

学校的理想装备

电子图书·学校专集

校园网上的最佳资源

统一之路



内容简介

揭开物质结构之谜一直是科学家孜孜以求的奋斗目标，也是现代物理学最基本的研究课题。经过近百年来几代物理学家特别是理论物理学家的不懈努力，人们对自然界存在的各种复杂的相互作用现象以及物质的最深层次结构初步建立起了一种简洁的、统一的图像。本书力求以一种通俗的语言，围绕攀登项目“90年代理论物理学重大前沿课题”所涉及的粒子物理、量子场论和凝聚态物理等三个方面的内容，简要地介绍了这一统一图像的发展历史、20世纪理论物理学取得的辉煌成就以及在世纪之交所面临的新的机遇与挑战。本书适用于广大的读者，对有关专业的大学生、研究生、教师、科研及管理工作者了解科学前沿和掌握科研动态有很大的价值。

代序

——在攀登计划实施大会上的讲话

国务委员、国家科委主任宋健

今天，“攀登计划实施大会”胜利召开，听了几位同志的发言我们大家都感到非常高兴。我们能够宣布攀登计划的启动和实施，清楚地说明，在科学技术的整体发展中，基础性研究工作受到党和国家的高度重视和关怀。

借此机会，我就当前科技工作的形势和进一步加强基础性研究工作的向题向大家作一简要说明。

自从小平同志视察南方的谈话发表以来，全国各族人民在党中央和国务院的领导下，进一步解放思想，加快改革开放的步伐，各条战线呈现出生机勃勃、欣欣向荣的景象。事实证明，经过几年治理整顿以后，我国已具备进一步加快发展的条件，正处在一个千载难逢的重要历史发展时期。我们面临的任务是：坚定不移地全面贯彻执行党的基本路线，牢牢把握经济建设这个中心，坚持四项基本原则，坚持深化改革和扩大开放，大力发展社会生产力。全体科技界的同志们将和全国人民一起，为实现这一战略方针做出历史性贡献。这是当代最高的战略目标，其他各方面工作都要服从这一目标。

党中央把发展科学技术列为我国经济和社会发展战略的重点，多次强调要把国民经济切实转移到依靠科技进步的轨道上来。在当前全国加快改革开放，尽快地把经济建设搞上去的新形势下，国家对科学技术工作提出了更高的要求。科技战线已经积极行动起来，加快科技体制改革的步伐，动员广大科技工作者更好地面向经济建设主战场，引导经济和社会发展；与此同时，要努力稳住和加强基础性和大力发展高新技术，攀登科学技术高峰，提高科技创新和自主开发的水平，增强科技实力和后劲，使科技在我国经济建设和社会发展中发挥更大的作用。简单地说就是两句话：一是“面向、依靠、攀高峰”，二是“加强一头，放开一片”。

“面向、依靠”，就是坚决贯彻党中央“经济建设必须依靠科学技术，科学技术工作必须面向经济建设”的战略方针，加快科技体制改革

的步伐，发展和提高全民族科学技术水平，促进科技与经济的紧密结合，鼓励广大科技工作者直接为推动企业科技进步，加快高技术产业开发区的建设、加强农业科技服务、合理开发资源、推动社会进步和保护生态环境等各方面贡献力量。

“攀高峰”，就是小平同志在视察南方的谈话中说的，“搞科技，越高越好，越新越好。越高越新，人民高兴，国家高兴。”

科学技术高峰，首先是我国社会主义现代化建设和社会进步的需要，只有创造出大量的、高水平的研究成果，才能更有力地推进国民经济发展和社会主义文明的建设。

攀登科学技术高峰也是科学技术自身发展的需要。近几十年来，科学技术以前所未有的速度向深度和广度迅猛发展，新思想、新概念、新发现层出不穷。加强知识产权保护和我国争取恢复在关贸总协定中的地位这种形势也对我国的科技工作的创新能力提出了更高的要求。只有树立雄心壮志，在各层次上勇攀科技高峰，才能不断提高科技工作水平，为中华民族在世界科技前沿领域占有一席之地建功立业。

总之，“攀登科学技术高峰”是新形势下为更好地贯彻“面向、依靠”的战略方针的一部分，是国家对科技工作提出的要求。这是指科技工作的各个方面包括攻关、高技术、基础性研究三个层次，研究、开发、建立产业等各项工作、理论、实验、工程技术各类专家，各个学科都要有自己的目标。努力创新、提高研究开发工作的水平。

在我国科学技术工作中，基础性研究的发展具有十分重要的意义。它不仅为人类利用和改造自然提供必要的知识基础，是新技术、新发明的先导和源泉；而且是培养高水平科技人才的摇篮。因此，确保基础性研究持续稳定地发展应是我国长期坚持的一项基本方针。建国以来，在毛主席、周总理等老一辈无产阶级革命家的关怀下，我们建立了一大批研究机构，形成了多种学科的基础科学研究体系，为社会主义建设作出了历史性的贡献。党的十一届三中全会召开以后，我国采取了一系列重要措施推动基础性研究的发展。例如，

1980年，中国科学院学部恢复了活动。

1981年，国家恢复了博士研究生制度。1985年7月，经国务院批准，我国开始办博士后科研流动站，试行博士后研究制度。

1986年2月，国务院决定成立国家自然科学基金委员会。

1987年，国家科委组织了对全国基础研究和应用基础研究现状及发展情况的调查研究。

1989年2月，经国务院批准，国家科委召开了全国基础研究和应用基础研究工作会议，会议明确指出，基础性研究是我国科技发展战略部署中的三个层次之一，对促进国民经济、社会发展和国防建设等具有重大意义。

近十年来我国先后建立的北京正负电子对撞机、兰州重离子加速器、合肥同步辐射加速器、高功率激光实验装置、受控热核反应HT—6H实验装置、长基线干涉仪、大型光学望远镜和毫米波射电望远镜、太阳磁场望远镜、中国环流器一号、低温核供热实验堆、串列静电加速器等十多项大型科学工程；建立了77个国家重点实验室和近100个部门开放实验室。增拨了实验室运行补助经费，为这些实验室更好地发挥作用和改善基础性研究的学术环境创造了条件。

今年，国家进一步增加对基础性研究的投入，国家自然科学基金有了较大幅度的增加，又拨出专款支持“攀登计划”的实施。这对于加强基础性研究，攀登科学技术高峰，具有重要的支撑意义。

随着对基础性研究重要性认识的提高，国家对基础性研究更加重视和支持，工作条件还将逐步得到改善。十多年来，我国科学家奋战在基础科学研究各前沿领域。学术水平不断提高，国际地位逐渐上升。如高临界温度超导体的研究，计算机科学的理论研究，若干人工功能晶体的理论和实验发现，微分动力系统数学理论研究，分子轨道图形理论方法，以酵母丙氨酸转移核糖核酸的人工全合成，蛋白质功能基因的修饰与其生物活性之间的定量关系等为代表的生物科学方面的新成就都处于世界先进行列；我国科学家对东亚大气环流、大地构造、成矿(成油)理论的研究，对新疆、青藏高原的考察，以及对雄性不育杂种优势、固氮基因的结构与调节，中国植物和动物分类编目等研究都取得了世界瞩目的成就。为我国资源勘探开发、自然灾害防治、环境保护提供了科学依据，为我国面临的与人口问题相关的农业、医药、卫生等重大社会问题的解决开辟了道路，也为世界科学技术的繁荣做出了贡献。“七五”期间，我国共有238项成果荣获国家自然科学奖，其中有一批荣获国际高层次科会和生产的速度大大加快。

事实表明，我国基础性研究工作已经进入了一个持续稳定发展的新

阶段。

但是我们也要看到当前我国基础性研究与国际前沿水平相比还有较大差距和面临着许多困难。例如，首创性不够，经费不足，仪器设备陈旧，课题分散，新秀培养不够等。还要预计到在加快改革开放的过程中，由于社会主义市场经济的发展和各方面的新变化，基础性研究还可能受到一定的冲击。因此，我们要采取果断措施，加强对基础性研究的指导，决心投入必要的人力物力，以求在力所能及的范围内，对重要的科学前沿和保证中华民族持续发展的最重要的基础科学方面，造就和保持一支有战斗力的精干队伍，务求取得第一流成就，为中国的发展和社会进步提供支撑，为世界科学事业作出我们应有的贡献。在增加基础性研究经费，逐步改变目前投资偏低状况的同时，要采取强有力的措施，发挥各方面的积极性，开创基础性研究的新局面。

在新的形势下加强基础性研究，我们特别要注意以下四个方面的问题：

第一，基础性研究要强调创新，强调高水平。基础性科学研究的主要目的是探索自然界的本质，发现自然界变化、发展的规律，提高我们对自然界的理性认识水平和深度。因此，对基础性研究的要求是要创新，要创造新思想、新方法，要争世界先进水平。有的科学家讲，基础研究只有第一，没有第二。事实证明，在当代科技发展中谁有更多的创新，有更高的水平，有更大的突破，谁就处于主动地位，就掌握了后来居上的机遇。基础性研究工作一定要坚持高标准，出突破性的成果，出高水平的人才，才能攀上新的高峰。这对提高一个国家科技创新和自主开发的能力是非常关键的。

第二，基础性研究项目要精选。在现代社会中，基础性研究已不仅仅是科学家的个人兴趣和行为问题，已成为一个有一定社会规模的科学事业。由政府或科学机构直接组织和推动基础性研究重大项目是当前世界科学技术发展的趋势。因此，国家在通过科学基金等方式支持科学家根据自己的特长和意向提出的自选课题和从学科发展的优先领域中选取的重点课题的同时，还要组织实施一些对国家的发展和科学技术的进步具有全局性和带动性、为我国科技界所公认的基础性研究重大关键课题。

根据这样的考虑，从 1989 年开始到现在，经科学界的认真讨论研究

后遴选了 30 个国家基础性研究重大关键课题，构成了“攀登计划”的首批项目，受到科技界广泛关注和欢迎。攀登计划的制订和实施，有利于完善我国科技工作的纵深部署，有利于吸引培养和造就新的学术带头人，形成高水平的科学家队伍；有利于集中力量，提高投资强度，在一些有优势和对国家长远发展有重大意义的重要领域取得突破，在世界科技发展的竞争中树立起中国的旗帜。对我国科技事业和国民经济的长远发展具有十分重大的意义。

第三，队伍要精干。中国是发展中国家，财力有限，不可能全面出击，要选择重点，组织一支精干的基础性研究队伍，集中兵力打歼灭战。我们衷心希望老科学家们的作用得到更多更好的发挥，更希望年轻一代科学家迅速成长。青年人思想活跃，精力充沛，容易冲破旧思想的束缚，始终是推动基础研究不断发展的生力军。要创造条件，让年轻的学术带头人在国际科学前沿拼搏；同时也要注意根据实际需要和个人特长，使优秀的人才及时开辟新的研究领域。要鼓励研究人员带着已经取得的成果参与应用研究和科技开发方面的工作。只有这样，才能使基础研究队伍始终充满活力，使基础性研究的成果和人才不断推向应用技术领域，为经济建设和社会进步服务。

第四，基础性研究工作要创造一个宽松、民主、有利于新思想产生和优秀人才脱颖而出的环境和条件。首先，我们要加强从事基础性研究的科研机构，建设和运行好已建成的各种大型科学工程、国家重点实验室和部门开放实验室，使它们在向国内外开放的大环境中，逐步形成具有自己特色的，在国际学术前沿努力拼搏的基础性研究的科学基地。我们还要加强科学数据积累工作，建立科学数据库，促进国际国内数据交换。为科学研究、经济建设和社会发展提供确切、快速、方便的信息服务。

我们要继续大力倡导尊重科学、尊重人才的社会风尚，要求科学界遵守科学道德规范。基础性研究是艰苦的探索性劳动，具有较大的不确定性。它需要科学家们数年、数十年、甚至毕生的探索，经过无数次失败后，才有可能取得进展。在少数成功者的后面，有很多科学家默默无闻的奉献。他们的工作应该得到人们的理解和支持，他们的成就应该得到政府和社会的鼓励和尊敬，这是尊重科学、尊重人才的社会风尚的一个重要组成部分。要尽可能为科学工作者创造较好的工作和生活条件，

以利于更好地发挥他们的聪明才智。

随着改革开放的深入进行，我国的基础性研究已经步入持续稳定发展的新阶段，展现出更加美好的前景。攀登计划的出台，在我国科学界受到了广泛重视和热情支持。这对稳定基础性研究队伍，落实“加强一头，放开一片”的方针，为基础性研究多出成果，多出人才，攀登科学高峰，必将产生深远的影响。我们希望有幸从事基础性研究的科学工作者，能密切关注我国经济建设和社会发展，从中抽象出新问题、新学科并组织力量加以研究。我们还希望从事基础性研究的同志们，能够尽力支持战斗在经济第一线的同行们，热诚地帮助他们，鼓励他们，分享他们的成就和喜悦。为了国家的未来，为了中华民族的前途，让我们一切从事基础性研究、技术开发和向产业转化的科技工作者，团结一致，相互支持，共同努力，为把我国科技工作提高到世界先进水平而努力奋斗。

（1992年7月22日）

编者的话

一个国家、一个民族要自立于世界之林，经济实力、发展速度固然十分重要，但经济长期稳定发展的后劲取决于科技进步，尤其是基础科学的发展水平。循因索果，是基础科学带动整个科技的发展，科技发展又带动社会经济的进步。这个道理人人都能懂。20 世纪的一个显著的时代特点就是科学的社会化与社会的科学化。科学技术的发展、进步与成果直接改变了我们周围的一切，影响着人们的生产、生活和思维方式。

科学活动大致可分为基础性研究和应用基础研究。基础性科学研究的目的，在于探索人类生存于其中的宇宙的本质，即自然界的本质。从这个意义说，基础性科学研究，是人类必需的理智行为之一。人类的每一次重大的基础性研究的突破，都对整个文明的进程带来从观念到现实的巨大震撼。然而，这种突破，在本世纪中叶以前，基本上只是以个别科学家的个人兴趣而存在着。今天，科学已不再只是少数科学精英在书斋和实验室中运作的个人追求，而是需要全社会都来关注、理解、支持的公众事业。因此，我们的科学工作者肩负着两个使命，使用着两种语言，当一项成功的科学成果，包括理论发现和技术发明诞生时，不仅要专家语言公诸同行，还必须用通俗语言告知大众，后者就是科学普及的使命。很难设想一个没有科学普及的时代会有科学技术的全面、高速发展。基于这种思考，我们在国家科委基础研究高技术司的支持与指导下策划了这套以大众语言反映我国基础科学研究顶尖水平的国家攀登计划的科普读物，恭请攀登计划的首席科学家主编了这套《攀登计划普及丛书》，每个项目一本。目的在于唤起社会，尤其是广大科技工作者及大、中学生对这一跨世纪伟大工程的了解与理解。

值得指出的是，本书的作者大部分是两院院士、国家级功勋科学家，他们惯常使用专家语言表达思想和智慧，此次他们改用通俗语言推介高技术知识及其进展，实不是一件轻而易举的事。为此，他们付出了极大的努力。在紧张繁忙的科研工作间隙，以十分严谨的态度构思写作计划，反复修改提纲，一遍遍为文稿润色，核实资料，推敲提法并雕凿文辞，力求深入浅出，生动活泼，可读易懂，同时又不失真诌与全貌。他们的这种求实笃诚的科学精神与严谨学风十分令人钦佩。当然，书中也有某些理论与技术问题由于本身的灰度，很难在不失原意的基础上做出通俗

的表述，只好留给读者一些灰色的谜。也许当读者在掌握某些学科的阶梯知识与背景材料之后，理解起来就会水到渠成。这就是说，科学普及，不能单方面强求我们的科学家走下高台，更重要的是，也要求我们的读者逐渐提高自己的科学知识，迈上台阶。理解与沟通永远是双向的。

科学普及工作是一项长期艰巨的社会文化工程，要做的事还很多，目标也将会越来越高。我们衷心希望一切关注社会进步与民族命运的同志和朋友都能与我们结伴而行，给我们批评和鼓励。科普的事业大有可为。

丛书自 1992 年组织起始，朱新民、丁乃刚、李文范、白松乾等同志在选题、组稿的策划中做了大量的工作，国家科委基础研究高技术司及湖南省新闻出版局给予了极大的关心和支持，谨此表示衷心的感谢。

《攀登计划普及丛书》编辑委员会

1995 年 3 月

前言

物理学是研究物质的基本结构和物质运动最一般规律的科学，是一切自然科学和工程技术的发展基础。对于这一点，公众通常均有或多或少的了解。但是对于理论物理学和理论物理学家，一般的公众恐怕知之甚少。

物理学本质上是实验科学。17 世纪到 19 世纪，正是在大量的观测和实验工作的基础上，牛顿建立了经典力学；焦耳、卡诺、开尔文、克劳修斯发展建立了热力学；克劳修斯、麦克斯韦、玻耳兹曼成功地发展了分子运动论；麦克斯韦在库仑定律和法拉第电磁感应定律的基础上建立了电磁学的系统理论。到 19 世纪末，人们普遍认为经典物理学理论已经系统、完整地建立了起来。但好景不长，光学和热学的一些实验就与这些被奉为经典的理论发生了尖锐的矛盾，从而引发了物理学的一场革命性的变革。

20 世纪初，几位卓越的物理学家，特别是一些刚刚进入物理学研究领域的杰出青年，对传统的物理观念提出了挑战，创立了相对论和量子力学这两个全新的物理学基本理论，实现了人类对物质结构和运动的基本规律的认识的又一次飞跃。正是在这两大理论基石之上，近代物理学建立了起来并得到了迅猛的发展。

如果说 20 世纪以前许多物理学家能身兼二任，既做实验又做理论研究，而且又都能做得很好，那么进入 20 世纪以后这种情况极为罕见。近代物理学以高速、微观和复杂系统为其研究对象，实验要求精度越来越高，设备越来越复杂，规模越来越大；而相应的理论日益抽象，日益艰深，对从事理论研究的人的物理和数学基础的要求越来越高。这使理论物理学与实验物理学的分化成为本世纪物理学发展的一种必然趋势，对整个自然科学有着极其深远的影响。

理论物理学是立足于实验物理学家提供的全部实验信息基础之上，对于未知的物质结构和运动基本规律进行理论探索的独立研究领域。理论物理学家并不亲自做实验，但他们对实验的进展都非常关心，其中有不少人对实验中的一些关键问题有着相当深入的了解。他们深知实验对于物理学理论发展的作用和意义。当然，也有不少理论物理学家的工

与实验没有直接的联系，或者很难看到与实验会有什么联系。对于一门基础学科，这一点也不奇怪。今天没有应用价值并不意味永远没有。很可能它们具有的潜在价值要经过相当长的时间才能被人们认识。也确有一些理论永远不会有任何实际的应用价值，但却是人类知识宝库中不可缺少的财富，具有特殊的意义。

回顾 20 世纪物理学的发展历史，人们不难看到在物理学取得的一个又一个重大的突破性进展过程中，理论物理学家起了突出的和关键的作用。诺贝尔物理奖从 1901 年开始授奖，到 1990 年共授奖 84 次。其中就有 27 次是授予理论物理学的成就，占了三分之一。共有 43 位理论物理学家先后获得此项物理学最高荣誉。这也反映了理论物理学的发展在物理学的前沿发展中的地位和作用。

物理学的每一次重大突破都会带来物理学新领域、新方向的发展，以至新分支学科的产生。理论物理的研究和发展为这些新兴学科、交叉学科的产生和发展准备了重要的不可缺少的条件。新兴学科的产生一般是不可能事先预知的，它们是在探索基本规律的前沿发展过程中生长出来的。许多新兴学科和交叉学科的产生和成长的早期阶段与理论物理的发展密切相关，理论物理的一部分精干的研究人员转入这些新兴学科和交叉学科成为主干，从事开拓性的研究，这是由理论物理学科性质所决定的自然的现象。例如量子力学和原子分子物理学的发展为化学周期律提供理论基础，促进了运用量子力学理论研究化学反应机理的量子化学的产生，使理论化学的发展有了质的飞跃。再如原子核物理理论的研究给出通过原子核裂变反应获取巨大能量的可能性，促进了原子核物理学和原子能科学技术的发展。二次大战中美国研制原子弹时，就集中了一批优秀的理论物理学家和原子核物理学家进行探索研究。由于核武器的研制属于高度军事机密，所有有关的实验和理论研究成果都是不公开的。各国在研制时都要重新集中一批优秀的理论物理学家和原子核物理学家进行研究，前苏联、法国、中国都经过这个过程。又如 50 年代初期，由于原子核物理学的理论研究给出通过原子核聚变反应获取巨大能量及高温等离子体可以提供实现可控制热核反应的可能性，促成一个新兴学科——等离子体物理学的形成和发展。当时美国和前苏联都集中了一批优秀的理论物理学家、原子分子物理学家、电真空物理学家和原子核物理学家进行等离子体物理和可控制热核反应的保密探索研究。5 年之后，

由于预计在可以预见的数十年内可控制热核反应不可能实现，两国先后解密，将等离子体物理领域实验和理论的系统研究成果公诸于世。由此可见，理论物理为物理学中新兴学科和交叉学科的产生、成长和发展既在学科准备方面又在研究队伍的建立和形成方面做出了不可缺少的重要贡献，起了“人才银行”或“人才库”的作用。

理论物理学研究的主战场始终是探索未知的物质结构和运动基本规律的最前沿。他们涉足于物理学各分支学科及许多交叉领域，内容非常广泛，其中最主要的领域为：

1. 探讨物质微观最深层次的结构和运动规律；
2. 探讨早期、极早期宇宙演化规律；
3. 研究复杂系统的凝聚态理论；
4. 探讨场的基本理论。

通过对这些重要领域的研究，理论物理学在微观、宇观和复杂系统这三个基本方向上把人类对自然界的认识推进到了前所未有的深度和广度，推动了整个自然科学体系的发展，带动技术革命从而促进了人类社会的进步和繁荣。

在第一个认识方向上人类探索的前沿已“深入”到空间范围仅有 10^{-15} 厘米，寿命仅有 10^{-23} 秒的亚原子粒子“内部”，建立了标准模型。由这个模型导出的结果和迄今为止积累的大量实验数据惊人地符合，这使人们相信，至少对于能量小于千亿电子伏特(100GeV)的物理过程，标准模型在弱、电、强三种基本作用方面是相当成功的。在此认识方向上发展起来的原子分子物理、核物理研究正推动着激光技术、核能利用和核技术的高速发展。

在第二个认识方向上，现代天文观测手段所及的空间范围已达到 150 亿光年，即 10^{28} 厘米，时间范围以这部分宇宙的寿命计算，为 150 亿年即 10^{17} 秒。在这样巨大的时空尺度内，描述物质相互作用和运动规律的引力理论和宇宙论取得了巨大成功，广义相对论的预言已为在太阳系中的实验观测所证实，描述宇宙演化规律的大爆炸宇宙模型得到了越来越多的观测支持。

在第三个认识方向上，即在对物质存在形式的多样化和物质运动复杂性的探索上，理论物理学的进展突出表现在以下三方面：

其一，以大量粒子组成的复杂多体系统为对象的凝聚态理论，围绕

着发现和研究凝聚态物质新基态和元激发以及探索对称性、对称破缺和标度现象这两条主线得到了丰硕的成果。80年代以来以分数量子霍尔效应和氧化物超导体的发现为代表，凝聚态理论研究又一次面临了对一系列新现象和新规律的探索，同时也由于新材料和新工艺的进展，又开辟了诸如介观系统理论研究等新的领域。

其二，确定性系统长时间行为可能对初值的微小变化极其敏感的混沌现象的研究揭示了非线性系统的长期不为人们所注意的重要规律，给人类知识体系以重大冲击。

其三，理论物理学研究推进了物理学与其它自然科学的交叉和结合，形成了大量新的交叉学科，其中特别有意义的是物理学与化学、生物学这两个研究物质更高层次运动规律学科的结合。理论物理学的发展促成了量子化学的形成和结构化学及化学动力学的改观，理论物理学方法已在化学研究中发挥重要作用。理论物理学和生物物理学两方面研究的发展正酝酿和准备着物理学与生命科学的实质性结合。

展望即将到来的21世纪，理论物理学正在前述三个方向上酝酿和准备着新的重大突破，充满了机遇和挑战。主要表现在：

1. 粒子物理标准模型中若干基本困难如希格斯(Higgs)粒子尚未发现、某些对称性的破缺尚未理解，正激励人们去探索非标准模型粒子和对称破缺的机理；现代量子场论的进展正促进粒子物理动力学规律的探索和统计模型求解的发展，并与现代数学的某些分支相互促进；非微扰量子色动力学理论及其对核物质和高能核碰撞的应用正加深人们对物质基本动力学规律的认识。

2. 由于天体物理观测技术的高度发展、实验数据的大量积累，近30年来形成了如量子引力、暗物质的本性、正反物质不对称性的缘由等一些困难而基本的理论物理问题。这些问题与物质深层次的结构和运动规律的探索有着深刻的内在联系。国际上这方面研究工作正通过与粒子物理和量子场论研究的相互促进、相互补充而迅速发展。

3. 在第三个认识方向上，强关联多电子系统的研究正揭示着占周期表几乎近一半左右元素（有f或d电子）所组成体系的未知属性；介观系统基本物理性质的理论研究正在探索传统理论未曾进入的新领域；有机导体、有机磁性和富勒烯系统的理论研究正与新技术密切结合；微结构系统的真空涨落、宏观仪器的作用和非平衡量子统计的研究正在一个

新的水平上重新考察人们对于自然规律的认识，也在为现代科技和工艺发展开辟道路；而诸如生命过程中的非平衡统计、遗传编码的信息分类、蛋白质折叠机理等理论生物物理研究课题也正在成为国际理论物理界共同关心的热点，有可能成为物理科学和生命科学本质性结合的突破点。

面对 20 世纪对理论物理学提出的挑战，科学地安排我国理论物理学发展的战略并采取切实有力措施优先支持一批优秀的理论物理学家的研究工作，对于我国未来基础科学的发展是至关重要的。这次“90 年代理论物理重大前沿课题”被列入“攀登计划”，正体现了对这一研究领域的重视。

本书主要目的在于介绍 20 世纪理论物理学取得的辉煌成就和它在世纪之交的今天所面临的新的挑战，希望能引起广大读者特别是有志于理论物理研究的年轻人的兴趣。由于当今的理论物理学已是一门极为庞大的学科，其涉及的范围之广，研究的层次之多，面对的问题之复杂，使任何一位或几位物理学家都不可能全面掌握其内涵和精髓。这样一本普及性的书更不可能面面俱到，我们只能挑选一些重要内容做一些非常基本的介绍。本书的第一篇简要地介绍了作为理论物理基础的量子力学和相对论的一些知识以及一些背景材料，第二、三、四篇分别介绍粒子物理、量子场论和凝聚态理论的进展情况、取得的成就及存在的挑战性问题。鉴于作者们学识浅陋且只工作于非常有限的狭小领域，这些介绍必定有很大的局限性。选择的内容未必都重要，有些重要的问题完全没有提到，对问题的讨论带有作者的偏见，说了不少外行话甚至错话等，所有这些敬祈各方面专家学者及广大读者批评指正。此外，尽管我们尽量使用了通俗的语言，注意使叙述尽可能深入浅出，但理论物理的基本概念过于抽象，本书仍然未能很好地完成使广大读者易读易懂的艰巨任务，这是一个极大的遗憾。

本书的写作是在攀登项目“90 年代理论物理重大前沿课题”的首席科学家苏肇冰主持下完成的。第一和第二篇由丁亦兵撰写，第三和第四篇的初稿分别由廖力和任德铭执笔，由丁亦兵改写并完成全书。在写作过程中得到了许多理论物理学家的热心帮助。几部分初稿曾经分别由汤拒非、宋行长和林宗涵审阅，提出了许多宝贵的意见。在修改时参考了高崇寿、张元仲和林宗涵为本项目撰写的有关材料，并请徐湛补写了一节关于 Seiberg-Witten 最新的工作简介。此外，黄涛和赵雅琴等在多方

面提供了支持和帮助，我们谨此一并表示谢意。

编著者 1996 . 8 .

1 理论物理学的两大基石

1.1 普朗克把“量子”引入物理学

19世纪末，经典物理学理论遇到了来自实验的严重挑战，其中黑体辐射实验是非常关键的一个。所谓的黑体是指一种能够把外来电磁辐射全部吸收的物体，例如炼钢炉。射入炉口的光线不会反射出来，所以一个没有加热的炉子，从炉口看上去完全是黑色的。加热以后会发射出光，熟练的工人从光的颜色可以估计出炉温。实验家们测量了不同温度下黑体发射光的强度随频率变化的曲线。几位著名的物理学家从经典物理学出发做了许多尝试，企图对这些曲线给出理论解释，但是都失败了。

德国柏林大学教授普朗克(M. Planck)另辟蹊径，从1897年开始，经过三年的努力，从已有的一些经验公式出发，终于在1900年找到了一个新的数学公式，相当准确地拟合了实验曲线。普朗克又用了两个月的时间，进一步去探索这个公式的理论根据。他发现不得不放弃经典物理学中的连续性观念，认为光的辐射是以分立的方式一份一份地发生，才能从理论上得到这个公式。他把每一小份能量称为“量子”。量子的能量正比于光的频率。比例常数是个很奇怪的、普适的小常数 h (人们后来称之为普朗克常数)。这种量子概念从经典物理学是绝对无法理解的，以至他的同事中几乎没有什么人接受它。

普朗克的工作实际上在物理学中引发了一场深刻的革命，最先注意到这一点的是爱因斯坦(A. Einstein)。在1905年发表的一篇使爱因斯坦获得诺贝尔奖金的论文中，他把普朗克的不连续的能量量子的概念和光电效应联系起来。

光电效应是指光照射某些金属表面打出电子的现象。这种现象的怪异之处在于，从金属中释放出来的电子的能量不取决于光的亮度，而取决于光的颜色，即光的频率。爱因斯坦在普朗克假设的基础上进一步假设：不仅光的发射过程，而且光的传播过程也是一份一份地进行的。于是作为电磁波的光，不仅有波动的性质，也具有粒子性。它们是由光量子，或光子组成。光量子的能量正比于频率，其动量反比于波长。比例系数都是普朗克常数 h 。利用这一假设，爱因斯坦相当好地说明了光电效应，并给出了定量预言，它引起了实验家的极大兴趣。爱因斯坦提出光

电效应公式时，实验家密立根(R . A . Millikan)是持怀疑态度的。在他的一段回忆中这样写道：“我花了 10 年来时间来检验爱因斯坦的 1905 年的方程，结果和我所有的预期完全相反。在 1915 年，我不得不宣布它的毫不含糊的鉴定，尽管它与我所知道的光的干涉的每件事都相违而不合乎常情。”1925 年康普顿(Compton)使单个光子与电子散射，高度精确地直接验证了爱因斯坦的光子假说。

1.2 玻尔的量子论挽救了原子有核模型

经典物理学在 19 世纪遇到的另一个挑战性的问题来自于原子结构的研究。原子作为物质最小的、不可分割的单元，最早是古希腊哲学家德谟克利特的一种猜测。它的真实存在，直到十九世纪初才被英国化学家道尔顿证实。在 1808 年道尔顿原子论发表之后，原子不可分割的观念又持续了半个多世纪。1895 年伦琴发现 X 射线，1896 年贝克勒尔发现了原子放射性，1897 年汤姆逊发现电子等，十九世纪末物理学家们完成的这一系列重大发现，使这一观念发生了根本的动摇。1903 年汤姆逊分析了大量实验，认识到带负电荷的电子应该是物质的基本组分。考虑到原子是电中性的，他提出了一个原子结构模型，认为电子嵌在连续分布的正电荷上构成原子，就象葡萄干粘在面包上。尽管这个模型不能对原子的其他性质特别是原子光谱给出解释，但它揭开了原子结构研究的新篇章。

1898 年，27 岁的卢瑟福(E. Rutherford)经过长时间的研究认识到铀放出的射线有两种不同类型。一种很容易被吸收，叫做 α 射线。另一种穿透性很强，叫做 β 射线，后来发现它就是电子。不久，维拉德又发现了第三种更有穿透力的射线叫做 γ 射线。1908 年卢瑟福证实了 α 粒子就是带正电的氦离子。

1909 年卢瑟福为了研究 α 粒子通过物质时的各种现象，设计了著名的“ α 粒子散射”实验。实验发现，一些 α 粒子打在金箔上，产生了 90° 或甚至更大角度的散射。卢瑟福对此大为惊奇，后来谈及此事说：“那确实是我一生中遇到的最难以置信的事。”“这几乎和你向一张纸发射一枚 15 英寸的炮弹，结果炮弹被弹回来击中你一样难以置信。考虑到粒子的高速度和很大的质量，我意识到这种反向散射必定是单次碰撞的结果。计算发现，除非设想一个原子的绝大部分质量聚集在微小的核心，否则不可能得到那种数量级的结果。”1911 年，卢瑟福公布了他的这一具有历史意义的研究成果。

卢瑟福通过实验建立了一个新的原子模型，它与汤姆逊模型完全不同，很象一个小的太阳系。原子的几乎全部质量都集中在中心，称为原子核。原子核带有正电荷，带负电的电子绕着它做轨道运动。进一步研究表明原子核的尺度只有 10^{-13} 厘米，称为一个“费米”，(记为：

$1\text{fm}=10^{-10}\text{cm}$)。而原子半径大约 10^{-8} 厘米，所以原子核的尺度只有原子尺度的十万分之一。如果把原子设想成一个几百米跑道的运动场，原子核就象在运动场中心处的一粒米。你可以想象原子内部有多么空荡了。

卢瑟福的原子有核模型是通过实验建立起来的。但奇怪的是：按照当时奉为经典的牛顿力学和麦克斯韦电磁学理论，这个模型完全行不通。沿轨道做圆周运动的电子，按照经典理论会由于不断地发射电磁波而损失能量，轨道半径会越来越小，直到最后落入原子核内，原子也就崩溃了。计算表明，这个过程在极短的瞬间即可完成，显然与自然界中原子稳定性的事实有着尖锐的矛盾。就在实验家们对此束手无策的时候，27 岁的丹麦理论物理学家玻尔(N. Bohr)于 1912 年来到了卢瑟福的实验室。他很快就开始了对卢瑟福原子模型不稳定性问题的研究。

当时的玻尔已经注意到了普朗克和爱因斯坦的工作，而且早在 1910 年，他在准备关于金属电子理论的博士论文时，就已经认识到他正在研究的一些现象不可能用经典力学来解释。于是，当他面对卢瑟福原子有核模型和经典物理学导致电子运动不稳定的尖锐矛盾时，很快想到：问题出在经典物理学不能描写这个系统。玻尔猜测：原子必定有一种“稳定态”。在这种状态下，电子的运动轨道是稳定的，它不辐射光，也不会跌入原子核。

正在玻尔为寻找新的实验证据而困扰时，他的一个老友汉森来访，问他能否用他的模型解释巴尔末线系。这是瑞士科学家巴尔末对氢原子的一些光谱线总结出的一个简单的公式，可以用来很容易地计算出这些谱线的波长。玻尔看到了这个公式马上就意识到，这正是他要寻找的东西。不久他就建立了一个氢原子结构的量子理论模型。在这一模型中，正常状态下，原子内的电子处在能量最低的一个稳定态，叫做基态。给原子增加能量，比如加热，电子则跳到较高能量的稳定轨道；停止增加能量，电子则跳回基态，同时放出一个光子。光子的频率为两个定态的能量之差除以普朗克常数 h ，刚好就是巴尔末线系公式。

1913 年他的这些成果以三篇论文形式发表，对量子物理学的发展起了划时代的作用，后人称之为伟大的三部曲。玻尔晚年回忆起这段历史时说：“很清楚，要解决卢瑟福原子的难题，除了来一个重大的变动，别无它法。”玻尔的文章在物理学界引起了很大的反响，许多物理学家认为它有着极重要的历史意义。例如，爱因斯坦认为：“凭这一小点既

不保险而又互相矛盾的理论基础，居然足以使玻尔这样一位具有独特直觉和洞察力的人发现了光谱线的主要规律，……，这对我来说真是一个奇迹。”但也有人持反对意见，斯特恩、劳厄等甚至表示，如果万一证明玻尔是对的，他们就不再干物理这一行了。

1922年由于玻尔这些开创性的工作，使他获得了诺贝尔奖。同一年，玻尔用原子结构理论对元素周期表作出了解释。他认为原子是由原子核及其周围作轨道运动的电子壳层组成，每个壳层有一定数量的电子。化学性质相似的元素，其最外壳层的电子数相同。据此，玻尔预言了当时一种新元素镧的性质，很快得到了实验的完全证实。

尽管玻尔的原子结构模型一再取得成功，但玻尔自己很清楚，他的理论有着严重的弱点。他企图把他的理论扩充，研究有两个电子的氦原子，但始终没有能成功。他感到理论的进一步发展单靠直觉是建立不起来的，它需要有更坚实的基础和需要一代新生力量。

1.3 量子力学的两种形式：矩阵力学与波动力学

1922年初夏，21岁的海森堡(W. Heisenberg)来听玻尔的演讲。海森堡原打算做一名数学家。玻尔的报告使他兴奋不已，会后他与玻尔讨论了很长时间。他被原子理论问题完全迷住了，决定改做理论物理。玻尔也喜欢上了这个敢于争论的学生，邀请他到哥本哈根工作。

1924年海森堡完成了博士论文之后，来到了哥本哈根，与玻尔开始了“别具一格”的协作，投入了量子理论的研究。海森堡对原子结构的看法与玻尔完全不同。他认为电子轨道近于幻想，因为没有人能看到它们。唯一可以从实验测量的，是原子内部发出的光的频率和振幅。海森堡决定放弃模型假设，直接对可以测量的这些数所构成的数组进行研究。

与玻尔合作一年之后，海森堡回到了德国，做玻恩(Born)的助手。1925年5月，由于花粉过敏症，海森堡到北海的赫戈兰岛上休假。岛上的清新空气诱发了他的灵感。经过一个昼夜的艰苦工作，他发现了那些数组满足的代数关系，它们的乘法与普遍的数的乘法不同，不能交换次序，或称不可对易。于是他写了一篇论文，建立起了一套新的体系。休假回来之后，他和玻恩进行了讨论。玻恩马上认出，他的这套体系是当时物理学家们还不熟悉的矩阵代数。海森堡与玻恩及他的学生约当三个人又经过了三个月的精心研究，发展了这一体系，建立起了处理量子理论的新的数学框架，人们称之为矩阵力学。海森堡的密友，与其同龄的青年人泡利(Pauli)解决了用矩阵力学计算氢原子光谱的问题，得到了与玻尔相同的结果。

与矩阵力学发展的同时，量子理论的另一形式也在发展。爱因斯坦的光量子理论和玻尔的原子模型引起了法国德布罗意(deBroglie)亲王的极大兴趣。他在分析了光量子说中共同存在的波粒二种性质的同时，研究了原子内电子定态轨道对应的分立的整数与波动现象中驻波的特征的类似，从而大胆地假设：既然光作为电磁波同时具有粒子性，电子作为一个物质粒子也应该有的时候表现为一个波。仿照爱因斯坦的做法，他导出了电子的能量与它相应的波的频率及动量与波长的关系，用它们成功地解释了玻尔的原子量子化条件。1924年德布罗意把这些结果写成了自己的博士论文。论文答辩委员会对他的结果持怀疑态度，不知

道怎样评价，决定征询爱因斯坦的意见。爱因斯坦表示了极大的兴趣，认为这篇论文的内容有一些极为重要的发现。1927年，贝尔实验室的戴维逊从实验上证实了德布罗意的想法。于是德布罗意的博士论文成了物理学史上的一篇最著名的博士论文，也得迄今唯一的一篇获诺贝尔奖的博士论文。

瑞士苏黎士的一些例行的学术讨论会使量子力学取得了另一个新的重大进展。在一次讨论会结束之前，主持人德拜环顾四周，想找一个下次讨论会的报告人。当时，德布罗意的惊人发现的消息刚刚传来，人们对其细节知之甚少。德拜的眼光停在了苏黎士大学薛定谔(E. Schrödinger)教授的身上说：“薛定谔教授，你是不是愿意讲一讲德布罗意的工作？”薛定谔点了点头，表示他愿意。

在下次例会上，薛定谔阐述了德布罗意如何用电子的物质波的驻波，解释玻尔原子模型中稳定的电子轨道。报告结束时，德拜对他说：“薛定谔教授，看来有点怪，你讲了半天波动但却没有给出任何波动方程。”这个评述促使薛定谔开始思考怎样去建立一个这样的方程。几个星期以后，当薛定谔就同一个题目给出第二次报告时，开场白就是：“德拜教授建议应当有一个波动方程。现在，我已经把它找到了。”

不久，人们就清楚地认识到薛定谔做出了一个重大的发现。薛定谔方程一经发表，立即受到了普遍的注意，迅速闻名于世。主要原因在于他使用的数学方法是微分方程，物理学家都非常熟悉。在解决具体问题时，比海森堡的方法容易得多。人们很快地把薛定谔的新理论应用到了各个方面。薛定谔自己计算了氢原子光谱，得到了与玻尔相同的结果。他还证明了矩阵力学与波动力学完全等价。1926年初他所发表的一系列文章，标志着量子力学的一种新的形式——波动力学的诞生。

至此，量子理论无疑已经取得了突破性的进展，但随之产生了许多新的疑难问题困扰着理论物理学家。德布罗意的物质波其实质究竟是什么？薛定谔方程中这个波用一个波函数来描写，究竟代表什么？对这个根本性问题，当时没有任何人能够给出准确的回答，其中包括薛定谔本人。薛定谔认为：波函数本身代表一个实在的和物理的可观测量，它描述物质的分布，粒子可以想象为一个波包。但这种看法有着难以克服的困难，因为即使原来的波包很窄，理论分析表明，它将随时间很快地弥散。也就是说电子会越来越“胖”，这是与实验尖锐矛盾的。

1926年玻恩仔细地研究了这个问题。他认为德布罗意的“物质波”或薛定谔方程中的波函数并不代表实在的物理量的波动，而是描绘粒子在空间分布几率的一种几率波。波函数绝对值的平方给出的是发现粒子在给定位置出现的几率。这是一种把粒子性与波动性统一起来得到的认识。量子力学的描写是一种统计性的，而不再是为人们所熟悉的决定论性的。经典物理学是满足决定论的。比如按照牛顿力学，只要知道了原子内某一时刻电子的位置和它的动量，就能完全确定以后任何时刻电子的位置及其动量。而量子力学只能告诉我们，某一时刻电子在原子内部各个不同位置的出现几率。知道了所有这些几率分布的知识，我们可以预知以后任何时刻电子的几率分布。

1927年海森堡发现的测不准关系进一步说明了量子力学这种怪异行为。按照测不准关系，在原子这样小的尺度内，粒子的位置和它的动量是不可能同时精确测定的。假设我们试图以任意高的精确度去测量一个电子的位置，你将会发现，它的动量的不确定程度会是无穷大，我们绝不会知道这个电子朝哪个方向飞，速度多大。反之，如果我们确知电子的动量，它的位置在哪里我们将一无所知。与此同时，玻尔对经典和量子物理学的矛盾也进行了认真的思考，提出在原子范围内应该放弃经典物理学的决定论，代之以波动性和粒子性两种互不相容而又互相补充的新观点。

量子力学带来的翻天覆地的变化，并不是人人都欢迎的。即使是曾经做出过重大贡献的爱因斯坦，对于几率解释和以玻尔为首的哥本哈根学派的观点也拒不接受，总想用决定论的说法代替它。在1927年布鲁塞尔召开的索尔维会议上，玻尔与爱因斯坦就量子力学的解释问题进行了激烈的论战。每天的早餐桌上，爱因斯坦总要向玻尔提出一个经过精心思考而认为足以使玻尔陷入矛盾的关于测不准关系的反例，玻尔和他的学生们立即加以研究，直到在爱因斯坦的批评中挑出毛病为止。但是这时爱因斯坦已经另外想好了一个事例。最后，爱因斯坦不得不承认他没有能够找到任何充足有力的反例。但是他在给玻恩的一封私人信件中说，他依然坚信“上帝是不掷骰子的。”

到二十年代末，新量子力学的解释逐渐完备，其应用研究极度兴盛起来。化学键、顺磁、铁磁等许多长时间悬而未决的现象都得到了很好的解释。许多新的研究领域在量子力学的基础之上蓬勃地发展起来。

1.4 爱因斯坦和他的相对论

除了量子理论以外，1905年刚刚得到博士学位的爱因斯坦发表的一篇题为《论动体的电动力学》的文章引发了二十世纪物理学的另一场革命。文章研究的是物体的运动对光学现象的影响，这是当时经典物理学面对的另一个难题。

十九世纪中叶，麦克斯韦建立了电磁场理论，并预言了以光速 C 传播的电磁波的存在。到十九世纪末，实验完全证实了麦克斯韦理论。电磁波是什么？它的传播速度 C 是对谁而言的呢？当时流行的看法是整个宇宙空间充满一种特殊物质叫做“以太”，电磁波是以太振动的传播。但人们发现，这是一个充满矛盾的理论。如果认为地球是在一个静止的以太中运动，那么根据速度迭加原理，在地球上沿不同方向传播的光的速度必定不一样，但是实验否定了这个结论。如果认为以太被地球带着走，又明显与天文学上的一些观测结果不符。

1887年迈克尔逊和莫雷利用光的干涉现象进行了非常精确的测量，仍没有发现地球有相对于以太的任何运动。对此，洛仑兹(H.A.Lorentz)提出了一个假设，认为一切在以太中运动的物体都要沿运动方向收缩。由此他证明了，即使地球相对以太有运动，迈克尔逊也不可能发现它。爱因斯坦从完全不同的思路研究了这一问题。他指出，只要摒弃牛顿所确立的绝对空间和绝对时间的概念，一切困难都可以解决，根本不需要什么以太。

爱因斯坦提出了两条基本原理作为讨论运动物体光学现象的基础。第一个叫做相对性原理。它是说：如果坐标系 K' 相对于坐标系 K 作匀速运动而没有转动，则相对于这两个坐标系所做的任何物理实验，都不可能区分哪个是坐标系 K ，哪个是坐标系 K' 。第二个原理叫光速不变原理，它是说光（在真空中）的速度 c 是恒定的，它不依赖于发光物体的运动速度。

从表面上看，光速不变似乎与相对性原理冲突。因为按照经典力学速度的合成法则，对于 K 和 K' 这两个做相对匀速运动的坐标系，光速应该不一样。爱因斯坦认为，要承认这两个原理没有抵触，就必须重新分析时间与空间的物理概念。

经典力学中的速度合成法则实际依赖于如下两个假设：1. 两个事件

发生的时间间隔与测量时间所用的钟的运动状态没有关系；2．两点的空间距离与测量距离所用的尺的运动状态无关。爱因斯坦发现，如果承认光速不变原理与相对性原理是相容的，那么这两条假设都必须摒弃。这时，对一个钟是同时发生的事件，对另一个钟不一定是同时的，同时性有了相对性。在两个有相对运动的坐标系中，测量两个特定点之间的距离得到的数值不再相等。距离也有了相对性。

如果设 K 坐标系中一个事件可以用三个空间坐标 x 、 y 、 z 和一个时间坐标 t 来确定，而 K' 坐标系中同一个事件由 x' 、 y' 、 z' 和 t' 来确定，则爱因斯坦发现， x' 、 y' 、 z' 和 t' 可以通过一组方程由 x 、 y 、 z 和 t 求出来。两个坐标系的相对运动速度和光速 c 是方程的唯一参数。这个方程最早是由洛仑兹得到的，所以称为洛仑兹变换。

利用洛仑兹变换很容易证明，钟会因为运动而变慢，尺在运动时要比静止时短，速度的相加满足一个新的法则。相对性原理也被表达为一个明确的数学条件，即在洛仑兹变换下，带撇的空时变量 x' 、 y' 、 z' 、 t' 将代替空时变量 x 、 y 、 z 、 t ，而任何自然定律的表达式仍取与原来完全相同的形式。人们称之为普遍的自然定律对于洛仑兹变换是协变的。这一点在我们探索普遍的自然定律方面具有非常重要的作用。

此外，在经典物理学中，时间是绝对的。它一直充当着不同于三个空间坐标的独立角色。爱因斯坦的相对论把时间与空间联系起来。认为物理的现实世界是各个事件组成的，每个事件由四个数来描述。这四个数就是它的时空坐标 t 和 x 、 y 、 z ，它们构成一个四维的连续空间，通常称为闵可夫斯基四维空间。在相对论中，用四维方式来考察物理的现实世界是很自然的。狭义相对论导致的另一个重要的结果是关于质量和能量的关系。在爱因斯坦以前，物理学家一直认为质量和能量是截然不同的，它们是分别守恒的量。爱因斯坦发现，在相对论中质量与能量密不可分，两个守恒定律结合为一个定律。他给出了一个著名的质量-能量公式： $E = mc^2$ ，其中 c 为光速。于是质量可以看作是它的能量的量度。计算表明，微小的质量蕴涵着巨大的能量。这个奇妙的公式为人类获取巨大的能量，制造原子弹和氢弹以及利用原子能发电等奠定了理论基础。

对爱因斯坦引入的这些全新的概念，大部分物理学家，其中包括相对论变换关系的奠基人洛仑兹，都觉得难以接受。旧的思想方法的障碍，

使这一新的物理理论直到一代人之后才为广大物理学家所熟悉，就连瑞典皇家科学院，1922年把诺贝尔奖金授予爱因斯坦时，也只是说“由于他对理论物理学的贡献，更由于他发现了光电效应的定律。”对于相对论只字未提。

爱因斯坦于1915年进一步建立起了广义相对论。狭义相对性原理还仅限于两个相对做匀速运动的坐标系，而在广义相对论性原理中匀速运动这个限制被取消了。他引入了一个等效原理，认为我们不可能区分引力效应和非匀速运动，即非匀速运动和引力是等效的。他进而分析了光线在靠近一个行星附近穿过时会受到引力而弯折的现象，认为引力的概念本身完全不必要。可以认为行星的质量使它附近的空间变成弯曲，光线走的是最短程线。基于这些讨论，爱因斯坦导出了一组方程，它们可以确定由物质的存在而产生的弯曲空间几何。利用这个方程，爱因斯坦计算了水星近日点的位移量，与实验观测值完全一致，解决了一个长期解释不了的困难问题，这使爱因斯坦激动不已。他在写给埃伦菲斯特的信中这样写道：“……方程给出了近日点的正确数值，你可以想象我有多高兴！有好几天，我高兴得不知怎样才好。”

1915年11月25日，爱因斯坦把题为“万有引力方程”的论文提交给了柏林的普鲁士科学院，完整地论述了广义相对论。在这篇文章中他不仅解释了天文观测中发现的水星轨道近日点移动之谜，而且还预言：星光经过太阳会发生偏折，偏折角度相当于牛顿理论所预言的数值的两倍。第一次世界大战延误了对这个数值的测定。1919年5月25日的日全食给人们提供了大战后的第一次观测机会。英国人爱丁顿奔赴非洲西海岸的普林西比岛，进行了这一观测。11月6日，汤姆逊在英国皇家学会和皇家天文学会联席会议上郑重宣布：得到证实的是爱因斯坦而不是牛顿所预言的结果。他称赞道“这是人类思想史上最伟大的成就之一。爱因斯坦发现的不是一个小岛，而是整整一个科学思想的新大陆。”泰晤士报以“科学上的革命”为题对这一重大新闻做了报道。消息传遍全世界，爱因斯坦成了举世瞩目的名人。广义相对论也被提高到神话般受人敬仰的宝座。

从那时以来，人们对广义相对论的实验检验表现出越来越浓厚的兴趣。但由于太阳系内部引力场非常弱，引力效应本身就非常小，广义相对论的理论结果与牛顿引力理论的偏离很小，观测非常困难。七十年代

以来，由于射电天文学的进展，观测的距离远远突破了太阳系，观测的精度随之大大提高。特别是 1974 年 9 月由麻省理工学院的泰勒和他的学生惠斯勒，用 305 米口径的大型射电望远镜进行观测时，发现了脉冲双星，它是一个中子星和它的伴星在引力作用下相互绕行，周期只有 0.323 天，它的表面的引力比太阳表面强十万倍，是地球上甚至太阳系内不可能获得的检验引力理论的实验室。经过长达十余年的观测，他们得到了与广义相对论的预言符合得非常好的结果。由于这一重大贡献，泰勒和惠斯勒获得了 1993 年诺贝尔物理学奖。

1.5 狄拉克把相对论和量子力学结合在一起

现在，我们可以回过头来，继续我们对量子力学的讨论。除了前文提到的由海森堡与薛定谔给出的量子力学的两种形式之外，还有一种形式是由比海森堡小一岁的年轻英国物理学家狄拉克(P. A. M. Dirac)建立起来的。1925年，年仅23岁的狄拉克收到了海森堡第一篇矩阵力学的清样稿，这是他初次与量子力学打交道。他一口气连看了10天。海森堡提出的非对易量，使狄拉克产生了浓厚的兴趣。

狄拉克有独自散步思考问题的习惯。正是在一次散步时，他突然想到这种非对易量的运算与经典力学中一种叫做“泊松括号”非常相似。他只是在阅读高深力学著作时看到过关于“泊松括号”的介绍，由于一直觉得它没有多大用处，很快就忘掉了。于是他匆匆忙忙赶回住所，翻遍了所有的书本和论文，却找不到有关的参考资料。狄拉克讲：“当天是星期日，不能到图书馆去，只得在躁急难忍中熬过了那一晚。第二天一清早，图书馆的门刚打开，我就跨了进去，赶忙去查看‘泊松括号’到底是怎么回事。终于发现，它和我的想法完全相符。”这样，狄拉克很快就建立了一种新的量子力学体系，在人们早已经熟悉的经典力学与海森堡新力学之间建立起了一种非常密切的联系。

狄拉克关心的另一个重大的理论问题是如何把量子力学与相对论协调起来。海森堡和薛定谔所建立的量子理论对原子尺度的微观现象给出了相当满意的描写，但他们给出的理论形式不符合爱因斯坦创立的狭义相对论的要求。确切地说，薛定谔方程在洛仑兹变换下不具有协变性，它是一种非相对论的方程。这意味着它不能正确地处理速度接近光速的微观客体的运动。

薛定谔后来告诉狄拉克，他最初建立的方程是一个相对论方程，因为德布罗意给出的关系式是相对论的。但是，当他用这个方程对氢原子进行定量计算时发现结果与实验不相符，他非常失望，开始怀疑自己所有的做法是否完全错了，于是把它们一古脑丢开没有发表。过了一段时间，他才重新捡起来，发现如果对这个方程取非相对论的近似，也就是说考虑粒子的速度远小于光速，得到的结果相当好，于是就把这个非相对论方程发表了，人们称之为薛定谔方程。对此，薛定谔非常惋惜。因为在那之后不久，他所丢弃的方程就由克莱因(Klein)和高登(Gordon)重

新得到而且发表了。但这个方程有一个很大的困难，即计算出来的几率有可能是负的。狄拉克对薛定谔方程和克莱因-高登方程存在的问题进行了认真的分析。发现关键在于他们没有考虑到电子有自旋。

电子可能有一种自旋角动量，即电子象一个陀螺不停地自转的想法，最早起源于人们发现原子能级“加倍”现象，即实际的原子可能允许的状态比起当时理论的计算结果要增加一倍。克朗宁最早提出来用电子的自旋解释这种现象。他给泡利写了一封信，把这一想法告诉了泡利。泡利的回答是：“不，这绝不可能。”一句话就把克朗宁的想法吓退了。因为，在当时物理学家中，泡利在判断和发现任何理论的弱点上，差不多具有传奇式的能力。因此除非得到泡利的赞同，很少人对自己的工作感到完全有把握。可惜的是这次泡利反对错了。

与此同时，1925年两位在埃伦费斯特手下工作的年轻的荷兰人，乌仑贝克和戈德斯密特，也就电子有自旋的想法，写了一篇文章给埃伦费斯特看。埃伦费斯特非常感兴趣并建议他们去找洛仑兹谈一谈。洛仑兹对他们说：“不，电子完全不可能有自旋。我自己也曾经想过这个问题。假如电子真的有自旋的话，电子表面的速度会比光速还大。所以，它是完全不可能的。”两位年轻人回到埃伦费斯特那里，对他说他们想撤回他们的文章。埃伦费斯特说：“太晚了，我已经把它寄出去发表了。”于是，电子自旋的想法就这样公诸于世了。世人也真该感谢埃伦费斯特对青年人的热情。遗憾的是，由于对自旋的本质一直得不到清楚的理解，乌仑贝克和戈德斯密特始终未能得到诺贝尔奖。

1927年，狄拉克经过几个月的努力，力图克服薛定谔方程和克莱因-高登方程存在的问题，而且要把电子的自旋纳入到理论中来。考虑到泡利曾用两个分量的波函数修改薛定谔方程成功地解释了与自旋有关的实验现象，依据相对论的协变性要求，狄拉克用四个分量的波函数描写电子，在1928年终于得到了一个新的方程。它满足相对论协变性，不存在负几率困难，同时在低速下回到了薛定谔和泡利方程的结果。这个方程还非常自然地导出了电子自旋角动量的量值，与实验完全一致，被人们看成是狄拉克方程的一个伟大的成就。

不过，这个方程也遇到了一个很大的难题，即它不可避免地预言存在有带负能量的电子。其实，只要是相对论理论，就可能出现负的能量，因为爱因斯坦的能量公式包含一个平方根，从数学上讲，能量值的计算

结果可正可负。对于经典理论，能量守恒可以保证负能量不会出现。然而，量子理论则不同，这时从正能态向负能态的跃迁不但是可以允许的，而且从能量减小看极为有利。这个问题使狄拉克感到十分困惑。狄拉克对于自己依照一般的相对论协变性及量子力学原理建立的新方程十分满意，它使许多困难都迎刃而解，唯独负能问题却依然故我，而且成为不得不解决的最基本的矛盾。

幸运的是，泡利的不相容原理帮了大忙。这个原理不允许任何一个量子状态容纳多于一个电子。据此，狄拉克大胆地假设，真空不是人们想象的那样空无一物，而是所有的负能态都被电子填满，所有的正能态都空着的状态。这就是说真空象一个没有底的负能电子的大海，每个态上都有一个电子占据着。带正能的电子在这个大海的海面上运动。它们不可能掉到海里去，因为里面没有它的位置。当真空由于某种原因，比如一个很高能量的 射线的光子闯入，而吸收了一份足够大的多余能量时，负能海中某一位置上的电子可以被激发从海中跳出来，这时我们看到的是海平面上多了一个正能量的电子，同时负能海中出现了一个空穴。这个空穴需要一份负能量和一个负电荷才能填平，所以它相当于一个带正能量和正电荷的粒子在真空中自由运动。

这个正电荷的粒子是什么呢？它的质量有多大？对于这个问题，狄拉克思考了很久。正如他后来所讲的：“当我开始产生这一想法时，似乎觉得在空穴与通常电子之间应该有一种对称性。但是，当时已知的唯一带正电的粒子是质子，于是我又好象觉得空穴必须是质子。我缺乏勇气去建议一种新粒子。当时有相当多的根据使人们相信只存在两种基本的带电粒子——电子和质子。由于电荷也只有两类，一类是正电荷，一类是负电荷，因此每一种电荷也只需要有一种粒子。人们对于存在新粒子的建议普遍持反对态度。当然，我也没有敢这样做。于是我把我的理论作为电子与质子的理论发表了出来。我认为在电子与质子之间很大的质量差或许可以用某种方法从电子之间的相互作用求得。但是，我承认这是极困难的，因为这个质量差实在太大了。”

没过多久，狄拉克的这种观点就受到了批评，而最坚定的反对者是一位数学家外尔。他不受物理现实的太多约束，更多的是从数学的对称性考虑问题，认为空穴形成的新的粒子应该与电子质量完全相同。

实验上拖后了几年才找到了这种粒子，其原因也与当时普遍流行的

只存在正电与负电两种粒子就足够的看法有关，约里奥夫妇在研究钋和铍产生的辐射时，明明看到了云室拍得的照片上，一些径迹恰好与带负电的电子运动方向相反，但他们以为那是向辐射源运动的电子，没有认真的对待，错过了发现正电子的绝好的机会。年仅 27 岁的博士后安德森(C. Anderson)，仅凭新建的带有磁场的云室拍到的第一张宇宙线的照片，就判断出存在一个与电子质量相同、电荷为正的粒子并命名为正电子。他很快地发表了这个结果，因此而于 1936 年获得了诺贝尔奖金。

1.6 量子场论在无穷大的困扰中诞生

早在 19 世纪，人们就已经认识到粒子和场是物质存在的两种形式。但本世纪二十年代中期建立起来的量子理论实际上只限于处理微观粒子。薛定谔方程和狄拉克方程也讨论了带电粒子与电磁场的作用，但没有把两者置于同等的地位。粒子是量子的，场是经典的。它们不能很好地处理光电效应、原子发射与吸收光子以及后来发现的各种粒子的产生和湮灭现象。所有这些现象都只能在场的量子理论，即量子场论中才能进行适当的定量研究。

量子场论的发展是从电磁场的量子化开始的，它是由狄拉克在 1927 年首先实现的。他把电磁场分解成无穷多种振动方式的迭加。然后把每一种振动方式仿照海森堡的做法进行量子化，使其能量取一系列分立的数值。频率为 ν 的振动方式受到激发，跳到高一个能级，就相应于产生了一个频率为 ν 的光子。激发消失时，该振动方式跳回到原来的能级意味着一个光子的湮灭。1928 年约当和维格纳(E. Wigner)引入了电子场的概念，认为狄拉克提出的电子的相对论量子力学方程，实际上是电子场的运动方程。他们仿照电磁场量子化的方式，建立起了电子场的量子化理论。电子场的激发相应于电子的产生、电子场激发的消失相应于电子的湮灭。

电磁场是矢量场，由它经过量子化得到的光子是自旋为 1 的粒子。而电子场是旋量场，量子化后得到的电子是自旋为 $1/2$ 的粒子。这两种粒子遵从很不相同的统计物理规律。光子是玻色子，而电子是费米子。此外，电子场的量子化还自然地导致两种粒子出现，即除了电子之外，还有它的反粒子-正电子。光子的反粒子就是它自己。

1929 年，海森堡和泡利进一步研究了电子与电磁场之间相互作用的量子理论。对应经典电动力学，通常人们把这种理论称之为量子电动力学或 QED。在 QED 中，电子以电流的形式与电磁场相耦合。电子之间的相互作用过程都可以看成电流之间通过电磁场为媒介发生耦合。遗憾的是，理论家们只会精确求解没有相互作用的自由电子场和电磁场方程。对于有相互作用的方程，考虑到电子之间的相互作用强度是一个很小的量，通常记为 $\alpha = 1/137$ (精细结构常数)，理论家们把相互作用看成一种对自由场的微弱的扰动，给出了一种相当好的近似求解方法。他们把

实验上所关心的散射截面和衰变宽度等物理量表示成为 α^n 的幂级数，然后，逐级求出它的近似解。这种方法称之为微扰论。

按照微扰论，电子与电子的相互作用过程的最低级近似是交换一个光子。光子与电子散射（康普顿散射）的过程，最低级近似是电子先把入射光子吸收掉，然后再把它放出来。而对于电子与正电子的相互作用过程，其最低级近似是电子与正电子先湮没，生成一个光子，然后这个光子又变成一对电子和正电子。除了初态和末态的那些真实的物理粒子外，所有中间交换的粒子都不是观测对象。它们的存活的时间都极短，其性质也与真实粒子有很大的差别，通常被称为虚粒子。虚粒子在量子场论中起着极其重要的作用。

所有上述的最低级近似计算都很简单，而且在当时的实验精度下计算结果与实验的符合是相当满意的。但是，不久人们发现，把计算的精度再提高一级这个理论就暴露出了严重的问题。

1930年，美国物理学家奥本海默计算了电子与它自己的场的相互作用，这是一个电子发射一个光子然后再把它吸收回去的过程。在这个过程中，光子不是做为真实粒子发射出来的，而是一个虚光子。按照 QED，这是一个完全可以发生的过程。奥本海默的计算涉及到一个对虚光子动量的积分，它的值是无穷大。电子与自己的场的这种相互作用称为电子的自能，也就是电子的质量。这个结果表明，在最低级近似下求得的电子质量是一个不可思议的无穷大。

在康普顿散射这一实际物理过程中，电子自能是二级微扰必然涉及到的过程。这就是说，一个实的电子先吸收掉入射的实光子变成一个虚的电子，然后这个虚电子放出一个虚光子再把它吸收掉，最后这个虚电子再放射出一个实的光子，它自己也重新变成了一个实的末态电子，完成了一个全过程。按照奥本海默的推论，这样的过程的截面的计算，给出的结果也必然是无穷大。它显然与实验完全不符。开始，人们不相信这个结果，以为奥本海默搞错了。后来海森堡和泡利证实了它。

在研究电子与电子或电子与正电子通过虚光子发生相互作用的过程中也遇到了类似的困难。例如电子与电子的相互作用过程中，电子放出的光子可以先变成一对虚的正负电子，然后这对虚的正负电子很快地湮没重新变成一个虚光子。这个虚过程称为真空极化，也是 QED 二级微扰近似过程中的一个部分。它的计算涉及到对虚电子的动量的积分，结果

也是无穷大。它相当于电子的电荷变成了一个无穷大的量。

实验上观测到的电子，其质量和电荷都是有限的。问题出在什么地方呢？1936年外斯科夫(V. Weisskopf)提出了一个建议，他认为电子总是被虚的光子和虚的电子—正电子对所包围，像是穿了身衣服。它的真实面目或裸的电子，我们是看不到的。很可能这些裸电子的电荷、质量本来都是无穷大，刚好与计算得到的无穷大相抵消，使我们所看到的真实的、物理的电子有有限的质量和能量。所以我们尽可以把那些计算得到的无穷大，吸收到电子的质量与电荷的重新定义中去，用观测得到的真实电子的质量和电荷做为理论计算所需的数值。这种技术称为“重整化”。实际上，要具体实现这一想法是很难的，外斯科夫自己并不知道怎样去做。当时有人觉得量子场论似乎大难临头了，但有的人却并不以为然。他们认为这些都属于QED高级效应，实验达不到这么高的精度。

第二次大战后的1945年，奥本海默战前的学生兰姆(W. Lamb)，利用新发展的微波技术，测量了氢原子最低的两个激发态的能量。一级QED微扰计算预言，它们是相等的。兰姆发现它们有百万分之0.4的差别，后人称之为兰姆位移。1947年6月在歇尔特岛召开的一次关于量子力学基础的讨论会上，兰姆报告了他的实验结果。当他的报告结束时，出席会议的理论家们一致认为，为了能解释这一实验结果，毫无疑问必须设法计算量子电动力学的更高级效应，把它们有意义的贡献找出来。于是，外斯科夫及后来的克拉末所建议的重整化方案，重新引起了人们的浓厚兴趣。

在返回康奈尔大学的火车上，贝特(H. Bethe)开始了电子自能的计算。为了减少些困难，他采用了比较简单的非相对论方法。回到康奈尔之后，他就完成了这个计算，对观察到的数值的百分之九十以上给出了解释。他确信，相对论的计算能够解释剩余的部分。人们听说了这个消息，都大吃一惊。

参加会议的人中有很多也都投入了这项计算工作。第一个得到结果的是兰姆和他的学生，但他们采用的减除无穷大的方法既笨拙又不可靠。其实，外斯科夫和富兰西比兰姆早得多对于兰姆位移进行了认真的计算，得到了与实验相当符合的结果。他们把使用的计算方法与得到的结果给了费曼(R. P. Feynman)和施温格(J. Schwinger)，请他们给看一看。施温格和费曼各自独立地重复这个计算，结果与外斯科夫差了一个

小的数值常数。而且问题还在于他们两个人的结果是一样的，这使外斯科夫极为担心。他和富兰西反复核算，怎么也找不出错来。几个月之后，费曼告诉外斯科夫，他和施温格都出了个错。这时外斯科夫才把文章发表，但已经落在了兰姆的后面。外斯科夫为之惋惜，他深有感触地说：“自信是一个成功的物理学家的重要的素质。”

第二年，在宾夕法尼亚，几乎是同样的人参加，召开了第二次讨论会。会上施温格是第一个报告人。他详细地介绍了由他精心设计的新的量子电动力学系统，列出了可能出现的每一个困难，指出了如何去克服它们。他用了上百页的计算完成了理论的重整化。甚至玻尔也对他的数学才能感到惊奇。但并非与会的每个人都象玻尔和奥本海默那样对施温格的方法心领神会，因为它太复杂了，用了大量人们很不熟悉的数学。奇怪的是，费曼在同一个会上给出了对这个问题的第二种更容易理解而且后来深受人们欢迎的解法。在会上却受到玻尔和奥本海默一再地批评。

费曼对重整化问题进行了大量的研究。在发展这一理论的同时，他发明了一套新的、非常有用的图形技术。紧接着施温格，他在会上报告了这一研究成果。他在黑板上画了几张图，然后用它们来阐述他所做的事情。但没有一个人能够跟得上他，主要是因为他所做的太直观了。玻尔如此气忿，以致他对费曼说：他应当回去学一学基本的量子力学。

实际上，费曼一直不喜欢通常的量子力学形式。他发展了自己的一套现在得到了广泛应用的路径积分量子化方法。对于会上的参加者，这种新的形式使他们迷惑不解。费曼企图说服他们，但未能成功。会后，费曼决心证明他的图形技术是可用的。不久，他把它们发表了。它们不仅立即被人们接受了，而且很快变成了进行这类计算的标准方法。

现在让我们来看一看费曼所引进的这些图形，它们被称为费曼图。他画一条实线表示电子，一条波线代表光子。例如两个电子的散射过程，通过交换一个虚光子而发生，画成费曼图如图 1.1 所示。

按照通常的约定，这个图形的外部有四条代表实电子的线，称为外线。线上箭头标出了电子的动量方向。中间一条代表虚光子的线称为内线。电子线与光子线的交点称为顶点，在顶点处，光子被发射或者吸收。费曼指出，他可以把电子与光子之间的任何一个可能的相互作用过程用这样的图表示出来。而且，他还可以利用这些图根据一些简单的规则写

出对这些相互作用过程进行计算所需的数学公式。这些图也可以用来讨论重整化。

上面给出的图称为树图，是在电子与电子散射过程的微扰展开中，贡献最大的部分，或称第一级近似。它的计算是很容易完成的。下一级近似包括了几种可能的贡献，即有几个图。其中之一是上图中的虚光子变成一对正、负电子，然后，它们又演变成一个光子。即我们前面提到的真空极化。图 1.2 画出了相应的费曼图。另一种可能是在两条相邻的电子外线之间，再交换一个虚光子，如图 1.3 所示。

这些图与前述的树图相比，都多了一个封闭的圈。图 1.2 多了一个电子与正电子构成的圈，而图 1.3 多了一个由电子和光子组成的圈。以第一个图为例，一个虚光子分裂成正负电子对，根据能量和动量守恒，正电子与负电子的四维动量之和应等于光子的动量。实际上有无穷多种可能性满足这个要求。这就是说闭合圈中的粒子的动量可以取无穷多种值，每一种取法都对这个过程有贡献。总的贡献是对圈动量的积分。不幸的是，这些积分都是发散的。为了能求得可以与实验比较的结果，必须象外斯科夫所建议的那样，进行重整化，把无穷大设法减除掉，抽取出第二级近似的真正的贡献。类似的做法还可以推广到更高级近似。微扰近似的级别越高，封闭的圈出现的越多。但人们发现无穷大的类型只有很少几种。通过重新定义质量、电荷等物理常数，可以把它们全都吸收掉，得到有限的、定量的结果。对于量子电动力学(QED)，这种重整化的手续可以在任意高级的近似下实现。人们称 QED 为可重整的理论。它给出的理论预言值可以达到与实验在惊人的精度上相符合。例如实验测得的兰姆位移值（即氢原子的 $2s_{1/2}$ 与 $P_{1/2}$ 的能级差，用频率表示）为 1057.862 ± 20 兆赫。而理论计算值为 1057.864 ± 14 兆赫。

量子电动力学的这段故事到此并没有结束。当奥本海默从宾夕法尼亚会议回到他的办公室时，发现桌子上有一份文稿正在等着他审阅。它是 1943 年日本物理学家朝永振一郎(S.Tomonaga)完成的关于同样的问题的第三种解法。由于战争，日本与美国的通讯实际上完全中断。这篇文稿拖延了几年才到达美国。

这个故事的最后一幕是，戴逊(F.Dyson)于 1949 年证明了这三位理论家给出的重整化的不同形式是完全等价的。戴逊是当时能完全看懂施

温格和费曼的方法的少数几个人之一。他还完成了 QED 直到所有阶近似都是可重整的证明。1965 年，施温格、费曼和朝永共享了诺贝尔物理奖。他们获奖的时间所以拖后了近 20 年，主要原因是玻尔一直持否定态度。出于对玻尔的尊敬，一直到玻尔去世，诺贝尔评奖委员会才做出了这个决定。消息发表时奥本海默给费曼发了一封电报，电文只有一个字：“终于”。

量子电动力学重整化取得的成功使物理学家们异常兴奋。他们还从来没有过任何一种理论，能与实验在如此高的精度上相一致。不少理论家相信，用不了多久就会完全地理解所有的相互作用。于是几乎所有的高能物理的理论家们都企图借助于 QED 中取得的成功经验，通过适当的量子场论方法处理强相互作用和弱相互作用。

首先我们来看弱相互作用。1934 年，在泡利提出的中微子假设的基础上，费米构造了弱相互作用的一个量子场论模型，处理有四个费米子参加的中子 β 衰变过程。他尽可能地模仿 QED，把这个过程看作是四个费米子组成的两个电流直接相互作用，通常称为四费米子相互作用。它成功地解释了当时得到的大量关于弱相互作用的实验数据。但当人们企图仿照 QED，讨论弱作用过程的高级近似时，发现它是不可重整的。随着微扰展开的级数增高，圈图引入的无穷大的种类越来越多。它们不能象 QED 那样，通过重新定义有限几个物理常数而消除掉。于是高阶的计算成为没有意义的，人们只好满足于最低级树图的计算结果。但尽管如此，人们发现还有问题。用这种四费米子相互作用算得的截面随能量成线性增长，到一定的能量时会破坏量子力学中最重要的定律之一：几率守恒。

强相互作用典型的量子场论是汤川秀树(H. Yukawa)于 1935 年建立的。他也是模仿量子电动力学，认为强作用是通过 π 介子传递的。如图 1.4 所示。

理论家们很容易地证明了，汤川所建立的这种强相互作用理论是可以重整化的。但同时他们也发现，这个理论实际上根本不能用。在量子电动力学中，由于相互作用强度正比于精细结构常数 α ，它是一个很小的量，因此级别越高的微扰图贡献越小。而强相互作用中与 α 类似的量，其数值为 15。因此，对强作用每高一级图其贡献要增强约 15 倍，根本不适用“微扰”的概念。就连最低级的计算，也是毫无意义的。

由于扩展 QED 的尝试接连碰壁，不少人对量子场论的信心从动摇到

崩溃，甚至想摒弃场论方法。量子场论发展中的这个低谷一直持续了将近 20 年。60 年代末，规范场理论取得一系列重大的突破性进展才使量子场论重新焕发了青春。

2 粒子物理学——探索物质最深层次 微观结构的科学

2.1 粒子物理学的前奏曲

粒子物理学是探索物质最深层次微观结构，研究物质的基本组元和它们之间的相互作用规律的一门前沿学科。人类对于物质基本组元的探求历史可以追溯到古老的年代，从朴素的猜测到进行科学的实验研究，认识不断深化。在一个时期被认为是基本的组元，随着科学的发展，一般总是发现它们有更深层次的结构，是由更小的组元构成的。因此，粒子物理学的研究对象和它的研究方法是在不断变化的。它形成一个独立的学科而且得到迅猛的发展是在本世纪中期宇宙线实验的发展和高能加速器的建造成功之后，所以粒子物理又称为高能物理。

本世纪 20 年代末，实验上得到证实的可以称之为“基本”的粒子有三种，它们是质子、电子和光子。质子和电子是实物粒子的最小结构单元，而光子作为电磁场的量子，是在带电粒子之间传递电磁相互作用的媒介粒子。

当时大多数物理学家都相信，一正一负两种粒子作为原子的基本成分就足够了。但是仔细地研究元素周期表就会发现新的矛盾。在周期表中元素所在的位置由原子序决定。当时已经认识到原子序数就是核外电子数，也就是原子核内包含的质子个数。每种元素又都有一定的原子量，它标志着以氢原子量为单位的核重量。除了氢原子，所有的元素原子量都比原子序大。越重的元素，这两个量的差异越大。例如，氦的原子序为 2，原子量为 4，而镭的原子序为 88，原子量为 226。如果说原子序表示核内质子的数量，那么多余的核的重量又如何解释呢？唯一的可能性是核中除质子外还有一些不带电的粒子。

1920 年卢瑟福在一次演讲中预言：除电子、质子之外还可能存在着第三种原子成分。他说：“可能存在一种核电荷为零的粒子”，它不会是一种新的基本粒子，而是由一个电子和一个质子紧密结合形成的一个中性粒子。“这种粒子会有新奇的特性，它的外部电场实际上可能等于零，所以能够自由穿透物质。它应该很容易进入原子核，”因而“将成为探测原子核的最有效的工具。”

卢瑟福的助手，29岁的查德威克(J. Chadwick)听了这个讲演。他说：“这些谈话使我确信中子一定存在。唯一的问题是到底怎样去找到它。”在那以后不久，他就开始了实验探寻，可惜这些实验仍遵循卢瑟福的思路。他们想出各种办法轰击氢原子，企图从氢原子核得到质子与电子紧密结合而成的中子。结果毫无所获。

1930年德国物理学家博特和贝克尔发现用 α 粒子轰击铍得到的辐射强度，比轰击其他一些元素几乎大十倍，他们把它称为铍辐射。1931年，约里奥·居里夫妇对这一现象进行了细致的研究，他们观察到铍辐射的穿透力很强，而且可以从石蜡及其它含氢物质打出高速质子。由于他们没有听过卢瑟福的讲演，所以在他们于1932年1月向法国科学院提交的报告中，仍认为铍辐射的这种高穿透力的物质是强 γ 射线。实际上，这种观点没有理论根据。 γ 光子可以从原子中打出电子，但质子比电子重1840倍，不可能被 γ 光子打出来。

查德威克在2月初读到居里夫妇的论文后立即开始重复他们的实验。他发现，这种铍辐射不仅很容易地穿透2厘米厚的铅，而且能够从他用作为检验的所有元素中把质子轰击出来。由此查德威克得出：这种辐射是“由质量与质子几乎相等的粒子组成。”……“这种粒子没有电荷，”它就是卢瑟福在1920年所预言的中子。从开始重复居里夫妇的工作到完成简要的报告《中子有可能存在》，前后总共花了十天。

中子的发现使原子结构得到了完满的解释：两个质子和两个中子组成一个氦核，7个质子和7个中子组成一个氮核，……。依次类推，原子序与原子量的矛盾被非常漂亮地解决了。一切都是那么完美，多彩的自然界只需要三个基本粒子：质子、中子和电子。不久，正电子的发现使人们相信，质子和中子还应该有自己的反粒子。

另外，人们在研究原子的放射性时，还发现了一个非常奇怪的现象。当时人们已经认识到 β 粒子就是电子。实验家测量了在这个过程中出射电子的能量，发现它不是一个确定的数值。衰变前后的原子都有确定的能量，为什么放出的电子能量是不确定的呢？面对这一困难问题，玻尔无计可施，甚至打算放弃核现象中的能量守恒定律。

泡利经过详细地研究发现，出射电子的能量分布有一个最大值，这个值是满足能量守恒的。但通常测得的出射电子能量都比它的最大能量小，很可能在发射过程中还有一个看不见的粒子，把能量带走了一部分。

此外，泡利还发现，这个看不见的粒子必须带有 $1/2$ 的自旋角动量，否则这个过程不仅不满足能量守恒，而且还破坏了另一个得到普遍证实的守恒定律：角动量守恒。于是 1930 年泡利建议，这个看不见的粒子是一个难以探测的，质量为零的，自旋为 $1/2$ 的中性粒子。1931 年泡利应邀在加利福尼亚美国物理学会的一次会议上，报告了他的这个发现。虽然会上有很多人不相信，但费米(E.Fermi)非常喜欢这一想法。他由此出发建立了 β 衰变的完整理论，而且把这个粒子命名为“小中子”或称中微子。

泡利的中微子假设提出以后，实验家们千方百计去寻找它。由于它与物质有极微弱的相互作用，穿透力极强，所以探测中微子的实验非常难做。1952 年，戴维斯根据中国物理学家王淦昌建议的方法用铍所做的俘获实验证实了中微子的存在。1953 年，洛斯阿拉莫斯实验室的科万(C.Cowan)和莱因斯(F.Reines)进行了探测反应堆释放的中微子实验，用了三年时间找到了它。实际上， β 衰变中放出的是反中微子，他们的实验所证实的也是反中微子。真正的中微子是 1968 年利用安放在意大利的一个废矿井中的探测器发现的。它们来自太阳内部的核反应。

中子被发现以后，原子核结构问题的研究立即引起了海森堡、伊凡宁科等人的重视，他们尝试引入一种强的相互作用力，把质子与中子紧紧束缚在原子核内。他们的一系列文章对当时正在京都大学攻读博士学位的 27 岁的汤川秀树产生了很大的影响。汤川联想到电子与质子之间的电磁作用力是由交换光子引起的，对海森堡关于核力可能是某种交换作用的观点产生了很大的兴趣。他考虑到，由于电磁作用力的范围可以延伸到无穷远，所以为光子没有质量，而核力起作用的范围很小，所以传递核力的粒子一定很重。正如传球游戏，球越重越难扔得远。汤川从量子力学出发，根据核力的力程只有 10^{-13} 厘米估算出这个新粒子的质量大约是电子的 200 倍。由于中子和质子都已确认为自旋 $1/2$ 的粒子，理论分析表明，这种新的粒子自旋可能是零。汤川于 1935 年发表了他的这一研究结果。这使他在 1949 年成为了第一位获得诺贝尔物理奖的日本物理学家。

汤川的预言发表不久，1937 年安德森就在宇宙线中发现了一个质量约为电子 200 倍的粒子，称为 μ 介子。开始，人们都认为这就是汤川预言的粒子。但详加考察之后，发现它的穿透力很强，与原子核几乎不发

生作用，不可能是核力的传媒。直到 1947 年鲍威尔利用核乳胶发现，有一种径迹对应于一种比 μ 介子更重的粒子，其质量约为电子的 270 倍，而且确实与核子有很强的相互作用。人们称之为 π 介子，它完全符合汤川的预言。

到此为止，基本粒子世界的主要角色似乎已经齐全了。质子与中子靠 π 介子传递的很强的吸引力束缚成原子核。原子核与电子靠光子传递的电磁力束缚成原子。核力比电磁力要强上千倍，人们称为强相互作用。实际上还发现，中子和 π 介子都是不稳定的粒子。自由中子只能存活 14 分钟左右，然后衰变成质子、电子和一个反中微子。 π 介子的寿命只有 10^{-8} 秒，即一亿分之一秒，它主要的衰变产物是 μ 介子和中微子。

μ 介子是一个很奇怪的角色。它除了质量比较大之外，其他方面的性质很象一个电子，而绝不象一个传递强相互作用的介子，因此更恰当地应称之为重电子或称 μ 子。 μ 子也是一个不稳定的粒子，它只能存活 10^{-6} 秒，即经过一百万分之一秒就要衰变成一个电子和两个中微子。1962 年哥伦比亚大学的莱德曼(L.M.Lederman)等在纽约长岛的布鲁克海文的国家实验室证明了， π 衰变成 μ 子所伴随的中微子与 π 衰变中伴随电子的中微子是两种不同类型的中微子。前者称为 μ 子中微子，后者称为电子中微子。由于这一重大发现，他们获得了 1988 年诺贝尔物理奖。

中子的衰变是原子核具有放射性的根源。实验上后来发现，中子、 π 介子和 μ 子的衰变有许多共同的性质，它们都是由一种很弱的相互作用引起的。这种相互作用的强度不到电磁相互作用的万分之一，人们称之为弱相互作用。

基本粒子世界的这种似乎完美的局面并没有维持多久。就在 π 介子发现的同一年，一些出乎预料的新粒子的发现重新打破了短暂的平静，给基本粒子世界带来了前所未有的繁荣。这一进展首先应当归功于实验家们为改进实验技术而做出的不懈的努力。

如前所述，正电子、 μ 子和 π 介子最初都是利用云室和原子核乳胶在宇宙线中发现的。宇宙线是由外层空间来的高能粒子与高层大气中的粒子碰撞而产生的。它们虽然给我们提供了大量信息，使我们找到一些预言的粒子，但是由于人们只能被动地记录所发生的过程，无法进行深入的研究。由此人们认识到，要想在基本粒子研究上进一步取得进展，必须设法在实验室中产生高能量粒子。

1937年，劳伦斯(E.O.Lawrence)设计建成了第一台回旋加速器。他本来希望能使粒子加速到一亿电子伏(100MeV)，但是不久发现，回旋加速器在25MeV以上就不能工作了。对于基本粒子的研究而言，这个能量是太低了。按照相对论质量能量关系式估算， μ 子的质量约为106MeV，而 π 介子为140MeV左右。要想用这样的加速器产生出 μ 子或 π 介子，显然是不可能的。为了克服这个限制，人们又设计了同步回旋加速器。六十年代以后人们开始建造对撞机。有的用正负电子对撞，有的用质子和质子或反质子对撞，从几十GeV到几千GeV，能量越来越高。它们的运行对粒子物理的研究起了至关重要的作用。

2.2 基本粒子家族“人口”大爆炸

1947年研究宇宙线的实验家罗彻斯特和巴特拉在云室中发现了两个V形径迹，它们记录的是一个中性粒子衰变成两个带电粒子的过程。其中，一个事例的衰变产物为质子和一个 π^- 介子，相应的中性粒子称为 Λ 粒子；另一个事例的衰变产物是一个 π^+ 和一个 π^- 介子，相应的中性粒子被称为 K^0 介子（右上角标代表粒子的电荷）。用粒子物理学家惯用的表示方法，这两个衰变过程为：

$$\Lambda \rightarrow p^+ + \pi^-$$

和

$$K^0 \rightarrow \pi^+ + \pi^-$$

这些粒子的行为很奇怪。它们的衰变产物都是参加强相互作用的粒子，如果是强相互作用导致它们衰变的话，它们的寿命应该只有 10^{-23} 秒，实验测到的都是 10^{-10} 秒，比预计值要长十万亿倍。这两个新粒子的发现引起了物理学家对宇宙线的新的兴趣。他们开始大量地探测这些天外来客。果然，不久就发现了一大批具有类似的奇怪行为的新粒子。一类与 Λ 相似，比质子和中子的质量大，自旋也是 $1/2$ 。其中包括 Σ^0 、 Σ^+ 、 Σ^- 和 Ξ^0 、 Ξ^- 等，人们称之为超子。另一类比 Λ 介子重，但比核子轻，类似于 K^0 、包括 K^+ 、 K^- 和 K^0 ，称之为重介子。它们与 Λ 介子一样，没有自旋。由于这些粒子的行为都很奇怪，人们把它们统称为奇异粒子。

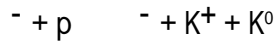
理论家们感到困惑的是已有的理论给不出关于这些粒子的任何解释。1952年派斯(A.Pais)认为，这些粒子通过强作用过程产生而通过弱作用过程衰变，产生时应该成对出现（协同产生）。他的这个预言不久就被用加速器所做的实验证实。在这个实验中，人们发现 Λ^- 和 K^0 的确是同时产生的，即

$$\Lambda^- + p \rightarrow \Sigma^- + K^0$$

与此同时，还观察到 Λ^- 超子和 K^+ 介子也是协同产生的：即：

$$\Lambda^- + p \rightarrow \Sigma^- + K^+$$

派斯还预言，奇异粒子的衰变过程只有奇数个奇异粒子参加。对于 Λ^- 和 Σ^- 超子，这个预言也是对的。但对随后发现的新的奇异粒子 Ξ^- ，这两个预言都不对。它的产生过程有三个奇异粒子出现，即：



而衰变过程包含两个奇异粒子：



就在奇异粒子困扰粒子物理学家的同时，五十年代起，在美国和欧洲相继投入运行的几个大型加速器在研究 π 介子与质子低能散射现象时，又发现了一大批短寿命的粒子，使基本粒子的数目急剧增加。

在散射实验中，人们通常测量的重要物理量叫“散射截面”。它的物理意义可以理解为入射粒子束流所看到的靶的实际面积。正如通常打靶，靶的面积越大，打中的机会越多。在散射实验中，散射截面越大，散射越容易发生，因而入射的束流中被散射的粒子数也就越多。由于基本粒子的尺度非常小，散射截面的单位“靶”是个很小的量：

$$1 \text{ 靶(b)} = 10^{-24} \text{ 厘米}^2$$

实际的截面通常采用的是单位是毫靶($1 \text{ mb} = 10^{-3} \text{ b}$)，微靶($1 \mu \text{ b} = 10^{-6} \text{ b}$)，纳靶($1 \text{ nb} = 10^{-9} \text{ b}$)和微微靶($1 \text{ pb} = 10^{-12} \text{ b}$)。相互作用越强的过程，散射截面就越大。如果都以 1GeV 能量的粒子入射，则 $\bar{\nu}$ 轰击质子的强相互作用过程，如 $\bar{\nu} + p \rightarrow n + \pi^0$ 的截面约为 10mb。而光子与质子的碰撞，出射仍为光子和质子的电磁相互作用弹性散射，截面为 $10 \mu \text{ b}$ 。反中微子轰击质子所发生的弱作用过程如 $\bar{\nu}_\mu + p \rightarrow n + \mu^+$ 截面只有 10pb。从这些典型实例大致可以看到，强作用过程的截面比电磁作用过程大一千倍，而电磁过程的截面比弱作用过程大一百万倍。

第一个短寿命的粒子是 1952 年在芝加哥大学的一台回旋加速器上发现的。当时费米和他的同事们正在研究 π^+ 介子轰击质子过程的截面随入射能量的变化关系。在描绘这一变化关系的曲线时，他们惊奇地发现，在入射能量为 1230MeV 附近，散射异常强烈。类比波动过程中的共振现象，物理学家们称这时质子与 π^+ 介子发生了共振。它们束缚在一起，形成了带两个单位正电荷的新的不稳定粒子，被命名为 π^+ 。其质量为共振能量 1230MeV。由峰的宽度可以推知，它的寿命只有 5.7×10^{-24} 秒，不到普通的基本粒子寿命的十万亿分之一。这样短寿命的状态通常称之为共振态。此后，人们用类似的方法又发现了 π^+ 、 π^0 和 π^- 粒子，它们与 π^+ 的质量差不多，自旋都是 3/2，因而都被命名为 ρ 粒子。

ρ 粒子的发现引起了人们用加速器寻找共振态的强烈兴趣。他们发

展了一些分析方法，从多粒子的复杂的末态中挑出二个或三个特定的粒子，测出它们的能量和动量，然后用相对论的质能关系计算出它们的总质量。这个总质量称为该子系统的不变质量。对多粒子终态的过程，不同的事例能量、动量的分布是不一样的，子系统的不变质量也会不同。以不变质量为横坐标，每个不变质量对应的事例数为纵坐标，画出一条曲线。如果所挑选的子系统有共振态存在，在这条曲线（称为不变质量谱）上就会出现一个峰。由这个峰可以估算出共振态的质量和寿命。正是利用这样的分析方法，五十年代末和六十年代初相继发现了 (550MeV) ， (770MeV) ， (780MeV) ， $K^*(890\text{MeV})$ 和 (1020MeV) 等几十个短寿命的粒子，使基本粒子的总数上升到二百个左右。难道这些粒子都是基本的吗？这个问题自然提到了理论物理学家的面前。

2.3 在混乱中寻找秩序

五十年代粒子物理理论家，很象上一世纪中叶的化学家。1869年俄罗斯化学家门捷列耶夫把几十种化学元素按照重量和化学性质分类排队，发现了元素周期律，为原子结构的研究奠定了基础。面对粒子物理世界的混乱局面，理论物理学家的首要任务也是设法整理出秩序来。

最早的尝试也是按质量把基本粒子分类。质子、中子以及比它们还重的超子和一些共振态，统称为重子。电子、 μ 子、中微子归类为轻子。质量介于两者之间的如介子、K介子以及 ρ 、 ω 等共振态称为介子。光子自成一类。后来又按照是否参加强相互作用，把重子和介子合称为强子。

这种分类过于粗糙，再进一步必须考虑基本粒子的其他物理性质。由于这些粒子的尺度极小，行为遵从量子力学规律，描写它们物理性质的质量、电荷、自旋等物理量的值都是量子化的，人们称之为量子数。理论家们发现，量子数与守恒定律相对应。这些守恒定律对基本粒子的产生、衰变以及其它各种反应过程，给出了许多严格的限制，在物理学家探索粒子物理世界奥秘的过程中起了关键作用。

基本粒子的第一个特征是它们的质量或者静止时的能量 M_0 。这个量子数是与能量-动量守恒相关的。用高能物理使用的单位，质子质量为938.3MeV，中子为939.6MeV。中子只比质子重1.3MeV。这一微小的差别使自由中子可以通过弱相互作用衰变为质子，而质子不可能通过任何相互作用衰变成一个真实的中子。后者为能量和动量守恒定律所禁戒。

理论家早就严格地证明了，能量和动量守恒定律是与时空的均匀性或系统具有时空平移不变性相联系的。例如中子的寿命取决于中子内部发生的衰变的几率。不管我们是在美国还是在中国（空间平移了一段距离），也不管是在泡利当年还是在今天（时间平移了一段间隔），测得的中子寿命都是14分钟。这就是所谓的时空平移不变性。能量和动量守恒正是由这种时空平移不变性所决定的。

与时空特性相关的另一个守恒定律是角动量守恒。在经典物理学中，角动量是描写粒子沿圆周做轨道运动时引入的物理量，它等于轨道半径与粒子的动量之积。如果系统有转动不变性或者说空间各向同性，则角动量一定守恒。非相对论的量子力学把这一概念推广，建立了描写

轨道运动的轨道角动量，而且同样证明了在空间各向同性时是守恒的。和经典力学不同的是，量子力学的角动量只能取一些分立的数值。

对于基本粒子，由于它们的速度通常接近于光速，必须用相对论量子力学或量子场论来描写。这时人们发现，空间各向同性时必须把粒子固有的自旋角动量与轨道角动量加在一起得到总角动量，才满足角动量守恒定律。于是，自旋角动量 S 成为基本粒子的又一个重要的量子数。

对于电子， s 为 $1/2$ 个角动量单位。虽然可以粗略地把电子想象为一个旋转陀螺来解释自旋角动量，但实际上完全行不通，因为迄今实验的观测结果仍表明电子在相当高的精度上是个点粒子。自旋角动量的真实来源至今仍是基本粒子理论中的一个谜。质子的自旋也是 $1/2$ 。与电子不同，质子是由夸克组成的。如何从夸克的角动量理解质子的自旋是当前很活跃的研究领域。从初步结果看来，问题也不象想象的那么简单。到目前为止，尚未能找到一个自洽的理论。

自旋不同的粒子行为有很大的不同，特别是当多个全同粒子构成一个多体系统时。自旋为半整数的粒子称为费米子，自旋为整数或零的粒子称为玻色子。描写全同费米子多体系统的波函数对任意两个粒子的交换要改变符号，这个性质被称为波函数具有反对称性。由此出发很容易导出著名的泡利不相容原理。它是说，不可能有两个或两个以上全同的费米子处于完全相同的状态。它是解释周期律的理论基础，也是统计物理学中所谓的费米 - 狄拉克统计的理论依据。玻色子系统不受这一原理的限制，它的波函数对任意两个粒子的交换不改变符号，称为对称波函数。因此，同一个态上可以允许有任意多个全同的玻色子，它们遵从玻色 - 爱因斯坦统计。

实验结果表明，重子的自旋都是半奇数，因而都是费米子。其中质子、中子和超子自旋是 $1/2$ 。 Λ 、 Σ^* 和 Σ^* 粒子等自旋都是 $3/2$ 。介子如 π 、 K 、 η 等自旋是零，而 ρ 、 ω 和 K^* 等自旋为 1。它们都是玻色子。所有的轻子的自旋都是 $1/2$ ，均为费米子。而光子的自旋为 1，是个玻色子。

基本粒子第三个重要的量子数是它们的电荷。质子和电子的电荷符号相反大小相等，它的量值被取为电荷单位 e 。其它粒子的电荷可正、可负，但总是 e 的整倍数。也有不少中性粒子如中子、光子等不带电荷。

现代基本粒子理论认为所有的相互作用都归结为场和场之间的相互

作用。带电粒子之间的电磁相互作用实际上是与带电粒子相应的场通过电磁场传递的相互作用。电荷的大小决定着电流与电磁场相互作用的强弱。带电粒子之间的电磁相互作用的强度

正比于它们的电荷的乘积。通常引入一个普适常数 $\alpha_A = \frac{e^2}{4\pi\hbar c}$ 来描写两

个基本电荷之间的相互作用强度。其中， \hbar 为普朗克常数除以 2π ， c 是光速。 α_A 的数值近似为 $1/137$ 。它最早是由原子光谱的精细结构的研究中测定的，故称为精细结构常数，是物理学中非常重要的一个基本常数。

电荷是一个守恒量，即任何基本粒子参加的过程电荷总量保持不变。这个守恒定律与时间、空间无关，是由基本粒子内部特性所决定的。在量子场论中，基本粒子用场函数描写。把场函数乘上一个与时空坐标没有关系的绝对值为1的因子，称为对场函数做整体规范变换。若在整体规范变换之后，场函数满足的运动方程保持不变，则可以导致电荷守恒。实际上，所有的基本粒子的场函数满足的运动方程都具有这种整体规范不变性，所以电荷守恒是个普遍成立的守恒定律。

类似于电荷，粒子物理学家分析了大量的实验，引入了重子数 B 和轻子数 L 。这两个重要的量子数的守恒定律也可以从类似的整体规范不变性推导出来。所有的重子有重子数+1，所有的反重子的重子数为-1，而所有的其他基本粒子的重子数为0。重子数守恒必然导致重子总是成对产生以及最轻的重子 - 质子绝对稳定的两个重要推论。实验在极高的精度上证实了这些结论。轻子数的情况要复杂得多。人们发现，必须认为电子和电子中微子有电子轻子数 L_e 为+1， μ 子和 μ 中微子有 μ 轻子数 L_μ 等于+1；它们的反粒子相应的 L_e 和 L_μ 取-1，其他粒子的 L_e 、 L_μ 均为0；而 L_e 和 L_μ 分别守恒才能与实验符合。也有一些理论模型对重子数和轻子数守恒提出了挑战，但都难做到与实验相符。

上述的这些守恒定律对各种相互作用过程都成立。还有一类守恒定律只对一部分相互作用过程成立，而对另一部分相互作用过程不成立。对这些有限制成立的守恒定律的深入研究是近几十年来粒子物理学取得突破性进展的重要原因之一。首先，我们来介绍宇称守恒问题。它是由物理规律的左右对称性或镜反射不变性所决定的守恒定律。

宇称是一个没有经典类比的概念。它的意义只能在量子力学的理论框架下才能理解。用量子力学的语言讲，描写粒子或粒子系统状态的波

函数如果在空间反射变换下或由左手坐标系变成右手坐标系时，不改变符号，则称其具有正宇称，或宇称量子数为+1；如果波函数改变符号，则称其具有负宇称，或宇称量子数为-1；每种基本粒子都具有自己的固有宇称。质子的宇称为+1， π 介子为-1，光子也是-1。其他粒子的宇称可以通过实验来确定。粒子系统的总宇称为各粒子的固有宇称与轨道运动所贡献的宇称之积。

左右对称性在我们的日常生活中随处可见，它带来了和谐与美。人们很自然地认为镜中世界与现实世界唯一的差别是左和右交换了位置，其他没有什么变化，一切现象都照常发生。物理学家的观念在1956年之前也正是如此，一直自然地承认了镜反射不会使物理规律发生变化，宇称守恒定律应该普适地成立。两位美籍中国物理学家李政道和杨振宁对奇异粒子 θ 与 τ 之谜的深入研究使这一观念发生了根本的动摇。

θ 和 τ 是当时人们发现的两个奇异介子。它们的自旋、质量和寿命完全相同，但通过弱相互作用 θ 衰变成三个 π 介子，而 τ 只能衰变成两个 π 介子。根据宇称守恒定律可知， θ 粒子的宇称应为-1， τ 粒子的宇称应为+1。这两个粒子太象了，很多人都觉得，它们很可能是同一个粒子。但宇称又截然相反，它们又不可能是同一个粒子。这就是所谓的 θ - τ 之谜，在当时引起了广泛的讨论。李政道和杨振宁经过认真地分析之后，确信只有认为弱相互作用过程的宇称不守恒才能解开这个谜。他们查阅了过去的文献，在1956年5月得到了下述结论： θ 过去做过的关于弱相互作用的实验实际上与宇称守恒问题并无关系。在电磁相互作用和强相互作用方面，确实有许多实验以高度准确性确立了宇称守恒定律，但弱相互作用过程中宇称守恒还是不守恒并没有做出准确的判断。

杨振宁教授回忆当时的想法时说：“在没有实验支持的情况下，长期以来，人们竟错误地相信弱相互作用中宇称守恒，这个事实本身是令人吃惊的。然而更令人吃惊的是，物理学家如此充分了解的一个空间时间对称定律可能面临破产。我们并不喜欢这种可能。我们是由于试图理解 θ - τ 之谜的各种努力都遭到挫折，而被迫考虑此可能的。”

为了明确检验在弱相互作用下宇称守恒与否，李政道和杨振宁在1956年夏提出了几个实验建议。1956年下半年，美籍中国物理学家吴健雄和她的实验组用钴60的 β 衰变做了这个检验。经过精密的实验测量，

于 1957 年证明了弱作用过程中宇称的确不守恒。就在同一年，李政道和杨振宁共同荣获了诺贝尔物理学奖。后来，实验家们对各种衰变过程进行了详细的考察，都证明了弱相互作用中宇称不守恒的理论预言，而在强相互作用和电磁相互作用中没有找到任何一个宇称不守恒的证据。

另一个局限守恒的量子数是在 30 年代引入的同位旋 I 。大量实验表明核力与电荷无关，即中子和质子、质子和质子、中子与中子之间的核力大小相同。1932 年海森堡建议，在研究核力时可以把中子和质子看成是同一种粒子（核子）的两种不同的状态。类似于自旋，可以把核子想象成在一个假想的空间旋转的陀螺，向右旋转对应于质子，向左旋转为中子。维格纳在同一年进一步完善了这一概念的理论形式，建议把这个假想空间命名为同位旋空间，核子具有类似于自旋矢量的同位旋矢量 \hat{Y} ，其长度为 $1/2$ 。它沿着同位旋空间的第三轴（相当于普通空间的 Z 轴）上的投影 $I_3=+1/2$ 时对应质子， $I_3=-1/2$ 对应中子。核力的电荷无关性意味着强相互作用具有同位旋空间的转动不变性，或者说强相互作用过程中同位旋守恒。1938 年科莫推广这一概念，认为在强相互作用过程中 $^+$ 、 0 和 $^-$ 也可以看成同一个粒子。它们的同位旋为 1 ，而 I_3 分别为 $+1$ ， 0 和 -1 。

弱相互作用如中子的 β 衰变使中子变成质子，同位旋改变了一个单位，因此同位旋不守恒。电磁相互作用是与电荷相关的，因此也不可能满足同位旋守恒。但同位旋做为一个新的量子数，可以使基本粒子按其取值分成许多小家族，对基本粒子分类很重要。例如，质子与中子构成了只有二个成员的小家族，称为同位旋二重态。它们有着近似相同的质量，相同的自旋、同位旋，只是电荷不同。三个 π 介子，三个 ρ 粒子等都是同位旋三重态。所有的强子都可以纳入这个分类方案中。

第三个局限守恒的量子数是盖尔曼 (M. Gell - Mann) 于 1952 年引入的奇异数 S 。当时盖尔曼只有 23 岁，他刚刚在 1951 年得到博士学位，然后到了芝加哥大学和费米一道工作。奇异粒子的怪异行为引起了极大的兴趣。针对派斯的“协同产生”理论在解释奇异粒子实验中遇到的困难，他建议引入一个新量子数称为奇异数 S 。对所有非奇异粒子，如质子、介子、电子等， $S=0$ 。大多数的奇异粒子如 Λ 和 Σ ， $S=-1$ ， Ξ 粒子的 $S=-2$ ， K 介子的 $S=+1$ 。

盖尔曼假设在强相互作用和电磁作用过程中奇异数守恒，而在弱作

用过程中奇异数不守恒。利用这个假设，他既解释了派斯所谓的协同产生，也解释了派斯理论所不能解释的一些产生过程。而且按照这种假设，奇异粒子衰变成非奇异粒子的过程，奇异数不守恒，因而一定是弱衰变，故寿命长。还有些衰变过程，衰变前后都有奇异粒子，如

$$\bar{K}^0 \rightarrow \pi^+ + \pi^-$$

这种过程的前后总奇异数改变了一个单位，于是奇异数也不守恒，因此也是弱衰变。这样一来，奇异粒子的怪异行为都得到了解释。

1953年日本物理学家西岛(K.Nishijima)独立地提出了同样的理论假设。盖尔曼和西岛还研究了强子的电荷 Q ，重子数 B ，奇异数 S 和同位旋第三分量 I_3 之间的关系。他们定义了一个新的超荷量子数 $Y = S + B$ ，发现：

$$Q = I_3 + Y/2$$

这个公式就是著名的盖尔曼-西岛公式。表一列出了到五十年代末发现的强子的名称和它们的主要量子数。

奇异数 S 和超荷 Y 引入之后不久，粒子物理学家们就着手在同位旋多重态的基础上，利用这些新量子数把强子归类成更大的家族。在这些研究工作中，群论方法得到了广泛的应用。群论是研究系统对称性的数学工具。强子分类的依据恰恰是对称性质，因此群论的应用是必然的。

例如，前面提到质子和中子交换或同位旋空间转动下，强相互作用不受任何影响。如果采用群论语言，同位旋空间的转动称为 $SU(2)$ 群变换，强相互作用对同位旋空间转动的不变性称为其具有 $SU(2)$ 对称性。质子和中子二重态，称为 $SU(2)$ 群的一个二维表示。由于同位旋 $SU(2)$ 群是在这个二维表示的基础上建立起来的，所以它被称为 $SU(2)$ 群的基础表示。 $SU(2)$ 群可以有一维、二维、三维以及任意整数维的表示。 π^+ 、 π^0 、 π^- 三重态就是它的一个三维表示。而 π^{++} 、 π^+ 、 π^0 、 π^- 粒子构成它的四维表示。

如果同位旋 $SU(2)$ 群是强子的严格对称群，则同一表示内的各强子除了同位旋第三分量 I_3 取不同值外，其他性质应该完全相同，实际上并非如此。各同位旋多重态内部的粒子，比如中子和质子，质量并不是完全相同的。这是因为强子除强相互作用外，还参加电磁相互作用。它产生了质量差使同位旋 $SU(2)$ 群对称性遭到了一些破坏。

引入新的量子数 Y ，寻找比同位旋多重态更大的强子家族，用群论的

语言就是要寻找比同位旋 $SU(2)$ 群更大的对称群,把同位旋与超荷同时包容在内。由于强相互作用过程中同位旋与超荷都是守恒量,更高的对称性确实存在。物理学家们尝试了各种可能性,其中 1961 年盖尔曼利用 $SU(3)$ 群建立的八重态方法取得了最大的成功。(与他同时,萨拉姆指导下的一位以色列物理学家涅曼(Y. Ne'eman)几乎循着同样思路,独立地建立起了八重态结构。)

$SU(3)$ 与 $SU(2)$ 是属于同一种类型的群,只不过 $SU(3)$ 比 $SU(2)$ 大,后者是前者的一个子群。 $SU(3)$ 群最低维的一些表示的维数为 1, 3, 8, 10, 27……等。三维表示是它的基础表示。这些表示可以用来填充自旋、宇称相同但同位旋和超荷不同的粒子,但这种填充决不是随意的,做为群的表示,每一个位置的量子数都是完全确定的。盖尔曼发现,最低质量的重子,恰好有八个,它们的同位旋和超荷量子数完全符合 8 维表示的要求。于是可以得到一个如图 2.1 所示的自旋为 $1/2$ 的重子八重态。

最低质量的自旋为 0 而宇称为负的介子(称为赝标介子)一共有九个。而自旋为 1,宇称为负的矢量介子也恰好有九个。 $SU(3)$ 群没有 9 维表示,但盖尔曼发现这两组介子的量子数与 1 维表示加上一个 8 维表示恰好相符。而且 8 维表示中有一个同位旋为 0,超荷也为 0 的态与 1 维表示量子数完全一致,它们可能会产生混合。于是这两组介子都做为单态与八重态的混合而填充,通常称之为介子九重态。它们的具体安排如图 2.2 和图 2.3 所示。

盖尔曼确实实现了把强子纳入了更大家族的目的,但是这种方案也有不尽人意之处。首先,当时的实验表明 超子和 超子宇称是相反的。把它们放在同一个多重态里面是不合理的。幸运的是,西欧中心(CERN)的实验家们重新进行了测量,1963 年他们宣布 超子与 超子的宇称相同,给了重子八重态的安排以很有力的支持。

其次,在同一个 $SU(3)$ 多重态的成员之间,质量差相当大。赝标介子九重态中,K 介子与 介子的质量差甚至达到 介子质量的三倍左右。这种质量差与同位旋多重态内微小的质量差不同,不可能解释为由电磁相互作用或弱相互作用所造成的。它表明 $SU(3)$ 群不是强相互作用的精确的对称性,而只是一种“破缺”的对称性。盖尔曼和日本物理学家大久保(S. Okubo)假定存在着强相互作用的一种破缺成分,推出一个 $SU(3)$ 多重态的成员之间的质量关系式。对于自旋 $1/2$ 的重子八重态和矢量介子九

重态，这个著名的盖尔曼 - 大久保质量公式符合得很好。

除了前述的八重态的重子和九重态的介子外，盖尔曼还讨论了自旋为 $3/2$ 的重子。当时已发现的有 $^{++}$ 、 $^{+}$ 、 0 和 $^{-}$ 以及三个 * ，它们的量子数符合 10 维表示的要求。但要填充这个十重态，如图 2.4 所示，还缺三个粒子。1962 年在西欧核子研究中心(CERN)召开的一次会议上，实验家们报告了新发现的两个粒子，称为 $^{*-}$ 和 *0 。使盖尔曼和涅曼惊奇的是这两个粒子刚好是在缺的三个粒子之中。从盖尔曼 - 大久保质量公式可以导出十重态的几个同位旋多重态之间满足等距关系，新发现的两个粒子完全符合这一质量关系。这样，十重态只剩下一个未知的粒子了。

盖尔曼当即在会上报告了他对这个粒子的理论预言。他称这个粒子为 $^{-}$ ，其自旋为 $3/2$ ，奇异数为 -3 ，质量为 1685MeV 。由于所有的奇异数守恒的强衰变方式都将导致总质量超过 $^{-}$ 的末态，所以 $^{-}$ 粒子可能发生的衰变只能是奇异数不守恒的弱作用过程。这使得 $^{-}$ 的寿命相当长，应当在 10^{-10} 秒量级。这些理论预言引起了实验家的强烈兴趣。

表 2.1 50 年代已知的基本粒子及其性质简表 (未包括反粒子)

分类	名称	符号	电荷	自旋	总同位旋	同位旋投影	奇异数	质量 (MeV)	寿命 (秒)
光子	光子	γ	0	1			0		
轻子	电子	e^-	-1	1/2			0	0.5	
	电子中微子	ν_e	0	1/2			0	~0	
	μ 子	μ^-	-1	1/2			0	105.7	2.2×10^{-6}
	μ 中微子	ν_μ	0	1/2			0	~0	
重子	质子	p	+1	1/2	1/2	+1/2	0	938.3	
	中子	n	0	1/2	1/2	-1/2	0	939.6	1.0×10^3
	超子		0	1/2	0	0	-1	1115.4	2.5×10^{-19}
	超子	+	+1	1/2	1	+1	-1	1180.5	0.8×10^{-19}
		0	0	1/2	1	0	-1	1192.5	$<10^{-14}$
		-	-1	1/2	1	-1	-1	1197.4	1.6×10^{-19}
	超子	-	-1	1/2	1/2	+1/2	-2	1321.3	1.7×10^{-10}
		0	0	1/2	1/2	-1/2	-2	1314.9	2.9×10^{-10}
介子	介子	+	+1	0	1	+1		139.4	2.6×10^{-8}
		0	0	0	1	0		135.0	0.9×10^{-16}
		-	-1	0	1	-1		139.6	2.6×10^{-8}
	K 介子	K^+	+1	0	1/2	+1/2	-1	493.8	1.2×10^{-8}
		K^0	0	0	1/2	-1/2	-1	497.8	$K_s^0 \ 0.9 \times 10^{-10}$ $K_L^0 \ 5.3 \times 10^{-12}$

布鲁克海文实验室从 1963 年 12 月开始实验。他们用 5000MeV 的 K 介子束轰击质子。到 1964 年 1 月底，从拍摄的五万张照片中终于发现了一组径迹是来自 Λ^- 的衰变。几个星期以后，第二个 Λ^- 粒子的证据也被找到了。1964 年 2 月，他们报告了这个实验结果，测得的质量为 $1686 \pm 12\text{MeV}$ ，与理论预言惊人相符。被认为是基于强相互作用 SU(3) 群对称性建立的八重态方法取得的最令人信服的成就。

2.4 夸克模型的建立

尽管八重态方法取得了很大的成功，但它毕竟只是一个类似周期表的分类方案。当时已发现的粒子有几百种，它们是不是都是“基本的”，它回答不了。回顾历史，人们经过几十年的努力才弄清元素的周期性是原子的壳层结构的反映。一百来种元素的真正基本组分只有质子、中子和电子三种。八重态方法是否也预示着强子有更深层次的结构呢？这个问题在 1964 年尖锐地摆在了理论物理学家的面前。

寻找更基本组分的努力实际上早就开始了。1949 年费米和他的学生杨振宁发表了一篇题为“介子是基本粒子吗？”的论文，建议介子是由中子、质子和它们的反粒子组成的复合粒子，即 $\pi^+ = p\bar{n}$, $\pi^0 = \frac{1}{\sqrt{2}}(p\bar{p} - n\bar{n})$, $\pi^- = n\bar{p}$ 。但难于理解的是：由非常重的核子和反核子结合成介子时，总质量减少到只有原来的两个组分质量和的十分之一左右。它意味着一定存在一种极强的相互作用把核子和反核子紧紧地束缚在一起。

奇异粒子发现以后，为了把它们包括进这种组合模型之中，1956 年日本物理学家坂田昌一(S. Sakata)推广了费米 - 杨模型，添加了一个带奇异数的 Λ 粒子，与 n、p 一起看成基本单元。这恰好相当于以 SU(3) 对称性代替 SU(2) 对称性。这三个基础粒子与它们的反粒子的所有各种组合，构成 9 个介子。除了当时已知的三种 π 介子和四种 K 介子外，还预言了两种新的中性介子。不久它们被相继发现，称为 η 和 η' 。这可以看作是坂田模型的成功。但对于重子，这个模型无法给出合理的解释。

实际上盖尔曼是在坂田模型的启发下提出八重态方法的，他认为坂田采用 SU(3) 对称性的做法可取，但以 p、n 和 Λ 做基础粒子有问题。盖尔曼的八重态方法是把 SU(3) 的 3 维表示扔在了一边，直接采用了 8 维和 10 维表示。他对自己的这种做法也并不完全满意，因为 3 维表示毕竟是 SU(3) 群的基础表示，在群论中起着极为重要的作用。SU(3) 群的所有其他表示都可用 3 维表示与反 3 维表示相乘得到。坂田正是利用了这种乘法得到了介子的八重态和单态。（用群论的语言，即 $3 \otimes \bar{3} = 1 \oplus 8$ 。其中的 \otimes 和 \oplus 都是表示的相乘与相加。）如果把坂田的做法照搬到八重态方案之中，也认为所有其他表示的粒子都是由 3 维表示的基础粒子所构成，则盖尔曼发现，它们将是三个非常奇怪的粒子。它们的电荷是基

本电荷的 $1/3$ 或 $2/3$ 。这样的带电粒子从来没有见过。这使盖尔曼处在进退两难的困境之中。

后来，盖尔曼重新进行了考虑，觉得分数电荷也许真的是可能的。只要假设这些基础粒子永远束缚在普通粒子内部，实验上的困难就可以避免。于是他写了一篇题为“重子与介子的示意性模型”的文章。在文章中他给这三个基础粒子起了一个很怪的名字，叫做“夸克”。其实夸克原是海鸟的叫声，出自詹·乔埃斯的小说，《芬尼根彻夜祭》。书中写道：“‘夸克……夸克……夸克……，’三只海鸟伸直脖子，一齐冲着绅士马克。除了三声夸克，马克一无所得。……”盖尔曼当时觉得不管起个什么名字，这个怪物一定会很快被人们忘掉，不如让它带点幽默。在文章结尾的地方，他一方面说：“即使夸克真是具有有限质量的物理粒子，推测它的行为方式也是荒唐可笑的”，因为“它们是永久禁闭在强子内部的”。另一方面，他又说：“寻找电荷为 $-1/3$ 和 $2/3$ 的稳定的夸克，会消除人们对于不存在真实夸克的疑虑。”这表明他对夸克的矛盾态度。

按照盖尔曼建议，夸克有三种，它们是上夸克(u)、下夸克(d)和奇夸克(s)。它们的反夸克为 (\bar{u}) 、 (\bar{d}) 和 (\bar{s}) 。夸克填充SU(3)的3维表示（如图2.5所示）。它们的量子数在表2.2中列了出来。

表 2.2 夸克的量子数，J：自旋，Q：电荷，S：奇异数，B：重子数，I：同位旋， I_3 同位旋第三分量。

	J	Q	S	B	I	I_3
u	1/2	2/3	0	1/3	1/2	1/2
d	1/2	-1/3	0	1/3	1/2	-1/2
s	1/2	-1/3	-1	1/3	0	0

所有的介子都是由一个夸克和一个反夸克组成，通常记为 $q\bar{q}$ 。所有的重子都由三个夸克组成，记为 qqq 。介子和重子的量子数由它们的组分的量子数相加得到。例如： 0^- 介子的组成为：

$$\begin{aligned} \pi^+ &= u\bar{d} & \pi^0 &= (u\bar{u} - d\bar{d})/\sqrt{2} & \pi^- &= d\bar{u} \\ K^+ &= u\bar{s} & K^0 &= d\bar{s} \\ K^- &= s\bar{u} & \bar{K}^0 &= s\bar{d} \\ \eta &= (u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6} & \eta' &= (u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3} \end{aligned}$$

其中的夸克与反夸克的自旋取向相反，以保证这些介子的自旋为0。 1^- 介子的成分与 0^- 介子相同，只是夸克与反夸克的自旋取向相同，总自旋为1。

又比如，自旋为1/2的重子八重态的组成为：

$$\begin{aligned} p &= uud & n &= udd \\ \Sigma^+ &= uus & \Sigma^0 &= (uds)_1 & \Sigma^- &= dds \\ \Xi^+ &= uss & \Xi^0 &= dss \\ \Lambda^0 &= (uds)_0 \end{aligned}$$

其中 Ξ^0 和 Λ^0 的基本组分都是u、d、s，但具体组合的方式不同。

自旋为3/2的重子 + 重态的组成为：

$$\begin{aligned} \Delta^{++} &= uuu & \Delta^+ &= uud & \Delta^0 &= udd & \Delta^- &= ddd \\ \Sigma^{*+} &= uus & \Sigma^{*0} &= uds & \Sigma^{*-} &= dds \\ \Xi^{*+} &= uss & \Xi^{*0} &= dss \\ \Omega^- &= sss \end{aligned}$$

在自旋 1/2 的八重态中三个夸克中有两个夸克自旋取向上方向，一个取向下方向，而在 3/2 自旋的重子 + 重态中，三个夸克的自旋是平行的，均取向上方向。

从重子的结构可以看到，要保证重子数均为 1，三个夸克的重子数必然都是 1/3。要保证 Δ^{++} 的电荷为 +2，三个 u 夸克的电荷必须都是 2/3。而要保证 Δ^- 和 Σ^- 的电荷均为 -1，则 d 夸克的和 s 夸克的电荷必须都是 1/3，这样刚好可以保证质子电荷为 +1，中子的电荷为 0。所以，如果盖尔曼所建立的模型是正确的，夸克有分数电荷是不可避免的。

分数电荷引起了物理学家们的极大兴趣，人们认为，如果它们真的存在，在实验上应该能找到一些证据。实验家们首先想到的是，在云室照片上分数电荷的径迹应该与整数电荷有明显的不同，于是重新考察了

从加速器和宇宙线得到的大量照片，结果没有任何新的发现。后来人们又推测会不会在地球的形成过程中有些分数电荷的粒子被保留下来，于是又进行了大量的地质学研究，考察了岩石、海水甚至大洋底部和月球表面，结果仍一无所获。1977年曾有人声称用铌球实验发现了一例 $1/3$ 电荷引起了一阵轰动。可惜至今没有人能够重复这个实验结果。

和盖尔曼同时，在 CERN 工作的一位年轻的博士后兹维格(G. Zweig)提出了几乎与盖尔曼完全相同的模型。象盖尔曼一样，他也认识到了SU(3)群的三维基础表示可以填充三个粒子，他称之为“爱司”。与盖尔曼不同，他认为这些粒子是强子的物理组分。从这种观点出发，他导出了很多预言。但他没有盖尔曼运气好，他的文章始终未能发表。兹维格的不幸还不止于此。正如他后来回忆中提到的，“理论物理学界对爱司模型的态度一般是不那么宽厚的。让 CERN 报告按照要求的形式发表是如此之难，以至我终于还是放弃了我的尝试。当美国的一所主要大学的物理系考虑我的工作申请时，他们的一位资深的理论家，深受理论物理学界所有的人尊敬的发言人，在教员大会上激烈地抨击爱司模型是‘骗子’的工作，因而反对我的申请。”

兹维格的遭遇其实也并不难理解。在六十年代，关于强相互作用有两种理论：量子场论和 S 矩阵。对这两方面的支持者，夸克模型都是难以接受的。为了解释为什么自由夸克在实验上观察不到，通常都是假设夸克非常重，加速器没有足够的能量产生它们。按当时的水平，这表明夸克质量至少为几个 GcV。然而由它们组成的强子，如介子，只有 140MeV。束缚态的质量比组分的总质量小这么多，意味着夸克被极强的引力束缚在一起。本来象核力那样的强作用，传统的量子场论已经束手无策，更不用说这种更强的相互作用了。至于当时正处于高峰时期的 S 矩阵理论与夸克模型更是格格不入。做为 S 矩阵理论的核心，根本不承认还有什么基本组分。对于他们，正如兹维格所说：“把强子这种核民主王国的居民，看成是由带分数量子数的基本粒子所组成的，这种思想简直是荒唐可笑的。”

中国北京粒子物理理论组把夸克看成是物质的更深层次结构，1966年提出了层子模型并对其进行了许多深入的研究，发展了相对论协变的处理办法。可惜由于一些历史原因，他们的工作当时未能被国际理论物理界所了解。

虽然在实验上和理论上，夸克模型都有一些需要进一步解决的问题，物理学家们对这个模型也还存在着不少疑问，但它也确实表现了诱人的应用前景和具有启发性的价值，解释了许多以前不能得到解释的现象。例如，它使同位旋和奇异数这两个量子数有了明确的意义。u 夸克带有 $+1/2$ 同位旋；d 夸克带有同位旋 $-1/2$ 。于是由 2 个 u 和一个 d 夸克组成的质子，同位旋为 $+1/2$ ，而一个 u 和 2 个 d 夸克组成的中子，同位旋为 $-1/2$ 。奇异数是由 s 夸克携带的。奇异粒子之所以奇异正因为包含有 s 夸克。

兹维格规则的建立是夸克模型的另一个重要的应用。1963 年，布鲁克海文的实验家们发 ϕ 介子主要衰变成 K 介子和反 K 介子而不是从能量看更为有利的 π 介子和 π 介子。有人对这个现象给出了一些解释，但很不自然。兹维格用夸克模型对这个衰变过程进行了分析， ϕ 介子是由 s 夸克和它的反夸克组成的。它的奇异数为 0，不是一个奇异粒子，但具有潜藏的奇异数。在衰变过程中，这一对 s 和 \bar{s} 夸克保留了下来，它们分别与真空中产生的 $u\bar{u}$ 或 $d\bar{d}$ 对结合在一起，自动导致包含奇异夸克的 K 介子和反 K 介子。这个过程可以画成图 2.6 所示的费曼图。如果衰变的产物是 π 介子和 π 介子，则衰变的图形如图 2.7 所示（其中两个来自 ϕ 介子）。这时 ϕ 介子的组分夸克 s 和 \bar{s} 首先湮灭掉，两对或三对非奇异夸克都从真空中产生。整个费曼图分成了不相通的两部分。兹维格把前一种图称为允许的衰变过程，而把后一种图称为禁戒的衰变过程。后者的宽度比前者有很大的压低。这个规则不仅对布鲁克海文的实验结果给出了很好的解释，在许多其他衰变过程的讨论中也起了很重要的作用。

夸克模型建立不久，人们就注意到夸克与轻子似乎存在某种对称性。夸克有三个，当时发现的轻子也只有三个，它们是电子、 μ 子和中微子。很多人觉得它表明自然界似乎有着完善的对称性存在。但是，哈佛大学的格拉肖 (S. Glashow) 考虑到 1962 年人们发现的中微子有两种不同类型 (μ 中微子和电子中微子)，即轻子实际上不是三个而是四个，1964 年与布约肯 (J. Bjorken) 一起建议还存在第四个夸克才能使夸克与轻子的对称性继续保持。他们把这第四个夸克命名为粲夸克 (c 夸克)。他们的文章对于夸克模型的进一步发展，起了重要的作用。

2.5 电子和质子的深度非弹性散射和部分子模型

60年代末，在美国斯坦福直线加速器中心（简称 SLAC）做的一些高能电子与质子非弹性散射实验，出人意料地提供了能够证明夸克存在的证据。

依赖于入射电子的能量，电子与质子的散射过程可以分成两大类。一类称为弹性散射，它发生在电子的能量比较低的时候。这时，入射电子和靶质子发生相互作用后，保持粒子类型不变而散射开来。第二类称为非弹性散射，它是在高得多的能量下发生的。这时，质子被打碎，产生出了新的粒子。能量越高，产生的粒子数目越多。

SLAC 的非弹性电子-质子散射实验是从 1967 年下半年开始的。他们测量的是所谓“单举”截面，可以用反应式 $ep \rightarrow eX$ 表示。由于非弹性散射产生的粒子很多，对每一种粒子都进行探测和鉴别（称为遍举测量）往往是很困难的。单举测量时只鉴别一种粒子（比如电子 e ），其他粒子（用 X 表示，通常是强子）不去理睬要简单得多。

对于低能电子的小角度散射，实验家们发现结果与预期的一致。但对大角度和比较高的束流能量，截面的数据使他们感到十分惊奇。他们预期尽管在非弹性散射过程大角度截面应该象弹性 ep 散射一样急剧下降，显示出衍射特征。实际却出乎预料，尽管个别的共振峰很快地消失了，但截面继续保持一个很大的值，与弹性 ep 散射截面的形状粗略相同。怎样解释这个现象呢？

实验家们首先想到的是会不会有些因素没有考虑进去，使得结果不准确。例如，在散射过程中，带电粒子会放射出一些光子而损失能量，给实验结果带来影响，通常称为辐射修正。但计算表明这部分修正加进来，仍解释不了实验结果。另一种可能是质子内部电荷分布不是均匀的，而是有一些类似于点粒子的结构（称为类点结构）。于是他们向理论家布约肯请教。

布约肯从 SLAC 理论组刚一成立就参加进来，从事非弹性散射的理论研究。他依据斯坦福大学德雷尔(S. D. Drell)的做法，引入两个结构函数描写质子。布约肯发现在适当的高能极限下（即深度非弹性散射），这两个函数的形式可以大为简化。原来每个函数都依赖两个变量。只要变换一下测量单位，在布约肯极限下，变成只依赖一个变量，这个现象

被称为“标度”现象。当实验家们把数据给了布约肯之后，他用这些数据，画了几条曲线，发现与他导出的“标度”现象完全一致。于是皆大欢喜，都觉得这是一个有意义的发现。但遗憾的是，它的意义究竟是什么，谁也不清楚。布约肯所用的量子场论中的流代数方法，实验家和大部分理论家很不熟悉，因此不能理解他究竟做了什么。

刚好这时费曼在对广义相对论做了几年研究之后，决定重新回到基本粒子上来。他向他的同事询问有什么有意思的问题，人们告诉了他关于 SLAC 的实验。于是，他决定访问 SLAC，看看他们正在做什么。费曼到达时，实验家们刚刚得到了许多新的数据，但没有人能解释它。费曼决定自己来试一试。他首先对电子与质子做了一番分析。由量子电动力学知道，电子要引起周围发生真空极化。它放出一些虚光子，这些虚光子又变成一对虚的正、负电子，正、负电子再演变成虚光子，于是电子实际上被电子云所包围。但由于电子是个类点粒子，在散射过程中，它的电子云一般不会表现出来。质子就不同了。它不仅有电磁过程产生的质子云，强作用过程使质子周围还要产生介子云。如果考虑质子的夸克模型，还要把夸克和反夸克产生的各种虚过程包括进去，图像非常复杂。费曼在分析时，没有逐个研究这些具体成份，而把它们统称为“部分子”。由此出发建立的模型称为“部分子模型”。

建立了模型之后，费曼进一步分析散射过程。由于能量很高，电子与质子的相对运动速度接近光速。电子是个点粒子，它的尺度不会受到速度的影响。电子所看到的质子，由于沿运动方向的收缩，变成了扁平的煎饼状了。又因为相对论引起的时间延缓，质子内部的部分子都冻结在自己的位置上。在这个基础上，费曼又进一步假设电子在靠近质子时发射的虚光子只与部分子中的某一个发生相互作用。每个部分子都被看做是自由的、没有结构的点粒子，它们与光子的相互作用时很象电子。于是，我们前面提到的深度非弹性 ep 散射与 ee 散射的散射截面对动量依赖关系的相似性，立即得到了很简单的解释。费曼还高兴地注意到，只要把布约肯的两个结构函数与质子内的部分子的动量分布联系起来，标度现象也自然地得到了。

费曼的部分子模型很容易理解，用起来也非常方便，受到物理学家的普遍欢迎。在很多人看来，部分子应该与夸克有某种关系。但按照费曼的理论，部分子又不象是夸克。理由之一是部分子是自由的，彼此之

间没有相互作用。而夸克之间有很强的相互作用。另外，质子中只有三个夸克，比解释标度现象所需的部分子的数目要少得多。

当时，有不少人致力于研究部分子的各种性质，试图确定部分子的自旋和电荷。1972年 SLAC 的理论家基尔曼评述有关的实验结果时指出部分子最可能的自旋值为 $1/2$ ，更加肯定了部分子与夸克的联系。后来人们认识到，部分子可能由两部分构成。一部分是质子的三个组分夸克，称之为价夸克。另一部分是不定数目的各种夸克与反夸克对，称为海夸克。部分子的这种图像，人们称为夸克-部分子模型。

1971年 MIT 的理论家外斯科夫和库提进一步发现，价夸克和海夸克的贡献仍不足以解释实验数据。他们引进了另一种能在夸克之间起胶合作用的粒子，称之为胶子，并假设它们是电中性的。胶子成分的贡献是一个自由参量，利用它可以使结构函数的理论值与实验相当好的符合。为了能更广泛地证明夸克-部分子模型，60年代末开始，一些理论家和实验家的注意力转向了中微子（或反中微子）与质子的深度非弹性散射。如果认为弱相互作用是由一个中间矢量介子 W 传递的，那么中微子与部分子之间的弱作用与电子与部分子之间的电磁作用极为相似（图 2.8 同时画出了这两种过程的费曼图，清晰地表明了这种相似性）。理论家确信，中微子与质子的深度非弹性散射也一定会有标度现象。但由于这种过程的反应截面太小，观察起来很困难。直到 1974 年，实验才肯定地证实这些深度非弹性中微子与质子的散射不仅存在标度现象，而且还证明了部分子带有分数电荷，支持从深度非弹电子与质子散射得到的类似结论。

到此为止，可以说夸克的存在已经得到了实验的证实。强子由夸克组成，夸克和轻子是基本粒子，这些概念也得到了广泛的承认。但夸克通过什么样的相互作用结合在一起呢？胶子是什么样的粒子？为什么在高能实验中可以把夸克看做是自由的部分子，但又不可能把它们释放出来？传递弱作用的 W 粒子又是什么？这一系列疑团仍然给理论家带来许多困惑。他们需要寻找更精确的理论。规范场论的深入研究终于出人意料地开创了新的局面。

2.6 奇妙的规范原理

在现代物理学特别是粒子物理学的发展过程中，对称性起着极其重要作用。所谓对称性，简而言之，就是一个物体或物体系统在某种变换下具有不变性。质地均匀的球绕其球心转动任意角度时，它的形状、大小、质量、密度分布等等，所有的性质都保持不变。我们称球对转动是对称的。方格子转动 90° 保持不变，五角星转动 72° 保持不变，人体外形对中心剖面有镜反射不变等等，在我们的周围，这样的几何对称性是随处可见的。它给自然界带来了和谐与美。

物理学家对于对称性的特殊兴趣，应该归功于本世纪初一位德国女数学家诺特(E. Noether)。她发现的著名的诺特定理告诉人们，物理量的守恒定律与对称性有着——对应关系。例如我们在前面指出的，能量 - 动量守恒是由时间 - 空间平移不变性决定的，角动量守恒可以从转动不变性推导出来。

多年来，物理学家们已经形成了一种习惯，即只要发现了一种新的对称性，就要去寻找相应的守恒定律。反之，只要发现了一条守恒定律，总要把藏在背后的对称性挖掘出来。对称性一旦确立，就会帮助人们推论出许多意想不到的结果，甚至导出惊人的发现。例如，我们在前面提到过电荷守恒。与之相关的对称性，用量子场论的语言来讲，叫做对整体规范变换具有不变性。近 20 年来，“规范不变”这个术语对粒子物理和量子场论的发展太重要了，我们不得不努力再给出一些粗浅的介绍。

在量子场论中，每一种粒子对应一个场，粒子是场的量子。粒子是用时间和空间坐标来描写的，场是用时间 - 空间坐标的函数如 $\psi(x)$ 描写，称为场函数。场函数满足一个运动方程，它决定着场的运动规律。例如电磁场的运动方程就是麦克斯韦方程，电子场的运动方程就是狄拉克方程等，这些方程在场量子化时决定着相应的粒子的行为。而场方程又可以从拉格朗日函数推导出来。用量子场论的语言讨论对称性，就是看一看在场函数做某种变换下拉格朗日函数是否具有不变性。拉格朗日函数不变，则由它导出的运动方程不变，场或粒子的运动规律保持不变。此时，还可根据诺特定理导出守恒量的表示式来。

整体规范变换指的是场函数乘上一个模为 1 的因子。它的一般形式是 $\psi(x)$ 变成为： $e^{iq\theta}\psi(x)$ 。其中， e 为自然对数的底， i 是虚数单位，

q 是粒子的特征量（如电荷）， θ 是描写该变换的参量，它很象空间转动的角度。在量子力学中这种模为 1 的因子称为相位因子。而用群论的语言讲，这种只依赖一个连续取值的参量 θ 的变换称为 $U(1)$ 变换。 q 称为该 $U(1)$ 变换的生成元。用诺特定理可以证明，在 $U(1)$ 变换下拉格朗日函数不变，则相应的守恒量就是该变换的生成元 q 。我们可以把它看做是系统的电荷。在整体规范变换中， θ 不依赖于时空坐标。这就是说整个场，每一个时空点对应的场函数统一地乘上一个相位因子。

整体规范变换要求空间各点的函数同时改变相同的相位，这是一个很特殊的变换。更一般的是每个时空点独立地做相位变换。这就是说，变换仍是 $U(1)$ 规范变换，但变换参量 θ 不是一个常数，而是时空坐标的一个任意函数，记为 $\theta(x)$ 。这种变换称为定域（或称局部） $U(1)$ 规范变换。

在研究定域 $U(1)$ 规范变换时，一个奇怪的现象出现了：单独一种物质粒子场（如电子场）存在时定域 $U(1)$ 规范不变性不可能成立，相应的拉格朗日函数在场函数做定域 $U(1)$ 规范变换之后不保持不变，而是多出了一些项。为了消掉这些项，从而保持定域 $U(1)$ 规范不变性，人们发现必须引入另外一种场，而且这种场还必须满足如下两个条件。第一个条件是在定域 $U(1)$ 规范变换使原来的物质场 $\psi(x)$ 变成了 $e^{iq\theta(x)}\psi(x)$ 的同时，新引入的场变成增加一个与 $\psi(x)$ 的四维微商（即 $\psi(x)$ 的四维梯度）成正比的项。第二个条件是这个新引入的场与原来的物质场之间存在着特定形式的相互作用。满足这两个条件的这个新引入的场称为 $U(1)$ 规范场。它是一个四个分量的四维矢量场。

有趣的是，为了使带电粒子相应的场有定域的 $U(1)$ 规范不变性而引入的规范场恰恰就是电磁场。四维矢量电磁势 $A_\mu(x)$ ($\mu=1, 2, 3, 4$) 在 $U(1)$ 规范变换下，增加了一个四维梯度项。实际上，在经典电磁理论中，人们早就发现与一定的电磁场相对应的电磁势不能唯一地确定，它们彼此之间可以相差一个任意函数的四维梯度。把这种相差为一个梯度的变换称为规范变换，起源于德国数学物理学家外尔。“规范”一词原意为量度的标准。1921 年外尔企图通过要求理论对测量的标度做一个与时间-空间相关的变换具有不变性，把引力场与电磁场统一起来，可惜未能成功。外尔把这种变换下的不变性称为规范不变性。后来，在他继续对电磁场进行研究时，仍然沿用了规范不变性这一术语，描写电磁场所固

有的对电磁势的梯度变换的不变性。

如上所述的定域 $U(1)$ 规范不变性不仅要求必须有电磁场作为 $U(1)$ 规范场存在，而且还限定了电子场与电磁场的耦合形式，并要求电磁场是零质量场。这两个要求在量子电动力学(QED)中都是满足的。此外，规范不变性还对物理量之间的关系以及计算的结果给出了很强的限制。总之，从量子电动力学这个例子，可以看出规范不变性这一原理的巨大威力。带电粒子的电磁相互作用理论完全可以靠这个原理建立起来。

既然规范不变性原理如此重要，能不能用它来建立其他的相互作用理论呢？五十年代初杨振宁决定试一试。他想通过与QED的类比，建立起具有规范不变性的强作用理论。他认为对强相互作用，同位旋守恒应是最好的出发点。他取了二分量场 (n, p) ，把 $U(1)$ 的规范变换推广成 $SU(2)$ 群相应的规范变换。但是在建立定域 $SU(2)$ 规范不变的拉格朗日函数形式时，遇到了一些困难。1953年他到布鲁克海文实验室访问。在那里，他与另一位理论家米尔斯(R. L. Mills)合用一间办公室。他们一起很快找到了解决方案。1954年发表了至今仍被场论和粒子物理学家奉为经典的著名的杨-米尔斯规范理论。

杨-米尔斯规范理论是一个具有定域 $SU(2)$ 规范不变性的理论。 $SU(2)$ 群比 $U(1)$ 群复杂得多。 $U(1)$ 群在数学上称为阿贝尔群，它的元素之间的乘法满足交换律。为了保持定域的 $U(1)$ 规范不变性，只须引入一个矢量规范场。 $SU(2)$ 群是一个不满足乘法交换律的群，为了使定域 $SU(2)$ 群规范不变性成立，必须引入三个规范场，它们都是矢量场。与电磁场不同的是，这种场有很强的自相互作用。它的量子化得到的粒子，可以发生三个粒子的自相互作用，也可以发生四个粒子的自相互作用，而光子没有自相互作用。这种自相互作用使与之相关的物理计算都变得极为复杂。

此外，规范不变性要求这些粒子没有质量，因而它们传递的相互作用必须是无穷大力程的。这一点与强相互作用的极短力程尖锐矛盾。而且实验上观察到的零质量矢量粒子只有光子，三个新的无质量粒子找不到任何实验的支持。为了解决这个问题，必须放进去质量项，但这会使规范不变性遭到破坏。

尽管存在着这些未能得到解决的问题，使杨和米尔斯当时未能用他们的理论成功地处理强相互作用，但是他们认为，他们这一工作的“整

个想法是相当有意义的”，所以他们仍决定把它发表。事实证明，他们的文章对其后几十年粒子物理学与场论的发展，起了极为重要的推动作用。

2.7 弱电统一理论

弱相互作用的第一个理论是费米在1934年建立起来的中子衰变理论。费米认为，在衰变过程中，中子变成了质子，同时中微子变成电子。由于弱作用的力程太短了，所以费米假定这四个粒子是在同一点发生相互作用的。其费曼里如图2.9所示。由于这四个粒子都是费米子，所以人们称这个理论为四费米子理论。虽然这个理论存在着一些难以克服的困难问题，但是用它可以在很高的精度下对弱作用现象进行定量研究，因此，很快就被人们接受了。

在费米理论中，中子和质子被认为形成一个与电流类似的带电的矢量流(记为 V 流)，中微子与电子形成了另一个矢量电流。四个费米子在一点的弱作用，可以看成是矢量流与矢量流的相互作用，它保持宇称不变，显然不满足1956年发现的弱作用宇称不守恒的要求。

1958年，费曼和盖尔曼与马尔萨克和苏达珊两组理论家几乎同时提出了“ $V-A$ ”理论，修改费米理论。

按照 $V-A$ 理论，中子与质子或中微子与电子不仅形成了矢量流(V)，而且还形成了一种轴矢量流(A)。流-流耦合还是对的，只不过现在的流是矢量流与轴矢量流的组合。 V 和 A 在空间反射变换下符号的变化刚好相反，所以“ $V-A$ ”理论中的拉格朗日函数包括的两项在空间反射变换下符号的变化相反，变换后的拉格朗日函数与变换以前不再相同，不变性不再存在。

$V-A$ 理论尽管有许多成功之处，但它没有改变四个粒子点作用的基本形式，因而费米理论的一些严重的困难问题仍然没有得到解决。其中包括不可重整化和在**高能下破坏几率守恒。理论家们继续探索克服这些困难的可能性。有的人建议弱作用中的这两个流不是直接接触，而是通过交换重的带电粒子发生作用。这些带电粒子自旋和光子一样，都是1，称为中间矢量玻色子，记为 W^+ 和 W^- 。还有的人认为弱作用和电磁作用既然这么象，它们应该有更深刻的联系。

最早给出弱作用与电磁作用统一形式的是施温格。他在1957年建议光子和弱作用中间玻色子 W^+ 和 W^- 是一个家族的三个成员，以非常直接的方式把弱作用与电磁作用统一了起来。但是 W 非常重，光子没有质量，它们怎么能看成一个多重态的成员，对此他不能给出令人满意的解释。 V

- A 理论发表后，施温格让他的学生格拉肖继续这项工作。格拉肖对规范不变性和重整化极为欣赏。他在规范理论的基础上重新思考这个问题。1959 年，格拉肖觉得自己已经建立了一个弱作用与电磁作用的统一理论，于是十分得意地把自己的想法告诉了当时正在做同样研究工作的萨拉姆(A. Salam)。谁知，令他十分沮丧的是，萨拉姆很快就发现了他的几处严重的数学错误。1961 年他终于发表了一个模型，成为后来居主导地位的理论。

格拉肖的模型把电磁相互作用 $U(1)$ 定域规范不变性与杨-米尔斯 $SU(2)$ 规范理论结合在一起，建立了一个 $SU(2) \otimes U(1)$ 的规范理论。规范粒子有一组形成 $SU(2)$ 三重态，另有一个是 $U(1)$ 群的单态。三重态的粒子中一个是正，一个是负，一个是中性的。 $U(1)$ 群的单态也是中性矢量粒子。通过单态的粒子与三重态的中性成员适当的混合，得到了一个重的中性粒子 Z^0 和一个零质量的光子。他把轻子也适当地填充在了 $SU(2) \otimes U(1)$ 群的表示中，保证了由光子传递的电磁相互作用是宇称守恒的，而由 W^+ 、 W^- 和 Z^0 传递的弱相互作用宇称不守恒。在这之后不久种萨拉姆也完成了完全类似的模型。

在格拉肖和萨拉姆的模型中都有一个很重的中性 Z^0 粒子，它传递一类完全新的弱相互作用。我们前面指出的如 $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$ ，是一个与中子衰变类似的过程。按照流-流相互作用的观点， μ^- 与 ν_μ 和 $\bar{\nu}_e$ 与 e^- 分别形成两个带电流，它们之间通过交换带电的规范粒子 W^+ 和 W^- 发生相互作用。费曼图如图 2.10 所示。中性 Z 粒子只能传递中性流的相互作用，如图 2.11 (β) 所示。这类过程从来没有被观察到过。

此外，在他们给出的模型中，粒子的质量问题没有能自然地解决，都是人为加上去的。这样做与规范不变性矛盾，因而成为这个模型的一大障碍。要克服这一障碍，看来需要有某种突破性进展，使得人们既能给规范粒子以质量，又能不破坏规范不变性。1961 年，芝加哥大学南部(Y. Nambu)的工作朝这个方向迈进了一步。

南部既是场论专家，又是一位超导理论专家。他熟悉高能物理与固体物理两大领域。他的研究工作侧重于量子场论在粒子物理和多体问题中的应用，很善于把一些概念从一个领域“翻译”到另一个领域。1961 年，他和一位博士后焦纳-拉西尼奥(G. Jona-Lasinio)一起，把在超导理论研究中起了重要作用的“对称性自发破缺”的概念引入到了粒子物理中。

正如我们在前面指出的，在粒子物理的各种对称性中，有些只是近似成立。如同位旋对称性只对强相互作用成立，电磁相互作用和弱相互作用使它发生破坏。这类近似的对称性称为明显破缺的对称性。南部的贡献在于，他根据与固体物理中的一些现象的类比，指出在粒子物理中还可能有一种对称性。它不是明显破缺的，其拉格朗日函数有着准确的对称性。但是由于系统的基态或真空态不具有这种对称性，使拉格朗日函数具有的对称性表现不出来。人们称为对称性被自发破缺了。萨拉姆举过一个很有趣的例子。假定人们已经在圆形餐桌前入座，准备进餐。每人面前有一个盘子，两个盘子之间有一块餐巾。于是每个人的面前有两块餐巾，一块在左，一块在右，但谁也不能肯定哪块餐巾属于自己。这就是说餐巾的摆放表现出左右对称性。这时，突然有一个人拿起了他右边的餐巾，大家也就随着他拿起了自己右边的餐巾。结果，左右的对称性被破缺了。

我们再来设想生活在一块大铁磁体的内部，其中有无数个小磁针。当温度很高，又没有任何外加磁场时，这些小磁针的取向完全杂乱无章，整个磁体有着绕空间任何方向的转动不变性，我们感受不到有任何特殊的方向性。但是，一旦温度降低到所谓的居里温度以下时，磁体内部的这些小磁针会自发地沿某一方向排列起来，出现了整块磁体的自发磁化。结果使绕任意轴的转动对称性遭到了破缺，只剩下了绕磁化方向的转动不变性。在磁体内部的我，凭着指南针可以觉察出空间出现了一个特殊的磁化方向。实际上描写磁体的拉格朗日函数并没有改变，它仍然具有绕任意轴的转动不变性，只是由于这时的最低能量态或基态变成了自发磁化态，它不具有绕任意轴的转动不变性。结果，对称性明显地降低了。人们把这种原因造成的对称性破缺称为对称性的自发破缺。由于原来的对称性被遮盖了，故也有人称之为潜藏的对称性。

南部的文章发表以后，很多理论家认为，自发破缺对称性很可能是解决规范粒子质量问题的关键。但是，就在人们开始进行尝试时，剑桥大学的戈德斯通(J. Goldstone)发表的一篇文章使这一工作遇到了意想不到的困难。在这篇文章中他给出了一个定理，人们称之为戈德斯通定理。它说：一个连续的对称性如果发生了自发破缺，则一定伴随出现一些零质量的粒子。人们称这些粒子为戈德斯通粒子。它们是自旋为零的玻色子。这些粒子所以不受欢迎是因为在自然界中人们发现的零质量玻

色子只有光子。我们不需要再多的零质量粒子。

为了解决规范粒子的质量问题，理论家们发现，既需要对称性的自发破缺，又必须消除零质量的戈德斯通玻色子，这一难题吸引了很多人的兴趣。美国的超导理论家安德森向粒子物理的同行们提过一个建议，他认为应当存在一种方法使这些无质量粒子不出现。因为在超导理论中，对称性的自发破缺没有带来任何无质量的粒子。找到了这一问题的一个相当吸引人的解决办法的，是一位英国理论物理学家希格斯 (P. W. Higgs)。

在 1964 年和 1966 年，希格斯在南部和戈德斯通的工作基础上，以一种最简单的场论模型为例，建议了一种所谓的希格斯机制。他讨论的是满足定域 $U(1)$ 规范不变性的一个复标量场与电磁场的相互作用。人们很早就认识到，由于光子的质量为零，所以它只有两个与其动量方向垂直的横极化方向。而一般的有质量的矢量粒子应该有三个极化方向，其中两个是横极化，一个是沿运动方向的纵极化。此外，在通常的情况下，复标量场的两个实分量，都具有非零的质量，场的真空态也是具有定域的 $U(1)$ 规范不变性的。但希格斯惊奇地发现，当理论中的参数取一种特殊选择使真空态的 $U(1)$ 规范不变性遭到破坏的同时，光子不再是零质量，它的纵极化分量出现了。而标量场只剩了一个有质量的分量，另一分量不见了。原来，由于对称性发生了自发破缺，标量场的一个分量转化为一个零质量的戈德斯通玻色子，它被光子“吃”掉了，变成了光子的纵分量，这就是希格斯机制。标量场剩下的有质量粒子被称为希格斯粒子。

希格斯的文章引起了温伯格的极大兴趣。开始他企图把希格斯机制用于强相互作用，但没有成功。到 1967 年，转而把希格斯机制与类似于格拉肖六年前提出的 $SU(2) \otimes U(1)$ 电磁与弱作用的统一模型结合了起来，果然得到了规范粒子所需要的质量。温伯格引入了由两个复标量场组成的二重态，它们一共包含四个实分量。通过对理论中的参数做一种特殊选择，可以使 $SU(2) \otimes U(1)$ 的对称性自发破缺成只剩下一个电磁的定域 $U(1)$ 规范对称性。相应的四个实标量场组成了三个零质量的戈德斯通玻色子和一个有质量的希格斯粒子。在未破缺之前有四个零质量的规范粒子，它们是光子和 W^+ 、 W^- 及 Z^0 粒子。自发破缺以后，光子仍保留了零质量，而 W^+ 、 W^- 和 Z^0 粒子分别吃掉了三个戈德斯通粒子，获得了非零质量。

温伯格的文章解决了弱电统一理论中规范粒子的质量难题，但从1967年这篇文章发表到七十年代初，极少有人对它感兴趣。原因之一是，人们认为它只是希格斯机制与拉格肖的弱电统一模型的一种结合，并没有对于一直困扰人们的这个理论是否可以重整化给出任何讨论。原因之二是温伯格的模型只涉及到轻子的弱相互作用。这方面的实验数据非常少。把这个模型简单的扩充，使其可以处理强子并不困难，但得到的结果与实验相矛盾。这样一来，这个模型似乎没有什么应用可言。萨拉姆几乎与温伯格同时完成了完全类似的工作，他也未能解决重整化以及中性流问题。

弱电统一规范理论的可重整性是由荷兰乌德勒支大学维特曼(M. Veltman)教授的博士生特·霍夫特(G. 'tHooft)于1971年给出证明的。维特曼从1968年起就研究规范理论的重整化问题。他发现：当令矢量粒子的质量等于零时，他从有质量规范理论导出的结果，不能回到已经确立的无质量理论。他想不出为什么会有这种怪现象。维特曼仔细地核对自己的计算，并更新了方法重新进行推导，最后他终于证明了人们通常的看法是对的，即一个靠手放进去规范粒子的质量的杨-米尔斯理论是不可重整的。而且有质量的矢量介子有三个极化自由度，无质量的只有两个极化自由度。多余的这个自由度引起了奇怪的现象。

就在这时，特·霍夫特成了他的研究生。他给特·霍夫特建议了几个选题，特·霍夫特都不感兴趣。特·霍夫特想要一个特别困难的、富有挑战性的题目。维特曼同他谈到了自己正在做的规范理论重整化的问题，但他觉得这个题目对特·霍夫特不合适。一个原因是特·霍夫特将会在一个没有人感兴趣的领域上工作。另一个理由是这问题太难，维特曼自己也还不会做。特·霍夫特听了这些介绍非常感兴趣，认为这正是自己要找的挑战性的工作。由于特·霍夫特有过关于对称性自发破缺方面的一些研究经验，他把标量多重态引入了杨-米尔斯理论，实际上重新发明了希格斯机制。他很快就证明了，通过对称性自发破缺使矢量玻色子获得质量的杨-米尔斯规范理论是可以重整的。维特曼听了他的结果之后，最初有些怀疑。刚好，维特曼设计了一套程序，可以用计算机进行复杂的解析运算。他验算了特·霍夫特的结果，发现它完全正确，这使维特曼极为惊讶。他一直想要的一个物理上现实的、有质量矢量玻色子的可重整化的理论终于找到了。

他安排了特·霍夫特于 1971 年夏天在阿姆斯特丹的一次高能物理会上给出了一个报告，但与会代表并没有完全相信这一结果。这是因为，特·霍夫特采用的所谓“路经积分”方法，人们认为数学上存在一些含混之处。此外，特·霍夫特是一个不知名的研究生，似乎不是能够解决困扰了许多主要的场论专家近二十年的难题的人。1972 年美籍韩国理论物理学家本杰明·李(B.W.Lee)以人们更熟悉的数学工具重新证明了这个结果。李一直从事规范场理论的研究，完成过一些很有影响的工作。他的权威性使人们相信了特·霍夫特的结论：规范玻色子通过希格斯机制获得质量的规范理论是可以重整化的，因此是一个物理的现实的理论。

在这以后，温伯格和萨拉姆的文章重新引起了人们极大的兴趣。温伯格的文章在 1967 到 1970 年之间只有一次被引用。而到 1973 年，一年被引用的次数高达 162 次。人们除了计算更高阶微扰修正外，许多还尝试构造替代模型。一般的做法是：选择更大的统一群，更复杂的对称性破缺方案以及不同的多重态的安排。怎样来判断一个模型优于另一个模型呢？温伯格在 1974 年的一篇评论文章中提出了两个标准：一个是自然，另一个是现实。所谓的自然是指理论中包含的参量不是手放进去的，而是可以计算的。所谓的现实，是指理论的预言能得到实验的支持。温伯格在文章中指出：“我们需要一个理论既是自然的又是现实的。迄今，这一点远没有做到。”

电弱统一模型的一个新奇之处是预言存在中性流。例如电子和中微子散射过程，在普通的弱相互作用 V-A 理论中只有如图 2.11(a)所示的由 W^+ 传递的带电流过程。而按照电弱统一模型，还应该同时存在图 2.11(b)所示的，由 Z^0 传递的中性流相互作用过程。在这个过程中，入射粒子和出射粒子的电荷没有发生任何变化。这两个图对散射过程的贡献是近似相同的。还有一类反应过程，只能由 Z^0 传递，它们在 V-A 理论中是不存在的。例如， $\nu_\mu e \rightarrow \nu_\mu e$ 的弹性散射，其费曼图如图 2.12 所示。如果实验上能找到这种类型的事例，无疑是对中性流存在给出最可靠的证据。1973 年在西欧中心(CERN)的实验家们分析了两年中拍摄的一百四十万张云室照片，终于发现了三个这样的事例，从而证明了纯轻子的弱相互作用中性流的存在。这使得电弱统一模型得到了充分的肯定，并促成了 1979 年温伯格、萨拉姆和格拉肖三位理论物理学家共享了诺贝尔物理奖。

当时在他们建立的模型中一些重要的角色特别是传递弱相互作用的媒介， W^+ 、 W^- 和 Z^0 三个规范粒子，还没有发现。理论家们相信它们一定存在，而且预言 W 粒子的质量大约为 80GeV ， Z 粒子约为 90GeV 。它们比质子重将近100倍。寻找这样重的粒子成为实验家面临的一个重大的难题。

实验家们深知，电子-正电子对撞机是寻找新粒子的最好的工具。可惜的是，到70年代末，已建成的 e^+e^- 对撞机中最大能量的是欧洲的PETRA，它能达到的最高能量只有 38GeV ，远远低于所需要的能量。当时运行的质子加速器，最大的有两台，一台是在美国费米实验室，最高能量是 500GeV ，另一个是西欧中心(CERN)的SPS，其最高能量为 450GeV 。表面上看，这两个能量值都比所预言的 W 和 Z 的质量高得多，似乎是可用的，但其实不然。原因是这两台机器都是把质子加速去轰击固定的靶中的质子。一方面，高速运动的质子束流与固定的靶质子碰撞时，很大一部分能量要转化成靶质子的功能；另一方面，质子与质子碰撞产生新粒子靠的是质子中的组分夸克之间的相互作用，它们仅能分得质子总能量的一部分、所以有效的能量其实是很小的。费米实验室的质子加速器的有效的能量最高只有 32GeV ，而SPS不超过 30GeV ，因此都远远低于所需要的能量。

西欧中心(CERN)当时正在设计能量为 100GeV 的电子正电子对撞机LEP，它无疑是最理想的。但它要到八十年代末才能建成，CERN的实验家鲁比亚不甘心等待。这是一场激烈的竞赛，拖延时间意味着会有更多的人能抢先得到发现这些重要粒子的荣誉。

1976年，鲁比亚(C. Rubbia)等开始探索一条捷径。到1981年他们成功地把超级质子加速器SPS改造成了 $270\text{GeV} \times 270\text{GeV}$ 的一台质子与反质子的对撞机。其后实验家们又经过一年多的努力，使亮度提高了一百多倍。1982年底，UA1组得到了140000个碰撞事例，经过计算机筛选，终于找到了5个事例可以确认为 W 粒子的产生。能量为 $81 \pm 5\text{GeV}$ ，与理论预言完全一致。 W 粒子终于被他们找到了。不久，UA2组也找到了四个事例，质量与UA1组的结果完全一致。1983年1月25日，他们正式发布了关于发现 W 粒子的消息。同年的5月份，6个 Z^0 粒子产生的事例也被鲁比亚发现。6月份，找到 Z^0 粒子的消息正式公布于众。它们的质量也完全符合理论预言。由于这两项重要的发现，鲁比亚于1984年得到了诺

贝尔物理奖

2.8 有色的夸克

夸克模型建立之后不久人们就发现它违背泡利不相容原理。按照这个原理，不能有两个全同费米子处在完全相同的状态。夸克的自旋为 $1/2$ 是费米子，所以也应满足这个要求，但实际上却并非如此。例如，在重子的自旋 $3/2$ 的十重态中有一个 Σ^- 粒子，它由三个 s 夸克组成。按照夸克模型，这三个粒子同处于轨道运动的基态。为了使总自旋角动量为 $3/2$ ，每个粒子的自旋必须都是 $1/2$ 。于是这三个夸克处在完全相同的状态，显然破坏了泡利原理。在同一多重态中的 Σ^{++} 粒子，由三个 u 夸克组成，它们的总自旋也是 $3/2$ ，因此也破坏了泡利原理。

1965 年南部研究了这个问题。他相信泡利原理不会错，问题出在夸克一定还有一种人们不知道的量子数，它可以把 Σ^- 的三个夸克区分开来，使它们仍然是处在不同状态而不违反泡利原理。他称这种新量子数为色荷(这里的色与普通的颜色没有任何联系)。色荷有三种，分别称为红(R)、绿(G)和蓝(B)。它们构成一个 SU(3) 群的重态。u、d 和 s 三种夸克也构成 SU(3) 群的重态。但这两种不是同一个 SU(3) 群。前者称为色 SU(3) 群，后者称为味 SU(3) 群。u、d 和 s 可以看成是三种“味”的夸克(这里的味也是借用做为量子数的名称)。

引入了色荷和色 SU(3) 群，南部能够对夸克模型的另一个谜给出简单的解释。按照夸克模型，介子由一个夸克和一个反夸克组成(即 $q\bar{q}$)。而重子由三个夸克组成(即 qqq)。为什么夸克的其他组合形式如： qq 和 $qq\bar{q}$ 等不会出现呢？南部借鉴原子的电中性提出了一个假设，认为只有色中性或无色的态才能被实验观测到。由色 SU(3) 群知道，只有颜色相反的夸克和反夸克才能构成无色的介子，也只有一个红、一个绿和一个蓝的三个夸克才能构成无色的重子。上述的一些其他的组合都是带色的。它们很可能具有极大的质量，因而绝不可能被观测到。

可惜的是，除了对上述问题给出了很好的解释之外，南部经过几年的努力，未能找到色荷在现象学方面的其他应用。不少人把它看成是理论家玩的一种技巧，没有什么特别的用途，因此没有引起人们的兴趣。直到 70 年代初两个重要现象的研究使色的概念终于有了用武之地。

第一个现象是，人们发现 π^0 衰变成两个光子的过程的理论结果对夸克色的数目有敏感的依赖关系，带三种色是不带色的 9 倍。而实验测量

值与夸克带三种色的理论值完全符合。另一个现象是电子与正电子湮灭的电磁相互作用过程中，可以产生一对 μ^+ 和 μ^- ，也可以产生一对夸克和反夸克然后再变成强子。实际上，夸克对产生与 μ 对产生过程完全类似，截面大小只与它们相应的电荷乘积的平方成正比。理论上计算的和实验上观测的都是强子对产生的截面与 μ 对产生截面的比值，称为R值。1972年，实验家们给出了对撞能量在1GeV到3GeV区域内R值的测量结果为2。理论上，如果夸克没有色荷，R值就等于u、d和s三种夸克的电荷平方和，这个值为 $2/3$ 。如果夸克有三种色，这个值要乘上三倍，刚好与实验完全符合。另外，南部曾建立了一个整数电荷的夸克模型，它给出的R值为4，与实验严重不符。所以这个过程不仅证明了色应该有三种，而且还肯定地有利于夸克具有分数电荷。

就在色荷得到了有利的证据的同时，1972年，盖尔曼着手扩展南部的思想，尝试把规范不变性纳入到色SU(3)群中，建立起强相互作用的动力学理论。但一开始，他们就遇到了一个难题，即解释不了布约肯的标度行为。众所周知，部分子是解释标度行为的关键。质子内部的部分子很象是自由粒子。而在盖尔曼等的理论里，带色的夸克在短距离情况下是否是自由的，他们不知道。

对夸克的这种所谓“渐近自由”认真进行了研究的是两名研究生。一个是哈佛大学的波利则(D. Politzer)，其导师是阔曼(S. Coleman)，另一个是普林斯顿的维尔切科(F. Wilczek)，导师是格罗斯(D. Gross)。这两组人研究同样的问题但出发点不同。波利则首先发现杨-米尔斯理论有“渐近自由”性质。这意味着强子内部的夸克当距离很近时，很象自由粒子。他觉得自己得到了一个重要发现，于是去找导师阔曼。阔曼说：“不对吧？我刚刚给格罗斯打了电话，他说他的学生维尔切科已经证明了杨-米尔斯理论没有渐近自由。你一定有一个符号弄错了。”波利则回去重新核算，结果仍然一样，这使他确信自己是对的，于是又去找阔曼。这时阔曼也听到了消息，维尔切科已经找到了在自己的结果中有一个错。波利则的结果是对的，杨-米尔斯理论的确是有渐近自由的。哈佛和普林斯顿的两个组在1973年的同一期“物理评论”上发表了他们的文章。与此同时，特·霍夫特也得到了同样的结果。

听到证明了杨-米尔斯场有渐近自由的消息，盖尔曼非常高兴。只要把渐近自由添加到他们的模型中，一切问题就迎刃而解。由此，盖尔

曼建立起了强作用的定域 SU(3) 规范理论。传递夸克之间的相互作用的新的场被引了进来。它相应的粒子称为胶子。通过它们，夸克被胶结在一起。

值得再一次强调指出的是，1964 年盖尔曼和兹维格建立夸克模型时引起的味 SU(3) 对称性不是规范理论，而且由于三种夸克的质量并不相等，所以只是一种近似的对称性。强作用的定域 SU(3) 色规范理论是以同一种夸克的三种色为基础建立起来的，是准确的对称性。仿照量子电动力学(QED)，盖尔曼把这种色荷之间相互作用的动力学理论称为量子色动力学(QCD)。

QED 是定域 U(1) 规范理论，它的规范粒子是光子。光子是电荷之间相互作用的媒介，但它本身不带电荷，所以光子与光子不会发生相互作用。带电粒子放出光子和吸收光子之后，电荷不会发生变化。QCD 是定域 SU(3) 规范理论，它的规范粒子是胶子。SU(3) 群是不可交换群，通常称为非阿贝尔群，它比 U(1) 群复杂得多。这些差别在光子与胶子的性质上有明显的反映。胶子很象光子，传递色荷之间的相互作用。不同的是胶子有八个，每个都带色荷。于是夸克放出胶子和吸收胶子之后，色荷要改变。而且胶子和胶子还会发生自相互作用。三个胶子、四个胶子都可以发生相互作用，正是胶子的这种自相互作用使 QCD 和 QED 中的夸克之间相互作用的行为完全不同。

前面曾提到过，按照 QED 理论电子周围总伴随着电子云存在。它们由虚的光子、电子和正电子对组成。由于排斥作用，虚电子被推到远离电子的外层区域。虚的正电子被吸引到电子的周围。电子的负电荷被包围的正电荷所屏蔽，所以真正的裸电子是看不到的，通常见到的总是被屏蔽的或“穿了衣服”的物理电子。本来按照库仑定律电荷之间的作用力随分开的距离增大而减小，屏蔽作用使这种趋势更明显些。

按照 QCD 理论，夸克的周围也有粒子云包围。它是由虚的夸克、反夸克和胶子组成。这个粒子云起的作用比电子云要复杂得多。主要是胶子之间存在很强的作用，它们也带色荷。于是粒子云不仅没有把夸克的色荷屏蔽掉，相反的是，反而使色荷增强。这使得夸克分离得越远，色相互作用力反而越强。通常称为反屏蔽效应。理论家证明，这是一切非阿贝尔群理论的共同特点。当夸克距离很近的时候，反屏蔽不起作用，夸克之间几乎没有胶子交换，于是很象一个自由粒子，这就是渐近自由。

当夸克距离很远的时候，交换的胶子数目大大增加，反屏蔽效应极强，夸克之间会发生非常强的吸引力，使它们紧紧地拉在一起。理论家们给出一个简单的估算，要想把夸克和反夸克拉开一个宏观的距离比如一米，需要的能量高达 10^{13}GeV ，比现在最大的加速器的能量大一百亿倍。即使能修建一个环绕地球赤道的大加速器，也远远达不到这样高的能量。这个例子告诉我们，QCD 理论能够对夸克被永远禁闭在强子内给出一些定性的解释。可惜的是，迄今，这个理论已经建立了二十年，仍未能精确地证明夸克的禁闭。理论家们仍然只能利用从实验得到的知识，建立一些模型来处理有关禁闭的问题。

例如南部建议用一种弦模型来描写介子和重子。如图 2.13(a)所示，介子是夸克和反夸克被橡皮筋一样的弦栓在一起，重子是三个夸克被如图 2.13(b)所示的弦束缚在一起形成的。弦可以随意拉伸。如果夸克和夸克(或反夸克)靠得非常近，弦变得很松，没有作用力存在，相当于渐近自由的状态。如果把夸克和夸克(或反夸克)拉开，弦就会绷紧。拉开得距离越大，绷得越紧，作用力越强。假如你有足够大的能量，弦是可以拉断的。不过断的地方出现了一对夸克和反夸克，它们栓在了断弦的末端。于是，我们得到的不是单独的夸克和反夸克，而是两个新的强子。

麻省理工学院(MIT)的约翰逊建议用一个口袋来描写强子，称为 MIT 口袋模型。这种模型把强子看成是好象一个橡皮球里面装着些气体分子。夸克就是这些分子，它们在强子口袋里自由运动，但不能从口袋里面穿出去。口袋可以变形，当变成了象飞船一样的长形时，很象一条弦。

各种模型都能对强子的某些性质给出一些解释，但没有一种模型能够解释所有的性质。理论家们坚持量子色动力学最终能够导出夸克禁闭，沿着这个方向也确实不断取得一些进展，但由于这时的相互作用非常强，场论专家熟悉的微扰论完全不适用，严格的解析求解又几乎是不可能的，所以要求得完全的和满意的答案仍是当今理论物理学家面临的重大难题之一。

2.9G.1.M. 使统一更完美

温伯格-萨拉姆模型建立之初是一个纯轻子的模型,它有 $SU(2) \otimes U(1)$ 群的定域规范不变性。这个群是由 $SU(2)$ 与 $U(1)$ 群相乘得到的。 $SU(2)$ 称为弱同位旋群,它相应的守恒量为弱同位旋 I 和它的第三个分量 I_3 , 规范粒子为 W^+ 、 W^0 和 W^- 。 $U(1)$ 群称为弱超荷群,它相应的守恒量为弱超荷 Y , 规范粒子为 B 。 W^0 和 B 量子数完全相同,它们按以下方式混合起来,得到光子 A 和中性的 Z^0 粒子,

$$A = \cos \theta_w B + \sin \theta_w W^0$$

$$Z^0 = \sin \theta_w B - \cos \theta_w W^0$$

其中混合角 θ_w 是该理论中极其重要的参量,称为温伯格角。

在温伯格—萨拉姆模型中,考虑到弱作用中宇称不守恒,以及零质量的中微子 ν_e 和 ν_μ 都是纯粹的左旋粒子(即它的自旋总是它的运动方向成左手螺旋关系,也称为左手粒子),把电子和 μ 子都分成了左手和右手两个分量,记为 e_L 、 e_R 和 μ_L 、 μ_R 。然后,取 ν_e 和电子的左手分量以及 ν_μ 和 μ 子的左手分量,分别构成弱同位旋 $SU(2)$ 群的二个二重态,即

$$\begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \text{ 与 } \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}$$

电子和 μ 子的右手分量均为弱同位旋 $SU(2)$ 群的单态。由此可见, μ 子及 μ 中微子几乎就是电子与电子中微子的复制品。通常称电子及其中微子为第一代轻子,而 μ 子及其中微子为第二代轻子。在此安排下,温伯格-萨拉姆模型的对称群记为 $SU(2)_L \otimes U(1)_Y$ 。

到 1973 年发现了弱中性轻子流过程为止,纯轻子的温伯格-萨拉姆模型在电磁过程、带电流弱相互作用过程以及中性流弱作用过程等几个方面都得到了实验的证实,电弱统一理论在纯轻子这一分支取得了很大的成功。它给出了温伯格角的近似值为: $\sin^2 \theta_w = 0.24 \pm 0.04$ 。

在此基础上人们发现,只要把当时已知的 u 、 d 和 s 三种夸克适当地安排,强子的电磁作用和带电弱流引起的弱作用也可以纳入这个模型。具体的做法是这样的。首先把 u 、 d 和 s 三种夸克也分成左手和右手分量 u_L 、 u_R 、 d_L 、 d_R 和 s_L 、 s_R 然后,为了符合 K 分子的弱衰变宽度与 π 介子弱衰变宽度比的实验值,利用 1963 年意大利物理学家卡比玻引入的卡比玻角 θ_c ,把 d_L 与 s_L 组成一个新的左手混合态 d'_L ,使 $d'_L = \cos \theta_c d_L + \sin \theta_c s_L$ 。最后,由 u_L 和 d'_L 组成一个弱同位旋二重态,它们与轻子二重态完全类似,即

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix} \leftrightarrow \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}$$

右手态 u_R 、 d_R 和 s_R 与轻子右手态类似，是弱同位旋单态。把夸克的这些二重态和单态仿照轻子，与规范粒子及希格斯粒子耦合起来，即可得到包括强子在内的弱电统一模型。

这样建立的模型可以很好地解释强子的电磁衰变以及由带电流引起的一些弱衰变过程，但是由 Z^0 粒子传递的中性流相互作用过程遇到了一些困难。这些过程可以分成两大类：一类过程只有 u 和 d 夸克参加，称为奇异数不变的中性流过程。另一类称为奇异数改变的中性流过程。我们先来看第一类。

奇异数不变的中性流过程的典型实例是 $\nu N \rightarrow \nu N$ ，即中微子与核子的弹性散射。与其类似的带电流弱过程是 $\nu N \rightarrow \mu^- X^+$ 。图 2.14 给出了这两个过程的费曼图。理论估算，这两个过程的截面应该相差不多。实验上，由于中微子的探测极为困难，直到 1973 年西欧中心从新建成的大型气泡室拍摄到的二十九万张照片中发现了二百例中性流过程 $\nu N \rightarrow \nu N$ ，同时大约发现了四百例带电流过程 $\nu N \rightarrow \mu^- p$ 。它们的比与理论的估算是相容的。1974 年，美国的费米实验室也找到了这种中性流的信号，进一步肯定了奇异数不变的中性流的存在。

奇异数改变的中性流遇到了更大的麻烦。最尖锐的矛盾发生在 K 介子的衰变。 Z^0 可以传递几种中性流的衰变模式，例如 $K^0 \rightarrow \mu^+ \mu^-$ 和 $K^+ \rightarrow e^+ e^-$ 。图 2.15 给出了这些过程的费曼图。理论的估算表明，它们与类似的带电流过程的几率是在同一个数量级。但实验上发现，这些中性流过程的几率极小。与带电流过程相比， $K^0 \rightarrow \mu^+ \mu^-$ 约为十亿分之一， $K^+ \rightarrow e^+ e^-$ 只有一千万分之一左右。它们不仅与中性流的理论预言严重不符，而且即使认为中性流不存在，这些过程也可以通过高阶带电流相互作用发生(如图 2.16 所示)，计算的结果也比实验大得多。

理论家们对这个异常现象非常感兴趣。1970 年哈佛大学的格拉肖、伊里奥普罗斯(J. Iliopoulos)和迈阿尼(L. Maiani)三个人建议了一种机制给出了最好的解释，它被称为 GIM 机制。这种机制的关键是格拉肖从夸克与轻子对称性出发建议存在的第四个夸克，即粲夸克 c 。它的引入增加了一个新的守恒量子数。味 $SU(3)$ 对称性扩大成味 $SU(4)$ 对称性。它预

言存在一个包含 c 夸克的强子家族，但从 1964 年起直到 1970 年，实验上没有任何迹象证明这一点。尽管如此，GIM 组仍然确信 c 夸克在解释奇异数改变中性流过程的异常压低的现象中起关键作用。他们认为，c 夸克非常象 u 夸克，它的电荷为 +2/3。它本应该与 s 夸克一起构成一个左手二重态，但由于弱相互作用的特殊性质，s 夸克必须与 d 夸克混合在一起。除了前述的那种混合，得到 d_L 之外，还有另一种混合： $s_L = -\sin\theta_c d_L + \cos\theta_c \delta_L$ 。c 的左手态 c_L 与 s_L 一起构成了弱同位旋的另一个二重态：

$$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$$

c 的右手态 c_R 是弱同位旋的另一个单态。按照同样的做法，很容易把这些新成员纳入到推广的温伯格-萨拉姆模型中。它们给理论带来了新的带电流和中性流。对于原来的理论影响最大的恰好是奇异数改变的中性流，它与原来的中性流符号相反，刚好抵消。它的带电流在高阶过程下也与原来的带电流的高阶过程的贡献相消。于是很好地解释了实验结果。

GIM 机制的成功使 c 夸克成了电弱统一理论不可缺少的重要角色。有了 c 夸克，两代轻子 $((\nu_e, e^-), \nu_\mu, \mu^-)$ 和两代夸克 $((u, d), (c, s))$ 之间具有了完美的对称性。SU(2)_L ⊗ U(1)_Y 电弱统一理论，对电磁相互作用和弱作用过程都给出了相当满意的定量解释。1973 年它的可重整性得到了证明后，人们对这个模型的兴趣急剧地上升。不久，c 夸克的质量估算为 1GeV 的结果发表了，它的宽度也估算出来了。人们急切盼望着实验家们能够给出明确的回答。1974 年冬，粒子物理学的一场革命终于爆发了。

2.10 十一月革命

1974年11月11日，美国斯坦福正负电子对撞机 SPEAR 的实验家里奇特(B.Richter)和布鲁克海文质子同步回旋加速器的实验家丁肇中同时宣布发现了一个质量为 3.1GeV 的粒子。里奇特称它为 J/ψ 粒子，丁肇中称为 J 粒子。于是，一个带有双重名字 J/ψ 的粒子，被载入了粒子物理学史册。十一天后，在 3.7GeV 附近，里奇特组又发现了第二个粒子，命名为 J/ψ' 粒子。经过大量的细致的研究，到 1976 年年中人们普遍地承认这些新粒子就是格拉肖预言的 c 夸克组成的粒子。这一重大发现标志着粒子物理一个新的时代的开始。关于强子由夸克组成的争论终告结束，规范理论取得了极大胜利，物理学家们称这一重大突破为“十一月革命”。

J/ψ 粒子比质子重两倍多。它不带电，自旋为 1，是个介子。这意味着它应该是夸克和反夸克组成的。但 u、d 和 s 夸克与反夸克组成的介子都早已存在了，因此它一定是一类新的夸克和反夸克组成的。此外，这个粒子的最不寻常的性质是它的长寿命。按照通常的规律，越重的粒子寿命越短。但 J/ψ 却比与其类似的轻的矢量介子 ρ 、 ω 等寿命长几千倍。在它被发现的消息刚一宣布，很多人就猜测它是由粲夸克(c 夸克)组成的。

关于c夸克，格拉肖已经谈论了许多年。1974年夏，盖拉尔德等指出在 e^+e^- 对撞中可以产生 χ 夸克与它的反夸克组成的束缚态或“隐粲”态。按照“兹维格规则”，如果明显具有粲量子数的粒子(称为“裸粲”粒子)比较重，而“隐粲”粒子的质量不足以产生两个“裸粲”粒子时，这些“隐粲”粒子的寿命就会比较长。它们只能衰变成破坏“兹维格规则”的普通强子。 J/ψ 刚好就是这种“隐粲”粒子。

哈佛大学的波利则与阿佩奎斯特(T.Applequist)注意到 J/ψ 粒子很象人们早就熟悉的电子与正电子组成的电子偶素。由于电子偶素只不过相当于氢原子中把质子换成正电子，所以两者的能级结构极为类似。 J/ψ 粒子由 c 夸克与其反夸克组成，可以称为粲偶素。由于 c 夸克很重，大约为 1.5GeV，粲偶素可以看成是一个非相对论的束缚态，它的能级结构可以照搬氢原子的处理方法求出来。粲偶素的能级图如图 2.17 所示。

这个图的横轴是 J^{PC} ，其中 J 表示粒子的总角动量， P 是空间宇称， C 是粒子与反粒子共扼变换下的宇称量子数，简称为 C 宇称。图的纵轴是粒子的质量(以 GeV 为单位)。 J/ψ 是轨道角动量为 0 自旋为 1 的粒子中能量最低的态， ψ' 是它的径向激发态。在图上点线之下还画出了六个态。它们可以分成两类，一类是 ψ_c 和 ψ_c' ，它们与 J/ψ 有相同的轨道角动量，只是总角动量 J 不同。这是因为它们的自旋不同造成的。粲夸克的自旋为 $1/2$ 。在 J/ψ 和 ψ' 中，粲夸克的自旋与反粲夸克的自旋平行，总自旋为 1，称为矢量粒子。而在 ψ_c 和 ψ_c' 中，粲夸克与其反夸克的自旋取向相反，总自旋为 0，称为赝标量粒子。这两类粒子总是结伴出现，它们的能级距离很近。例如 J/ψ 和 ψ_c 相差为 110MeV 左右，不到其质量的 4%。这个能级差是由于自旋与自旋的相互作用引起的，通常叫做能级的超精细分裂。

另外有四个都是轨道运动的激发态，它们的轨道角动量为 1，称为 P 态粒子。其中 x_0 、 x_1 和 x_2 的自旋为 1，总角动量为 0、1 和 2 三个值。这三个态之间的能级差也很小，主要来自于组分夸克的自旋与轨道角动量耦合，通常称为能级的精细分裂。最后一个 h 粒子总自旋为 0，总角动量与轨道角动量一样都是 1。

图中用实线标出的是实验上已确认的粒子，而虚线标出的是尚待确认的粒子。粒子质量的理论预言值最早采用与电磁作用类似的库仑力描写夸克与反夸克的相互作用来计算。但库仑力作用下，夸克绝不可能是禁闭的，因此这个模型过于简单了。后来，康奈尔大学的艾可藤 (E.Eichten) 和颜东茂等，在库仑势的基础上加上了一项与夸克之间的距离成正比能对禁闭结出最简单的定量描写的线性势(即著名的康奈尔势)，在势参数的适当选择下求得的能级与实验结果极相接近。

在 J/ψ 之上还有许多能级。在发现了带有粲量子数的粒子(如 D^0 、 D^*)之后，人们才知道，比 J/ψ 高一点的能量对应于一个分界线，在这个分界之上的能级都可以通过“兹维格允许”的衰变过程很快的变成一对“裸粲”粒子，因而寿命极短。只有 J/ψ 及其以下的态是有比较长寿命的稳定的粒子，它们的强衰变过程都是“兹维格禁戒”的。通常把能够发生“兹维格允许”的强衰变过程的最低能量值称为阈值。对于 $c\bar{c}$ 系统，阈值刚好超过 J/ψ (在图上用点线表示)。

有了理论家给出的能级图，实验家们就有了具体的目标。1975 年和

1976年，在正负电子对撞机上分别找到了三个 χ 态和 χ_c 粒子的证据。1992年初， h 粒子也在费米实验室得到了确证。目前只有 χ_c 还没有最后被证实。

此外，按照理论估算，“裸粲”粒子 $c\bar{q}$ 或者 $\bar{c}q$ (其中 q 表示 u 、 d 和 s 夸克) 质量应该为 2GeV 左右。1976年4月，SPEAR的戈德哈伯(G. Goldhaber)与皮埃尔(F. Pierre)一起在质量为 1.87GeV 处鉴别出“裸粲”粒子 $c\bar{u}$ ，他们称之为 D^0 。紧接着他们又发现了 D^+ 介子，它对应于 $c\bar{d}$ 。不久 $c\bar{s}$ 粒子 D_s^- 也被发现了。

1976年7月在前苏联第比利斯召开的高能物理会议，对于从1974年开始的十一月革命是一次很好的总结。会上肯定了粲夸克存在的事实。理论家们面对四个夸克与四个轻子完善的对称方案感到十分欣慰，但实际上，就在人们欣赏这种美妙与和谐的同时，打破这种佳境的尝试正在加紧进行。刺激这种努力的动机之一仍是 R 值的危机。考虑四种味 (u, d, s, c) 和三种色之后，理论预言值为 $10/3$ ，而实验值停在了 5 左右，这种偏差引起了很多人的注意，其中就有人想到重的轻子存在的可能性。

到1974年，已知的轻子只有四个，两个带电的轻子 e^- 和 μ^- ，它们各有一个中性的、无质量的伙伴： ν_e 和 ν_μ 。在电弱标准模型中，这四个轻子形成了两代完全对称的结构。但没有任何理论根据使人们相信只能存在两代、四个轻子。

1974年，意大利的电子-正电子对撞机 ADONE 报告了带电重轻子质量下限为 1GeV ，SPEAR的佩尔(M.L. Perl)对此

很感兴趣。佩尔认为，假如在 e^+e^- 对撞中产生一对重轻子，比如称之为 τ^- 和 τ^+ 。那么最容易观测的是 τ^- 衰变成 μ^- 加中微子， τ^+ 衰变成 e^+ 和中微子，导致探测器上形成“ $e\mu$ ”事例。到1974年底，佩尔探测到了大约20个 $e\mu$ 事例。1975年他发表了这个结果。称之为 τ^- 。到1978年， τ^- 的存在被普遍承认了。由于这一重大发现，佩尔和1953年发现反中微子的莱因斯一起，共享了1995年诺贝尔奖。

对产生对于 R 的测量值贡献为 1 ，而四种味和三种色的夸克对 R 值的贡献为 $10/3$ ，于是等于 5 的 R 值基本上得到了解释。但新的矛盾又产生了。原来的四个轻子与四个夸克的对称性再一次被打破。轻子加中微子又构成了新一代轻子，使轻子总数变成了六个。人们希望能

再找到两个夸克，使对称性重新得到恢复。

SPEAR 和 DORIS 的实验家们在宣布了轻子被发现之后，马上就开始了寻找新夸克的行动。他们仔细地从小能量开始逐步向高能区推移，直到 8GeV 左右，达到了机器的极限，结果什么也没有发现。这意味着新的夸克要么根本不存在，要么就很重，这两台对撞机的能量都不足以产生它们。于是人们的注意力转向了更高能量的机器。刚好，在美国费米实验室的 500GeV 质子加速器投入运行，它的能量是足够用的。

在费米实验室的这台机器上从事搜寻的是由莱德曼领导的一个合作组。从 1975 年到 1976 年，在 5.5GeV 到 10.0GeV 之间他们发现了许多 e^+e^- 对产生，在 5.8GeV 和 6.2GeV 之间挤在一起的有十几个事例。他们认为这是在 6GeV 附近存在一个长寿命粒子的证据，于是把这个结果发表了，而且命名为 Υ 。1977 年，他们改进了探测器后，事例增多了，但 6GeV 处的聚集现象却不见了。在 9.5GeV 附近出现了一个新的峰。于是他们重新报告了新的数据，认为这个新粒子的质量应该是 9.5GeV。不久他们又发现，在这附近的能区内，至少有三个粒子。 Υ 的质量进一步被确定为 9.4GeV，另外两个窄态分别为 10.0(Υ')和 10.4(Υ'')GeV。1978 年 DESY 和 1979 年美国康奈尔的 CESR 相继证实了这些结果。

Υ 被解释为带一种新味的重夸克与其反夸克组合而成的束缚态。前面我们曾提到，待寻找的新夸克应该有两个，通常称它们为顶夸克和底夸克。前者简记为 t 夸克，它的电荷为 +2/3，很象 u 或 c 夸克，弱同位旋为 +1/2。后者记为 b 夸克，电荷为 -1/3，很象 d 或者 s 夸克，弱同位旋为 -1/2。实验上发现的矢量介子 Υ 粒子有利于解释为 $b\bar{b}$ 束缚态而不利于 $t\bar{t}$ 。

1978 年奎格(Quigg)和罗斯纳(Rosner)对这些态给出了理论估算。

Υ 、 Υ' 和 Υ'' 分别被记为 $(1S)$ 、 $(2S)$ 和 $(3S)$ 。 $(1S)$ 是自旋为 1 的轨道角动量为 0 的基态，后两个是它的激发态。实验测得的质量和轻子衰变的宽度与理论预言符合得很好。按照理论估算，S 态的更高激发态，如 $(4S)$ 应当在“裸 b”介子 B 的阈之上，可以通过“兹维格允许”过程

衰变成 $B\bar{B}$ 末态，因此宽度比阈以下的三个粒子大得多。 $\Upsilon(4S)$ 、 $\Upsilon(5S)$ 和 $\Upsilon(6S)$ 果然被康奈尔的 CESR 先后在 1980 年发现，它们的性质与理论

的预言是一致的。

$\gamma(4S)$ 衰变产生一对 $B\bar{B}$ 介子是1983年在CESR确认的。B是b和u、d
B是b和u、d等轻夸克组合而成的介子。它的质量约为5.28GeV，后来b
和s夸克组成的 B_s 介子也在加速器上找到了。对这些B介子的细致研究
至今仍是粒子物理理论与实验的重要课题。

家族的发现使已知的夸克数增加到五个。从对称性考虑，第六个
夸克即t夸克(顶夸克)应该存在。在当时，人们认为德国DESY的电子-
正电子对撞机PETRA最有希望能够找到它。因为它的总能量可以达到
38GeV。从1979年PETRA的实验家们开始在整个可以允许的能量范围内
搜寻。1980年在斯坦福直线加速器中心(SLAC)建成了一台新的 e^+e^- 对撞机
PEP(最大能量为36GeV)，也投入了搜寻工作。可惜的是，，他们都未
能成功。八十年代中期日本建造了一台最高能量75GeV的 e^+e^- 对撞机
TAISTAN。他们缓慢地提高能量，扫过整个能区仍然毫无所获。其后西欧
中心建成了LEP,SLAC建成了SLC,这两个 e^+e^- 对撞机都可以达到100GeV
的最高能量，也一直没有进展。1994年4月，在美国费米实验能量为
1800GeV质子反质子对撞机TEVATRON工作的CDF和D0两个组中，CDF组
找到了几个t夸克事例但D0组却未能证实。为慎重起见，CDF组宣布观
察到了t存在的实验证据。直到1995年3月2日，CDF和D0两个组才同
时宣布，他们肯定地证实了t夸克的存在。CDF组给出的质量值为 $176 \pm$
 13GeV ，D0组给出的为 $199 \pm 30\text{GeV}$ 。在误差范围内，这两个值是一致的。
与理论家的间接估算也完全相符。这一发现无疑是近年来粒子物理取得
的最重要的结果。

2.11 强相互作用的规范理论——QCD

70年代末和80年代初，作为弱电统一理论的温伯格-萨拉姆模型，其核心部分得到了实验的证实，取得了非凡的成就。与此同时，在色SU(3)群定域规范不变性基础上建立起来的量子色动力学(QCD)，也经受了許多实验检验，确立了作为强相互作用理论唯一候选者的地位。

“十一月革命”带给粒子物理世界许多新粒子，它们提供了检验QCD理论的绝好机会。如前所述，受QCD理论的启发，理论家们建立起的各种模型对新粒子的性质给出了很好的解释，其理论预言给了实验家们很大的帮助。

按照QCD理论，不仅 $q\bar{q}$ 和 qqq 可以形成无色的介子和重子，夸克、反夸克和胶子的其他一些可能的无色组合，也应该对应于物理上可以观测到的束缚态。这些态统称为奇特态。例如四夸克态 $qq\bar{q}\bar{q}$ ，混杂态 $q\bar{q}g$ 等。特别令人感兴趣的是，由于胶子有色荷，它们的自相互作用可以使二个胶子或三个胶子形成无色的束缚态 gg 或 ggg ，人们称之为胶球。

早在1975年盖尔曼等就预言了胶球的存在。人们估算了它的质量，讨论了它的自旋、宇称的可能取值。由于胶子不带电，只参加强相互作用，所以胶球与光子没有直接的耦合，只能通过由夸克与反夸克形成的封闭的圈图衰变成光子。表现在实验上，胶球衰变成两个光子的几率远远小于衰变成二个胶子再强子化为强子的几率。这两个几率之比称为粘度，是鉴别胶球的一个重要的依据。另一个依据是由于胶子与各种味的夸克对的耦合强度与味无关，胶球衰变成普通强子的几率与衰变成奇异强子的几率应该差不多。

根据这些理论预言，实验家们进行了认真的搜寻。80年代以来，在J/ψ辐射衰变过程中的确发现了一些有希望的候选者，其中包括(1440)，(1720)，(2230)(括号内为质量值，单位是MeV)。这些粒子很难看成普通的介子，因为它们在介子对称性分类中的相应位置已经被占据了。但它们的观测行为与关于胶球的预言又不完全一致，不能断定它们就是胶球。

从目前情况看，由于理论上和实验上都存在一些问题，使寻找胶球的努力遇到了很大的困难。理论上，胶球的质量和宽度缺乏明确的预言，

给不出一个强有力的判据。实验上， J/ψ 的辐射衰变在最有希望观测到胶球的 1GeV 到 2.5GeV 能区，结构非常复杂，很难分辨得十分清楚。最近，对在北京对撞机上关于 J/ψ (2230) 数据所做的最新分析，提供了有利于这个粒子是胶球的一些证据。如果实验上关于这个粒子的性质辨认得更清楚一些，能够最后确证它真的是胶球的话，将是了不起的成果。

与强子谱学相比，利用高能加速器进行的一些实验对 QCD 理论可以给出更直接的检验。前面我们曾介绍过，60 年代末，对电子与质子的深度非弹性散射过程的研究发现了布约肯的标度规律，它导致费曼建立了部分子模型。1973 年人们发现，这个模型的理论基础正是 QCD。QCD 具有的渐进自由特点，可以给出标度规律。但是随着加速器能量的提高，ep 散射的能量范围逐渐扩大，数据大量增加。到 1975 年 SLAC 的许多实验结果表明标度规律并不是精确成立的。理论家们详细地研究了标度破坏的现象，他们发现只考虑 QCD 渐进自由的朴素的部分子模型不足以解释越来越复杂的实验现象。他们利用 QCD 理论的最低级微扰近似进行了定量计算，结果与标度破坏的实验曲线符合得相当满意。实验家们很高兴，认为他们很好地检验了微扰 QCD 理论。理论家们却有些担心。因为这毕竟只是最低级微扰论的结果。更高级的微扰会带来多大修正，理论家并没有很大把握。

对 QCD 理论的另一种很有力的检验来自高能 e^+e^- 对撞实验。如前所述，1978 年在德国的 DESY 开始运行的 PETRA (38GeV) 和 1980 年在斯坦福投入运行的 PEP (36GeV) 寻找 t 夸克的努力毫无所获，寻找 W 和 Z 粒子又无能为力，那么还能寻找什么呢？QCD 的理论家们建议他们来研究“喷注”现象。

首先让我们来解释一下什么是喷注。电子与正电子对撞时先湮灭成一个虚光子。如果对撞能量比较低，则虚光子将变成一对正负电子或一对 μ 子，它们将被探测器所接收。如果能量很高，虚光子会变成一对正、反夸克。当能量刚好达到某个矢量粒子的质量附近(称为该矢量粒子的产生阈)时，正、反夸克对会形成束缚态，在散射截面的曲线上会有一个窄的共振峰。实验家们正是根据这个峰，发现 J/ψ 和 Υ 粒子的。

如果能量再高则共振态不会形成，正反夸克对背对背地飞离开来。由于色禁闭，裸的夸克或反夸克不可能象自由粒子那样被探测器接收到。按弦模型，产生的这一对夸克与反夸克被一条弦紧紧的拉着。当它

们飞离开时，弦随着离开距离的增大而越来越紧，到一定程度终于断裂成两条或几条新的弦，它们都是一些物理的强子。其中的一半沿原来夸克的产生方向飞出，另一半沿着反夸克的方向飞出。探测器看到的是两个由强子组成的背对背的锥体。能量越高，锥的顶角越小。这两个锥形出射的粒子束就称为两个喷注。喷注的轴的方向实际上就是母夸克或反夸克飞出的方向。喷注中的强子，其动量都靠近轴的方向，垂直于轴方向的动量分量(横动量)都很小。

1975年，汉森(G.Hanson)等在斯坦福的 e^+e^- 对撞机 SPEAR 上第一次观察到背对背的两个喷注现象，而且发现喷注的产生截面对于喷注轴与束流夹角的依赖关系只能用产生喷注的母粒子有 $1/2$ 的自旋才能解释，它刚好与喷注产生于夸克与反夸克的理论是一致的。

1976年，CERN的理论家：艾利思(J.Ellis)等用QCD预言除了双喷注现象之外，还可能有三喷注现象出现。这是因为虚光子变成夸克、反夸克对后，按照QCD理论，高能的夸克或反夸克会发射一个高能胶子(称为硬胶子)。由于色禁闭，这个胶子也不能孤立地释放出来，它也要碎裂成一些强子，因此也要形成一个喷注。于是寻找三喷注来检验QCD就成了 e^+e^- 对撞机的实验家们一个新的重要的研究课题。

实验上寻找三喷注现象比双喷注现象困难得多。双喷注是背对背沿着一个轴发生的，轴与束流方向的夹角理论上可以预言。三喷注中有三个轴，它们的方向依赖于硬胶子发射的能量与动量，理论上无法确定。此外，硬胶子发射过程的几率与不发射过程几率之比等于QCD的耦合常数 g_s 。这个耦合常数在胶子很硬时远远小于1，三喷注的几率远小于双喷注，所以直到1979年，实验家们才在PETRA对撞机上找到了第一个三喷注的证据。而且证明了这些事例的几率和整体特征与QCD的预言一致。三喷注事例的发现有力地支持了QCD硬胶子发射的图象，提供了胶子存在的证据。

在强子-强子高能碰撞过程中硬胶子发射也带来明显的修正。70年代末，西欧中心及费米实验室的质子-反质子对撞机对轻子对产生即著名的德雷尔-颜过程进行了详细的研究。从夸克-部分子模型来看，这个过程刚好与轻子对湮灭产生夸克对相反。它是通过作为质子组分的夸克与另一个质子(或反质子)中的反夸克湮灭生成虚光子，然后虚光子产生一对轻子。当质子中的组分夸克之间交换硬光子时，发生湮灭的夸克带有了

很大的横向动量，结果使产生的轻子对相对于束流有很大的横动量。这个预言引起了实验家们的注意，因为它可以给 QCD 提供新的检验。1979 年报告的实验结果与理论相当符合。

但是无论是强子谱学，还是高能碰撞，都只能对 QCD 理论给出很不完全的检验。它们都涉及到 QCD 的低能和长距离的行为。这时 QCD 的耦合常数会变得很大，微扰论方法不再适用。而对非微扰问题的处理，理论家的能力非常有限。1974 年，年轻的理论物理学家威尔逊想出了一个好的主意。他把建立在连续时空的量子场论与处理不连续的晶格的经典统计力学详细地比较，发现基本的计算公式惊人的相似。只要把时空也看成一种不连续的格子，量子场论的公式就可以很容易地翻译成统计力学的公式。经典统计力学经过上百年的研究，积累了大量的经验，有许多成熟的计算方法略加改造就可以用来计算量子场论中的物理量。关于这部分内容我们将在下一篇中介绍。

2.12 标准模型的建立

至此我们已经看到粒子物理有两个很成功的理论。其一是统一描写弱电相互作用的温伯格-萨拉姆模型。它是 $SU(2) \otimes U(1)_Y$ 规范理论。其二是描写U(强作用的量子色动力学QCD。它是 $SU(3)_C$ 规范理论，C代表色。通常人们把这两个理论组合成 $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ 规范理论，统一描写强、弱、电三种相互作用，称为基本粒子的标准模型。

标准模型包含的基本粒子有三类。一类是物质粒子，包括夸克、轻子及它们的反粒子，它们都是自旋为1/2的费米子。另一类是传递相互作用的媒介粒子，它们都是自旋为1的玻色子。还有一类是为了使 $SU(2)_L \otimes U(1)_Y$ 规范理论自发破缺而引入的自旋为0的希格斯粒子。

夸克有六种味，它们是u、d、s、c、b、t。每一种味有三种色，因此夸克总数为18。轻子共有六个，它们是e、 ν_e 、 μ 、 ν_μ 、 τ 和 ν_τ 。于是，夸克加上轻子一共24个。它们的反粒子也有24个。于是物质粒子(包括反物质)一共48个。)传递相互作用的媒介粒子也有三类共12个。一类是传递电磁作用的光子，第二类是弱作用的媒介 W^+ 、 W^- 和 Z^0 粒子，第三类是传递强作用的八个胶子。

希格斯粒子的数目决定于为破缺而引入的希格斯场的数目。引入一个复希格斯场二重态的模型称为最小标准模型。它只有一个中性的希格斯粒子H。它的最小扩充是引入二个复希格斯场二重态，破缺之后剩下五个希格斯粒子，其中两个带电，三个是中性的。

在标准模型中，物质粒子分成了性质几乎完全相同的三代，他们的唯一不同是粒子的质量。每一代包括两个夸克和两个轻子。第一代最轻，第三代最重。实际上，只有第一代是构成我们周围复杂纷纭的物质世界的基本组分。第二代与第三代构成的物质寿命都极短，只有用极高能量的加速器才能使它们产生出来。表 2.3 列出了三代的粒子的安排，并列出了相应的近似质量。

表 2.3 标准模型中的三代物质粒子

	第一代	第二代	第三代
夸克	u ~ 5MeV	c ~ 1.5GeV	t ~ 176GeV
	d ~ 8MeV	s ~ 0.15GeV	b ~ 5GeV
轻子	$\nu_e < 5\text{eV}$	$\nu_\mu < 0.16\text{MeV}$	$\nu_\tau < 23.1\text{MeV}$
	e ~ 0.5MeV	$\mu \sim 0.1\text{GeV}$	$\tau \sim 1.78\text{GeV}$

应该指出的是，表中所列的质量，并不是从标准模型计算出来的。标准模型中物质粒子的质量虽然是通过希格斯机制产生的，但这种产生是在物质场与希格斯场存在着汤川型耦合的基础上发生的。实际上，对应每一个非零质量的物质粒子都必须引入一项汤川耦合，耦合系数与粒子质量成正比。因此，若如上表所示，这十二个物质粒子质量都不为零，就必须有十二个汤川耦合系数，它们都是标准模型中待定的参数。

对于带电轻子 e^- 、 μ^- 和 τ^- ，列出的是实验测量值。 e 和 μ 的实验值精度很高，达到 8 位和 9 位的有效数字。轻子的质量在 1993 年以前的测量值为 1784MeV 左右。它与 τ 寿命的测量值一起作为讨论弱作用对三个轻子普适性的判断标志，曾经带来一些困难。1993 年，北京正负电子对撞机 (BEPC) 给出了新的精确值 $1776.9 \pm 0.4 \pm 0.3\text{MeV}$ ，比原来值减小了 7.2MeV，测量精度提高了 5~6 倍，使普适性的困难基本上得到解决，因此引起了国际上的重视。

中微子质量是极难精确测量的。尽管大量实验表明，它们的质量可能为零，但迄今并没有精确的证明。表中列出的是实验的上限。前苏联的一个实验组 80 年代初曾宣布，他们测量到 $17\text{eV} < m_{\nu_e} < 40\text{eV}$ ，在国际上引起过一些轰动。很多国家企图重复这个实验，均未能获得成功。后来发现在他们的数据分析过程中有一些问题，这一争论才告结束。

宇宙学家对于中微子质量问题极为重视。按照标准的宇宙学模型，宇宙是均匀而且各向同性的，它起源于一百多亿年前的一次大爆炸，至今仍然以极其巨大的速度膨胀着。它是永远地膨胀下去，还是膨胀到一定程度后开始收缩，或者说宇宙是开放的还是封闭的，完全取决于宇宙的物质密度。标准宇宙学模型预言，宇宙的临界密度为 $\rho_c = 5 \times 10^{-30}$ 克/厘米³。按照这个值估算，每个立方厘米平均只有半个质子。如果密度小于这个值，宇宙就是开放的。反之，如果大于这个值，就是封闭的。根

据大量事实，理论家们推算，我们的宇宙密度比 ρ_c 应该略小一些。它意味着我们的宇宙是一个开放的宇宙，它会永远的膨胀下去。然而，大量的观测表明，宇宙中可见的物质密度远远小于这个临界值。大约只有 ρ_c 的十分之一。这就是说，宇宙中一定存在着大量的看不见的物质，人们称为暗物质。寻找这些暗物质的合理解释，当前已经成为宇宙学家和粒子物理学家共同感兴趣的热门课题。初步分析表明，由于宇宙中重子物质密度很小，它们不会是暗物质的主要成分。而中微子却是一个重要的候选者。

理论家们估计，宇宙中中微子很多。它们的数目与光子差不多。根据著名的 2.7°K 宇宙背景辐射估算，每立方厘米的体积中大约有 400 个光子。由于光子质量是零，它们对宇宙密度的贡献完全可以忽略。如果中微子质量是零，那么它也不会有任何贡献。但如果它的质量不是零，结果就大不相同了。据估算，只要中微子有 30 到 40eV 的质量，它们对宇宙密度的贡献就和临界密度差不多。一些实验表明，不能排除中微子的质量是在这个范围内。因此中微子质量问题的研究是有很重要意义的。在标准模型中，实际上认为中微子是零质量的，它们只有左手分量，没有右手分量。

由于色禁闭，自由的夸克不可能存在，所以它的质量的确切含义是什么，是很难回答的问题。在标准模型中做为参数出现的质量值在各种反应过程的计算中使用，称为流夸克质量。表中所列的就是这些值。在研究强子谱学时，通过使理论与实验的质量谱拟合所确定的夸克质量不同于流夸克质量，称为组分夸克质量。其数值与模型有很大的依赖关系。实际上，组分夸克质量是考虑了强子内部的各种复杂的虚过程之后得到的等效质量。对于轻夸克，它与流夸克质量差别很大，例如 u、d 夸克的组分夸克质量大约在 300MeV 以上。而流夸克质量只有几个 MeV。而对于重夸克，两者的差别不是太大。

标准模型的规范群是 $SU(3)_c$ 、 $SU(2)_L$ 和 $U(1)_Y$ 三个群的乘积，每一个群对应一个耦合常数，用 g_s 、 g 和 g' 表示。在自发破缺以后，由 $SU(2)_L \otimes U(1)_Y$ 部分得到弱相互作用和电磁相互作用。由此可以导出电荷 e 与 g 和 g' 的关系是：

$$g \sin \theta_w = g' \cos \theta_w = e$$

其中 θ_w 为温伯格角。显然：

$$\tan \theta_w = g' / g$$

费米的四费米子耦合的弱相互作用理论曾给出一个普适的弱相互作用常数 G_F 它与 g 的关系为：

$$G_F / \sqrt{2} = g^2 / 8m^2W$$

其中 mW 为 W 粒子的质量。利用上式和早就由实验上已知的 e 和 G_F 值，可似把 mW 表示为

$$mW = \frac{37.3}{\sin \theta_w} \text{ GeV}$$

理论上， Z^0 粒子的质量 m_z 与 W 粒子的质量 m_w 有下列关系式：

$$m_z = \frac{m_w}{\cos \theta_w}$$

由上述分析可见，温伯格角 θ_w 起着至关重要的作用。由于 e 是一个相当精确地知道的量，只要从实验上测得 θ_w ，则 g 和 g' 即可求出， m_w 和 m_z 就可以预言出来。幸运的是 70 年代为检验中性流过程而进行的深度非弹性散射实验，可以给出 θ_w ，结果为 $\sin^2 \theta_w = 0.22$ 。利用这个值预言了 $m_w = 80\text{GeV}$ 和 $m_z = 90\text{GeV}$ 。1983 年这两个粒子在实验上发现时，给出的质量值在误差范围内与上述预言值一致，使标准模型得到了令人信服的证明。

在标准模型中希格斯场起着非常重要的作用。规范粒子和物质粒子的质量都是靠它产生的。在最小标准模型中，希格斯场部分引入了两个参数。一个是希格斯场的真空平均值 v ，另一个是希格斯场的自相互作用的耦合常数 λ 。破缺之后希格斯部分只剩下一个希格斯粒子 H 。它的质量 m_H 为 $m_H = \sqrt{2}\lambda v$ 。理论上可以推知 $v = 246\text{GeV}$ ，但由于 λ 不知道，所以 m_H 没有任何明确的理论预言，成为标准模型中最不确定的因素。如果 λ 很大，比如 $\lambda > 4$ ，则 $m_H > 1\text{TeV}$ 。这时它的宽度接近它的质量，很难再看作一个基本粒子。所以通常认为 m_H 不会太大。

标准模型中还有一些参数来源于弱电统一部分所要求的夸克之间的混合。我们在表 2.3 中所列的物质粒子的三代结构，称为质量本征态形式。在没有发生破缺的原始的标准模型中，它们并不是以这种形式出现的。为了保证规范不变性和弱作用的 V-A 流-流耦合形式以及 GIM 机制的要求，第一代夸克都以一个左手的二重态和二个右手的单态出现，而且，三代夸克的“下类”分量左手态 d_L 、 s_L 与 b_L 通过一个经 3×3 的矩阵混

合在一起。这种形式的态是规范群的本征态。它们构成各种流以及与希格斯场的汤川型的耦合。在自发破缺的过程中，这些态要重新组合，才能由汤川型耦合得到费米子的质量项。规范群的本征态转化成了质量本征态。两者靠一个混合矩阵联系起来。

这个混合矩阵其实并不是标准模型给出的一个新的结果，它早在1975年就由两位日本理论物理学家川石(J.Kobayashi)和增川(K.Maskawa)发现了。他们把卡比玻为解释奇异数改变的弱衰变几率压低而引入的d夸克与s夸克的混合推广到三代夸克，建立了一个 3×3 的混合矩阵，它正是标准模型中的混合矩阵。人们取他们两个人英文名字的字头命名这个矩阵，简称为KM矩阵。或者加上卡比玻的名字，称为CKM矩阵。

KM矩阵有9个矩阵元，它把u、c、t与d、s、b联系起来。最简单的写法，是把这9个矩阵元记为： V_{ud} 、 V_{us} ...等。它们也是相应的九个带电流的系数，因此它们的数值大小决定了相应的带电流的强弱，对各种反应过程的计算都是十分重要的。

在最普遍形式下，KM矩阵的每个矩阵元都是复数，9个矩阵元需18个实参数才能完全确定。但由于KM矩阵为么正矩阵，再考虑到一些相因子，独立的参数降为4个。它们可以有多种选取方法，称为不同的参数化方法。川石与增川类比于卡比玻角，取了三个实参数为三个混合角，另一个参数为绝对值等于1的复数，称为KM相角。

对KM矩阵的理论分析与实验测定可以采用任何一种参数化方法。由不同的实验过程确定参数数值，从而可以检验标准模型。也可以对每一个矩阵元单独进行测定，看看所得的结果是不是满足么正矩阵的条件。例如，由中子的衰变与 μ 子的衰变宽度相比较，可以定出： $V_{ud} = 0.9734 \pm 0.0017$ 。而由K介子的弱衰变定出： $V_{us} = 0.2205 \pm 0.0018$ 。由D介子可以定出： $V_{cd} = 0.215 \pm 0.0016$ 。其他的矩阵元也可以类似地估算出来。KM相角是至今未能给出明确结果的一个参数。这个相角的存在使KM矩阵成为一个复的矩阵，它导致标准模型在破缺以后的拉格朗日函数成为复的，理论家们认为，其结果会使弱相互作用过程出现CP对称性的破坏。这里的P即空间反射变换，C是正、反粒子互换的变换，也称电荷共轭变换。CP是这两种变换的联合作用。对于电磁相互作用，P对称性、C对称性及CP对称性都是成立的，而对于弱相互作用，P对称不成立即宇称不守恒。人们发现，C对称性也不成立，但CP对称性是可

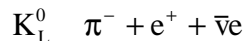
能成立的。让我们来看一个例子。

图 2.18 对比了弱衰变过程 $\mu^+ \rightarrow \pi^+ \nu_\mu$ 及其在 P 变换、C 变换和 CP 变换下的相应的过程。由于实验发现，自然界只存在左旋的零质量中微子和右旋的反中微子。所以在原来的衰变过程中(左上角) ν_μ 是左旋的。由于 μ^+ 自旋为 0， μ^+ 一定是右旋的。P 变换下(右上角) ν_μ 变成右旋，C变换下 $\bar{\nu}_\mu$ 变成左旋，这两者都是不允许的，所以这个衰变过程中P和C的对称性分别都不成立。只有右下角CP联合变换，得到 $\pi^- \mu^+ \bar{\nu}_\mu$ ，其中的 $\bar{\nu}_\mu$ 确实为右旋才是可以允许的。所以，这个反应过程，CP对称性是成立的。

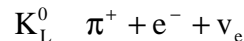
CP 破坏的现象是 1964 年克朗宁(J.W.Cronin)和费奇(V.J.Fitch)在中性 K 介子衰变过程的研究中发现的。中性 K 介子通过强作用产生时，以 K^0 和 \bar{K}^0 一对正反粒子出现。 K^0 的夸克结构为 $s\bar{d}$ ，而 \bar{K}^0 的结构为 $\bar{s}d$ 。但在弱作用中，它们组合成两个新的中性粒子出现，一个叫 K_L^0 ，另一个是 K_S^0 ；

K_L^0 主要通过 3π 衰变， K_S^0 主要衰变成 2π ，实验上测得两者寿命相差将近一百倍。用CP对称性可以对这个现象给一个很好的解释。因为，K介子和 π 介子的宇称都是负的，而在C变换下， K^0 变成 \bar{K}^0 ，所以，CP不变性要求 K_L^0 只能衰变成 3π ， K_S^0 只能衰变成 2π 。由于 3π 的质量和比 2π 质量和大得多，所以 K_L^0 比 K_S^0 衰变几率小很多，寿命长得多。但克朗宁和费奇的实验却发现每一千个 K_L^0 粒子，大约有3个衰变成二个 π 介子。这表明CP对称性在中性弱衰变过程中并不是严格成立的。

1967年，本纳特(S.Bennett)研究了 K_L^0 介子的半轻子衰变过程：



和与之相应的 CP 变换后的过程



如果没有CP破坏，则 K_L^0 通过这两个过程放出的 e^- 和 e^+ 粒子数 $N(e^-)$ 和 $N(e^+)$ ，应该相同。实验结果表明并非如此。用一个电荷不对称性的量 给出的结果为：

$$x = \frac{N(e^+) - N(e^-)}{N(e^+) + N(e^-)} = (0.32 \pm 0.012)\%$$

这说明 CP 破坏的比例约为千分之三。

CP 破坏为什么会发生？理论家们进行了多年的研究，建立了一些模型，但都没有能给出很好的解释。标准模型建立以后，川石和增川就曾指出，KM 相角是 CP 破坏的可能来源。在 KM 矩阵中，通常的参数化方法表明，KM 相角主要出现在与重夸克 b、t 有关的矩阵元中，如 V_{td} 和 V_{ub} 。利用标准模型对于中性 K 介子 CP 破坏的测量参数进行的计算发现，在 K 介子的衰变过程中，这些重夸克是通过复杂的高级费曼图起作用。例如图 2.19 所示的箱式图和图 2.20 所示的企鹅图。

计算结果表明，只要 KM 相角不为 0 或 π ，CP 破坏的参数一定不为 0。但理论的计算结果还有很大的不确定性。实验的精度也有待进一步提高。要得到关于标准模型能否完全解释中性 K 介子的 CP 破坏效应的明确回答，还需要做出更多的努力。

中性的 B^0 和 \bar{B}^0 介子在很多方面与 K^0 、 \bar{K}^0 介子类似。由于 B 介子比 K 介子重得多，衰变方式 K 介子多，而且它本身是由重夸克组成的，KM 相角与它的联系更为直接。这一切都使理论家相信，B 介子一定会有比 K 介子丰富得多的 CP 破坏现象。目前实验家们在不断的努力，但毕竟这些效应的量级太小，测量很困难。因此，迄今尚未能得到明确的结果。有人估计，如果 B 事例数能达到一亿，明确的结论是可以做出的。由于正在设计的一些产额比较大的 B 介子工厂，以检验 CP 破坏为其主要运行目标，所以，对 CP 破坏的理论研究正在引起人们更大的兴趣。

2.13 标准模型真的标准吗？

至此我们已经看到，在标准模型中有 19 个参数，它们是：耦合常数 3 个，物质粒子质量 9 个(中微子被认为质量为 0，所以省掉了三个质量参数)，希格斯场部分 2 个，KM 矩阵 4 个。还有一个与 QCD 真空结构有关的参数，我们没有介绍。尽管参数很多，但由于这个模型原则上可以处理除了引力作用之外的一切基本粒子过程，因此仍有着极强的预言能力。

自从 1983 年 W 和 Z⁰ 粒子相继发现，标准模型基本确立之后，从各方面检验标准模型引起了理论家与实验家的越来越大的兴趣。80 年代末，西欧中心的大型正负电子对撞机 LEP、美国 SLAC 的 SLC、费米实验室的质子与反质子超级对撞机 TEVATRON 等先后投入运行，它们的能量大于或远大于 Z⁰ 粒子的质量，为检验标准模型提供了极为有力的工具。特别是 LEP，使检验的精度提高到了惊人的水平，在一些方面甚至超过了理论所能达到的精确度。

理论家把精确检验标准模型看成了压倒一切的中心课题。他们很快把最低级微扰论(树图近似)所能计算的过程几乎全部计算过了，对于计算中遇到的不能严格求解的非微扰(或长距离行为)问题，建立了许多模型和发展了一些巧妙的计算技巧绘出了很好的处理。继而进行了更高级微扰(单圈图、甚至双圈图)的计算，使这种检验能够达到更高的精度。

观测的量	实验结果	理论预言	偏离程度
()			
$M_r(\text{GeV})$	91.1884 ± 0.0022	91.192	-1.64
$m_x(\text{GeV})$	2.4963 ± 0.0032	2.4987	-0.74
$\sigma_h^0(\text{nb})$	41.488 ± 0.078	41.462	0.33
$R_{hl} = \Gamma(h \rightarrow l\bar{l})$	20.788 ± 0.032	20.756	0.99
$R_{bh} = \Gamma(h \rightarrow b\bar{b})$	0.2219 ± 0.0017	0.2157	3.67
$R_{ch} = \Gamma(h \rightarrow c\bar{c})$	0.1540 ± 0.0074	0.1722	-2.46
A_{FB}^l	0.0172 ± 0.012	0.0167	0.41

A_{FB}^1	0.1418 ± 0.0975	0.1482	-0.85
A_τ	0.1390 ± 0.0089	0.1482	-1.03
A_e	0.0997 ± 0.0031	0.1039	-1.35
A_{FB}^b	0.0729 ± 0.0058	0.0744	-0.26
$\sin^2 \theta_w$	0.2325 ± 0.0013	0.2314	0.87
M_W	80.26 ± 0.16	80.40	-0.88

检验的结果怎么样呢？十几年来，世界范围的粒子物理学会议名目繁多、数不胜数，其中最大规模的有国际高能物理会和轻子光子会，它们交替地隔一年举行一次。每次会议都有一位著名的物理学家总结关于标准模型的检验状况。在他们的结语中你总可以听到这样一句话：“标准模型与实验数据惊人地符合”。下面，我们列出 1995 年 8 月在中国北京召开的国际轻子-光子会议上，报告人对 LEP 对撞机上得到的一些重要的量的测量结果与理论预言值所做的比较情况。

对这个表我们首先给出一些简单的说明。观测的量中， M_Z 为 Z^0 粒子的质量， Γ_Z 为 Z^0 粒子的总的衰变宽度， $R_{b\bar{b}}$ 、 $R_{c\bar{c}}$ 、 $R_{\tau\bar{\tau}}$ 、 $R_{\nu\bar{\nu}}$ 四个量分别是 Z^0 粒子衰变成强子、一对 b 夸克、一对 c 夸克以及一对轻子的宽度。 A_{FB} 和 A_b 、 A_e 等分别表示各种衰变过程中测得的粒子分布的一些不对称性大小， $\sin^2 \theta_w$ 为温伯格角 θ_w 的正弦值的平方。 M_W 为 W 粒子的质量。偏离程度等于测量结果减去理论预言值之后与测量误差之比值，通常以为单位。

从上述所列的实验结果看，除了几个不对称性测量之外，所有的量都达到了千分之五以下的精度。从实验值与理论值的偏离程度看，只有 $R_{b\bar{b}}$ 和 $R_{c\bar{c}}$ 超过了两个 σ ，因此理论与实验的符合程度之好是相当惊人的。

对于这两个量的偏差，实验家们并不认为有多么严重。他们认为不能排除统计涨落的可能性。但不少理论家对此极为重视，他们认为这是迄今为止标准模型精确检验中唯一重要的偏差。它也许是一个苗头，表现出某种超出标准模型的新的物理规律的可能性，因此可以大做文章。于是，几个理论组分别计算了超对称粒子、人工色(Technicolor)粒子等对 Z

粒子衰变到 $b\bar{b}$ 及 $c\bar{c}$ 等过程的影响，企图得到对这些可能的新的物理模型的支持。但究竟孰是孰非，还要看理论与实验的继续发展。

有趣的是，时间刚刚过去一年，情况就发生了明显的变化。在 1996 年 8 月，于波兰华沙召开的新的一届国际高能物理大会上，实验家们报

告了他们测得的最新结果。其中，

$$R_b = 0.2178 \pm 0.0011$$

$$R_c = 0.1715 \pm 0.0056$$

这两个值与理论预言值相当接近，已经不存在大的偏差。理论家们又一次感到失望了。

标准模型受到普遍欢迎的重要原因在于它的可计算性。比如依据实验事实，标准模型自然地纳入了三代结构。虽然模型本身

回答不了为什么会出现现代的结构，为什么会有三代，还有没有可能出现更多的代等问题，但是理论家们根据 LEP 提供的精确实

验数据，从丢失的宽度(衰变成中微子因而测量不到)对 Z^0 粒

子总宽度的贡献，计算出质量不超过 Z^0 粒子质量一半的轻的(或零质量的)中微子的种类数 N_{ν_0} 结果为

$$N_{\nu} = 2.980 \pm 0.027$$

它恰好与三代中微子的事实完全相符。

另一个例子是关于顶夸克(t)的寻找。标准模型的建立意味着 t 夸克一定要存在。如果说，一开始这还只是理论家的一种信念，那么没有经过多久，许多实验迹象表明它的存在是不可避免的。但是顶夸克质量有多大，并没有任何直接的理论约束。从 1977 年底 b 夸克被发现，证实了 b 夸克存在的那个时候起，寻找 t 夸克就成了一种强大的动力，激励着实验家和理论家。十几年来，实验家们争得了巨大的财政支持，建造了能量越来越高的加速器和越来越精密的探测器。理论家们依据不断提高精度的各种实验数据，进行了精心的计算，不断更新结果，给出越来越准确的间接理论估计。由于希格斯粒子的质量一点线索也没有， m_t 的估算准确度受到了极大的影响。尽管如此，理论家们经过对许多过程的复杂的高阶修正的计算，使 m_t 的理论值的不确定程度越来越小。到 1994 年，预言值为(m_H 取 300GeV)：

$$m_t = 172_{-18}^{+16} \pm 19 \text{ GeV}$$

如前所述，1995 年 3 月 2 日，费米实验室的 CDF 组和 D0 组同时在新闻发布会上宣布的 t 夸克质量的可能值分别为： $m_t = 176 \pm 13 \text{ GeV}$ 和 $m_t = 199 \pm 30 \text{ GeV}$ 。两者相比，符合的程度是十分惊人的。1996 年这两个实验组测得的数值更加接近，误差也进一步减少，它们给出的最新的实验平均值为

$$m_t = 175 \pm 6 \text{ GeV}$$

与理论预言值的符合程度更加惊人。

时至今日，再想用实验打倒这个理论，看来已是绝不可能的。尽管人们对它的精确检验仍然热情不减，仍期待着积累更多的实验数据使统计量进一步改善，结果的精度进一步提高，但出发点已经从前几年的证实标准模型，转向寻找可能的偏差。标准模型解释不了的那些偏差，无疑将成为超出标准模型的新物理学的突破口。遗憾的是：到现在为止，除了前面提到的 R_{bh} 之外，人们没有发现任何现象明显偏离标准模型。尽管还有一些实验，由于其精度不够，不能确切判断是否存在偏差，但普遍认为它们留给超出标准模型的新物理学的余地是很小的。

尽管如此，寻找超出标准模型的新物理仍是今天粒子物理学家的一个奋斗目标。究其原因，主要是标准模型本身在很多方面不能令人满意，存在着许多解答不了的疑难问题。说是标准模型，其实很不“标准”。

首先，这个模型中的基本组分至少有 61 个，其中物质粒子及其反粒子 48 个，规范粒子 12 个，希格斯粒子 1 个。理论本身并不能排除第四代甚至更多代的粒子存在的可能性，只要它们相应的中微子的质量超过 Z^0 粒子质量的一半即可。如此众多的粒子都是基本的吗？这个老问题又一次提了出来。

其次是这个模型参数太多。如果中微子质量不为零，至少要有 22 个参数。如果中微子质量为零，参数可减少 3 个。理论家们不相信，有这么多参数的理论会是一个真正的基本理论。标准模型并没有真正把强、弱和电磁相互作用统一起来，这是它的又一个缺陷。它的对称群是 $SU(3)_C$ 、 $SU(2)_L$ 和 $U(1)_Y$ 三个群之积，因此才会有三个独立的耦合常数 g_s 、 g 和 g' 。或许有人会说，由于电弱统一， g 和 g' 与电荷 e 有下列关系式： $g \sin \theta_w = g' \cos \theta_w$ ，因此 g 和 g' 并不独立。其实，这个关系只不过表明，可以用电荷 e 和温伯格角 θ_w 二个参数代替 g 和 g' 。如果从理论上可以把 θ_w 计算出来，参数确实可以减少。但可惜的是，标准模型做不到这一点。它的 θ_w 是由实验测定的。此外，基本粒子除了参加强作用、弱作用和电磁作用外，还要参加引力相互作用。虽然，引力的相互作用强度极弱，比弱作用要弱几十个量级，但它在构造物质世界方面起着极为重要的作用。特别是在宇宙学研究中，引力是最主要的角色。标准模型没有包括这种相互作用，自然也是它的一个不足之处。

标准模型有三代物质粒子，假如它们都没有质量，代和代之间存在严格的对称性，彼此好象是复制品。为什么会有这种奇妙的对称性？为什么只有三代？标准模型解释不了。实际上这些粒子都是有质量的(除中微子外)，质量的大小以及彼此之间的关系也不能由标准模型给出来。标准模型解释不了为什么三代粒子中第三代比第二代重，第二代比第一代重，也解释不了为什么同一代中夸克比轻子重，为什么第一代 u 和 d 夸克质量相差不大，第二代中 C 夸克比 S 夸克重 10 倍，而第三代中 t 夸克比 b 夸克重近 40 倍。

至于希格斯部分，它对温伯格-萨拉姆弱电统一理论起了极其重要的作用。严格的规范对称性要求规范粒子都是零质量的，希格斯机制导致对称性自发破缺。破缺之后，规范粒子获得了质量，同时剩下了一个有质量的希格斯粒子。此外，在弱电统一理论中对每一种费米子都引入了一项费米场与希格斯场的汤川耦合项，破缺之后它使费米子获得所需要的质量。所有这些粒子的质量都正比于希格斯场的非零真空期望值 v 。所以 v 成了一个质量单位，或说给出了一个质量标度。理论上要求 V 必须等于 246GeV，才能解释实验上得到的 W 和 Z^0 粒子的质量。但是，理论家们发现在标准模型中 V 是一个极不稳定的量，虚粒子的交换可以使它发生极大的变化，为了保持住 246GeV，原则上必须使理论的一些原始参量有着极高的精度，需要做不可想象的精细调整。标准模型中的这个困难，常称为“自然性”问题。

由于希格斯机制对温伯格—萨拉姆模型起着如此重要的作用，寻找希格斯粒子就成为粒子物理实验家的迫切的任务，也引起了理论家们的极大的兴趣。但是困难在于理论家对于希格斯粒子的质量究竟有多大给不出任何明确的回答。只能对质量的上限和下限给出一些间接的限制。这些限制与 t 夸克质量有着密不可分的联系。一些理论家把关于 t 夸克质量的最新实验结果做为输入，发现如果要使自发破缺希格斯场等效势的最小值是稳定的，则标准模型的希格斯粒子质量应满足如下的约束。若标准模型的破缺发生在很高的能量标度(如 10^{16} GeV 以上)则希格斯粒子质量下限为 135GeV，若这个破缺发生在低得多的能量标度(如 10^3 GeV)则希格斯粒子质量下限降低为 90 ~ 100GeV。超对称模型可以使希格斯粒子质量下限更低。实验上最新的 LEP 数据表明，希格斯粒子质量不小于 64GeV。

希格斯粒子质量的上限，理论上同样也说不清楚。通常认为，如果它比 1TeV 还高，会给理论带来一系列严重的问题。许多反应过程的几率守恒这一根本原理将会受到威胁(通常称为么正性条件将被破坏)。而且这时希格斯粒子的宽度差不多等于它的质量，是否可以再把它看成基本粒子或者共振态将成为问题。于是人们在讨论希格斯粒子参加的各种反应过程时，都把它的质量上限取为 1TeV。最近人们利用 LEP 和费米实验室的实验结果综合分析，认为希格斯粒子质量可能不大于 510GeV。

在 t 夸克已经被发现的今天，希格斯粒子成为弱电统一理论中唯一尚未找到的重要成员。探讨可能的各种途径寻找希格斯粒子，无疑是摆在粒子物理理论家和实验家面前的首要任务。如果希格斯粒子始终找不到，希格斯机制就会垮台，但对标准模型本身不会带来严重的影响。其实，从温伯格-萨拉姆模型建立之初，人们对希格斯机制就存在着争议。它的引入毕竟缺少严格的理论基础，含有过多的人为因素。不少人认为希格斯机制可能只是一种等效理论形式。他们努力寻求更深刻的基础理论，用动力学的破缺机制来代替希格斯机制。

对标准模型全面地、深入细致地分析，使理论家们确信，尽管它有极强的预言能力，极好地符合了所有的实验，但它绝不是一个终极的理论。超出标准模型的新物理一定存在，一个比标准模型更好的理论一定能够建立起来。理论家们一直在朝着实现这个目标努力，不断地尝试各种超出标准模型的可能性。

2.14 统一之梦

早在标准模型建立不久，一些理论家就开始考虑能否把统一的思想再向前推进一步。弱电相互作用与强相互作用会不会是从一种统一的相互作用经自发破缺而分道扬镳的呢？夸克和轻子会不会是由原始的一个更大的家族分出来的呢？如果真是这样，自然界的规律将会是更简单、更美妙。它只有一种作用力而不是强、弱、电三种。它只有一种物质粒子而不是夸克和轻子两大类。这种统一实际上是理论物理学家多年来追求的梦想。标准模型的建立使这种梦更接近于实现。

从标准模型可以得到关于可能的更大的统一理论的许多重要启示。首先，它必须是一个规范理论。这是因为 QCD 和弱电统一模型都是规范理论。其次，规范群必须是一个更大的群 G ， $SU(3)$ 、 $SU(2)$ 和 $U(1)$ 三个群都是它的子群。在这个新的统一模型中，耦合常数只有一个而不再是三个。

1973 年，美国的马里兰大学的派提(J.Pati)和萨拉姆合作进行了最早的尝试。他们认为轻子是带第四种色的夸克，建议了一种 $SU(4) \otimes SU(4)_L$ 模型。在这种新模型中轻子和夸克属于同一个家族，所以夸克可以变成轻子，轻子也可以变成夸克。它的一个最严重的后果就是质子会衰变。因为质子里面有三个夸克，只要有一个夸克变成轻子，质子就不复存在。尽管派提和萨拉姆很不喜欢这一点，但他们无法回避。

第二年，格拉肖与哈佛大学的一名研究生乔治(H.Georgi)一起也讨论了这个问题，他们想找一个单一的更大的规范群。乔治尝试了几种可能性，发现 $SU(5)$ 群是满足统一理论要求的最小的群。它的基础表示是 5 维，接着有一个 10 维表示。三个带色的 d 夸克和一个正电子，一个反中微子，正好与 5 维表示的要求相符。其余的粒子安排在 10 维表示，量子数完全对。在这个群中传递相互作用的媒介粒子，包括光子、 W^\pm 、 Z^0 和胶子都是一个家族的平等成员。同时还引入了十二个新成员，称为 X 粒子。乔治与格拉肖对这种安排都很满意。但在这个模型中，X 粒子传递的相互作用使夸克与轻子相互转换，因此也必然导致质子衰变。格拉肖查到了当时对质子寿命的估计大于 10^{27} 年。据此，他估算 X 粒子非常重，比已发现的最重的粒子还要重一千万亿倍以上。现代任何加速器也产生

不了这样重的粒子，只有象太阳系那么大的加速器才会有这样的能力。于是他们不再担心，于 1974 年发表了他们的文章。

读过这篇文章的人都认为这是一篇经过精心设计的杰作，但一接触到质子衰变问题，人们的热情马上就冷了下来。除此之外，在这篇文章中，还有些问题没有能解决。其一是尽管预言了质子会衰变，但不会定量计算。其二是这个模型把强、弱、电三种相互作用统一成为一种单一的相互作用，但解释不了为什么破缺之后三种相互作用强度会如此不同。

基本粒子之间的四种相互作用中，若以强作用的强度为 1~10，则电磁作用为 1/137，弱作用约为十万分之一 (10^{-5})，而引力作用大概是 10^{-39} 。在标准模型中，与 $SU(3)_C$ 、 $SU(2)_L$ 和 $U(1)_Y$ 三个群对应的描写耦合强弱的参数分别为 g_s 、 g 和 g' 。但实际描写强弱电三种相互作用过程强弱的耦合常数分别为 α_s 、 α_w 和 α_e ，它们的定义为： $\alpha_s = g_s^2 / 4\pi$ ， $\alpha_w = g^2 / 4\pi$ 和 $\alpha_e = g'^2 \cos^2 \theta_w / 4\pi = e^2 / 4\pi$ 。最后一个就是精细结构常数。

由于 $SU(3)_C$ 、 $SU(2)_L$ 和 $U(1)_Y$ 都是定域规范群，理论家们证明， α_s 、 α_w 和 α_e 都是随相互作用过程的能量大小变化的跑动耦合常数。强作用的耦合常数 α_s 在能量很低或夸克之间的距离拉开比较大时取很大的值，引起夸克禁闭。当能量很高，或夸克之间距离很小时， α_s 很小甚至接近于零，称为渐近自由。 α_w 的变化趋势与 α_s 很像，但变化范围很小。 α_e 是与 $U(1)_Y$ 群相联系的，这个群是一个可交换群，与 $SU(2)$ 或 $SU(3)$ 很不相同。 α_e 的跑动特点与 α_s 和 α_w 完全相反。能量越低、距离越大时， α_e 越小，不会造成禁闭；能量越高、距离越近时， α_e 越大，因此没有渐近自由。

通常把这三个跑动耦合常数随能量的变化关系画成三条曲线，称为耦合常数的演化曲线，研究它们是否交于一点以及交点相应的能量是什么。如果能交于一点，则意味着三个耦合常数可以统一成一个。由交点的能量可以定出实现统一所需的能量。不同模型给出的这三条演化曲线形状不同。乔治·温伯格和奎恩(H. Quinn)根据当时的测量值画出的 $SU(5)$ 模型中的三条曲线果然交于一点，相应的能量为 $M_G = 10^{15}\text{GeV}$ (相当于距离为 10^{-29} 厘米)。它表明能量达到 M_G 时 $SU(5)$ 群可以使强、弱、电三种相互作用统一为一种相互作用，称为大统一。这时耦合常数为 α_G 。图 2.21

示意地画出了当年 SU(5) 大统一模型中耦合常数的演化曲线。

大统一能标 10^{15}GeV 是一个极高的能量。迄今为止，人类制造的加速器的最高能量约为 10^3GeV ，与之相比简直微不足道。如果这个能量靠热运动产生，温度应为 10^{27}K 。只有宇宙大爆炸的最初 10^{-36} 秒的瞬间才能达到这么高的温度。这就是说，只有在如此特定的环境下，电、弱和强三种相互作用的耦合常数才完全相同，大统一才能实现。由此推算，在最小 SU(5) 大统一理论中，X 粒子质量 M_X 约为 $10^{14} \sim 10^{15}\text{GeV}$ ，相应的质子寿命约为 $10^{27} \sim 10^{31}$ 年。如果实验证实质子的寿命超过这个范围，则至少最小 SU(5) 模型将会被排除。

大统一模型最成功的预言之一是可以从理论上计算出温伯格角。在乔治和格拉肖的文章发表不久，有人就算了这个值，但结果与实验不符。乔治与温伯格发现这些计算中没有考虑耦合常数的跑动。把这个因素考虑进来重新计算的结果， $\sin^2 \theta_w$ 为 0.21 左右，与实验结果很接近。

SU(5) 模型的其它成功之处还有：导出 d 夸克的电荷为电子电荷的 $1/3$ ，预言了 b 夸克质量与轻子质量之比为 3，没有树图的味改变中性流，不存在曾使理论家们大伤脑筋的反常等等。但标准模型中未能解释的代的结构问题，希格斯机制问题，以及由于大的标度 M_X 的存在使希格斯场的重整化带来的精细调节问题(即前曾指出的自然性或也称规范等级问题)等等都仍然未能解决。而且还带来了一些新问题。比如按照 SU(5) 模型在 200GeV 以上直到大统一能标 10^{15}GeV ，不存在任何新粒子，很像是一个一望无际的大“沙漠”。如果情况果真如此，将是十分扫兴的。

幸运的是，目前已经得到了一些实验结果对 SU(5) 大统一模型十分不利。我们先来介绍质子衰变实验。由于理论预言质子的寿命极长，所以似乎这类实验不会引起人们的兴趣。开始萨拉姆曾试图说服一些实验家使他们发生兴趣，但未能成功。在 SU(5) 模型取得了一些很好的结果后，实验家们才开始重视起来。实验家们关心的一个问题是质子在衰变之后，变成了什么。按照 SU(5) 大统一理论，衰变是由 X 粒子引起的。一种可能性如图 2.22 所示，衰变末态是 $\bar{0}$ 和 e^+ 。还有一些与其类似的过程。

另一个问题是如何实现这一实验。按照乔治和温伯格的计算，质子的寿命为 10^{31} 年。如果我们能够有 10^{33} 个质子则每三到四天就可以有一个质子衰变。 10^{33} 个质子并不难得到，大概有一间大房子那么大体积的物质就够了。

为了防止宇宙线带来无法区分的背景，实验必须在地下很深的地方进行。最早报告观察到一个衰变事例的，是由印度和日本物理学家合作，在印度库拉金矿废矿井中，用 150 吨铁安放在地下 12000 英尺处进行的实验。欧洲、美国也都在进行类似的实验，但都没有得到进一步的肯定结果。1983 年，美国的 IMB 合作组，利用 6000 吨水，放在 600 米深的地下，也没有观测到任何事例。他们给出的质子寿命实验的下限为 10^{32} 年，比 SU(5) 模型给出的 10^{31} 年大。这使最小 SU(5) 模型遇到了很大的麻烦。

SU(5) 大统一模型遇到的另一些困难来自于 LEP 实验给出的精确度越来越高的结果。例如， $\sin^2 \theta_w$ 的最新结果是 0.2322 ± 0.0007 ，它比由 SU(5) 模型定出的 0.21 左右大得多。更严重的是依据最新实验数据计算出的耦合常数的演化曲线，对于 SU(5) 大统一模型不再交于一个点。这个结果实际上已经能够把 SU(5) 大统一模型排除掉了。

SU(5) 大统一模型只是实现大统一的一种最简单、最经济的可能性，类似的可能性还有许多种，通常采用了更大的群，有更多的参数。其实，SU(5) 模型虽然使耦合常数减少了两个，但同时引进了更多的希格斯粒子，总的参数个数并没有减少。其它的大统一模型比 SU(5) 模型参数更多。理论家发现，当前的实验已经把这一类简单大统一模型基本上排除了。得到实验的新的支持因而成为新的热点的是在更高的对称性基础上建立的统一理论——超对称性大统一理论。

超对称理论是七十年代初由前苏联、美国和欧洲几组理论物理学家从不同的出发点独立地建立起来的。我们前面曾经指出，基本粒子按其自旋为半奇数还是整数分成费米子和玻色子两类，它们表现出两种截然不同的“性格”。玻色子比较随和，容易聚在一起，相同的状态可以容纳的玻色子数不受限制。费米子个性极强，绝对憎恶自己的同类占据相同的状态。泡利最早发现了这一点，称为不相容原理。所谓的超对称性是指在费米子与玻色子之间存在着一种新的对称性，使它们成为超空间中同一个家庭的成员。很象同位旋空间中同位旋向上代表质子，向下代表中子，你也可以设想超空间中有一个矢量，它指向上代表费米子，向下代表玻色子。在超对称变换下，玻色子会变成费米子，费米子也会变成玻色子。

最简单的超对称模型，就是在原有的一些模型的基础之上添加超对称性。它是一个整体的规范变换。标准模型经过这样的改造(称为超化)

得到超对称标准模型。大统一模型，比如 SU(5)，超化得到超对称大统一或超对称 SU(5) 模型。在这些模型中，玻色子和费米子成了同一个家族的成员，超对称变换也确实把玻色子变成费米子。把费米子变成玻色子，但它不是把原来模型中的玻色子和费米子互相变换。原模型中的费米子是夸克和轻子等实物粒子，而玻色子是传递相互作用的规范粒子和希格斯粒子，这两种粒子之间不存在超对称性，这是实验事实。超对称性的引入实际上使原来的粒子都多出了一个自己的超对称伙伴，粒子的数目加了一倍。这些伙伴粒子与原来的粒子相比，除了自旋和质量不同之外其它量子数完全相同，参加相似的相互作用过程。

引入超对称性的一个最大的好处是，由于费米子和玻色子一对伙伴对于闭合的圈图给出大小完全相同的贡献，但符号恰巧相反，因此严格抵消，使原来难以解决的发散问题很容易得到解决。例如，由标量希格斯粒子的重整化带来的自然性问题，只有在超对称理论下才能克服。当然，这个结论是在严格超对称性下成立的，那时一对伙伴粒子均具有相等质量。实际上，超对称性只是一种近似的或破缺的对称性。我们观察到的只是伙伴中之一个，另一个质量要重得多，至少在目前还观测不到，但由不存在自然性问题或发散基本相消可以得到一对伙伴的两个粒子质量差不应很大。通常认为它们的差应小于 1TeV。

此外，为了保证重子数和轻子数守恒，超对称性理论中还引入了一个叫做 R 宇称守恒的分立对称性。非超对称的普通粒子 R 宇称为+1，它们的超对称伙伴 R 宇称为-1。R 宇称守恒导致两个重要的结论，即超对称伙伴粒子必须成对产生和最轻的超对称粒子是稳定的粒子。通常认为，这个最轻的超对称粒子是中性粒子，因而很难探测。

近年来最受关注的是最小超对称标准模型(MSSM)，它是标准模型的最小扩充。标准模型的每个粒子在这种扩充下，都只增加一个超对称伙伴粒子。如夸克对应的有标夸克，轻子对应的是标轻子，规范粒子对应的伴侣叫规范微子，例如光微子、胶微子等。变化比较大的是希格斯粒子。超对称性限制了粒子之间的耦合形式，使希格斯场增加为二个复的二重态。二个复二重态有八个实分量，被规范粒子吃掉三个，还剩 5 个做为希格斯粒子保留了下来。正因为如此，他们的超对称伙伴也是五个，称为希格斯微子。

标夸克和标轻子自旋都是零，所以称为“标”粒子。“微子”都是

费米子,例如 r 、 g 和 W 的费米子自旋都是 $1/2$ 。超粒子的质量究竟有多大? 它们之中最轻的粒子能否在现有的加速器或者下一代加速器(如 LHC 等)找到? 这是当前理论家和实验家最关心问题之一。

比超对称标准模型更进一步的理论是超对称大统一理论,特别是超对称 SU(5) 模型。由于这个模型与简单的 SU(5) 大统一模型相比增加了普通粒子的超对称伙伴粒子,它们的质量可能在 1TeV 以下,使原来 SU(5) 模型中的大“沙漠”有了一些生气。

1993 年,一个新的重要进展是,利用 LEP 的精确实验数据,在简单的 SU(5) 大统一模型下重新计算三条耦合常数演化曲线,发现它们不再交于一点(如图 2.23 所示,其中 α_1 正比于 α_e 而 α_2 和 α_3 分别为弱和强作用耦合常数)。这使简单大统一理论受到了致命打击。纳入了最小个数的超对称粒子之后,如果这些超对称粒子比 Z^0 粒子重不了多少,比如它们的质量不超过 1TeV,则三条演化曲线的确可以交于一点。如图 2.24 所示。这个结果被人们看成是对超对称大统一模型的最有力的支持。

由耦合常数演化曲线的交点定出的超对称大统一能标及在夸克与轻子之间传递相互作用的粒子 X 的质量为 $M_X \sim 10^{16} \text{GeV}$ 。由此求得 $\sin^2 \theta_w = 0.23$, 极好地符合目前的实验值,比简单大统一有很大的改进。此外质子衰变几率也有所压低,大于目前实验下限 10^{33} 年。

无论是简单大统一还是前述的超对称大统一都没有包括引力。强、弱、电和引力是基本粒子同时参加的四种相互作用,如果只能把三种相互作用统一,当然美中不足。实际上,理论物理学家们把建立一个包括所有的相互作用的统一理论做为最终目标,不是今天才产生的想法。早在十九世纪中叶,对电磁感应现象进行了实验研究的法拉第(M. Faraday)就曾说过:“我很久以来一直有个看法,简直成了信念,我相信这是很多自然知识爱好者所共有的,即在多种形式之下的物质的作用力能被表明出自同一根源。”其后,麦克斯韦成了法拉第思想的热烈追随者。他朝着统一的目标迈出了第一步,把磁力与电力用单一的电磁场理论统一了起来。在本世纪二十年代,广义相对论建立不久,爱因斯坦、外尔、海森堡和克莱因等都进行了许多尝试,企图把的电磁场与引力场统一起来,建立一个统一场论。

1915 年到 1916 年期间,爱因斯坦完成了他的最伟大的理论工作——广义相对论。他给出了一组引力场方程,把空间、时间、物质、能量、

引力与几何学极巧妙地结合成一个整体。理论的优美形式强烈地把爱因斯坦吸引住了。此外，对经典的决定论的坚定信念又使他成了几率解释的量子理论的顽固反对者。从那以后，他脱离了理论物理发展的主流，直到逝世为止的将近 40 年中把全部心血倾注在统一场论的研究中，企图把经典的电磁场与引力场统一到一起。爱因斯坦认为：“只要没有统一的场，对我来说就没有物理学。”他顽强地努力建立一个没有内在不完备性的物理图像。非常可惜的是，他没有成功。今天看来，正如盖尔曼指出的，他的失败的根本原因之一是他对于近代的量子力学所持的否定态度，尽管对早期的量子理论的发展他曾经做出过巨大的贡献。另一个原因是受到了时代的局限。直到 60 年代，理论与实验两个方面才为统一理论的建立奠定了初步的基础。今天，新的统一理论之所以能取得很大的进展，实际上是几代物理学家智慧和辛勤劳动的结晶。其中爱因斯坦的理论及其深刻的思想，包括他未能成真的梦，无疑做出了最伟大的贡献。

标准模型是量子理论，要把引力统一进来也必须把引力场量子化。但引力场方程做为一种规范场方程，是极为特殊的。标准模型中的三个规范场的耦合常数都没有量纲，是一些普通的数。而引力耦合常数 G_N 有量纲，正比于质量的负二次方，即 $G_N = 6.707 \times 10^{-39} \text{GeV}^{-2}$ ，它给引力场的量子化带来了无法克服的发散困难。迄今为止，建立量子引力理论的尝试都未能获得成功。

70 年代超对称性理论的建立给解决引力量子化问题带来了新的希望。当人们把整体超对称性扩充成局域超对称性时，很自然地把量子引力纳入了进来。理论中出现了引力的媒介粒子—引力子。它的自旋为 2，没有质量，满足爱因斯坦方程。引力子是玻色子，它的超对称伙伴粒子叫引力微子，是自旋 3/2 的费米子。由于这两种粒子对发散的贡献符号相反，提供了相消的机制，使人们对克服量子引力的发散困难看到了新的希望。这种新的引力理论称为超引力理论。

超引力与大统一理论的结合，即局域化超对称大统一理论称为超引力大统一理论。它在一个更高的能标—普朗克质量 M_G 之上，把四种相互作用力统一了起来。 M_G 实际上就是引力常数倒数的平方根，它大约为 10_{19}GeV 。在这样高的能量下引力效应变得很强，与强、弱、电相互作用不再可区分。

普朗克质量相应的温度为 10^{32}K ，这个温度只有在大爆炸产生宇宙的最初 10^{-44} 秒的极短瞬间内才能达到。只有在那时，才能真正的实现四种相互作用力的统一。过了 10^{-44} 秒，引力首先分离了出来，到 10^{-36} 秒大统一也破缺了，强作用与弱电统一的相互作用分离开来，再到 10^{-10} 秒时弱电统一的 $SU(2) \otimes U(1)_Y$ 对称性也破缺了，弱作用与电磁作用分家，得到我们在普通的低能下见到的情况。

从定性上看，这种统一的图像似乎很诱人，四种相互作用可以用一种单一的相互作用来代替，不管它们的实现所要求的能量有多么高，存活的时间多么短暂，实现这个目标似乎已经不再是一个梦想。然而这毕竟只是一种定性的图像，回到我们的今天的、低能的定量的理论，情况就不那么简单了。超引力大统一理论实际上相当复杂。理论的自由参数不但没有减少，实际上反倒增多了。理论预言了大量的超对称粒子，它们的质量究竟有多大，理论家只能给出一个大致的估计。下一世纪初预计将会有能量为 500GeV 左右的 e^+e^- 直线对撞机 NLC 投入建造。西欧中心(CERN)将建成 $7\text{TeV} \times 7\text{TeV}$ 的质子-质子对撞机 LHC。人们普遍认为，如果在这些机器上仍然找不到这些超对称粒子，则超对称性的理论将会受到严重的威胁。

普通超引力大统一理论除了参数仍比较多以外，还存在一些解决不了的问题。比如，为什么存在三代仍无法理解，自发破缺的机制非常复杂，引力的发散也并没有完全克服，在质子寿命的计算中也还存在一些有可能与实验冲突的大的贡献的项等等。于是理论家们积极地寻找更好的理论模型，其中基于超弦理论发展起来的低能超引力大统一模型目前吸引了许多理论家的兴趣。

我们前面曾提到过，夸克模型建立不久，为解释禁闭，南部等建立了强子的弦模型，把介子看成是夸克与反夸克由一根弦拉在了一起。它能解释强子谱和一些反应过程，得到了许多好的结果，因而一直为人们所重视。70 年代初，一些理论家们把这种思想进一步推广，形成了对物理世界的一种新的基本观点，认为时空的最基本结构可能不是点状的，而是尺度极小的弦，比如说它只有普朗克长度 10^{-33}cm 这么小的尺度。在这样小的尺度下，时空与我们所生活的普通四维时空(三维空间加一维时间)完全不同，它有很高的维数，比如 26 维。多余的那些维数，我们观测不到，它们都蜷缩在上述的极微小的尺度里了。这就好像一条水龙带，

本来是三维的物体，远远看去只能看到一维的线。分辨不出它还有一个横向的面积。在这种弦理论中，各种粒子被看做是弦的各种振动模式。正如一根琴弦，可以有各种振动方式，从而发出了不同的音调。弦的振动模式可以有无穷多种，因此原则上粒子的种类也会有无穷多。我们所需要的粒子只是最低的几种振动模式而已。

在最初发展的弦理论中，同时解释费米子与玻色子的存在曾是很困难的问题。70年代末、80年代初，施瓦兹(J.Schwarz)、格林(M.Green)和威藤(E.Witten)等几位理论家把弦理论与超引力巧妙地结合在一起，构造了一种新的理论，称为超弦理论。他们不仅使时空的维数降低到了十维，而且找到了使多余的六维蜷缩起来的所谓自发紧致化的技巧。在这个理论中，四种相互作用自然地起源于唯一的一类弦之间的相互作用。强、弱和电磁作用大统一的能标可以提高到 10^{18}GeV ，质子衰变问题可以安全地解决。同时容纳多代的粒子在这个理论框架内是不成问题的，只是还不能限定一定是三代。超弦理论是一个原则上不存在发散困难的理论，长期以来困扰着理论物理学家的重整化，在这个理论中可能完全不需要。总之，这个理论几乎是能自治地解决量子引力并同时又足够包容所有其它相互作用的非常美妙的理论。

从80年代中期开始，在理论物理学界掀起了世界范围的热潮。超弦理论被看成可以解决所有物理学理论问题的一种可能的终极理论，是继量子力学和相对论之后，20世纪理论物理学达到的又一顶峰。它吸引了大量优秀人才。到90年代初，这个热潮逐渐平静了下来。人们认识到事情远非如此简单，超弦理论的发展很快遇到了一些新的难题。人们发现从一个超弦理论可以有无穷多种可能性通向描述真实世界的理论，无法判断出哪种是正确的。微扰真空态的定义也是一个难以解决的困难问题。此外，在超弦理论的破缺中起重要作用的是非微扰的贡献，理论家对非微扰的处理办法极少。此外，也有些问题涉及到当前数学工具的不足。正因为如此，迄今为止还没有能建立起任何一种可供实用的超弦模型。

近来，人们对四维超费米弦理论重新表现出了极大的兴趣。由于它不存在多余的时空维数，不需要蜷缩和紧致化等。更重要的是利用近几年关于超共形代数的研究成果，对于可能的超对称大统一模型给出了许多严格的限制，使参数可以大大减少。人们由此出发建立了一些预言能

力很强的模型，称为超弦诱导模型。例如翻转 $SU(5)$ 超引力模型， $SU(5) \otimes U(1)$ 超引力大统一模型等引起了人们很大的兴趣。

2.15 机遇和挑战

如果以 1895 年伦琴发现 X 射线做为对物质深层结构进行探索的开端，至今刚好过去 100 年。100 年前经典物理学遇到的严重挑战造成了前所未有的机遇，给物理学乃至整个自然科学带来了空前的繁荣。在这 100 年中在量子力学和相对论两大基石之上，近代物理学迅猛地发展起来，物质结构的研究无疑取得了非常辉煌的成就。而粒子物理学正是在这一进程中发生与发展并且成为最活跃、最富成果的研究领域。如今我们又进入了世纪之交，90 年代还有最后的 5 年。回顾粒子物理迄今的发展历史，展望新的世纪，我们发现理论物理学家与实验物理学家一道面临着更多的严重挑战。

首先，当前粒子物理理论的状况显得非常奇怪。一方面，经过几代理论物理学家半个多世纪的努力所建立的标准模型，无疑是可以与法拉第、麦克斯韦的电磁理论相比的划时代的辉煌成就，是理论物理学家对人类科学发展做出的又一伟大贡献。除了希格斯粒子没有找到以外，标准模型给出的所有理论预言都与精确的实验数据符合得非常之好。它有很强的预言能力，基本粒子的一切强、弱和电磁过程都可以用它定量的计算。实验检验的不仅是那些最低级的所谓树图过程，而且还包括单圈，甚至一些双圈的高级过程。随着实验精度的不断提高，给超出标准模型的理论留的余地越来越小。但是另一方面，这个理论的确又有许多难解之谜。爱因斯坦曾经讲过：“一个理论越深刻，它的前提就越简明，它描述的事物就越多，它适用的范围就越广泛。”以此来衡量，标准模型很难认为是一个真正“深刻”的理论。为了解释几十种粒子的质量，在这个模型中不得不引入十几个参数，它们都与一个捉摸不透的希格斯粒子有关。做为这个理论框架的基础之一，它的三代结构尽管与实验相符，却给不出合理的解释。凡此种种，可以开列出一个很长的清单，我们在前文已经详细的做了介绍。这么多的“未知”存在着，使理论物理学家们一致认为标准模型充其量不过是一个更好的理论的某种低能的等效形式。所以理论家宁可看到某一个实验的检验显示出偏离，使超出标准模型的努力能有较大活动余地。但迄今为止，除了前述的 Z^0 到 $b\bar{b}$ 衰变宽度有微小偏离以外，没有发现任何明显偏离。究竟问题出在哪里？没有人能说得清楚。突破口肯定存在，粒子物理学家们正在从各个方面积极探

索。

在 t 夸克于最近被正式发现以后，寻找希格斯粒子成为当前理论家与实验家最关注的焦点。在现有实验可达到的能量范围内，这种粒子存在的任何迹象都没有发现，人们十分迷惑不解。人们正在非常细致地搜寻对它的质量的各种可能的约束条件，研究它与费米子、规范粒子的相互作用的各种观测效应，对更高能量下寻找希格斯粒子的可能性做仔细的分析。期望不久会有突破性的进展。

另一个普遍关注的大问题是强作用 QCD 理论的非微扰问题，它是许多计算的最大的不确定性的来源。原则上，QCD 理论是夸克与胶子的一个精确的、完备的理论，应该能够对所有能量下的强相互作用问题做定量的处理，实际上远远做不到这一点。主要障碍是夸克和胶子不可能做为自由粒子独立的存在。从 QCD 第一原理出发，对这种禁闭现象给出精确的、定量的描写，一直是理论家们的梦寐以求的目标。今天看来这个问题太困难。量子场论发展了半个多世纪，至今除极个别的低维问题可以精确求解之外，理论家们仍只会用微扰论做近似计算。尽管已经证明 QCD 是一个渐近自由的理论，当能量很高或距离很近时相互作用变得很弱使微扰论可用，但实际上几乎所有的过程都包含有微扰论不适用的低能的、长距离的部分。

例如，对高能深度非弹性散射过程，我们已经介绍过，部分子模型给出了很好的描述，是检验微扰 QCD 的最好的实例。实际上，为了能与实验观测的结果相比较，必须知道反应之初夸克或胶子在强子内部是如何分布的，反应终了被撞击出来的夸克和胶子如何构成强子，或者说原来的强子在高能粒子的撞击下是如何碎裂成新的强子的。这些问题不可能用微扰论处理，它们属于非微扰 QCD 的对象。人们通常用所谓因子化的方法把微扰部分与非微扰部分分离开来，分别进行处理。分离出来的非微扰的贡献至今并没有找到可靠的计算方法。通常的做法是，引入一些参量用实验数据拟合，或者是建立可以用来计算的模型。这些做法本身都给理论的预言带来很多不确定性。而且，因子化本身的立论根据其实并不完全清楚。有人甚至讲，微扰 QCD 的检验只不过是检验给出这些结果的人对于问题的不可计算部分使用的技巧是不是聪明。这种说法其实是不无道理的。

高能过程的状况尚且如此，对于大量的低能衰变和反应过程非微扰

QCD 带来的不确定性影响更大。尽管理论家们建立了许多唯象的模型对强子静态性质给出了很好的描写，也使衰变或反应过程的计算成为可能，但由于各种模型都不能直接从 QCD 第一原理建立起来，所以它们的可靠程度很难给出恰当的估计。它们的应用能力都是很有限的。此外，理论家们还发展了许多近似的方法或等效理论如 QCD 求和规则、重夸克近似和大 N 展开等，应用的范围也都有限。

格点 QCD 理论力求直接从第一原理出发解决非微扰计算问题，确实已经得到了许多好的结果，但是由于在格点上如何处理费米子问题始终没有找到满意的解决办法，它的应用能力受到了很大的限制。此外，它过多地依靠计算机，其结果必然受到计算机能力的制约，它们在多大程度上逼近真实的、连续的 QCD 理论，目前还无法做出令人满意的判断。

总之，理论家们对于涉及非微扰、长程、禁闭现象等问题的研究已经做出了相当大的努力，但存在的问题仍非常严重，随着实验数据的大量积累，这里正孕育着出现新物理观点和理论的突破口。我们的最终目标是，要从一个基本理论出发正确地处理非微扰现象，自然地得到符合实验数据的结果。实现这一目标将会给理论物理学带来新的飞跃。

粒子物理理论中的还一类问题是关于大量粒子的聚集状态性质和结构问题。重的原子核可以看作是以大量质子和中子作为组元聚集成的高温为数百亿度的强子物质相，这个温度按宏观现象的标准来看是极高的，但从粒子物理世界的标准来看是相当低的。理论预言，如果把强子物质相的温度提高到数万亿度，就可能发生微观物质的相变，从以重子、介子为组元聚集成的高温强子物质相转化为以夸克、胶子为组元聚集成的高温夸克物质相。夸克物质和强子物质的组元不同、聚集方式和相互作用性质和强度也不同，表现出来的物理性质和行为特征完全不同。大爆炸宇宙模型认为随着时间的进展，宇宙的尺度不断增大，宇宙的温度不断降低。温度降到一定程度时，宇宙中物质演化表现为夸克物质相。温度继续下降到数万亿度时，又发生相变，夸克物质相转化为强子物质相。高能物理实验的一个重要领域是相对论性重离子碰撞，其目标之一就是通过超高能原子核的对碰，实现在重原子核大小的体积范围内，达到高温高密度，从而产生从强子物质相到夸克物质相的相变。如果实现了这种相变，就从大量粒子的聚集状态方面检验了标准模型，看到了新物理，同时也是在实验室条件下，在重原子核的体积范围内，模拟研究宇宙演化的某

一阶段。用 QCD 理论处理这类问题的研究工作还非常初步，需要有更大的发展。

至于谈到超出标准模型的理论，存在的问题就更多。从目前精确的实验数据看，我们在前面详细地介绍过的超引力大统一理论也许是最有希望的候选者，为当前理论家所关注。近年来这一理论模型取得了很大的进展。但总的来看，给出的理论预言还包含了很大的不确定性，有许多基本的困难问题有待解决。其它的模型，如用新的动力学机制代替希格斯机制的所谓人工色理论，把夸克和轻子等看成是更深层次的粒子组成的复合模型等超出标准模型的尝试，相比之下更不成熟。尽管目前的实验数据并没有完全排除这些可能性，但理论本身存在更多困难问题，使它们达到能为实验精确检验的地步，需要做的工作更多。

从实验方面看，1993 年遭受了一次意外的沉重打击。为了寻找希格斯粒子，完成对标准模型的最重要的检验并对超出标准模型的所有的理论做出判断，粒子物理学家满怀希望地期待着美国的 20TeV × 20TeV 超级超导质子对撞机 SSC 的建造成功。这台预算达上百亿美元的超大规模工程，1989 年在美国德克萨斯州达拉斯附近破土动工，计划用六年时间就可建成。由于美国经济的不景气以及政治上、管理上、舆论上种种复杂因素的共同影响，在工程已经完工将近五分之一的情况下，于 1993 年秋被迫停工下马。它不仅导致二千名高能物理方面的理论家、实验家和工程技术人员失业，而且余波所及，其它一些已建成的大实验室的经费也受到了影响。

当然，事物的利与弊毕竟不是绝对的。SSC 的垮台无疑使粒子物理的研究进程严重受挫，但实验家们的精力却可以更多地集中到如何充分利用现有的设备和设计一些投资少、见效快而实际价值并不见得差的设备上。

目前，世界上有几个大型的高能加速器在运行。其中最引人注目的是西欧核子研究中心(CERN)于 1989 年建成的 e^+e^- 对撞机 LEP。对撞束的质心能量为 100GeV 左右，刚好在 Z^0 能区。五年来取得的大量数据把对标准模型的检验提高到了一个新水平。温伯格对此给予的评价为：“由于 LEP 实验的成功，弱电理论已经达到了可以和量子电动力学的精度相匹敌的程度。”几乎与 LEP 的建造同时，美国在 SLAC 也建造了一台 100GeV 的 e^+ 、 e^- 对撞机 SLC。这是一台半直线对撞机，也工作在 Z^0 能区，但亮

度要比 LEP 小。近年来也提供了大量 Z^0 粒子的实验数据，特别是关于极化现象的实验结果，受到了人们的重视。

美国芝加哥的费米国家实验室于 1987 年建成了质心能量 1.8TeV 的质子-反质子对撞机 Tevatron。它远远超过 LEP 能区，因此可以得到 t 夸克这样重的粒子，以及搜寻更重的希格斯粒子和超对称粒子。

除此之外，1992 年在联邦德国的 DESY 还建成了世界上第一个高能电子-质子对撞机 HERA。它以 30GeV 的电子与 820GeV 的质子对撞，提供了研究轻子对核子深度非弹性散射的最前沿阵地。它可以探测 10^{-17} 厘米之内的核子结构，有非常重要的意义。

从 1996 年开始 LEP 将实施更新计划，能量将提高一倍或更多，完成后改称为 LEP II。然后在本世纪末到下世纪初，欧洲还将建成更大的质子-质子对撞机 LHC，能量达到 14TeV。费米实验室的 Tevatron 更新计划也已执行，其能量与亮度也将有很大的改善。

带电粒子沿圆周高速运动时，会因同步辐射使带电粒子本身的能量不断损失，因而为维持其圆周运动就必须不断补充能量。理论分析表明，粒子越轻、能量越高、半径越小能量损失越严重，对于环形加速器，LEP II 已经达到了建造水平的极限。因此近年来，美国、欧洲、日本的科学家们对建造 300 ~ 500GeV 的 e^+ 、 e^- 直线对撞机 NLC 表现了极大的兴趣。有的实验室已经开始了预研制。人们认为它将是精确测量 t 夸克的性质，寻找希格斯粒子，研究带电的弱中间玻色子 W 的最有力的机器。到下一世纪初建造成功大有希望。

从历史上看，粒子物理实验的发展有两条互为补充的路线。一条是努力向更高的能量前沿攀登，上面介绍的那些正在运行或计划建造的大型加速器都属于这种发展路线的产物。由于能量越高规模越大，投资以惊人的速度增加，这种方式的发展前途是不容乐观的。另一条发展路线是向高精度方向努力。CP 破坏的发现、电子与 μ 子反常磁矩的测量及强子谱和稀有衰变过程的研究等，都属于这类工作。近年来美国、欧洲、日本和中国等热烈地讨论和积极筹划的各类所谓的粒子工厂正是这后一种发展路线的重要组成部分。

粒子工厂实际上就是一些加速器实验室。它们在选定的能区以高于一般加速器几百倍的束流强度下运行，产生特定粒子，产额比普通加速器大得多。而且设计的能量使所需要的粒子刚刚能够产生出来(称为近阈

产生), 背景非常干净, 可以进行最佳研究。由于其重大的物理意义和造价相对适中, 目前国际上提出的计划相当多。例如, 加拿大的 K 介子工厂、意大利 Υ 介子工厂、美国 SLAC 和日本 KEK 的 B 介子工厂, 中国北京的 Υ -c 工厂等。其中除了 K 介子工厂是进行固定靶实验的质子加速器以外, 其它粒子工厂都是 e^+e^- 对撞机。

在 B 介子工厂的物理目标中, CP 破坏现象的研究居于首要地位, 其结果不仅会对于粒子物理学对称性的研究有重要的意义, 对理解宇宙中物质与反物质不对称性的起源也有极其重要的作用。为此, B 介子工厂多设计成两个能量不相等的束流对撞。使产生的 B 介子速度比较大。根据相对论时间延缓原理, B 介子在其存活的时间内, 能飞行更长的距离, 于是 CP 破坏效应更显著, 实验观测更容易。

Υ -c 工厂以 Υ 轻子和含 c 夸克的介子与重子为生产对象。它工作在两个能区, 一个是 3.6GeV 左右, 刚好能产生一对 Υ 轻子。另一个能区是 4GeV 左右, 通过 Υ 的各种激发态的衰变, 研究与之相关的丰富的物理现象。特别是寻找胶球吸引着人们极大的兴趣。

由上述可以看到, 虽然 SSC 的停建犹如一场暴风雪, 使粒子物理学家向更高能量前沿推进的道路暂时受阻, 但是不会对这一领域的发展进程带来根本性的影响。在最近一二十年内, 粒子物理学的理论与实验两个方面都有大量的工作要做。本世纪末、下世纪初的这段时期更是难得的黄金时代。人们期望着 t 夸克的进一步结果, 期待着希格斯粒子及超对称粒子的发现, 胶球、杂化态以及尚未露面的重夸克偶素家族成员被找到, 期待着 CP 破坏现象的研究有较大进展, 期待着各种超出标准模型的预言得到判定性检验。

一、二十年在人类历史的长河中只不过是短暂的瞬间, 对它期待得太多也许是不太实际的。但今天理论的发展水平以及实验技术的高超能力又确实是前所未有的。庞大的、训练有素的理论物理学家的队伍, 特别是那些年轻的硕士、博士、博士后中不乏优秀人才。众多的大型、先进的实验设备为理论的发展提供了极其良好的条件。新思想、新理论、新方法、新现象、新的物理学随时都有可能诞生出来。重大的突破往往在最预料不到的时候被得到。因此, 对于粒子物理学的未来发展, 做出精确的预测实际上是不可能做到的。

你可以有非常乐观的估计, 随着这些大加速器和粒子工厂的改造、

建造计划的完成，希格斯粒子和一些预言的超对称粒子相继被发现了，稀有过程的成功观测提供了重要的信息。超弦、超对称大统一、超引力、非微扰 QCD 以及 CP 破坏等在理论方面也取得了突破性进展。所有这些都或许可以使我们朝着一些物理学家津津乐道的“终极理论”或“包罗万象”的理论的目标大大向前迈进了一步。那时，理论物理学家们是否就无事可做了呢？实验家们是否就可以偃旗息鼓，不必为更高的能量的加速器的建造而奔波忙碌了呢？

情况也许完全相反，出乎人们的预料，该改造的改造了，能建造的建成了，而要找的东西却依然没有找到。处境仍然是标准模型经受了更高精度的检验，存在的问题仍然解决不了。或许仍然可以找到理由认为，希格斯粒子和超对称粒子质量太大超出了建成的加速器的能力。但建造更高能量的机器需要巨额投资将无法实现。理论本身由于遇到一些无法克服的困难，因而也成效甚微。在这种形势下，物理学家们会不会束手无策，以至放弃他们奋斗百年的目标了呢？

当然，极有可能的情况是在这两种极端之间。但不管怎样，历史告诉我们，物理学家是绝不会满足于已有成就的。他们不会停止思考，也不会停止行动。一个问题解决了，他们还会发现有更多的问题需要解决。一条路不通，他们还会开辟更多的路。物理学的总的趋势是在不断地发展，只不过有时快、有时慢。粒子物理学只不过有半个多世纪的历史，当前正处在富有成果的蓬勃发展的时期，它的发展前景仍然是十分诱人的。严重的基本问题摆在理论物理学家面前，呼唤着，孕育着新的物理观点和理论的出现。这正是 21 世纪理论物理学家面临的挑战和机遇。他们会为人类的文明作出新的贡献。

在这世纪之交的激动人心的时刻，回顾这一百年的辉煌历程是非常有意义的。19 世纪末物理学曾遇到了两大疑难，一个是以迈克尔逊—莫雷实验为代表的关于运动物体中光的传播现象的研究，另一个是从黑体辐射实验规律总结的普朗克公式。它们从根本上与经典物理学发生了尖锐冲突。前者导致爱因斯坦建立了相对论，后者导致量子力学的诞生。理论物理学家所做出这两项革命性的成就，开辟了现代物理学的全新的时代。物理学的各分支也都随之迅速地发展起来。老的学科焕发了青春，新的领域迅速崛起，核能、激光、半导体、超导、超级计算机等等高新技术的兴起及发展无不依赖理论物理的这两大基石。

历史会不会重演？什么是今天物理学的最大疑难？下一世纪初会不会也有革命性变革？这是引人深思的问题。美籍物理学家李政道教授对此有独特的见解。他认为，今天的物理学也有两大疑难，它们都存在于粒子物理之中。一个是对称性的丢失，比如宇称不守恒、CP 破坏及时间反演不变性的破坏等。另一个是夸克看不到，它们永远被囚禁在强子内部。这两大疑难，他认为都与真空的性质有关。物理规律本身，所有的对称性都是对的。由于真空的一些特殊性使对称性破坏了。规范对称性的自发破缺，正是真空的特性造成的，这是一个很好的例证。夸克的禁闭也可用真空是一个理想的抗色电媒质来解释。正如超导体是一个理想的抗磁体，把所有的磁场都排斥在外一样，真空把所有的色电力线都压挤在强子内部，不允许夸克自由地释放出来。于是如果能想办法使真空激发，就有可能观测到真空的物理效应。

李政道认为，本世纪末将投入运行的，布鲁克海文实验室正在兴建的，每核子 $100\text{GeV} \times 100\text{GeV}$ 的相对论重离子对撞机(RHIC)有可能提供这个机会。重的原子核，如金原子核，以极高的速度对撞，会使真空形成夸克-胶子等离子体(QGP)。在这种状态下，夸克和胶子的禁闭被解除了。但它只能存活极短的时间，然后，很快地碎裂成极大数量的强子。观测这种相变过程，可以对真空有深刻的了解，从而带来物理学的重大的突破。这个预言是否正确，下一世纪初是可以见分晓的。如果是这样，那么下一世纪的粒子物理学会不会以挖掘真空做为新的出发点呢？

参考文献

[1] 基本粒子物理学专门小组(美), 90 年代物理学：基本粒子物理学, 科学出版社, 1993.

[2] Y.Nambu, Quarks: Frontiers in Elementary Particle Physics, World Scientific, 1981.

[3] Andrew Pickering, Constructing Quarks: A Sociological History of Particle Physics, Edinburgh University Press, 1984.

[4] Berry Parker, Search for a Supertheory: From Atom to Superstring, Plenum Pub.Corp., 1987.

[5] Fayyazuddin and Riazuddin. A Modern Introduction to Particle

Physics, World Scientific, 1992.

[6] Sheldon Glashow, The Charm of Physics, The American Institute of

Physics, 1991.

[7] 李政道, 对称, 不对称和粒子世界, 北京大学出版社, 1992.

3 场论

量子场论或场的量子理论是把相对论和量子力学结合起来统一研究实物粒子与场的一门理论物理前沿学科，是从场的观点研究粒子的性质、粒子的产生与湮灭过程以及自然界存在的四种基本相互作用的理论。它始于本世纪 20 年代末，经过半个多世纪的发展，迄今硕果累累。作为微观现象的基本理论它已经广泛地应用于近代物理学的各个分支。

粒子物理学与量子场论有着密不可分的联系。粒子物理不断提出场论研究的新课题，而量子场论的研究成果极大地促进粒子物理学的发展，自 70 年代以来标准模型的巨大成功是最具说服力的例证。在统计物理、凝聚态理论和核理论中广泛地采用量子场论的格林函数与费曼微扰论方法，已成为这些物理学分支的基本理论工具，场论方法对温度不为零的统计物理学以及超导和量子液体、量子霍尔效应等现象的理论发展起了非常重要的作用。实际上，量子场论已经成为近代物理学各分支的共同基础理论，它的任何一个重要进展都会对不只一个分支的发展有重要的推动作用。

数学是物理学的语言，物理学的应用会反过来推动数学的发展。物理和数学的这种奇妙关系在场论的发展史上表现得尤为突出。对称性研究广泛地应用了群论。广义相对论(即引力场理论)与黎曼几何的关系是众所周知的，描述空间几何性质的数学被用来研究引力场的行为。杨振宁在 1974 年指出的规范场就是纤维丛上的联络，这种关系的发现大大地推进了关于场的拓扑性质的研究。此外，反常是与场在空间的整体性质相关的，研究这类性质的数学分支叫拓扑学。超弦理论的发展与数学的代数拓扑学和研究黎曼面的代数几何学密切相关。从这些例子可以看到数学在这门学科中的重要作用。所以，脱离严格数学就很难对量子场论这门学科作简明而精确的阐述。同时值得指出的是，有些数学分支正是通过在量子场论中的应用得到了蓬勃的发展和不断的完善。

数学的简和美对于基础物理概念的形成起着越来越大的作用，但是物理学和数学毕竟各有各的目标和爱憎，有明显不同的价值观和不同的传统。虽然二者令人惊异地具有某些共同的概念，然而如杨振宁指出的，“即使在这些方面，二者的生命力也向着不同的方向奔驰。”

3.1 量子电动力学

电磁作用是自然界中人们最早认识到的一种基本作用力。关于电磁作用的量子理论称为量子电动力学，是一个最简单的规范场理论，也是迄今量子场论最为成功的典范。量子场论的许多基本概念产生于此，它为人们寻找关于其它基本相互作用理论以及四种相互作用统一理论提供了依据。本节第一篇已经介绍了它的发展历史，这里再深入地阐述几个重要问题。

3.1.1 场和它的量子化

场的概念是在解释粒子之间非接触作用时建立的。熟知的例子是带电粒子周围存在着电场，处在电场中的其它带电粒子不需要同原来的粒子发生接触就会受到作用力，其大小与粒子间距离的平方成反比(即库仑定律)。如果场仅仅是依附于产生它的粒子之上，粒子之间通过场的相互作用就是瞬时的或称超距的。我们相信自然界的每一个基本理论必须与狭义相对论原理一致，而这种瞬时相互作用同这些原理是明显矛盾的。根据狭义相对论，信号的传播速度不可能大于光速，因此非接触力的作用不可能是瞬时的。由于一个粒子的位置或运动状态的突然改变而产生的场的分布的改变，必须经历一定的时间才能到达空间的其它位置从而被处在那里的粒子察觉到。这段必须的时间就是光信号在两粒子间传播所经历的时间。这样，场就首次获得了其真实性：场可以在空间传播，但传播速度永远不能大于光速。

另外，场还具有能量。如上述粒子 1 运动状态的改变将通过在粒子 2 处产生的场的改变而使粒子 2 的运动状态改变。这样，两个粒子之间就有能量的交换。一个孤立的体系的总能量是守恒的。那末我们就可以问，从粒子 1 的运动状态改变的瞬时到这一改变在粒子 2 处首次被感觉到的瞬时之间，到哪里去找那些最终要传送给粒子 2 的能量呢？答案只能是能量存在于场中。在迈克尔逊·莫雷实验否定了真空中存在以太介质之后，人们认识到电磁场不是以太的振动波而是具有真实的物理实在性：场可以脱离粒子单独地在空间以波的形式传播并携带有能量。

从上面的讨论看出，在相对论理论中，不但要考虑实物粒子，而且要考虑场。物质的存在形式被推广了。场可以想象为空间一定范围内弥漫开来的、连续的某种分布。用数学的语言表达，场是空间坐标的函数，比如记作 $\phi(x)$ 。当讨论场随时间的变化时，它同时还是时间 t 的函数，记作 $\phi(x, t)$ 。由于场是在空间连续分布的，描述场的状态就必须给定它在空间各处的值，这需要无穷多的参数，于是体系不再是一个有限自由度的体系。这是相对论性和非相对论性理论之间的深刻区别。随着量子理论的建立，物质的这两种存在形式——粒子和场进一步得到了统一。

我们在第一篇曾指出，狄拉克于 1928 年把量子力学与相对论结合建立了单电子满足的相对论性量子力学方程并提出了“负能电子海”的假说，认为所有的负能态已被电子占满，它是体系的最低能量态，也就是真空态。这个假说给出了一个全新的真空概念，真空变得不空了！自从爱因斯坦的狭义相对论打破了以太存在的假说之后，人们认为真空就是什么都不存在的绝对真空。相对论与量子理论的结合改变了这一局面。在相对论情况下，由于真空态电子海的出现，所谓的单电子实际上是一个无穷多粒子的系统。所以狄拉克方程是描写无穷多电子的集合——电子场的波动方程，如同麦克斯韦方程是电磁场的波动方程。

光子是电磁场的量子，同电子一样遵从量子力学规律。为描写大量光子的体系，需要将电磁场量子化。通常的作法是把电磁场的运动分解为一系列基本的简单的振动，各种基本振动方式代表电磁场的各个自由度。电磁场的量子化被转化为各个基本振动自由度的量子化。由量子力学可知，微观领域中的谐振子量子化的结果能量取值不连续。这样就解释了电磁场的粒子性，即解释了光子的存在。某一振动自由度被激发就相应于产生了一个光子，振动激发的消失就相应于光子的消失。电磁场既可以看作是一个光子的集合，又可以看作是一系列的电磁波。这一想法是狄拉克于 1927 年提出的。按照这一想法可以同样地处理电子场，并由此引入了量子场的概念。

量子场是微观物质存在的形式之一。可以用不同的量子场来描述不同种类的基本粒子。量子场的各种受激发态代表处在各种状态下的基本粒子系统。量子场的激发代表粒子的产生，激发的消失代表粒子的湮灭，用量子力学语言，这种产生和湮灭过程用产生和湮灭算符作用于真空态来实现。粒子间的相互作用表现为相应的场之间的作用。海森堡和泡利(1929)在此基础上建立了量子场论的普遍数学形式，用来处理各种基本粒子现象。

用前述把场的运动分解为一系列基本振动的办法，我们把作为多粒子系统的场的任一个态与振子的一个激发态联系起来。一个多粒子系统和一个振子的集合是同一物理实体的两种不同的数学描述。场是一个无穷多自由度的体系，它的基本动力学变量是场在各个时空点的取值，称为场量，相当于体系的“坐标”。而时间和空间坐标则做为参数标记这无穷多的“坐标”。场的量子化则是把场量对应为在相应的时空点产

生和湮灭粒子的算符。与体系的“动量”相当的量称为共轭场量，量子化后也成为算符。量子化要求同一时刻不同空间点的场量和动量对易而同一时刻同一空间点的场量和动量不对易，这称为正则量子化条件。物理上，同一时刻空间的距离不为零的两点即使用光信号也来不及联系，在这些点上进行的测量彼此不相干扰，因而我们预期属于这样的两点的物理可观测量彼此对易，这一要求称为微观因果律。所以，量子场论的这一基本要求是相对论性的必然后果。

上述作法意味着处于同一运动状态的光子的数目是不受限制的，这样的粒子称为玻色子。自然界中还存在另外一种粒子，它们遵从泡利不相容原理，也就是说，不能有两个全同的这种粒子处在完全相同的状态，这种粒子称为费米子。

在量子力学中，由两个全同粒子组成的体系，用波函数 (x_1, x_2) 描述其状态。 x_1, x_2 分别代表两个粒子。把两个粒子交换，波函数变为 (x_2, x_1) 。因为两个粒子的属性完全相同，体系只“关心”有两个粒子处在什么样的运动状态，至于哪个粒子处在这些运动态，体系并不“关心”，事实上即使探测到某一运动态上有一个粒子也无法判断究竟是两个粒子中的哪一个。所以只能认为 (x_1, x_2) 与 (x_2, x_1) 描写的是同一量子态。这种交换对称性会给波函数的形式带来很强的限制。它们只能相差一个取正 1 和负 1 两个可能值的常数因子。于是交换后的波函数要么与原波函数相等，要么乘上一个负号。前者称为对称波函数，描写玻色子系统；后者称为反对称波函数，描写费米子体系。

前面曾指出狄拉克方程描写的是一个无穷多电子的系统，即电子场。它的量子化意味着，象电磁场一样将场量当做算符。由于电子的自旋等于 $1/2$ ，人们发现它的量子化条件不同于电磁场，它的场量满足所谓的反对易关系，在算符交换次序时多了一个负号。更一般地可以证明，所有自旋为半奇数的粒子都是费米子，对应的场量满足反对易关系；而自旋为整数的粒子都是玻色子，应该由满足对易关系的场量来描写。

3.1.2 相互作用量子场

从物理学发展之初起，对称性一直是一个强有力的工具。如今，对称性观念以及研究方法已经渗透到物理学的各个领域，变成了物理规律理论表述的支柱，也是迄今发展起来的各种相互作用场论的基础。

一切对称性的根源在于某些基本量的不可观测性，这些基本量称为“不可观测量”。“不可观测量”的存在会使某些“状态”等价。例如，我们通常有这样的经验，如果洗出来的照片上一个人的手表是戴在右手上的，我们会有理由说照片洗反了，因为通常人们总是把手表戴在左手上的。如果照片上的人没有戴手表，我们会去看他的衬衣纽扣或是口袋。总之，我们试图通过区分左和右来判断照片是否洗反了。如果找不到这种左和右的差别，我们就无法区分同一底片正反洗出的两张照片。这时，由于左右的“不可观测”，正反洗出的两张照片看上去一样，或者说是“等价”的。这个例子说明一个对称性原理的三个方面的相互关联性，即不可观测量的假设，相关的变换(这里表现为正着和反着冲洗照片)以及所蕴含的不变性(这里表现为洗出的照片无法区分正反)。

基础物理学中一个重要的对称性原理是洛仑兹不变性。这是作为麦克斯韦方程的数学性质而被发现的，而麦克斯韦方程则是在电磁学实验定律的基础上建立起来的。后来闵科夫斯基(Minkowski)倒转了这一过程，从洛仑兹不变性入手要求场方程不变。爱因斯坦对此大加赞赏。他的广义坐标不变的想法加上等价原理导致了广义相对论的产生。

规范对称性是另一个极为重要的对称性原理。它的最初想法是由外尔提出的。受爱因斯坦将引力与电磁力统一起来的思想启发，自 1918 年起，外尔致力于将电磁学与引力统一的研究。既然坐标变换下的不变性导致了引力，外尔认为电磁学也应该可以从某种新的几何对称性得到。如果 x^μ 和 $x^\mu + dx^\mu$ 是相邻的两个时空点， f 是物理量，它在 x^μ 为 f ，外尔认为随时空变化的重新标度。从 x^μ 到 $x^\mu + dx^\mu$ 点时，它的“长度”被乘上了一个放大倍数，这个倍数与两时空点之间的距离成正比，记作 $1 + S_\mu dx^\mu$ 。关于这个标度因子，外尔观察到两点。第一， S_μ 有着与电磁势 A_μ 同样数目的分量。第二，如果要求这个理论在标度的变化下保持不变，那末只有 S_μ 的旋度，而不是 S_μ 自身有物理意义，而这也正是电磁势的特点。由于以上的性质，外尔认为标度因子就是电磁势，至多相差一个常

数因子。这样，电磁学就可以作为标度不变性而导出。可是这一想法没有成功，爱因斯坦证明了外尔的理论不可能描述电磁学。外尔于是放弃了他的想法。

一九二七年福克以及其后的伦敦从量子力学出发得到的新的结论是， S_μ 不和 A_μ 等同，而是等同于 $(-ie/hc)A_\mu$ (其中， e 是电荷， h 是普朗克常数除以 2π ， c 是光速)。这和外尔最初的想法所不同的仅仅是差了一个因子 i 。可是这个因子 i 影响深远，它使标度因子变成了

$$1 - (ie/hc)A_\mu dx_\mu \quad \exp[- (ie/hc)A_\mu dx^\mu]$$

这是位相的改变而不是标度的改变。因此，局部的位相不变是电磁现象的量子力学特点。外尔自己开头曾经将这一概念称为

“MasstabInvarianz”，后来又改称“Eich-Invarianz”。20年代初，这一名称翻译成英语，叫做“GaugeInvariance”，以后中译为规范不变性。

常识告诉我们，一个自由粒子在空间运动的轨迹是一条直线，直线是联结两点的最短距离。实际上这里蕴藏了一个非常重要的原理，即最小作用量原理。作用量是一个依赖于运动轨道的量，它是系统拉格朗日量对运动轨道的积分，随轨道改变而取不同的值，粒子实际经过的经典运动轨道总是使得这个量最小值。通过对作用量在轨道的各种假想改变下的响应进行研究可以得到粒子运动轨道应该满足的方程，也就是运动方程。这种办法完全可以推广用来研究场论，从而得到场方程。这样做的一个最大的优点就是可以很容易地使理论满足相对论性原理以及其它对称性要求。电子场的作用量可以从系统的拉氏量计算出来。在拉氏量的适当选取下，从作用量最小原理得到的运动方程正是狄拉克方程。

人们注意到在电子场做如下相因子的改变

$$\psi(x) \rightarrow e^{iQ} \psi(x), \quad \bar{\psi}(x) \rightarrow e^{-iQ} \bar{\psi}(x)$$

时，体系的拉氏量保持不变(其中 $\bar{\psi}$ 代表正电子的场)。这种不变性在体系的物理规律的反映即为系统的电荷守恒。所以说，电荷守恒这一由实验总结出的定律在理论上则解释为电子场作用量的一种对称性。事实上，在作用量这个框架内，所有的守恒量都是作用量存在某种对称性的必然结果，这正是我们在第一篇提到的诺特定理。

在上面关于场的位相变换中，变换是对场在同一时空点上进行的，称为内禀变换，以区别于那些因坐标系改变而导致的场变换。理论的相

对论不变性要求作用量在时坐标改变时保持不变，相应地得到动量、能量和角动量等各种守恒量，这是我们所熟悉的。对于内禀对称性，除了上面已经提到的场的电荷位相变换不变性外，还有如本书基本粒子物理部分介绍的重子数和轻子数守恒等，它们是电荷概念的某种推广。由于这种内禀变换对于场在各个时空点的改变完全相同，所以称为整体变换。如果场在各个时空点的位相改变并不相同而是各自独立，就称为定域变换。显然，整体变换是定域变换的一种特殊情况。

对整体变换的讨论可知，作用量之所以在电荷位相变换下不变，是因为场和反粒子的场以特定的组合形式出现，它们的整体位相改变刚好抵消的缘故。由于定域对称变换中，场的位相的改变与所在的时空点有关，场关于时空的微商的位相变换与场的位相改变不再相等，于是原来具有内部整体对称性的作用量并不具有相应的定域对称性。研究结果表明，保证原来的作用量具有相应的内部定域对称性的途径就是引入一个新的矢量场 $A_\mu(x)$ ，与前面关于外尔规范理论的讨论进行比较得出这个新的场量正是电磁势。这种定域位相变换下的不变性称为定域规范对称性，新引进的矢量场称为规范场。所以，要求理论有定域规范对称性就必然要引入相应的规范场，它起着将时空各点的位相变化联系起来的作用，具体到我们现在的例子，它实际上意味着电磁场传递带电粒子之间的相互作用。

3.1.3 非阿贝尔规范相互作用

象电荷这样简单的复数位相变换在数学上构成所谓的阿贝尔群。具体讲，将场乘上一个复数相因子就是对场进行了一次位相变换，再乘上另一个复数相因子则相当于做了两次位相变换。两次所乘的相因子本身相乘等于一个新的复数相因子，所以，两次变换完全等价于用新的相因子做一次变换。这就是说，两个变换相“乘”得到的还是一个变换。把所有的变换集合起来，数学上称为一个群。所以，群就是一些对称变换的集合，这些变换互相“乘”起来还是这个集合中的一员。用数学的语言就是说，群的元素在“乘”法运算下是封闭的。很容易证明这个群的乘法是可交换的。这样的群称做阿贝尔群；反之，则称为非阿贝尔群。

人们认识较早的非阿贝尔内部对称群是本书基本粒子部分介绍的强相互作用的同位旋对称性。它是从实验上发现的核力的电荷无关性建立起来的一种对称性。它使得只要问题仅限于讨论强核力，我们就无须区分质子还是中子；换句话说，在描写强核力的作用量中，我们可以将质子场与中子场用一个两行两列的系数矩阵进行线性组合，用新的场表达的作用量在形式上应该与原来的作用量完全一样。例如若取组合系数为：

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

则相当于把质子和中子对调，根据同位旋对称性，动力学规律应该不受任何影响。所有这样的系数矩阵在矩阵的乘法下封闭构成一个群，而且由于矩阵 A 乘矩阵 B 一般并不等于矩阵 B 乘矩阵 A，所以这个群是一个非阿贝尔群。出于物理上的考虑，我们只限于那些满足特定条件的系数矩阵，它们构成所谓的二阶么模么正群，记作 SU(2)。

根据群的一般理论，我们可以将系数矩阵写成一个复相因子的形式。只不过这时指数上不再是简单的纯虚数而是一个对角元素之和为零的反厄密矩阵，它在某种意义上是模为 1 的相因子的推广。所有这些 2 乘 2 的反厄密矩阵都可以写成 3 个独立的反厄密矩阵的线性组合，这 3 个独立的反厄密矩阵

$$r_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$r_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

$$r_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

称为群的生成元。非阿贝尔群的特点反映在群元素乘法的不可交换性，这种不可交换的程度可以通过对群的生成元满足的对易关系

$$r_i r_j - r_j r_i = 2 \epsilon_{ijk} r_k, \quad i, j, k=1, 2, 3$$

来描写，其中 ϵ_{ijk} 不恒为零反映了这种不可交换性，它称为结构常数。不同的群结构常数不同。同一个群可以有不同的表示。例如本书基本粒子部分曾提到， π 介子的三种状态构成同位旋 3 维表示，相应的生成元表示成为 3 乘 3 的矩阵。它们也满足与上式同样的对易关系。

1954 年杨振宁和米尔斯类比电磁相互作用的规范原理，将同位旋对称性定域化希望得到强作用的理论。他们发现，要构造一个作用量在定域的同位旋变换下不变就必须对应于同位旋的三个生成元引进三个规范场，记做 $(W_\mu^+, W_\mu^0, W_\mu^-)$ 。这三个矢量粒子和光子类似，也称为规范粒子，它们起着传递强力的作用。当进一步构造规范场自身的作用量时发现，要得到具有这种定域 SU(2) 对称性的作用量除了要有规范场的二次项外，还要有三次项和四次项。它意味着规范粒子带有一种“荷”（在这里是同位旋荷），使规范粒子之间存在自相互作用。非阿贝尔规范场的这一特点对于量子场论有深远的意义，现在人们已认识到它是强作用存在渐近自由行为的根源，但在当时它所带来的计算上的巨大复杂性却使得理论的发展足足推迟了十多年之久。杨振宁后来回忆这段发展时说：“我们完全陷入困境（指得到自恰的计算法则），规范场的非线性自作用项如此复杂以致我们看不出摆脱这种局面的途径。无论如何，由于对这个课题本身的足够兴趣，我们还是发表了一篇文章。”事实上，尽管后来的研究表明强作用的规范群不是同位旋 SU(2) 而是色 SU(3)，但这篇发表于 1954 年的经典之作却揭开了人们认识自然界基本相互作用的一个历史新篇章。

3.1.4 路径积分量子化

量子化是微观客体的特性，如何正确地构造量子理论始终是一个重要问题。一般而言，从经典理论得到相应的量子理论可以按照正则量子化的标准方法进行。但正则对易关系只对独立的共轭正则变量才成立。如果系统存在约束，那末首先应从约束方程中解得独立的变量，然后再进行量子化。规范场理论的量子化就属于这种情况。

电磁场的作用量在规范变换下保持不变，它表明并非电磁势的四个分量都是独立的场变量。物理的场满足一定的规范条件，它对电磁势的独立分量加了很强的限制。这时如果简单地将四个分量同时做为独立变量进行正则量子化，将会因无视它所满足的规范条件而得到矛盾的结果。从规范条件中解出独立的变量然后再进行量子化可以避免上述困难。但是，直接这样做并不方便，而且会牺牲整个理论的明显洛伦兹协变性，特别是对于非阿贝尔规范场的情形，规范场方程的非线性将导致实际求解的巨大困难。

对非阿贝尔规范场的量子化最先是 1967 年由前苏联学者法捷耶夫 (L. Faddeev) 和波波夫 (V. Popov) 用路径积分的办法实现的。路径积分量子化的最初想法来自狄拉克。他在寻找一种能够平等对待时间和空间的量子力学表述方式时发现，在经典理论的两种表达方式中，哈密顿形式的时间具有特别的地位，而拉格朗日最小作用量原理的时间和空间地位是平等的。他相信拉格朗日作用量也一定会在量子理论的表述中发挥作用。费曼接受并发展了这种想法，得到了路径积分量子化方法。

作为量子力学的基本原理之一的测不准原理表明，由于位置和速度这对相互共轭的量不能同时测准，所以经典的轨道概念在微观领域没有意义。但费曼认为，在量子力学中可以保留轨道概念，只不过一个微观粒子从初始时刻的位置到终了时刻的位置之间的所有轨道都是可能的，粒子沿哪条轨道走并不确定，每一轨道都有一定的几率幅。量子力学的特征正反映在轨道的不确定性上。这与经典力学的情况相反。在经典力学中粒子是按确定的轨道运动的，这个轨道称为经典轨道，它是使作用量 S 为最小的那条轨道。在量子力学中，费曼假定每个轨道对总几率幅贡献的大小相等，相角不同，即等于 $e^{iS(a,b)/\hbar}$ ，其中 $S(a, b)$ 是该路径的作用量值。总的几率幅为

$$\sum_{\text{路径}} e^{iS(a, \text{路径}, b)/\hbar}$$

这种路径积分量子化的特点是，它的表达式中所用到的量都是经典的而不是算符，但是由它所表达的却是量子力学的几率幅。

既然所有的路径都有同样大小的贡献，只是相角有所变化，那末在经典极限下某个特定的路径是怎样变成最重要的路径的呢？经典近似相应于尺寸、质量、时间等都很大，以致 S 与 \hbar 相比是巨大的，所贡献的相角 S/\hbar 是个很大的值。对一般的情况，路径的微小改变会造成相角的巨大改变，它们的贡献由于振荡而相互抵消。但对于 S 为极值的特殊路径，它的微小变化不会使 S 发生变化。所以，这个区域内各路径的贡献几乎都是同相的，不会相互抵消。于是当所讨论的尺度远大于 \hbar 时，这种表达形式自动地挑出了满足最小作用量原理的经典轨道。

路径积分量子化方法在概念上是简单的。但是对这样的求和给出更精确的数学上的定义是相当复杂的。这种求和的极限自然地应该表达成积分，由于这时的积分变量——路径——本身是时空的函数，所以它是推广了的黎曼积分，即所谓的泛函积分。关于泛函积分的存在性在数学上仍是尚未完全解决的问题，好在对于一类特殊的称为高斯型的泛函积分其存在性是得到严格证明的。而常见的几种场，如标量场、旋量场和规范场，它们的自由场拉氏作用量的路径积分都是属于高斯型的。

当考虑了场的相互作用后，拉氏量中往往包含场的三次或更高次幂的单项式。它们不属于高斯型积分，在具体计算中不得不借助于微扰展开。由此发展起来的理论与正则量子化的微扰论异曲同工。路径积分量子化方法有许多优越的地方。例如，由它可以发展一些形式场论的方法，对作用量的具体表达式尚不清楚的情况尤为有用；路径积分形式有利于做半经典近似；它的表达式与统计物理中配分函数的相似导致了有限温度场理论的产生和发展等。特别是由于规范对称性使规范场的自由度多于真实的物理自由度，即规范场是一个约束系统，而量子化是对于真实的物理自由度进行的，这在操作上给正则量子化带来困难，但在路径积分量子化方案中比较容易处理。

规范对称性意味着作用量在规范变换下不变，由规范变换所联系起来的场称为规范等价，它们的集合称为一个规范等价类(或规范轨道)。不加区分地对所有的场位形(路径在场论中的推广)进行积分将会导致对

于规范等价的场位形重复积分。由于场是无穷多维的系统，所以这种的重复积分的结果是无穷大。费曼路径积分应该只对于那些不在一个规范等价类的真正物理自由度进行积分。路径积分形式通过适当地修改积分的测度定义使积分只沿与规范轨道相交且仅交一次的路径进行，从而做到这一点。

这种办法是由法捷耶夫和波波夫首先得到的，他们对测度的修改通过引进辅助的场量完成的。这些辅助场在时空变换下表现得象个标量但却满足反对易的交换关系，由于这种奇特的性质，它被称为鬼场。它们只出现在费曼图的内线，从而保证了 S 矩阵的么正性。使积分路径只与规范轨道相交一次且仅相交一次是通过选取规范条件来实现的，如洛伦兹规范条件在场的位形空间就是给出这样的积分路径所满足的方程。格瑞伯夫(Gribov)发现，由于位形空间的拓扑性质，并不是总可以找到上述的积分路径，这种情况称为存在格瑞伯夫不定性。一般说来，非阿贝尔规范场的库仑规范条件就存在格瑞伯夫不定性，所以库仑规范不是一个好的规范条件。现在人们通过研究相信，格瑞伯夫不定性可能是非阿贝尔规范场的普遍性质，与其拓扑性质密切相关。格瑞伯夫不定性表现在规范条件确定的积分路径与规范轨道相交不只一次。不过这种重复相交一般发生在场量较大的位形，它并不影响微扰论计算。

路径积分量子化方案对于旋量场存在一定的困难。前面说过，路径积分的表达式中所用的量都是些经典的量，由普通的复数表示。但是，满足反对易关系的旋量场并没有经典的对应，换句话说，没有满足反对易乘法的数。所以要处理旋量场的量子化就必须推广关于数的定义。满足反对易乘法的“数”称为格拉斯曼数(Grassman 数)。Brezin 讨论了对于格拉斯曼数的函数进行微分和积分的问题。最近关于量子群在场论中的应用，特别是非对易几何的研究再次引起人们对于这些形式讨论的兴趣。

量子场论的路径积分形式由于前面提到的一些特点使得它成为现代量子场论的主要理论构架。量子场论的很多丰富内涵都通过路径积分测度的定义得到反映，鬼粒子就是一个著名例子。80 年代初，利用路径积分形式对于量子场论中的反常现象的研究表明反常与场在大范围的拓扑性质密切相关。

3.1.5 微扰、发散和重整化

描写电子(适合于其它自旋为 $1/2$ 的带电粒子)与光子体系的作用量实际上包括三个部份;电子场部分,电磁场(或者说光子)部分以及电子和光子耦合部分。前两部分分别称为电子场和光子场的自由部分;第三部分是因为要求作用量在定域规范变换下不变而引入的相互作用项。它的出现使得由最小作用量原理得出的场运动方程不再是自由的波动方程,电子场的运动方程中出现光子场,光子的运动方程中有电子场。对于这种耦合的运动方程人们尚未找到精确的解法,只能依靠一些近似方法。

现在我们可以通过具体讨论一对电子散射,看一下量子场论对于这样一个过程的处理方法。在初始时刻,两个相距无穷远的自由电子相向而来,随着时间的演化,两个电子逐渐靠近、发生相互作用然后再飞离开去直到无穷远。对于电子这样的微观粒子,人们无法观察碰撞的细节。按照量子理论的原理,体系的状态随时间的改变应由薛定谔方程决定。但由于人们关心的只是由初态的两个自由电子经相互作用后变成末态两个自由电子飞离开的几率,当相互作用如电磁作用这样很弱时,人们发现这个几率振幅的精确表达式(称为 S 矩阵元)可以对相互作用的耦合常数展开成幂级数,然后对这个无穷多项求和采用逐级近似方法(称为逐级微扰方法)计算。

先考虑零级(即第一项,因其中含耦合常数的零次幂而得名)贡献。它意味着没有相互作用,两个电子各自独立地运动并保持动量的守恒。这个过程可以用图 3.1 来表示。图中的黑点代表电子,两条线分别代表两个电子从初态到末态的过程,它们各不相干。人们更为关心的是有相互作用时的情况。考虑一级(也就是求和的第二项)贡献,发现其结果为零。因为它不能同时满足能量和动量守恒。用图来示如图 3.2 所示,其中直线代表电子,波浪线代表光子,这个图称为相互作用顶角。对于两个电子散射的过程,相互作用的最低级贡献为二级,如图 3.3 所示。

可以看出,电子通过交换一个光子而发生相互作用,而这个光子并未在初末态中出现,它的四动量可以不满足质能关系,故而称为虚粒子。虚光子从一个顶角到另一个顶角所对应的代数表达式称为传播子。在每个相互作用顶角,三个粒子的四维动量满足守恒定律。只要保证顶角动

量守恒，对虚光子的动量则没有什么限制。上面的图示称为费曼图。费曼在用量子场论进行各种计算时发现，用图来表示不仅有助于记忆而且在物理上更直观明晰。把图和一定的代数表达式对应起来的关系称为费曼规则。

继续上面的计算，把交换各种可能的虚光子对散射的贡献加起来应得到最后的结果。由于费曼图的拓扑结构，电磁相互作用的所有非零贡献费曼图一定含有偶数的顶角。而一对偶数顶角贡献一个由电荷和普朗克常数构成的无量纲的因子 e^2/ch ，其值为 $1/137$ ，即为精细结构常数。这个因子是实际耦合常数。对上面散射过程的计算所得到的级数解中，第一项正比于 α^0 的零次幂，第二项(二级微扰)正比于 α^1 的一次幂，第三项(四级微扰)正比于 α^2 的平方，依次递增。由于 α 比 1 小得多，通常只须考虑求和的头几项就可以达到所需的精度。

然而，从理论本身的严格性出发，人们希望能够处理微扰展开中的所有项而不仅仅只是头几项；从实际计算的精度考虑，对于某些过程，也需要把更高级的修正包括进来。这意味着对应的费曼图中含有更多的相互作用顶角，例如，

由于初末态是给定的，比如上面两个电子的散射初态只有两个电子，末态也只有两个电子，多出的顶角在图中增加了由内线构成的圈(虚的光子和电子)。由于这些圈并不直接和外线(即初末态粒子)相连，所以整个过程的动量守恒对圈上的虚粒子的四动量没有限制，其后果是在对所有的贡献求和时会出现令人意外的无穷大，从而使得结果毫无意义。导致这种令人困惑局面的原因是深刻的。表面上看是由于费曼图圈上虚粒子的四动量可以取正负无穷大之间的任何值，而所有交换这些虚粒子的过程都对散射过程有贡献。仔细分析就会发现，这种无穷大的出现是与点粒子的观念直接相关的。

通常人们把基本粒子看做是没有大小的几何点，它是合乎相对论要求的。在相对论的时空观中，有所谓的长度缩短概念，所以不存在刚体，因为刚体是和两点间距离保持不变的概念相联系的。于是把基本粒子看做有限大小的刚体的想法是行不通的。把基本粒子看作大小有限形状不定的物体也存在困难，因为这意味着其内部有相互运动，与基本粒子没有内部结构的观念相违背。但把基本粒子当做点粒子在经典理论中就已经带来无穷大的困难。例如电子的电磁能问题。把电子当做带有一定电

荷的小球，它的静电能将正比于电荷密度。当电子的半径趋于零时，其电荷密度就会变得无穷大，相应的静电能也将成为无穷大，这与实际情况是不符的。

在量子理论里，粒子具有波粒二象性，粒子的动量与对应的物质波的波长成反比。要探测越微小的结构就要用越精细的“尺子”即波长越短的波，也就是用能量很高的粒子。要探明一个粒子是点粒子就要用波长无穷短或动量无穷大的波。对于两个电子的散射，高阶修正带来的圈图上的虚粒子的动量正是可以到无穷大。所以，量子场论中高阶费曼图中存在的无穷大(或称发散)是与点粒子的假设直接相关的。

在量子场论中，描写粒子是采用定域的场量。所谓定域场量指场量是连续时空的连续函数，它在每一点上的值只和该点有关。所以连续的定域场实际上认为基本粒子是理想的几何点。人们可能会问，为什么要用定域场，以及它的微分方程来描述微观客体的运动呢。原因不只一个，但重要的是利用这种理论已经取得了与实验相当符合的结果，并且人们尚没有找到可以避免微分方程的理论形式。

研究已经表明，在对物理的量子场论模型进行微扰论的计算时都会出现上面谈到的无穷大。重整化理论就是为了要发展一套办法处理这种发散从而得到合理的、有限大小的高阶修正值，使微扰论在理论和实际应用两方面完善起来而建立的。重整化理论是复杂的，计算十分繁难，概念上也存在一定的问题，但由于在实际应用方面所取得的重大成果，迄今仍是量子场论的主干。

我们通过一个例子来阐明重整化理论的主要思路。在相对论量子场论中粒子的产生和湮灭现象是基本特征，所以量子场论从一开始就是一个多粒子体系。如前所述，在量子场论中，真空不空(狄拉克的电子海)，带电粒子可以使其产生极化效应，所以我们测量到的粒子的电荷不是它本身带有的电荷值(称为裸电荷)，而是把极化电荷的屏蔽效应包括在内的“穿了衣服”的电荷(重整化电荷)。例如，一个真空中的电子有如费曼图 3.5 所示的过程：

前面我们已经提到，圈图修正将会导致无穷大，它意味着真空极化带来的“屏蔽”效应是无穷大。那末如何解释我们实际测到的有限大小的电荷值呢？一个自然的解释就是，裸电荷本来就是无穷大，它与同样是无穷大的极化电荷相消使得测量到的电荷(重整化电荷)具有有限大

小。

上面的讨论提示人们要仔细对待理论中引入的参数和参量，如耦合常数、质量和场量本身的大小等，这样一些本身没有直接物理测量意义或无法直接测量的量在理论的重整化中有着特殊的作用。所谓的重整化就是把微扰计算中出现的无穷大吸收进这些参数中，象对裸电荷那样假定它们具有适当(往往是无穷大)的裸值以抵消圈图修正的发散从而给出与实验测量相符合的有限结果。它相当于把物理的振幅用测量到的质量和耦合常数来表达。并不是任何一个量子场的理论都可以把发散吸收到参数中去的，能够做到这一点的称为可重整化的理论，否则称为不可重整的。能否重整化已经成为一个量子场理论正确与否的重要判据。具体讨论一个理论能否重整化，首先要分析理论中费曼图的结构，把基本的发散图找出来。例如，对于量子电动力学，基本发散图只有图 3.6 所示的三种，它们分别对应为耦合常数的重整化，电子质量的重整化和光子场(波函数)的重整化。其它的发散图要末在结构上等同于它们，要末包含了它们。所以量子电动力学是可以重整化的理论。

重整化方案具体的实施首先要设法处理那些对圈上虚粒子动量的积分，这些积分由于动量值可以(从零)取到无穷大而发散，称为紫外发散(因为大的动量相应于短的波长，在光谱上对应于紫光)。将这种积分在形式上积出来进行处理，称为规制化(或正规化)。最简单的规制化办法是将动量积分“截断”，即假设动量最大取某个有限的，把积分算出来。然后把随 趋于无穷大的项通过重新定义裸质量和裸耦合常数吸收掉。

实际的作法是在作用量中相应地引入一些抵消项，这些项的系数是可与那些随 趋于无穷大的量相消的无穷大。如果所有这样的抵消项的场量组成形式在原来的作用量中都已经存在，那末就可以把相同的项合并在一起，最后的作用量在形式上和原来完全一样，只不过现在的作用量中的参数和参量是包含了无穷大的，它们刚好抵消了微扰计算圈图带来的发散。所以，重整化相当于对理论中的参数和参量进行重新定义。如果理论是可重整化的话，发散最终将能够被抵消掉，具体截断的值或办法并不影响也不应该影响重整化后的结果。事实上，“截断”只是处理问题的手段，理论本身不应依赖“截断”。

一个量子场理论主要的发散来自于紫外发散，这是根本性的困难。

有时，当虚粒子是质量为零的粒子时，积分会因为粒子动量取零值而趋向无穷大，称为红外发散，通常它与电子不断发射的很低能量的光子有关。由于能量小于仪器分辨率的光子探测不出来，当这样的光子的产生振幅与红外发散的散射振幅合在一起时，红外发散可以消除。

前面已经指出，规范对称性决定了体系的相互作用方式，这样建立的理论都是可重整的，这是规范理论成为基本相互作用理论的一个重要原因。规范对称性对于可重整化至关重要。由于规范对称性，发散积分之间存在一定的关系称为沃德(Ward)恒等式。通常重整化导致的参数重新定义不应该改变理论的形式，这是可重整化的含义所在。然而也会出现这样的情况，理论原来具有的对称性在重整化之后被破坏掉了，这就是所谓的反常。它通常意味着为消除发散而引入的抵消项形式并不都是原来作用量中已有的，如果被破坏的对称性是一般的整体内部对称性，反常并不会给理论带来特别的困难，有时这种反常甚至是必要的，如介子到两个光子的衰变过程。而如果这样的抵消项破坏了重整化的作用量的规范对称性，就会相应地导致沃德恒等式不再成立，从而使得理论不可重整化。

80年代以来，人们通过对场在大范围的性质的研究发现，一个量子场理论是否存在反常与场作为时空的连续函数在进行某些连续变换时的性质有关。这种连续变换称为拓扑变换，场在拓扑变换下的性质称为拓扑性质。研究还发现，这种拓扑性质与理论中费米场的情况有关，适当地安排费米场的数目以及它们在规范变换下的变换性质可以改善场的拓扑性质从而消除反常。目前描述基本粒子及其相互作用的量子场理论即标准模型对于夸克(它们是费米子)的数目有相应的限制，通过实验寻找夸克，看它的数目是否与这种限制一致是对标准模型正确与否的一个重要检验。顶夸克的发现使标准模型所需要的6个夸克全部找到了，这是对标准模型的巨大支持。

3.2 统一理论

自然界的现象丰富多彩、千变万化，但是物理学家一直怀有一种信念，认为这些复杂现象背后的规律是简单而优美的。自从人们认识到自然界存在四种基本相互作用以来就发现了关于这些相互作用的理论存在某些共同之处，于是建立一个能统一地描述四种相互作用的理论成为人们努力追求的梦想。虽然能否最终实现这个梦想尚属未知，但是目前沿这个方向的研究已经取得了非常令人鼓舞的成果，使寻找统一理论的信念更加坚定了。

3.2.1 对称性的自发破坏

规范相互作用原理的主要优点除了可重整外，还在于提供了把各种基本相互作用的机制统一起来的数学框架。在杨-米尔斯场基础之上建立的弱电统一模型就是很好的例子。但在建立这个理论时人们遇到了一个很大的困难。由于规范对称性的要求，规范场的作用量中不可能出现规范场的质量项(或规范场的平方项)，它将破坏规范对称性。如果电磁作用和弱作用都是以规范相互作用原理为基础的话，传递它们的规范粒子就应该都是零质量的。然而实验表明，光子确实无质量，而弱作用是一种力程很短的力，传递这样的力的中间玻色子的质量非但不为零而且必定很重。要把这种中间玻色子当成是规范粒子必存在如何既给规范粒子以质量又不破坏规范对称性的困难问题。

在前面的讨论中我们提到对称性总是指作用量在某种对称变换下保持不变。其实，整个理论的对称性还和它的状态，特别是基态有关。通常体系基态(也就是能量最低态)与作用量有相同的对称性。但是，存在这样的情况：体系基态具有的对称性比作用量要少，于是整个体系表现的对称性降低了。我们说作用量的一些对称性被自发地破坏了。这种情况往往发生在体系的最低能量态是一些简并态，基态只是其中之一。图 3.7 所示的一个简单例子可以形象地说明这一关系。

一个小球很小心地放在“墨西哥”帽的顶上。这时整个系统与作用在它上面的重力具有相同的绕过帽子中轴的转动对称性。但是这样的结构显然是不稳定的，如果小球沿帽顶滑下最后落在帽沿的某个位置，其重力势能最低，是能量的最低态，达到稳定平衡。然而这种位置破坏了原来的转动对称性。小球落在帽沿的具体什么位置纯粹是随机的，任何位置都是势能最低(是简并的)。重要的它总是使作用在体系上的重力所具有的转动对称性被破坏了。这种情况就是所谓的对称性自发破坏。

正是基于这种考虑，温伯格和萨拉姆认为弱中间玻色子的质量是通过自发地破坏规范对称性而获得的。理论的规范对称性只不过是隐藏起来了。为了实现对称性的自发破坏他们引入了额外的希格斯场，它通过与规范场的恰当耦合而给出类似“墨西哥”帽那样的相互作用势能。系统的基态对应于弱玻色子获得质量。它不仅克服了原来理论的不可重整化等困难，还很好地解释了低能下弱作用与电磁作用在强度上的差别。

关于弱电统一理论的粒子物理方面的内容及与实验的比较请参看本书粒子物理部分。这里要提到的是，从量子场论的角度看尽管弱电统一理论对这两种相互作用给出了同样的作用机制——规范作用原理，但是理论上没有给出温伯格角的值。或许对于这个理论更准确的说法是把弱电拼在一起而不是统一在一起。不满意于理论的这种局面促使人们去寻找进一步统一的理论。

3.2.2 重整化群及标度变换

重整化理论告诉我们，在构造理论时作用量中出现的参数或参量并不就是实际测量到的物理值，这是一个非常重要的观念。重整化理论是为消除量子场论的微扰计算中存在发散而发展起来的，但是现在看来，重整化是将理论计算用于同实验测量相联系的必要环节，而并非仅仅是因为理论计算中存在发散才引入的。举个固体物理的例子可以说明这一点。考虑在固体晶格中运动的一个电子。由于受到带正电的晶格的背景作用，电子的有效质量 m 与它在固体之外的质量 m_0 有所不同。换句话说，由于固体背景的作用，电子的质量从 m_0 重整化为 m^* 。通过将电子拿出或放入固体中，我们原则上可以测量到这两个质量。显然，因为 m_0 和 m^* 都是可测的并且是有限的，所以它们之差也是有限的。在相对论量子场论中，情况与此非常类似但存在两个基本的差别。首先，裸量(对应 m_0)与物理量(对应 m^*)之间的差是无穷大，这表现在圈图发散上；其次，量子场论的相互作用从一开始就存在，我们没有办法象把电子拿到固体之外那样把它除去，这就是说，相对于 m_0 的裸量是无法测量的。量子场论中背景作用来自真实，真空不空而且象是一个介质(相当于例子中的固体)无处不在，我们不能设想将电子“拿”到真空之外去！于是我们完全有理由假设裸的质量和电荷为任何可能的值(包括无穷大)，真空的背景作用(相应于微扰论中的圈图修正)刚好抵消了值为无穷大的裸量从而得到有限的测量值。

由于重整化并不因为存在发散才变得必要，所以重整化的修正并不一定都是无穷大。事实上，存在有限大小的重整化修正。仍以电子在固体中的运动为例。晶格的背景作用给电子对外力的响应带来的影响是同电子的运动状态有关的，所以重整化的质量或其它量依赖于粒子的运动状态，即依赖于粒子运动的能量标度。不同标度下重整化所得到的量(有效的量)是有限大小的量，它们之间的差别显然也是大小有限的。从一个标度下的有效量到另一个标度下的有效量通过重整化联系起来。所以，重整化给理论参量带来的改变可以看成是对它们的一种变换，所有这种变换的全体构成群，称为重整化群。它反映了体系在标度变换下的行为。

标度变换简单说就是尺度变换。在粒子物理中，很早就有这样的想法，即如果理论中粒子的质量为零并且不存在其它的有量纲参数的话，

那末标度变换下，量子场论的作用量保持不变。标度变换下不变对于理论是一个重要的性质，其中著名的是关于凝聚态物质相变的临界行为理论。在临界点附近，涨落的关联长度趋于无穷大，由有限晶格常数引起的所有效应都被抹平，这里不再有特征的尺度，于是热力学函数将具有标度不变性。

但是，通常的情况下粒子的质量可能并不为零而且也不能被忽略掉，这时标度对称性实际上是被破坏的。换句话说，简单的标度对称性所对应的沃德恒等式不再成立，这也是一种反常。取而代之的是重整化群的对称性，这时反常对场量在标度变换下的行为的影响(表现为反常量纲)已经被考虑在内，相应的等式就是重整化群方程。重整化群所反映的对称性其根源在于前面关于重整化理论的讨论中存在的任意性。裸量是不可测量的，关于它的假定带有任意性。如果微扰计算中的发散被某种重整化方案 R 抵消掉，那末存在另外的方案 R' 同样可以消除发散，两种方案之间相差有限大小的重整化。重整化方案的不同导致重整化后理论中的物理参数不同，它刚好补偿因重整化方案不同给物理结果(如散射截面等)的计算带来的改变从而保证了对物理结果的预言不受重整化方案改变的影响。

重整化方案的任意性具体表现在两个方面。除上述重整化参数与物理参数联系起来的关系上存在的不定性(所谓归一化条件的不定性)之外，另一个是减除点的任意性，这个减除点就是将发散积分做规制化处理时的截断点。物理的结果不应依赖于截断点的位置是重整化群方程的实质所在。重整化群方程给出了当截断点变动时重整化的参数(如重整化耦合常数，重整化质量)的相应变化规律以保证用它们表达的物理结果(如散射截面等)保持不变。对于两个粒子散射的典型例子，截断点是交换虚粒子的最大动量值，相应于两个散射粒子在散射过程中靠得最近时的距离。重整化群方程的解给出了耦合常数随这个距离的变化的函数关系，这种关系反映了真空作为“介质”对耦合常数(即荷)的屏蔽效应。非阿贝尔规范场的反屏蔽效应使得耦合常数随着夸克间的相互靠近而变小，这种性质叫做渐近自由。理论是否渐近自由是由重整化群方程的紫外不动点来决定的。

一个定域的量子场论能否自治，从重整化群角度看完全由其紫外稳定不动点所决定。但有时即使存在这样的不动点从而足以保证自治性，

却并不意味着理论在通常的幂律或渐近自由的意义上是微扰可重整的。例如，在 $2 + \epsilon$ 维的 Gross—Neveu 模型中，当 $\epsilon > 0$ 时理论是不可微扰重整化的，但是只要 ϵ 足够小，理论是自洽的。另一方面，微扰可重整有时并不足以保证理论是截断无关的。有一个截断无关的理论在数学上自然是美妙的，有时甚至是值得庆幸的。但是，这不等于说那些依赖于能标的有效理论就应该摒弃。实际上，有效理论也是极为有用的，至少可以预言新的物理将会出现的阈能。衰变的四费米子理论就是一个很好的例子。另一个重要的例子是量子色动力学在低能情况下的有效理论。如何构造一个有效理论在很多时候并非是轻而易举的事，一旦得到合适的有效理论将会给人们很多很好的启发。

重整化群，有效拉氏量以及渐近自由等概念的发现和引入使得人们对重整化的物理认识进一步深化，重整化不再仅仅是消除计算发散的一种技巧，而是物理相互作用随所研究现象的标度变化的体现。重整化群方法已经成为量子场论中的重要方法，它不仅能给出耦合常数的渐近行为，而且结合算符乘积展开对于一些深度非弹性散射过程的计算十分方便和重要。当然，重整化群方法还在统计物理中关于相变临界现象的研究方面发挥重要作用。重整化群理论的研究对于构造大统一场论也是非常重要的。

3.2.3 超对称与统一理论

规范理论的成功鼓舞人们在这个方向上继续探索。如本书粒子物理部分所述，在夸克模型的基础上发展起来的强作用理论也是一种规范理论，称为量子色动力学。相应的规范群是 $SU(3)$ 。强作用的规范粒子称为胶子。胶子本身带有颜色，并且是双色的。三种颜色两两组合共有 9 种双色，由于真正改变颜色的有 8 种，对应地有 8 种胶子。既然胶子本身也带有颜色，所以胶子之间也发生强相互作用。胶子象光子一样没有静质量，因此量子色动力学的 $SU(3)$ 规范对称性是没有自发破坏的。将 $SU(2) \otimes U(1)$ 电弱理论和 $SU(3)$ 强作用理论合在一起构成了我们现有的关于微观基本粒子物理的理论——标准模型。

量子色动力学理论的特点是存在渐近自由。胶子带有色荷使真空在强作用之下发生“反屏蔽”极化，导致等效色荷随夸克之间距离增大而增大。于是夸克之间的强作用随距离增大而变强，随距离减小而变弱。我们知道，小距离对应大的能动量。所以，耦合强度随距离的这种改变可以更方便地表示成能量的函数。由于 $SU(3)$ 规范理论中有较多的带有色的规范粒子，所以强相互作用的渐近自由效应要比其它相互作用更明显，它超过了夸克的屏蔽效应。对于 $SU(2)$ 弱相互作用，两种效应旗鼓相当，而 $U(1)$ 的电磁作用中没有规范粒子的反屏蔽效应，所以只表现出通常的屏蔽效应。这样一来，强作用的强度随能量减弱，弱作用的强度基本不变，而电磁作用的强度增强，于是有可能在某个能量尺度上三者汇合在一起！（见本书第二篇）从重整化群的计算可以得出耦合常数随能量标度的跑动是对数的关系，这种变化是比较缓慢的。所以，如果它们确实能够汇合的话，这个能量标度也远远大于目前加速器所能达到的尺度。但是无论如何，规范作用原理给这三种基本相互作用提供了相同的作用机制，并且耦合常数的跑动方向也是与存在统一理论的假设相一致的。

人们对于一个统一理论的基本要求包含了这样几个方面。首先，它应该是一个规范理论，相应的规范群应该是单群（而不是几个群乘在一起）以保证有单一的耦合常数；其次，它必须给出与实验结果相符合的低能行为。在大统一理论的诸多模型中最有影响同时也最简单的是 $SU(5)$ 最小耦合模型， $SU(5)$ 群是能够包含 $SU(3) \otimes SU(2) \otimes U(1)$ 的最小单群。三种

相互作用在这个理论中统一为一种，由 SU(5)群的规范粒子传递。从量子场论的角度讲，除了群更大更复杂外，这个理论与标准模型的场论并没有什么本质的不同。如果说有什么特殊之处的话，就是存在所谓的规范等级问题和关于质子衰变的预言。事实上，最近的实验结果已经宣告SU(5)大统一的失败。

寻找更大的规范群并不能使困难得到解决，必须引入新的思想，这就是超对称。超对称的基本变换操作就是把一个粒子或它相应的场转变成另一个与之自旋相差 $1/2$ 的粒子或场。也就是说，超对称把费米子变成玻色子，把玻色子变为费米子。到目前为止并没有迹象表明自然界存在超对称，所以，超对称如果存在也必定是被破坏掉的。但是我们已经清楚，一种对称性即便是被破坏掉了仍然可以有不可忽略的作用，只要这种破坏是自发的。当把超对称规范定域化之后，人们发现爱因斯坦的广义相对论就成为理论的必然结果，从而第一次有了将自然界的四种基本相互作用统一到一个理论框架中的可能性。这样发展起来的理论称为超统一场论。

超对称引入了大量额外粒子，粗略地讲，它至少使粒子数目增加一倍；超对称把任何粒子转变为相应的伙伴粒子，这些粒子和原来的粒子自旋相差 $1/2$ 但却有着相近的耦合方式和强度。这些超对称伙伴粒子的主要作用是提高了跑动耦合常数交汇点的能量标度而又不破坏大统一理论所要求的关系。另一方面，由于统一的能量标度提高了，质子的寿命变长，可以满足质子寿命的要求。

在弱电统一的理论中必须引进带弱荷的复双重希格斯场。但唯象分析表明不可以引进带色荷的希格斯场。因为它们可以导致质子衰变，除非这些粒子的质量足够大否则相应的质子寿命会短于实验定出的下限。没有什么先验的理由能说明为什么带强色荷的希格斯粒子比带弱荷的希格斯粒子重那么多，这个问题是所谓的规范等级问题的一个方面，它给建立超对称的希格斯场的表示时增加了困难。此外由于技术的原因，在最小耦合的超对称模型中必须引入两个弱的双重希格斯场。尽管所有的希格斯粒子以及它们的超对称伙伴粒子对跑动的贡献并不及夸克、轻子和规范粒子的贡献大，但最近的实验精度已经足以分辨出希格斯粒子对跑动的修正。结果表明最小耦合的超对称模型与实验符合得非常好而普通非超对称的模型则无法与实验吻合。

最小耦合的超对称模型肯定不是实现统一的唯一出路，更复杂的统一方案也能作到这一点。特别是通过引入新的带有奇怪量子数的粒子和更加复杂的对称破坏机制。然而从理论的简和美两方面看，最小耦合的超对称模型是比较受青睐的。

3.2.4 引力场论

到目前为止我们一直没有提到引力，这一方面是因为引力在微观粒子领域的作用非常弱以致于完全可以忽略，另一方面也是因为构造自洽合理的量子引力理论特别困难。迄今为止还没有找到好的解决办法。继超弦理论之后，人们更多地转向讨论一些简单的非物理模型，如 2 维引力理论，希望从中得到一些启发为最终构造一个物理的理论做必要的准备。

现在我们首先简单介绍经典的引力理论。狭义相对论关于时空变换关系的预言揭示出时空是整个物理的一个有机部分而不仅仅是物理过程发生的舞台，它不仅要求放弃牛顿的时空观以及牛顿力学规律而且还要求放弃牛顿的引力理论。牛顿的万有引力理论是关于静止物体的引力理论，根据这个理论，一个物体产生的引力瞬间作用到所有其它物体上。这种超距作用是与狭义相对论相违背的。可以说，牛顿的万有引力定律只相当于电磁理论中的库仑定律。而与动体的电动力学——狭义相对论相当的动体的引力理论就是广义相对论。所以广义相对论实际上是一个满足相对论性要求的引力理论。

引力场的概念可以象电磁场一样自然的建立，问题在于如何得到正确的引力场方程。引力与电磁力的不同之处在于两点。首先，引力更普遍：所有的物体都会产生引力场而同时所有的物体都受引力场的作用。其次，引力作用的“荷”是所谓的引力质量，而实验表明物体的引力质量与它对外力响应时表现出的惯性质量大小相等。这两点赋予了引力特殊的地位。由于引力的普遍存在，撇开引力去谈论时空就变得毫无意义，也就是说整个时空中本来就已遍布了引力线(电力线概念的自然推广，是一种形象的说法)，而由于引力质量与惯性质量相等从而可以从动力学方程中消去，于是所有的物体在引力场中的运动就与这个物体的特性(即它的质量)无关，这样一来，引力线的分布就完全等同于时空的结构了。引力线一般讲来象电力线一样是弯曲的，因而可以完全等价地说时空是弯曲的。两个离得很近的物体在自由下落过程中不断地慢慢靠近趋向地心。(如图 3.8 所示)可以将这种情况归因于地球的辐射状引力线，但也可以等效地认为地球以及它产生的引力场并不存在，两个物体的靠近是由于空间并非平直的所造成的。这称为等价原理。

我们通常讨论的时空坐标是直角坐标，坐标轴是两端无限的直线。用这样的—个坐标系来代表弯曲的时空显然是不行的，正如—个二维的平面直角坐标无法代表—个球面—样。这说明在时空弯曲的情况下整体直角坐标做不到。在考虑引力场之前，理论是在惯性系中表达的，它们要满足狭义相对论就是要在不同的惯性系中表现出相同的规律。原则上讲，那时的坐标总是可以取做整体的，即代表整个时空。由于引力的存在，所有的物体都会受到引力的作用，所以，惯性系并不存在。换句话说，无法用—个直角坐标系来代表整个时空。但是，经验告诉我们，当空间弯曲并不严重时，我们仍可以在—定的有限大小的区域内利用直角坐标。比如地球是圆的，但在较小的范围内使用直角坐标系仍是一个很好的近似。

直角坐标的特点是，两点之间的最短距离是直线，距离的平方等于两点的坐标差的平方和，即

$$ds^2 = \sum_{i,j=1}^3 dx^i dx^j g_{ij}$$

其中 g_{ij} 称为度规，在直角坐标系下它就是个对角元素为 1 的单位矩阵。当空间弯曲时，两点间可能并没有直线，比如在球面上。空间的弯曲反映在度规上就是，它不再是单位矩阵。因为各时空点的弯曲情况不同，所以度规的元素是时空点的函数。尽管如此，在某个有限(严格讲应是无穷小)的局部仍有可能通过选取适当的坐标系使度规在这个局部内是单位矩阵。这就是说，弯曲空间在局部仍可以看成是平直的，这和通常的经验—致的。

空间的弯曲代表了引力线的分布情况，也就描写了引力场的分布，所以我们首先找到了描述引力场的参量-时空的度规-引力势。它相当于电磁场中的电磁势。余下的工作是找到决定引力分布的“麦克斯韦”方程。描述弯曲空间以及其上函数性质的数学称为黎曼几何。黎曼几何告诉人们如何处理弯曲空间上的函数及它们对空间坐标的微商在坐标变换下的性质。利用黎曼几何提供的手段，爱因斯坦得到了引力场的“麦克斯韦”方程。这个方程就是著名的爱因斯坦场方程。

$$R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} = k T^{\mu\nu}$$

方程的左边是引力场在时空的分布，由度规函数对时空坐标的协变微商

代表；右边是产生引力场的物质分布情况。这个方程是满足相对论要求的，即在(局部)坐标系的变换下方程的形式保持不变。这时的协变性具有了新的含义。

广义相对论是不是一个最终正确的理论，还是一个有待回答的问题。已经明确的是，它在概念方面是有用的，它能够给出一些定量的判断，诸如水星近日点的进动、光线在引力场中的弯曲等。除了定量方面的成就外，广义相对论还预言了一些值得注意的现象，它们在宇宙的结构和动力学演化上可能会起一个关键的作用。其中最为突出的是预言了黑洞的存在。这是说，当一个物体的密度足够大以至其引力半径超过了物质半径时，物体的外部空间存在这样的区域，在那里光和其它粒子都只能单向地落向引力源，而不可能静止或是向远离源的方向运动。人们将这种特殊的时空区域叫做黑洞。尽管人们迄今尚未直接找到黑洞，但有许多理论上的原因和观测上的迹象使得人们相信它的存在。

由上述可见广义相对论事实上是一个引力理论，也是一种把运动方程表示成与坐标无关(即协变)形式的理论。从等效原理的观点来看，相对论与引力论是两个不可分割的概念。因此一个满意的理论必须能同时体现它们。这是引力理论的基本意义所在。然而，在大多数的天体物理和天文问题中，物质密度从大范围来看是如此之小，引力是如此之弱，以致可以忽略它的存在。在对原子尺度以下另外三种相互作用的讨论中我们也是这样做的。但从理论的角度来看，我们必须考虑将引力包括进来后现有的整个理论的框架是否还靠得住。那末，首先遇到的就是引力理论的量子化问题。

引力量子化的一个基本问题在于缺乏作为背景的几何空间；时空本身的几何在这里已经具有了动力学的含义。这一点正包含在理论的广义协变性之中。在标准模型中，时空作为一个固定的几何结构存在着对称性，这种对称性由一个有限维的群——彭加勒群所描写。而在广义相对论中这样的群却是无穷维的——时空的微分同胚群——它是所有满足爱因斯坦场方程的时空结构的几何变换群。这一特点将给理论的量子化带来两大难题。一方面，在构造希尔伯特态空间的内积时通常的定义不再适用，这被称为“内积问题”。另一方面，动力学不再由态空间上的几何对称群作用量所决定，这被称为“时间问题”。此外，由于引力子的高自旋将导致理论重整化的困难。80年代以来，一些看来互不相关问题的进展

为找寻引力量子化数学框架的前景带来了希望。这些进展主要来自称为拓扑量子场论和二维共形场论等方面。

3.2.5 微分几何，纤维丛和规范理论

在继续统一场论的讨论之前，我们从数学的角度来简单地看一下引力与其它三种力在描述上存在哪些共同之处，这对于后面讨论它们的统一是有好处的。它们的共同之处首先表现在都是用场来描述的。场是时空坐标的连续函数(在这里我们只讨论经典的场而不是量子化之后的场)。这些连续函数的图像在数学上用流形这个概念来表达。流形是普通的曲面概念的推广。研究流形几何性质的数学理论就是微分几何。人们发现，数学和物理这两个不同的学科在这个领域的诸多概念上有惊人的共同之处。杨振宁曾写到：“我惊奇地发现规范场正是纤维丛上的联络，而它却是由数学家在完全没有参照物理世界的情况下发展起来的。”

微分几何的出发点是微积分。微分对应于曲线的切线，而封闭曲线所包围的面积的理论就是积分论。微积分在几何上的应用演变成曲线论和曲面论。高斯关于内蕴几何的重要发现揭示出，如果要判断一个曲面是否弯曲无须从更高维的空间里去整体地观察而只须在局部通过适当地测量角度和线段长度即可得出结论。换句话说，高斯的曲率概念是一种局域的概念。比如，人们不需要环地球一圈或是从太空拍摄的照片就可以知道地球是圆的。这种环地球一圈或是从太空观测得到的是地球的整体面貌，是曲面(地球表面)的整体结构，是拓扑学所研究的范畴。球面在局部凹陷或隆起一些并不会改变它在整体上是球面的性质。从拓扑角度看，即使是在局部形变的情况下，它仍不同于平面或轮胎面。拓扑学与微分几何的联系是所谓的整体微分几何，主要是通过空间各处局部的测量得到关于空间整体的拓扑性质。这方面最著名的一个例子是 Gauss-Bonnet 定理，这个定理是说，将曲面在各处的高斯曲率积分得到的量是一个拓扑不变量，事实上，这个量是 2π 的整数倍。对于一个球面来说，无论如何凹陷或隆起，这个量都等于 4π ，对于轮胎面，它等于零；而对于图 3.9 的双孔洞的胎面，则为 -4π 。

辅助空间在对流形的研究中起着重要的作用。一个熟习的例子是曲面上各点切平面构成的空间；对于整个空间平直的情况，我们可以在全空间取唯一的直角坐标，空间里的任何点到原点的有向线段构成一个向量，向量的坐标分量在坐标系的转动下按照一定的规律变换以保证向量本身在空间不变。对于空间是弯曲的情况，比如曲面，不可能在全空间

取到一个直角坐标；但曲面上每一点的切平面仍是一个维数与曲面相同的平直空间，这个空间无限延展开去，当然已不再同于要研究的曲面，但在与曲面相切点无穷小的区域内，它与曲面吻合得足够好。对于流形，坐标只能是局部坐标，它本身并没有意义，倒是对于那些在局部坐标的变换下保持不变的性质(如切向量)的研究揭示了空间本身的几何与拓扑性质。流形上每一点的切空间是一个向量空间，从流形的每一点的向量空间里按一定的法则取一个向量，它们整体上就构成了一个向量场。除了向量空间外，还可以有其它的辅助空间，比如切平面上过切点的半径为 1 的单位圆，或是过切点的法线等。辅助空间与原来的空间(底空间)一起称为纤维丛。纤维指的是这些辅助空间——切平面、单位圆、法线或是其它的什么，而丛指的是把各点的纤维联在一块儿。直到 30 年代微分几何还主要是研究切空间构成的纤维——称做切丛，以及由它通过张量乘法得到的张量丛。后来才发展为讨论更一般的与曲面完全无关的纤维。

至少在局部看来，纤维丛是由底空间与纤维空间相乘得到的(图 3.10)。比如，当底流形是一个圆环，而纤维也是一个圆环时，相乘得到的纤维丛就是一个轮胎面。这样的乘法在全空间是一样的，称为平凡的。通过考查乘法的实现过程可以看到存在得到非平凡丛的可能。胎面的构造过程是将做为纤维的圆环沿底环移动一周。在最后相接的地方存在各种可能性。其中一种就是把最后的一个纤维环与头一个纤维环不经任何转动地点对点合在一起，这样得到的纤维丛就是所熟习的轮胎面。另一种连接的办法是将最后的纤维环翻转 180 度然后再和头一个纤维环合起来，就象把一个纸带翻转一次对接成莫比乌斯(Mobius)带那样。这样得到的纤维丛是一个克莱因瓶(见图 3.11)。

关于纤维丛以及它们的不变量的分类问题在 30 ~ 40 年代由 Whitney, Stiefel 和陈省身等人大大向前推进了。他们发现了一些非常有趣的积分表达式，这些表达式是高斯-Bonnet 公式的推广。到了 70 年代，这些整体不变量在物理学中不断地涌现。为了说清楚其中的原因，我们回过头来进一步看看关于弯曲的本质含义。

由于弯曲，流形上的坐标架只有局部的意义。流形上某一点的切空间里的向量因此也只对于该空间而言具有向量行为。不同点的切空间相互没有关系，是各自独立的。前面提到流形上的向量场，在流形的每一

点的切空间中取一个向量。不同点上的向量之间也是无关的。相对论性要求物理的方程在坐标变换下保持协变的形式，所以物理的量通常由向量描写，物理方程是这些向量的微分方程。在考虑引力之前，时空是平直的，各时空点的切空间全部重合在一起，它们中的向量有相同的变换关系。考虑引力之后，时空不再是平直的，时空各点的切空间不再重合，而是原则上相互独立的。这样就带来一个问题，即如何定义向量场关于时空坐标的微商。微商涉及两邻近点上向量的差，但由于两点的向量空间各自独立，两点的向量之差不再是向量，所以简单地按通常办法得到的微商不再具有向量的性质，这会破坏方程在坐标变换下的协变性。为此，必须设法使不同点的向量发生联系，比如把一点的向量“平行”移动到另一点，然后和那一点的向量进行比较。把不同点的向量联系起来的关系称为联络，在流形上各点的联络构成流形上的函数。有了联络就可以定义协变的微商。联络赋予了流形一定的几何结构。在流形上给出度规函数会自然诱导联络，同时，流形的形状就更为“固定”了。于是我们就可以讨论流形的弯曲情况。

一个观察者携带坐标架和时钟沿一个闭合的围道平行移动一周回到出发点与留在那里的另一个观察者的坐标架和时钟进行比较，坐标架之间的差别代表了由围道包围起来的曲面的曲率。曲率是由联络的微商构成的。按照广义相对论，时空的曲率就是代表了引力场的场强。引力场的这种几何解释从美学的角度看非常诱人，人们希望电磁场也存在类似的情况。电磁场的确也可以解释为曲率，它是四维时空上一个环丛的曲率。电子场是四维时空上的旋量(一种特别的向量)场。电子场在每一点上做相因子变换对应为该点上的一个纤维环(模为1的复相因子就取值在单位圆环上)。定域规范变换允许各点的相因子不同，因而纤维丛是非平凡的(弯曲的)。规范不变性要求引入规范场——电磁场。这里，规范场就是一种联络。而场强是由规范场的微商构成的，代表了纤维丛的曲率。

更复杂的规范理论对应为更复杂的相因子，也就是更复杂的纤维空间。这时曲率张量或场强成为矢势的非线性函数，非阿贝尔规范场的特点就源于这个非线性项。规范群描写了这些纤维丛的结构——它们的纤维在首尾相联时转过多少“角度”。于是，纤维丛非常自然地构成了规范理论的数学框架。而四种相互作用也因此具有了统一的几何基础。

为了得到量子场论，首先将经典的场论在拉格朗日作用量形式下表

述出来。对于没有物质场存在的情况，规范场的拉格朗日作用量是电磁场情况的直接推广。由于非线性，规范场分成若干不连接的拓扑类，它们由一个称为拓扑荷的整数标志。在此之前，量子理论中的离散现象来源于算子的本征值，由代表对称群表示的量子数标记。现在，上面的讨论提供了另一种可能，即有些量子数或守恒定律本质上是由于场在结构上有非平凡拓扑性质。物理上进行计算时是对各种场的构形求和的。场分为不同的拓扑类反映出理论的真空结构是非平凡的。量子场论的许多量来自费曼路径积分，积分是对所有的场构形进行的。事实上，因为人们并不十分清楚所谓的对各种可能的场构形做积分究竟意味着什么，所以量子场论的一个基本问题就是如何赋予这个路径积分适当的意义。

一般地，对积分的主要贡献来自作用量的稳定点附近。鞍点法可以用来确定稳定点并且进行近似计算。具体的办法是，将作用量写成场的二次项和剩余项之和。二次项做为高斯型积分可以完全求解，剩余项的贡献做为微扰修正。二次项的结果正是经典场的解，而场论的量子化效应恰由高次剩余项给出，它们相应于有闭合圈的费曼图。规范场作用量的鞍点由欧拉-拉格朗日方程定出，它与自动成立的卞奇(Bianchi)恒等式一起就构成自由的麦克斯韦方程组。对于非阿贝尔规范场，这组方程当然是非线性的。用鞍点法做微扰计算就应先找出作用量的稳定点。人们在四维欧几里德空间找到了这组方程的特殊解，它们对应于一些局部极小值点。这些解具有奇异性，称为瞬子解。通过适当的相因子变换，奇异性可以被“规范”掉；变换后的矢势不再是奇异的了。但是，变换前后的矢势在围绕原奇异点的三维球面上无法衔接。为消除断点，需要在四维球(四维欧氏空间加无穷远点)上构造具有 $SU(2)$ 对称性的纤维丛。通过规范变换使纤维绕过三维球面，这样可以在整个纤维丛上得到没有奇异性的矢势。这些使作用量获得局部极小值的解统称为自对偶解。

自对偶解多于瞬子解。寻找自对偶解可以归结为一个代数几何的问题。整体微分几何与代数几何帮助我们在经典水平上找到了自对偶解，但是这些解以及相应的构形如何影响规范理论的非微扰性质仍是没有完全解决的问题。最近，研究表明 4 维自对偶杨-米尔斯理论的量子化会自然出现所谓的量子群结构，这时取值在规范群上的场构成量子群的矢量表示空间。两个场量乘积的无穷小展开给出了非微扰的豪斯道夫量纲，

这方面的研究尚在进之中。毫无疑问，人们希望纤维丛的几何与拓扑性质与规范理论存在的这种密切关系能够为解决非微扰问题提供有效的帮助。

3.2.6 超弦理论

80年代中期，超弦理论的研究取得了突破性进展。众多的理论物理学家加入到了超弦理论研究的行列。不少人在不同程度上相信，自从发现自然界存在四种相互作用以来终于首次获得了一种量子理论能统一描述所有的基本粒子及它们之间的这四种相互作用，而且不存在通常的大统一场论所无法避免的各种困难。但是也有一些著名的物理学家告诫人们，超弦理论的方向严重地背离了物理学的宗旨从而有可能贻误年轻一代学生，因为尚不存在任何关于超弦理论的实验证据，而且他们认为永远也不会有这方面的实验证据。目前，尽管这种争论仍在继续，但人们更多地采取务实的态度，希望让事实来说话。无论如何，超弦理论做为一种理论模型为我们探索自然界提供了一种新的思路，从这个意义上讲，它的影响也许会比所预料的要深远的多。

这里我们只能对超弦理论做一个非常简短的介绍并指出尚待解决的困难。关于超弦理论的详细论述需要高深的现代数学。事实上，超弦理论的发展带动了这些新的数学分支的发展并且使得粒子物理理论、量子引力理论和某些新兴现代数学理论有机地融合在一起。

在超弦理论之前，人们按照传统的点粒子场论来处理引力的量子化理论，这种方法不可避免地要产生无法重整化的发散。问题主要来自引力理论在非常小的尺度下的奇异量子行为。在 10^{-35} 米的小尺度上，粒子的引力(Schwarzschild)半径变得与康普顿(Compton)半径相当的时候，真空态中充满了虚的黑洞。所以，在关于引力的点粒子场论方案中，时空本身在普朗克尺度将开始塌陷。超引力理论的引入并没有能够扭转局面。将超对称定域规范化，就会自动地得到一个包含引力在内的理论。这是因为洛伦兹群的定域化势必导致坐标架的局域化及广义坐标变换下的协变性，而按照爱因斯坦的等效原理，这完全等效于考虑了引力的效应。一般而言，超对称的引入会减弱点粒子量子场论的无穷大发散行为。在定域化之前，超对称使得人们头一次有可能构造一个无需重整化就已经是有限的规范理论。但是希望通过定域规范化同样得到一个有限的量子引力理论的设想并没有能够实现。由于普朗克尺度下真空的涨落，点粒子的超引力理论存在无法重整化的发散。

弦理论之不同于传统的量子场论在于假定物质的基本结构不是点粒

子而是弦——一条一维的曲线。它的特征尺度是由普朗克长度 L_p 和普朗克质量 M_p 代表

$$L_p = \left(\frac{hG}{c^3} \right)^{1/2} \cong 1.6 \times 10^{-33} \text{ cm},$$

$$M_p = \left(\frac{hc}{G} \right)^{1/2} \cong 1.2 \times 10^{19} \text{ GeV} / c^2$$

在这样的尺度下，弦理论与点粒子理论明显不同，但在较大的尺度或更低的能量标度下，由于人们“感觉”不到物质基本单元的这种弦结构，于是可以近似地用点粒子理论。正是由于这种弦结构存在于如此小的尺度(或等价于极高的能量标度)，所以人们担心或许永远无法证实弦理论。按现在的水平即使每 10 年增加一个数量级我们也至少要等上 200 年才可能在实验室的加速器上达到普朗克能量标度。但是，普朗克能标是与任何包含引力的统一理论相联系的，所以人们相信超弦理论的特点一定在低得多的能量尺度下有所反映，正如 SU(5) 大统一理论会有相应的低能有效理论一样。超弦理论的低能有效理论应该和目前人们在这个能标下的正确理论——标准模型有基本相同的行为并且可以克服标准模型现有的困难。事实上，下面我们将看到，超弦理论距离达到这个目标还有相当漫长的路程。

正如本书粒子物理部分指出的，标准模型取得了巨大的成功，但是也存在不少的问题。人们期望这些问题将由一些更基本的原理来回答。所以对于标准模型的扩充一直是人们关注的事情，何况标准模型并不完全——它没有包括引力。大统一、超对称以及额外的空间自由度这些概念是人们对标准模型的扩充中发展起来的。尽管这些概念能够在通常的规范场论中实现，但是它们全部可以自然地吸收进弦理论中。

弦理论最初出现在粒子物理中是在 60 年代后期。当时维涅吉阿诺 (G. Veneziano)，南部等人提出弦模型，用来解释实验上发现的强子共振态质量和自旋的如图 3.12 的关系。它称为雷吉 (Regge) 轨迹。该模型认为强子如介子由一个夸克和一个反夸克通过色力束缚而成，可以形象地用夸克和反夸克之间的色力线代表。由于色力很强，这些色力线在胶子-胶子作用下形成管状并且携带有能量。夸克的质量与之相比可以忽略，于是强子的质量主要来自这个色力线构成的管。好比橡皮筋两端系着小球，如果橡皮筋粗壮有力，而小球很轻的话，人们就容易只感觉到

橡皮筋而忽略了小球的存在。将强子的质量分布在弦上可以算出质量与自旋的关系。刚好解释雷吉轨迹。但是同时预言存在一个质量为零自旋等于 2 的粒子。强子谱中并没有这样的粒子。70 年代中期标准模型的成功，弦模型已被遗弃时，舍科和施瓦兹指出若把这个自旋 2 的零质量粒子用来描写引力场的量子—引力子，则弦的延展性或许将有助于克服以往的引力量子理论中存在的发散。1984 年格林和施瓦兹沿这个方向推进了一大步，构造了一种特殊的弦模型，它具有时空的超对称，因而称为超弦理论。当时空维数等于 10，内部对称群为 $SO(32)$ 时，这个理论不存在反常。

人们或许会问为什么只考虑弦这样的延展客体而不考虑维数更高的如二维膜的理论，事实上，构造延展客体的量子理论，使之满足所要求的么正性、微观因果性并不是轻而易举的事。对于一维弦人们找到了构造量子理论的方案，而且其经典解的粒子谱包含有一个自旋为 2 的零质量粒子，它们在相互作用中的确满足广义协变性从而自然地包含了引力。所有这些都是值得十分庆幸的。但可惜的是这些特性并不能推广到高维的客体。

弦有两种基本的拓扑结构：开弦和闭弦。开弦是两端自由的线段而闭弦是首尾相接的闭合环。弦运动的各种简正模式的量子激发给出了基本粒子谱。这些激发可以有弦的振动和转动自由度，对应到粒子谱上，反映为粒子存在各种内部自由度。在弦理论中，所有的基本粒子都是一个基本弦的不同运动模式而已。弦的运动态中低于普朗克能量的态数目是有限的，对应为可观测的粒子。那些质量与普朗克能量相当或是高于普朗克能量的模式有无穷多，它们很可能是不可观测的。一般说来，它们是不稳定的，会衰变为更轻的模式。

目前知道的有三类自洽的超弦理论。第一类是关于不定向的开弦或闭弦的理论。另外两类基于定向的闭弦，它们的区别在于内部对称群不同。其中一种称为第二类超弦理论；而剩下的那种称为异常弦理论。这三类超弦理论中不存在任何可调的无量纲参数或是其它的任意性。所以，除了这三种选择外，几乎得到了一个完全唯一的包含引力在内的量子理论。现在人们希望进一步缩小候选者的范围。如果能够证明其中两类是等价的而另一类是不自洽的，那末就只有唯一的理论来解释整个基础物理了。这是当前弦理论研究的任务之一。另外，只有正确的方程是

不够的。是方程的解提供了对自然现象的数学描述。一般地，人们先研究最低能量和较低能量的量子态。一个理论很有可能存在不只一个基态，在这种情况下就要根据实验数据来做出选择。对于弦理论情况正是如此。

弦理论的构造与点粒子的情况是类似的。量子场论的微扰展开处理中点粒子的相互作用是由费曼图表示的。点粒子的运动在时空中划出的轨迹称为世界线。世界线的交汇或分开代表了粒子的相互作用。在给定初末态的情况下把所有可能的图的贡献加起来就得到过程的完全几率振幅。可以按照拓扑性质对图进行分类。某个特定拓扑性质的贡献由一个有限维的积分给出。这个积分通常是发散的，但是对于可重整化的理论有办法从中得出有限的结果。弦在时空中运动的轨迹是一个二维面，称为世界面。相应的费曼图是有确定出态和进态的二维曲面。由于相对论的要求，弦的形状不可能是固定的(刚体是不存在的)，因而世界面的拓扑性质才是有意义的。对于第二类超弦理论和异常弦理论，存在单一的基本相互作用顶角。它的费曼图如图 3.13 所示。这类图称为裤子图。代表 T_1 时刻的平面与裤子相交，可以看到有两个闭弦。 T_2 时刻的平面与裤子相交就只有一个闭弦了。很显然，在中间某时刻两个闭弦彼此靠近最后合二为一了。由一个闭弦分解为两个闭弦的逆过程也是允许的。

裤子图所描写的相互作用与点粒子理论中的相互作用在基本的方面存在差异。图 3.14 中画有点粒子相互作用顶角和弦作用的裤子图。在什么时刻两个粒子变成一个粒子了呢？我们用两根虚线分别代表两个不同的观测者所处的洛伦兹坐标系中的等时线 t 和 t' 。可以看到相互作用发生的时空点对于不同的观测者是重合的，换句话说，有唯一的相互作用点。相反地，在弦理论的裤子图中相互作用点为等时面与世界面的切点，它因观测者的不同而不同。对于点粒子的相互作用顶角，作用点的唯一性是“流形”奇异性的体现。各种选择都是允许的，这也是为什么普通的量子场论在构造上存在如此多的自由度的部分原因。弦理论的世界面是平滑的流形，并不存在特殊的点。相互作用完全是曲面的拓扑性质的原因正在于此。这样一来，相互作用的性质就完全由自由理论的结构所决定了。

弦的世界面可以看做为黎曼面，利用复分析的技术来处理。弦理论的基本特点在于由共形映射 $z \rightarrow f(z)$ 联系起来的世界面是等价的，其中 z

是复坐标。所以在对各种构形的贡献求和时只考虑那些共形变换下不等价的世界面。幸运的是，对于每种拓扑构形，其共形不等价类由有限的参数标志，相应的费曼积分是有限维的。对于第二类和异常弦理论而言，费曼图的拓扑分类特别简单。它们由一个称为亏格数的整数来区别。亏格数代表了从球面上连出来的手柄个数(见图 3.15)。初末态弦由黎曼面上的黑点代表，它们共形等价于延伸到无穷远的管。可以看到，亏格数对应为微扰展开中的圈数。弦理论的一个显著特点是，对于微扰展开的每一阶只有唯一的费曼图，而在普通的量子场论中往往会有很多的费曼图。由于这个性质，人们可以利用弦理论的方法更方便地将量子色动力学中的复杂计算大大简化。不但如此，弦理论中与图相对应的积分的收敛性质也要好得多。亏格大于 1 的费曼图的计算仍未解决，这涉及到黎曼面理论、代数几何甚至数论的前沿研究。但是，有迹象表明所有的发散都是我们已经了解并能够处理的类型。例如，任何含零质量粒子的理论都会存在红外发散。我们已经知道如何处理它们。更为重要的是，弦理论中不存在那些导致理论无法重整化的发散。

为什么对于广义相对论进行的量子修正会导致无法重整化的发散，而在低能情况下与广义相对论等效的弦理论却是可重整化的呢？这似乎是矛盾的。事实上，弦理论可重整的主要原因在于普朗克尺度下其行为不同于点粒子场论。特别是弦的那些质量等于或大于普朗克能量的运动模式，它们对应粒子谱中无穷多的粒子。这些粒子做为虚粒子对于散射过程中间态的贡献很巧妙地相互抵消从而改善了费曼积分的紫外行为。这种情况类似于中间玻色子对于四费米子理论的作用。

超弦理论的基本原理的发展以及对于它们的更为几何化的描述已经吸引了一大批人为之付出了不懈的努力。把它的发展与当初广义相对论的情形做个比较将会是富有启发性的。爱因斯坦建立了具有深远意义的等价原理和广义协变性，随后又找到了实现这些原理的恰当的数学语言——黎曼几何。这导致了相应的动力学方程以及对实验的预言。这些预言许多已得到证实。在弦理论中，可以说我们得到了方程却还没有找到做为广义坐标变换和等效原理的推广的那些基本原则。这样的原则一定是存在的，因为广义相对论是弦理论的低能等效理论。无论这些基本原则究竟会是什么样的，它们很有可能要求一种新的几何，或许是黎曼几何在无穷维的推广。当然现在要弄清这些还为时尚早。对于超弦理论的

研究仍在继续，但期望很快就能得到所要的基本原理是不现实的。或许要经过几代人的艰辛工作才可能得到满意的答案。毫无疑问这方面的研究需要数学的进展，同时今后一、二十年里的实验结果也将有助于我们新观念的形成。

3.3 非微扰理论

非微扰理论是量子场论中人们感兴趣的另一个主要课题。一个物理理论即使精确可解，它能否完全代表客观世界仍是个很难回答的问题。但是如果理论只在微扰近似展开的意义上可解无论如何是令人遗憾的。所以，探索非微扰的理论具有多方面的意义。

3.3.1 精确可解模型

微扰论在量子场论的研究过程中始终是基本重要的方法但也因此带来许多令人困扰的难题。所以，精确求解量子场论就成了人们长期以来梦寐以求的愿望。毫无疑问，这个问题的解决对量子场论的研究将起巨大的推进作用。低维场论及精确可解模型的研究正是在这方面的尝试。它们还不是物理的理论和模型，但从中得出的方法甚至结果很有可能对构造和研究物理的理论和模型提供必要的线索和基础。例如，非阿贝尔规范场的瞬子解就被成功地用来解释量子色动力学中的 $U(1)_A$ 问题。

在弦理论中对耦合常数的微扰展开与量子力学中的展开在精神上是一致的。但这样的展开应用到量子场论中的时候，在展开参数的给定阶，微扰的项数比阶数的增加还要快，于是就导致发散。特别是当展开的相邻项符号并不交替变换时，这种发散通常是非微扰效应的表现，如隧道效应。发散还有可能表明真空的选取存在问题。此外在弦理论中决定粒子谱和规范群的真空态往往具有过多的对称性，例如超对称，必须有适当的机制来破坏它以保证理论的低能行为与实验一致，而微扰展开并没有提供这种机制。普林斯顿大学的格罗斯指出，缺乏非微扰的方法是理论进一步发展的巨大障碍。我们生活的宇宙很可能是众多可能的宇宙中的一个。量子引力理论允许不同的宇宙之间交流信息。联系不同的宇宙的时空构形—虫洞物理或许正向人们揭示这样的事实。但只有在非微扰的量子引力理论中才能做到。量子引力的虫洞及其它非微扰效应对于理解为什么宇宙常数趋于零同样可能是必需的。

弦模型和量子引力的精确解问题的研究可以追溯到70年代中关于量子色动力学的研究。特·霍夫特指出，将理论的对称群从 $SU(3)$ 推广到 $SU(N)$ ， $1/N$ 就可以做为一个有用的展开参数。强相互作用的许多突出特点在这种展开的领头项中就已经表现出来。 N 趋向无穷大时领头项的贡献称为平面近似。进一步发展平面近似，导致对一些简单模型的研究。在这些模型中，类似于量子色动力学的 $SU(N)$ 推广，时空的每一点有一个 $N \times N$ 的矩阵。70年代末，人们对于只含一个时空点的简单情况在平面极限下成功地解出了一类矩阵模型。这些解提供了费曼图分类的方便办法。当然它们的意义不止于此。80年代中期，人们进一步指出了可以用这些能够精确求解的矩阵模型来研究弦理论和2维量子引力的可能性。

1988年这一猜测得到证实。普林斯顿大学的理论物理学家们得到了球形时空下2-维量子引力的精确解，这个解与1985年用矩阵模型得到的结果是一致的。

二维时空中引力的量子力学描述表现为对二维时空面的求和。同经典点粒子动力学的最小作用量原理一样，爱因斯坦的场方程对于给定边界条件的解决定了时空的形状或者说拓扑，这样得到的只是引力的经典描述。一般而言，关于引力的量子理论涉及对各种可能“路径”求和。所以有关二维引力和弦理论的基本数学问题是枚举不同形状大小的曲面以及将它们对适当定义的费曼路径积分的贡献求和。

两端自由且没有自身相交的一条曲线可以平滑地变成一条直线，或者说它们拓扑等价。一般地，可以用一条曲线与自身相交的次数来作为拓扑结构的标志。同样，没有边界的曲面拓扑等价于一个带有若干手柄的球面。前面提过，手柄数又称亏格数。最新的精确解将各种亏格曲面的贡献同时加起来。在这种求和过程中，光滑的曲面首先近似地用多面体代替，这叫做剖分。当多面体面数增多边长缩短时就可以无限地逼近原来的曲面。将给定亏格的曲面做不同的剖分对应于曲面的不同构形。所以为求解量子引力或弦理论模型而对各曲面进行的求和需要了解曲面的剖分情况。

前面提到的简单矩阵模型在这里派上了用场。这些矩阵模型的费曼图的对偶正好相应于曲面的剖分。一个图的对偶是这样定义的：将图中的每个面用一个点来代替再把这些点用线连起来，在这个过程中每条线只能与原图中的一条边相交。例如，将一个四边形剖分为二个四边形有两种方式，它们的对偶分别为两个费曼图(图3.16)。一般地，将 n 边形剖分则对偶为有 n 条外线的费曼图；将 n 边形剖分为 m 边形对偶为含 m 个顶角的费曼图。这种对应对于立体图形也是成立的。这种剖分问题早在十八世纪中叶就引起大数学家欧拉(Euler)的注意，到了本世纪60年代数学家们利用图论的方法对其中许多问题给出了解答。现在，通过对简单的矩阵模型的研究使得我们可以用新的、相当普遍而优美的方法得到任何曲面的剖分。

如果矩阵模型中存在 p 点相互作用——相互作用拉氏量中含矩阵场的 p 次单项式，这个理论的费曼图的对偶就是将曲面剖分为 p 面体。而且，正如特·霍夫特1974年证明的那样，模型的 $1/N$ 展开就是按曲面的

亏格展开。对于展开的某个给定阶而言，其贡献来自于在亏格与阶数相同的曲面上笔尖不离开曲面能画出的所有费曼图。展开的领头项来自球面上费曼图的贡献，而球面拓扑等价于一个平面，这就是为什么特·霍夫特将 $1/N$ 展开的领头项近似称为平面近似的原因。一个矩阵模型的 $1/N$ 展开的 $(1/N^2)$ 阶中如果含有 p 次相互作用的话，将会得到把一个亏格为 r 的面剖分成 p 面体的各种可能剖分。

矩阵模型的平面极限只是在一个时空点的情形下进行求解的，最近的一些推广也是这样的。用弦理论的语言来说，这意味着弦张成的面所嵌入的时空只包含一个点。对于早期的解已有了许多的推广。最新的结果已将时空推广到 1 维。象数学物理的许多发展一样，新的解是通过求解微分方程得到的。微分方程来自于将曲面剖分求和的过程取连续的极限。在弦模型的方程中，自变量就是弦的耦合常数，所以方程所表达的是弦的关联函数或真空能量如何依赖耦合常数的关系。幸运的是，所有遇到的方程都属于称做 KdV 方程的一类可以积分的方程。普林斯顿大学的威藤(Witten)认为这些可以精确求解的模型可能与拓扑场论有关。格罗斯则提醒人们注意那些新的解，它们或许与包括十维超弦理论在内的所谓临界弦理论有关。

在过去的几年中，这些不同的方面交互发展。在临界弦(描写 26 维时空的玻色弦或 10 维时空中的超弦)理论中，二维的世界面度规场退耦和；而另一方面，如果试图建立一个非临界弦(在物理时空中运动的)理论，就不得不面对二维量子引力问题。前面的讨论表明，矩阵模型对于处理高亏格曲面的贡献的确非常有效，特别要指出的是，矩阵模型事实上提供了二维量子引力的一种非微扰定义。除此之外，处理 2 维量子引力还可以采用量子刘维(Liouville)方法。为此要用到二维连续时空中共形场论的技术。刘维理论以及它与黎曼面统一描述的关系问题已经有百余年的历史。尽管如此，充分利用刘维理论以便研究临界点之外的弦理论仍是非常有意义的。这些研究向超对称情况的推广也肯定是值得关注的方向。当然这些想法是否有助于理解相对论性非阿贝尔规范场理论在大 N 极限下的性质也还是悬而未决的问题。

量子场论的路径积分形式与统计物理中关于有限温度系统的配分函数有相同的形式。所以，场论的许多方法和概念已经渗透到凝聚态理论的研究中去了。例如，重整化群方法处理临界现象等。同时，统计模型

的研究也促进了场论的发展。

精确可解的统计模型由来已久，如著名的伊辛(Ising)模型。80年代初，对于二维平面格点系统顶角模型的量子反散射方法的研究导致了量子群概念的产生。一个典型的例子是6顶角模型(图3.17)。平面上水平直线和与之正交的纵向直线交点组成格点系统，满足周期边界条件。每个格点的状态由与其相邻的四个有方向的键决定。量子反散射方法(或称代数Bethe-Ansatz)的研究表明，这类模型精确可解的充分条件是要满足称为杨—巴克斯特方程的代数关系。这个方程最早在1967年由杨振宁教授研究有排斥作用势的一维薛定谔多体系统的精确解而得到。一般地，不带谱参数的杨—巴克斯特方程写成

$$R_{12}R_{13}R_{23}=R_{23}R_{13}R_{12}$$

其中R是 $N^2 \times N^2$ 的矩阵，N可以是任何给定的自然数。当N=2时得到的最简单常见的解为

$$R = \begin{bmatrix} q & 0 & 0 & 0 \\ 0 & 1 & q & -q^{-1} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & q \end{bmatrix}$$

由杨—巴克斯特方程的解可以构造相应的能够精确求解的顶角模型。杨—巴克斯特方程所代表的代数关系是李群的结构常数所满足的雅可比等式的推广，它的解所决定的代数是李代数的非平凡的推广，在数学上叫做Hopf代数，物理上称为量子群。它的特点是做为矩阵群的话，矩阵的元素不再是普通的数，而是属于某个代数的“算符”，它们之间的乘法通常是不可对易的。例如，对于上面的解，可以得到量子群 $SU_q(2)$ 的2乘2矩阵

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

满足的代数关系。这种不可对易性由q来标志，q称为形变参数。当形变参数等于1就回到普通的乘法： $ab=ba$ 。群是对称性的语言，所以量子群也应该代表了某种对称性，看来统计模型具有量子群所描述的对称性与模型可以精确求解之间存在着必然的联系。这方面的研究仍在进行。

现在，量子群的结构在许多地方出现，与数学物理的很多分支存在着密切的联系。它同辫子群的关系使得人们可以通过统计模型的解找到

用来区分环和结的拓扑结构的多项式。这种办法得到的多项式比数学家们原来的表达式简并度要低。数学上研究的辫子是用 N 条线去联接两根平行的直线上的 N 个点构成的，这 N 条线在三维空间内彼此不相交，任一辫子是由若干对线的交叉构成。每个交叉称为基本辫子。图 3.18 中的辫子是三个基本辫子的乘积 $b_i b_{i+1} b_i$ 。辫子群 B_N 是 N 个独立的基本辫子 $b_i, i=1, 2, \dots, N-1$ 生成的。它们满足的关系为

$$b_i b_{i \pm 1} b_i = b_{i \pm 1} b_i b_{i \pm 1}$$

它的解与杨—巴克斯特方程的解只相差一个置换。而上式是辫子的拓扑性质的反映，它表示图 3.18 的两个辫子拓扑上是等价的。

在不扯断线的情况下可以移动线的位置把左边的辫子拓扑地变换成右边的辫子。将辫子两端对接就得到环和结，如图 3.19。反过来，任何环或结都是某个辫子的对接，这称为亚历山大定理。然而这种对应关系并非唯一，所以如何辨别哪些辫子给出拓扑等价的环或结是十分重要的。

杨—巴克斯特方程的解在共形场论和拓扑场论中也有应用。量子群的代数满足不可对易的乘法，这使得人们对于它可能会在微观领域物理的描述方面有所作为而寄予了很大期望。事实上，海森堡的测不准原理已经向人们揭示出，不可交换是微观世界的基本属性。目前的场论所存在的基本困难有可能是与连续统一的概念联系在一起。在普朗克尺度，如果有一个自然的“截断”，那末发散的问题也许就不再是问题了。这种自然的截断可以表现为时空的量子化，换句话说，在普朗克尺度下，时空的坐标不再是普通的数，而是满足一定代数关系的“算符”了。在非对易的时空上推广了微积分的运算，讨论非对易的微分几何以及由此建立起来的量子场论是目前量子场论发展中的一个方面，人们期待它能为解决量子场论现在面临的一些基本困难投进一线曙光。

3.3.2 格点规范理论

夸克模型取得了巨大的成功。然而实验上迄今为止尚未找到自由存在的夸克。对此人们的解释是存在着所谓的禁闭效应。那末关于夸克之间强相互作用的量子色动力学能否从理论上给出夸克禁闭的结果呢？目前关于这个问题的最好解答来自格点规范理论。

格点规范理论于 1974 年由威尔逊首先提出。所谓格点规范理论就是将连续的时空简化为分离的格子，在其上建立规范场理论。格点化是一种数学上的技巧，它提供了一种自然的截断从而使得连续时空中一般的量子场论所固有的紫外发散不出现。象所有其他的规范化方法一样，在重整化之后必须将人为的截断去除，也就是说，只有在格子变得无穷细密的连续极限下才能得到原来的物理，因为我们真实的物理时空是连续的(至少在人类目前的认识水平上是如此)。

事实上，格点化这种规制化方案从量子场论的费曼路径积分角度看是更为自然的，因为积分本身就是无穷分割求和的极限过程。当然，格点理论因此会失去明显的相对论协变性，不仅如此，格点化使得人们难以利用费曼图展开的办法进行微扰计算。也正因为这样，格点理论为人们提供了一种新的近似计算方法—非微扰展开的办法。由于这个近似办法不是按耦合常数做展开，所以特别适合于讨论强相互作用，尤其是夸克禁闭这样一些本质上属于非微扰的现象(因为根据重整化群理论关于耦合常数跑动行为的讨论，当夸克间距离增大时，有效耦合常数也随之增大从而使得微扰展开不再适用)。

另外，格点理论同时还揭示了量子场论与统计力学之间的密切关系。事实上，从路径积分表达式可以看到，强耦合相应于高温展开，在统计力学中是容易处理的。在欧氏空间中，路径积分和一个相应格点系统的统计配分函数完全相同，耦合常数的平方相应于系统的温度。于是可以利用统计力学的方法进行研究。尽管格点规范理论的目的是为了研究量子色动力学的禁闭现象以及强子谱的问题，但是关于格点理论本身也存在许多有趣的现象。其中许多模型表现出的非平凡相结构在连续极限下的相应场论中并不存在。

分离时空格点上规范场论的构造最早由威尔逊提出。其关键性的一点在于注意到规范场是一个与路径相关的相因子。于是，基本的变量是

定义在连接两个最近邻格点的键上的群元素。与任意路径对应的群元素是那些构成路径的键所对应的群元的乘积。考虑一个一般的规范群 G ，相应于连接两个最近邻格点 (i, j) 的键，存在群元素 $U_{ij} \in G$ 。这里规范场定义为

$$U_{ij} = e^{ig \cdot A_{\mu} a}$$

其中 a 是格子的基本尺度。为了决定格点系统场变量的动力学行为，威尔逊给出了格点理论的作用量，并证明了其拉氏量在 $a \rightarrow 0$ 的连续极限下回到我们已经熟习的纯非阿贝尔规范场理论的作用量。

为了得到量子理论，将作用量做为权重因子对所有可能的规范变量进行积分，由于 U 是紧致群的元素，所以很自然地利用群上的不变测度作为积分的测度，在它的路径积分表达式中并未包括规范固定项，而在通常连续时空的理论中这样的规范固定项将用来消除由于对所有的规范变量积分而导致的发散。然而在格点规范理论中，变量是紧致群的元素，规范轨道也是紧致的。对于规范不变的量进行积分不会带来发散，因而无须引入规范固定项以便得到为进行微扰展开用的有效作用量。从运动学的角度看，格点理论将非阿贝尔规范场理论作为它的适当的连续经典极限。然而，即便是在取连续极限之前，模型也保留了一个规范理论的许多性质，尽管如此，在格点规范理论中适当地引进费米场仍然存在某些尚待解决的问题。

对于不包含费米场的纯规范理论，威尔逊的格点规范理论与统计力学中磁理论的模型有着许多极为类似之处。 U_{ij} 就象是晶格键上的自旋，在威尔逊的作用量中通过 4 -自旋耦和发生相互作用。于是可能会提出这样的问题，即格点规范理论能否产生自发磁化现象。在铁磁体中，自旋在特定的方向产生不为零的期望值。可以证明，在威尔逊的理论中不会出现这种情况。其原因在于这样将会破坏定域规范对称性。自旋模型中磁化提供了一个非常有用的序参量，所以人们期望在格点规范理论中也存在类似的序参量。既然问题与规范对称性相关，那末可以预期这个序参量应该是规范不变的。的确，由于路径积分走遍所有的规范，所以任何算子的非规范不变部分将会从期望值中抹去。在纯规范理论中，最简单的不变量是沿一个长方形的边得到的群元素的迹。它的期望值代表了相应热力学体系的内能 P ，可以由配分函数的微商得到，人们发现它可以作为所需的序参量。 P 作为序参量存在一个缺点，即它不是一个定域的

量。尽管如此，它在数值计算中起着重要的作用并且是最容易估算的量。

对于纯规范理论，威尔逊给出了另一个非定域的序参量。可以证明沿一个闭合圈上的群元素乘积的迹是规范不变量，它的期望值

$$W(C) = \left\langle \text{Tr}_{i,j \in C} U_{ij} \right\rangle$$

称为威尔逊圈，其中 C 代表闭合圈。最简单的威尔逊圈正是前面提到的小四方。为了考察 QCD 是否存在禁闭现象，可以计算由一对正反夸克组成的体系的能量 $E(R)$ ， R 是两个夸克之间的距离。如果不存在禁闭，那末可以预期当 $R \rightarrow \infty$ 时， $E(R) \rightarrow 2m$ ，其中 m 是夸克的质量。这表明当正反夸克相距足够远时它们成为两个独立的自由粒子，体系的能量就是正反夸克质量之和。相反，如果存在禁闭则夸克间的势能将是无界的，即

$$E(R) \rightarrow \infty, \text{ 当 } R \rightarrow \infty$$

威尔逊证明对于宽 R 长 T 的类时回路有 $W(R, T) \sim e^{-E(R)T}$ ，当 $T \rightarrow \infty$

其中 $E(R)$ 正是与相距为 R 的正反夸克体系相联系的规范势能。而在另一方面，可以证明在强耦合的情况下对于足够大的长方形回路，威尔逊圈满足所谓的面积定律，即

$$W(R, T) \sim e^{-KRT}$$

其中 K 是比例常数， RT 是长方形回路的面积。比较上面两式就可以得出结论：

$$E(R) \xrightarrow{R \rightarrow \infty} KR$$

正反夸克间的势能为线性的，这与雷吉轨迹的结果是符合的。面积定律中的比例常数为格点规范理论提供了又一个序参量。在非禁闭相，这个常数为零；而当夸克间存在禁闭时则不为零。由于它的直观意义，这个常数成为物理上重要的参数。特别要指出的是，即使在纯规范理论的连续极限下它也不为零。

禁闭现象的这种面积定律判据在引入夸克作为动力学变量后就失去了作用，分得足够开的源将通过真空涨落产生新的夸克对而使体系的能量降低。换句话说，这时一个大的威尔逊圈所衡量的是两个介子之间的势能而非两个裸夸克之间的势能。当然，如果我们知道如何计算整个理论的话，那末就无需关于禁闭的判据，我们所需要做的就是计算理论的质量谱并和实验结果进行比较。这方面的研究仍在进行中。目前关于质量谱和一些低能过程的性质方面虽然已经得到了许多好的结果。但在格

点理论中如何正确处理费米子仍是一个引人注目的难题。

关于弱作用规范理论中的序参量是另一个非常有趣的问题。在这种情况下，希格斯机制将给规范粒子带来质量。在这方面的研究中，格点规范理论几乎没有发挥什么作用。原因主要在于这时微扰展开的办法更为适用。尽管如此，如何通过序参量来区分希格斯相与禁闭相是一个有趣的问题。的确，在这两种相变下都可以得到能隙。富雷肯(Fradkin)和辛克(Hencker)于1979年已经证明这两种相并没有明显的区别，它们可以相互解析地转变。

格点理论将规范场论的费曼路径积分表达式简化为普通的多重积分。于是至少在有限的尺度下，人们可以预期对理论的配分函数及其它的观测量进行数值计算。然而稍加分析就会发现，即便是对于 10^4 大小的格点系统，积分的重数已经大得惊人。这样的体系有400000个键，以最简单的规范群如 Z_2 为例，配分函数简化为普通的求和，但是这个求和仍有 $2^{400000}=1.58 \times 10^{12041}$ 项。处理这样大的数目只好用统计方法。目前人们采用蒙特卡罗方法来对格点规范理论进行计算。这种方法的精髓就是通过随机地提供较少数目的典型构型来模拟整个求和。

蒙特卡罗模拟技术在统计物理中有相当长的历史。这一方法提供了在虚拟的“晶格”上进行实验的可能性。人们可以通过编制程序来选择相互作用的哈密顿量的形式。所以原则上这将能够把各种动力学特征孤立起来研究，将有助于弄清它们各自在诸如相变这样一些过程中的作用。这一技术在高温和低温区域下有很好的收敛性，尤其在中间区域吻合得很好，这一点对于粒子物理来说特别重要，因为人们急于弄清楚威尔逊理论在强耦和条件下关于禁闭的证明与弱耦和极限下的连续时空量子场论的关系。紫外截断的办法是各式各样的，即使在格点理论的框架内其变化仍是难以计数的。尽管对于格点理论这种非微扰方法而言，是否存在四维的连续极限理论尚不能严格地知道，但蒙特卡罗模拟技术的确提供了非微扰情况下对截断依赖关系进行研究的可能。

3.3.3 拓扑场论

相对论量子场论的最终目标是要建立一个包罗万象的统一理论，所以这样的理论难以严格求解就不足为奇了。为此，人们必须发展近似又足够准确的计算方法。长期以来，量子场论中这样的方法就是所谓的戴逊-费曼-施温格微扰展开法。然而量子场论的很多特性无法在这个近似展开中得到反映，所以人们不得不尝试其它的近似方法。半经典展开就是这种尝试的成果，它被成功地用来揭示量子场论的许多新特性。

半经典近似方法以量子场算符满足的海森堡运动方程为出发点，在经典极限下忽略所有的量子效应把场方程看成纯粹的经典非线性偏微分方程。当然这样的方程仍是非常难解的。物理学家借助各种各样的数学工具进行分析。这些工具是数学家按照他们自己的想法发展起来的，双方从这种结合中相互受益，这样的情况在科学的发展史中多次出现。利用数学上的拓扑分析，人们可以超出半经典近似从而得到一些量子场论的精确解。这些结果主要有三个方面。第一个方面是认识到可能存在一些标志量子场论特性的参数，这样的参数并不都出现在理论的拉氏作用量中；第二，对于自洽的量子场论而言，拉氏量中的参数如耦合常数、质量等，不可以随意选取；第三，规范场与物质场(指费米子场)的耦合形式受到很强的约束。除非恰当地安排费米子，那些看上去规范不变的自洽理论很可能在量子化之后变得不自洽，如出现反常。人们希望对于拓扑场论的进一步研究为量子场论的精确解提供更多的线索。

我们通过简单的例子来看看拓扑性质在这里的重要作用。空间的拓扑性质是指空间在做连续变换下不变的那些性质。这里空间的定义也是十分广泛的。为形象起见。不妨将空间看成是橡皮做成的曲面，只要不发生撕破或粘合的情况，任意地改变曲面的形状就对应为我们所说的连续变换。通过连续变换得到的那些曲面相互之间称为拓扑等价的，而那些保持不变的性质就称为拓扑性质。

求证两个空间拓扑等价是一个几何问题，将涉及怎样造出两个空间之间具体的同胚变换，所用的技巧则随问题的不同而互异。求证两个空间不同胚。则是性质完全不同的另一个问题。因为不可能枚举所有的变换来检验是否同胚，所以这时多采用拓扑不变量。如果两个空间的拓扑不变量不同，那末它们一定不可能拓扑同胚。连通性是拓扑不变量。利

用这个拓扑不变性可以证明一维直线和一个二维平面不同胚。从直线上挖去一点(如原点),则空间分成了两块(相应于正实轴和负实轴),这就是一个不连通的空间。假如有一个同胚映射把直线变成平面,那末,挖掉一点的直线就相应地变成挖掉一点的平面。但是,挖掉一个点的平面仍是个连通空间(即还是一整块)。所以一维直线和二维平面拓扑不等价。

更一般地研究拓扑不变量的理论是所谓的同伦论和同调论。彭加勒提出基本群的想法,认为每个拓扑空间对应以一个群,使得同胚的空间具有同构的群。如果这些群不同,那末空间就是不同的。比如图 3.20 的两个空间它们显然是不同胚的。因为两圆盘一个有孔一个没有。这个孔洞的影响可由图 3.21 内的环道很好地反映出来。正是由于有这个洞,使得环道不能连续地缩成一点。彭加勒就是用象这样的环道来产生一个群,称基本群。环道和一样也能使我们识别洞的存在,因为可以作不经过洞区的连续形变而变成。它们给出基本群的同个非平凡元素。有孔圆盘的基本群为整数所构成的自由循环群,对应为环道绕孔洞的圈数(也称绕数)。图 3.22 就是这个群的几个元素。

无孔圆盘的基本群是平凡群,因为任何环道可连续地缩成一点。如果得到的是相同的群,就得寻求更为精细的不变量来区别这两个空间。比如对于实心球和空心球,它们的基本群都是平庸的,但它们显然是拓扑不等价的。为了区别实心球和空心球,把基本群的环道相应地推广到二维球面。这样得到的群称为(二阶)同伦群,那些类似环道和一样可以连续转变的环道称为同伦的。空心球面的二阶同伦群是非平凡的,而实心球的则是非阿贝尔规范场具有非平庸的拓扑结构,这表现在它具有非平凡的同伦群。以 $SU(2)$ 纯规范场为例。在规范场路径积分量子化过程中,首先要挑出真正独立

的动力学变量,这是通过选择规范来实现的。比如库仑规范: $\nabla \cdot A = 0$ 。众所周知,规落条件不能完全将场确定下来,一般还会存在剩余的规范自由度。对于纯规范,仍可以相差一个只依赖空间坐标的规范变换 $U(r)$, 有

$$A_i(r) = -U^{-1}(r)\partial_i U(r)$$

如果库仑规范是一个好的规范条件的话,应能在上面的规范变换 $U(r)$ 下得以保持。也就是说,上面的方程与库仑规范合在一起只应有平庸解

$A_i=0$ 。事实上，由于规范场的非平庸拓扑性质，上面的方程不仅有平庸解还有非平庸解。

就规范群元素 $U(r)$ 而言，它是将三维空间映射到群上的函数。三维空间在把无穷远作为一点的情况下拓扑同胚于四维欧氏空间里的三维球面。另一方面， $SU(2)$ 群有三个生成元，它的参数空间也是一个三维球面。这样，规范变换 $U(r)$ 就是从一个三维球面到另一个三维球面的连续映射。这种映射对应于三维球面的三阶同伦群。 $A_i=0$ 相当于绕数为零的同伦环道。因为球面中有洞的事实，所以同伦群是非平庸的。因此还存在绕数 n 不为零的同伦映射：

$$n = -(1/24\pi^2) \int d^3x \text{Tr} \{ \epsilon_{ijk} A_i A_j A_k \}$$

规范场真空的这种非平庸拓扑结构与量子力学中的周期位势类似，实际的真空为一个确定的 所标志，称为 真空。将不同的真空联系起来的是瞬子解。相当于隧道效应。这些问题的研究对于规范理论，特别是理论的量子化，是非常重要的。

场的这种非平庸拓扑结构在很多方面将会给理论的构造带来一定的约束。非线性 模型在加上一个依赖于紫外截断的外斯—组米诺项之后就构成量子色动力学在低能下的有效理论。它能够正确地描写诸如手征对称破缺。低能定理和介子衰变等。更令人惊奇的是，考虑到场的拓扑性质后，外斯—组米诺项前的因子—耦合常数——一定是整数，并且就是夸克颜色的数目。这类似于狄拉克磁单极子的量子化条件。事实上，非线性 模型具有拓扑稳定的孤立子解的根源在于其规范群的非平庸拓扑结构。这些拓扑稳定的孤立子可以用来解释量子色动力学中的重子。

量子陈—塞蒙斯(Simons)场论在过去的几年中一直是拓扑场论研究的主题。在量子陈—塞蒙斯模型中，拓扑研究的手段不再局限于半经典近似的范围。事实上，无须再作任何近似，量子陈—塞蒙斯场论是一个“真正”的拓扑场理论，在这个理论中得到的任何观测量和观测结果都有其拓扑上的根源与含义。做为一个完全的量子理论，陈—塞蒙斯模型不仅在 R^3 和 S^3 中可严格求解，而且在任何一个闭的、三维连通定向流形上都是严格可解的。

象对所有的量子场论一样，人们感兴趣算符乘积的真空期望值。在拓扑场论中，特别典型的是考虑与一些 1 维环路相联系的、称为威尔逊圈算子的期望值，它定义在一个定向的环结(link)上，是一组有用的规

范不变的观测量。如果一般地考虑一个有 m 个分量 $\{C_1, \dots, C_m\}$ 的环结 (link) L ，则得到相应的环结 (link) 多项式。从纯数学的角度来研究环结 (link) 和纽结 (knot) 问题是一个古老的话题。只是当琼斯 (Jones) 发现关于环结 (link) 和纽结 (knot) 的新的不变多项式后，才重新引起人们对这个问题的兴趣，特别是对它们与物理问题的联系。人们认识到这些多项式在很多方面与 2 维系统的物理有着极为密切的联系。然而，更具挑战性的工作是找出这些多项式在 3 维空间的定义。威腾解决了这个难题。事实上正如前面已经看到的，这些多项式在陈—塞蒙斯理论中自然地出现；它们就是陈—塞蒙斯理论中威尔逊圈的真空期望值，它们的拓扑不变性由理论满足广义协变性而得到保证。

一旦求得环结多项式，接下来的主要问题就是如何识别它们。前面已经提到，从陈—塞蒙斯理论得到的环结不变量是与准三角型准 Hopf 代数的辫子群表示相联系的。另一方面，这种联系同时有利于揭示陈—塞蒙斯作用量所描写的模型的奇特物理行为。考虑 3 维空间的一个 $(2+1)$ 维分解，等时面可能与给定环结相交数次。这情形就好比点粒子的世界线与等时面相交，随着时间的演化，这些粒子将在空间运动，而且它们当中的每两个一对还有可能一起湮灭或产生 (图 3.23)。于是期望值 $\langle W(L) \rangle$ 可以做为由环结 L 代表的全部过程的量子力学振幅。对于 k 取一般的值，每个粒子与等时面的交点由规范群不等价不可约表示提供的量子数来标志；而当 k 为整数时，不同种类的粒子的数目是有限的。例如，对于 $SU(2)$ 规范群， $k > 2$ 的情况，只有 $|k| - 2$ 种不同的粒子。当 $|k| = 2$ 时，态空间是 1 维的，任何态矢量都正比于真空态。当 $|k| = 1$ 时，态空间是 2 维的，即除真空态外只有一个非平庸的粒子态。

此外，用威尔逊圈有可能描写引力理论。这时，遵循狄拉克处理约束系统量子化的一般程序，那些湮灭态空间的约束可以表达为环结和纽结不变量。而从上面的图可以看到，环结不变量为微分同胚约束所湮灭是极为自然的。人们发现 3 维陈—塞蒙斯理论与 4 维量子引力有很强的联系。当然量子引力本身不可能是一个拓扑场论，因为在引力理论中存在着定域的激发。尽管如此，广义相对论在圈方程的表示中与拓扑量子场论的联系是极富启发性的。用圈表示，在陈—塞蒙斯理论中出现的环结不变量同时也是量子引力中的一个态。有关这一结论的可靠性正在进一步研究中。在威腾最初从陈—塞蒙斯理论中得到琼斯的环结多项式

时,2 维共形场论起了关键作用。关于共形场论与 3 维拓扑量子场论的关系得到人们的极大重视。然而从路径积分的角度如何在数学严格意义下建立两者的联系仍然是没有解决的问题。

拓扑量子场论是一个比较新的方向,这方面的发展也非常快,已经取得了许多重要的结果。80 年代中期以来利用微分几何的方法对于量子场论中反常问题的研究表明反常是与场在大范围拓扑性质密切相关的。由此而建立了一些拓扑场论模型,如外斯—祖米诺—威腾模型,陈—塞蒙斯规范理论。由于 4 维时空中量子场论的巨大困难,这些模型通常是建立在较低维空间,特别是在两维空间和一维时间,因为 1+1 维时空中的纯规范场是平庸的。2+1 维非阿贝尔规范场是超可重整的理论,为研究禁闭现象以及 3+1 维规范场的高温性质提供了合适的模型。尽管对于反常的研究已经取得了很大的进展,但是仍然存在一些很基本的问题有待回答。例如,反常即使对于那些拓扑平庸的阿贝尔规范场也存在,另外,拓扑研究自然地涉及到一些积分的量,而反常却在本质上是定域的量。总之,拓扑量子场论中还存在许多有意义的重要课题等待进一步的研究。

3.3.4 共形场论

共形场论是指在共形变换下不变的量子场论，主要用于讨论场算符乘积在小距离下的行为。二维共形场论是人们主要感兴趣的领域。二维共形场论的非阿贝尔推广是在主丛模型的作用量中加进外斯-祖米诺项，称为外斯-祖米诺-威腾模型。维数大于3的共形场论的一般分析并没有多大进展，只是在一些具体模型的讨论中取得了一定的成功。

共形场论最早出现于60年代，目前有关研究的兴趣主要来自于它在两个完全不同的物理领域的应用。其中之一就是弦理论，在那里共形不变性给理论提供了较强的约束。另一个领域就是二维情况的临界问题的研究。临界现象中的标度不变性与彭加勒不变性一起意味着理论具有共形不变性，这种不变性在二维情况下有着特别丰富的物理结论，它对关联函数构成的约束足以确定所有的这些关联函数从而使理论可以精确求解。

我们首先来介绍一下共形不变性的概念。共形变换是D维欧氏空间里的保角变换，在维数大于3的情况下，共形群由平移、转动、标度变换和特殊共形变换构成，这是一个有15个参数的连续群。在2维情况下，共形群包含更多的共形变换。场论的拉氏作用量S在坐标变换下的改变由能量动量张量描述：

$$\delta S = \int d^D x T^{\mu\nu} \epsilon^\nu$$

从中可以看出，平移、转动以及标度改变下的不变性意味着共形不变性。这种不变性会给出各种关联函数(场论的格林函数在统计问题中的名称)所满足的约束——沃德恒等式。在2维情况下，有两个新的特点。一是可以利用复数坐标 $\bar{z} = x + iy$ 和 $z = x - iy$ 来代表x和y，这样能量动量张量自然满足 $T_{z\bar{z}} = T_{\bar{z}z} = 0$ ；另一个特点是，坐标的任何解析变换都构成一个共形变换。这意味着二维共形变换构成无穷维代数。可以想象，这样庞大的对称性对于理论是一个非常强的约束，因而使得人们可以用非微扰的方法(主要是李代数的方法)来研究这些共形理论。关联函数满足的沃德恒等式用算子的语言来表达将导致所谓的维拉索沃代数，它的生

成元出现在能量动量张量对复坐标的洛朗展开中，相应于z和 \bar{z} ，共形场理论中自然地存在两套维拉索沃代数。在物理上，人们感兴趣的是代

数的最高权表示，其原因在于这样的体系能量存在下界。这些表示的重要特性反映在它的中心荷以及结构常数上。因而存在共形场的分类问题。特别感兴趣的是所谓的“最小”理论，它具有一些显著的优点。在“最小”理论中，所有的反常量纲都是知道的，并且作为系统的线性微分方程的解，定域场的关联函数是可以计算的。从数学的角度看，这样的理论是与共形代数的简并表示相联系的。可以证明，最小理论的一个最简单的例子能够用来描述二维伊辛模型的临界行为。另外的最小理论描述什么样的统计模型的临界行为是共形场论研究的一个方向。

在弦理论的应用方面，其主要意义还在于共形场理论提供了弦理论的经典解。为此，人们特别关心 2 维的量子共形场理论以及将这样的理论与曲面上的动力学度规相耦合的情形。一个渐近自由的量子场论可以描写点粒子在经典时空中的运动。另一方面，按照塞曼则克(Symanzik)发展的用无规布朗路径讨论量子场论的想法，相互作用的量子场论可以看成是一个统计理论，描写一些相互作用着的无规路径构成的“气体”，这些无规路径(random loop)相应于场论中虚粒子在欧氏时空中的世界线。所以可以设想，在欧氏时空中弦理论就是一些无规面(random surfaces)的统计理论。但是，这种看法不可能是弦理论的一个基本解释，因为弦理论应该包括时空本身的动力学——引力，而构造一个经典时空中运动的弦的世界面的理论是事倍功半的。所以人们希望了解一个在散射理论的意义下渐近自由的弦理论。在量子场论中，渐近自由粒子的态是希尔伯特空间中的矢量，整个希尔伯特空间负载了彭加勒群的一个不可约表示，这个表示由粒子的质量和自旋来识别。在弦理论中，一个弦的态对应为一个 2 维共形场论希尔伯特空间中的矢量。这个 2 维共形场论的维拉索沃代数具有 $c=10$ 的中心荷，而其表示的最高权与弦的能量-动量张量、自旋等物理参数有关。在弦理论中，时空的性质自然地表现为共形场论的内在对称性。一个完善的弦理论应该能够告诉人们如何由一系列维拉索沃代数中心荷为 10 的共形场论态的张量积来得到散射理论量子力学振幅。进一步将能够计算不同的经典时空几何之间的隧道效应等。

在应用到统计物理的方面，共形场理论的主要问题是计算各个有关的定域场的反常标度量纲，因为这些量决定了系统在相变临界点附近热力学函数的奇异行为。最近的一些发展表明共形场理论与经典的可积系

统(KdV 型)、W 代数存在一定的关系。W 代数首先是为研究维拉索沃代数在高自旋情况下的推广而引进的。后来发现它的经典形式(即泊松括号代替对易子)与 KdV 理论有关并且可由哈密顿约化方法得到。另一种得到 W 代数的办法是用所谓的卡斯-莫迪代数。这两种办法是否等价是一个非平庸的问题。另外,关于 W 变换的几何意义仍是尚待阐明的问题。当然,现在人们越来越多的从量子场论本身的角度来研究共形场理论,希望由此得到对二维量子场论的非微扰处理办法,以及与数学,特别是拓扑、环结理论的引人入胜的关系。同时,关于四维情况的研究也不断地受到关注。

3.3.5 塞伯-威藤(Seiberg-Witten)理论 :四维空间量子场论的新飞跃

从 1994 年初开始，由塞伯、威藤和他们的合作者所完成的一系列工作震动了理论物理学界和数学界。从物理的角度讲，这些工作涉及如何精确地研究四维空间量子场论的非微扰行为(例如禁闭、手征对称性破缺和真空凝聚等)，也就是如何解析地处理强耦合的相互作用。从数学的角度讲，它们涉及如何简单地算出光滑 4 维流形的拓扑特征。以上这些问题在物理学和数学中都是极为重要的，然而二十多年来却一直困扰着物理学家和数学家。现在，人们终于看到了解决这些问题的曙光。

塞伯-威藤理论主要研究的对象是 4 维空间中的超对称非阿贝尔规范场论(包括 $N=1, 2$ 和 4)。它的主要内容包括以下几个方面。

塞伯首先注意到了超对称理论对其中的耦合常数是全纯的，这就强烈地限制了理论随耦合常数的变化。基于这种性质，人们一旦研究了弱耦合区域(在这里微扰论是可以应用的)，那么许多结果就可以推广到强耦合情形。他还利用这个性质研究了 $N=1$ 的规范模型的真空态(基态)，证明了其真空是简并的但不是等价的，即不同的真空将产生不同的激发谱。随后，塞伯和威藤一起研究了 $N=2$ 的情形，特别是仔细研究了 $SU(2)$ 规范模型。以后又有人研究了更广泛的模型。

所有这些模型有一个共同的特征，就是能够自然地产生磁单极，也就是基本场的集体激发(孤子)。这种磁单极不同于特·霍夫特和波利雅阔夫(Polyakov)在 70 年代提出的那种质量很重的磁单极。在这里，磁单极的质量随真空态而改变，在某些真空态下会变成无质量的，因而可以产生磁单极的真空凝聚。这种情形非常象超导体中带电的库伯对的凝聚。在那里，这种凝聚使得超导体完全排除磁场(迈斯纳(Meissner)效应)，类似地，在这里，带磁的粒子的凝聚使得真空完全排除电场。如果这里所说的“电”和“磁”就是 QCD 中的“色电”和“色磁”，那么这种效应正是“夸克禁闭”。他们也通过磁单极凝聚发现了手征对称性的破缺。关于磁单极凝聚导致夸克禁闭的思想在七十年代曾由特·霍夫特和其它人提出过，但他们无法进行实际的计算。现在塞伯-威藤理论把它具体地实现了，从而证明了这种想法是正确的。这无疑是对禁闭现象的更深的理解。

由于有了磁单极，理论也就有了一种非常重要的“对偶性”——电和磁的对偶性。在一定的意义上讲，这种对偶性是塞伯-威藤理论的核心，而且它还不只涉及电与磁。比如，对于 $N=2$ 的 $SU(2)$ 超对称规范模型这个具体的情形，他们证明了：这个模型对于电荷 \leftrightarrow 磁荷、粒子 \leftrightarrow 单极子(或者说基本场 \leftrightarrow 集体激发)同时 $g \leftrightarrow \frac{L}{g}$ 的互相

交换是对偶的。也就是说，从一个方面的计算结果可以推断出另一个方面的计算结果。这样一来，物理学家所面临的一个艰巨任务——处理“电”的粒子在强耦合下(而且是非阿贝尔的)相互作用，可以化为处理它的对偶情形：“磁”的粒子在弱耦合下(而且是阿贝尔的)相互作用，而后者微扰论可以用，因此容易得多。类似的对偶性在这个理论研究过的许多其它模型中也存在，而且事实上是该理论的主要应用所在。无疑地，这种“电-磁，粒子-孤子、强耦合-弱耦合”的对偶性质如果存在于真实的物理世界，那么它将成为物理学的主要基本特征。

在塞伯-威藤理论的数学意义方面，首先要谈一谈计算 4 维流形拓扑特征的督纳尔逊(Donaldson)理论。人们早已知道，与任何其它维数的流形不同，4 维流形是解析结构最为丰富(理论上说有无穷多种)的流形。所以，如何完全地计算出一个给定 4 维流形的拓扑特征，就成为人们关注的问题。1982 年，督纳尔逊提出了一个办法：用自对偶的非阿贝尔杨-米尔斯方程来研究 4 维流形的拓扑不变量(后来即被称为督纳尔逊不变量)，开辟了这方面研究的新途径，并逐渐成为 4 维流形拓扑学的基础。但可惜的是，督纳尔逊理论的计算量极大。塞伯-威藤理论提出以后，督纳尔逊不变量的计算就化为数出“单极子方程”(即单极子与光子相互作用的方程)经典解个数的问题，从而把这个计算成千倍地简化了。利用这种方法，人们已经轻易地证明了一些以前只是猜测的定理，而且新的定理还在不断地涌现。

总而言之，把塞伯-威藤理论称为 4 维空间量子场论研究的一次新飞跃是不过份的，它的意义还将在今后不断地被揭示出来。但是说到它与实在的物理世界的联系，比如能否回答在粒子物理的标准模型中人们关心的夸克禁闭、电弱对称性破缺、质量产生等问题，还存在一定的障碍。这主要是由于它是一种超对称的理论，而现在的标准模型是非超对称的，而且实验到目前为止也还没有发现自然界中存在着超对称性。这个

縫隙如何填补？谁向谁靠拢？仍然众说纷坛。这个问题只有靠理论的更深入的研究和实验的证据来回答了。

参考文献

- 〔1〕中国大百科全书，物理学卷，量子场论有关条目。
- 〔2〕狄拉克，物理学的进展，科学出版社。
- 〔3〕杨振宁，文集 1945—1980(英文)。
- 〔4〕超弦，中国对外翻译出版公司，1994。
- 〔5〕Weisskopf, V. F. , “量子场论在过去五十年的发展”，今日物理，1981(英文)。
- 〔6〕Frohlich, J., “量子场论的新进展”，第 24 届国际高能物理会议文集，1989(英文)。

4 凝聚态物理

汉语中“凝聚”一词是由“凝”字双音演化而来的。“凝”在东汉许慎的“说文解字”一书中同“冰”，指的是水结成冰的过程。可见我们的祖先最初对凝聚现象的注意可能始于对水的观察，特别是水从液态到固态的现象。英语的 condense 来源于法语，后者又来源于拉丁文，指的是密度变大，从气或蒸汽变液体。看来西方人对凝聚现象的注意可能始于对气体的观察，特别是水汽从气态到液态的现象。这是很有意思的差别，大概与各自的古代自然生活环境和生活习惯有关。不过东西方二者原始意义的结合，恰恰就是今天凝聚态物理主要研究的对象—液态和固态。当然从科学的含义上来说，二者不是截然分开的。所以凝聚态物理还研究介于这二者之间的态。例如液晶等。液态和固态物质一般都是由量级为 10^{23} 的极大数量微观粒子组成的非常复杂的系统。凝聚态物理正是从微观角度出发，研究这些相互作用多粒子系统组成的物质的结构、动力学过程及其与宏观物理性质之间关系的一门学科。

众所周知，复杂多样的物质形态基本上分成三类：气态、液态和固态，在这三种物态中，凝聚态物理研究的对象就占了二个，这就决定了这门学科的每一步进展都与我们人类的生活休戚相关。从传统的各种金属、合金到新型的各种半导体、超导材料，从玻璃、陶瓷到各种聚合物和复合材料，从各种光学晶体到各种液晶材料等等；所有这些材料所涉及到的声、光、电、磁、热等特性都是建立在凝聚态物理研究的基础上的。凝聚态物理研究还直接为许多高科学技术本身提供了基础。当今正蓬勃发展着的微电子技术、激光技术、光电子技术和光纤通讯技术等等都密切联系着凝聚态物理的研究和发展。

凝聚态物理以万物皆成于原子为宗旨，以量子力学为基础研究各种凝聚态，这是一个非常雄心勃勃的举措。凝聚态物理这个学科名称的诞生仅仅是最近几十年的事。如果追寻一下它的渊源。应该说出自于对固态中晶态固体的研究和对液态中量子液体的研究。在对这二种特殊态的长期研究中，人们积累了一些经验，也建立起了一些信心，并逐步把一些已有的方法推广用于非晶态和液晶乃至液态的研究，从而大大拓宽了视野，逐步形成了凝聚态物理。

今天，凝聚态物理的视野还在继续开拓。然而作为渊源的二种凝聚

态即晶态固体和量子液体，时至今日仍然是它主要的研究对象，内容当然越来越丰富了，考虑的问题也越来越深入了。毕竟我们面临的是同一个自然界，许多现象和规律是普适的。人们正是通过对一系列特殊态的深入研究来逐步认识和掌握那些普适的规律。

4.1 晶格和能带

人们对晶体(晶态固体)有记载的最早的定量研究,或许应该算是意大利的斯丹诺(N. Steno),1669年他发现了晶面角守恒定律。该定律说:同一物质的不同晶面大小、形状和个数可以不同,但是相应的晶面之间的夹角是固定不变的。1784年法国人阿羽依(R. J. Hauy)假设晶体可能是由相同的单元-晶胞堆砌起来的,定量地证明了这个定律。

在现代原子学说的基础上,1855年布拉伐(A. Bravais)提出了晶体的空间点阵学说。按照这个学说,晶体是由原子或原子团周期性地在空间中规则排列而成的。我们可以想象有一个三维空间的点阵,在每一点附近有一个或一团原子,它们在各点的布局都是一模一样的。组成点阵的最小单元连同它所拥有的原子或原子团便成为整个晶体的组成单元。晶体点阵就是这些单元被许多空隙隔开的几何形体。通常是一个平行六面体。布拉伐不仅提出了晶体的点阵学说,还证明了简单晶格只有14种类型。所谓简单晶格是指在点阵每一点附近只有一个原子的晶体。一般的情况称为复式晶格。

对于复式晶格,1890年俄国的费多洛夫和1891年德国的熊夫利分别独立地证明了只有230个对称群,又称空间群。这里的对称群是指,由旋转、平移和反演(或镜像)这些基本操作(运算)所组成(生成)的保持复式晶格不变的群。我们在本书的前面部分已看到对称及描绘它的数学工具群论在粒子物理理论发展中的重要性。然而,在历史上,晶体学可能是最早运用群论工具的学科。虽然晶体的空间点阵学说在19世纪末还只是一种假设,但是以群论为基础的晶体几何学在那时却已完全建立起来了。

由点阵结构,再考虑晶体的密度、晶体物质的原子量,人们不难大致估算点阵的间隔-晶格常数。它大约是 1Å ,也即 10^{-8} 厘米的量级。1895年伦琴发现了X射线。根据它产生的方式,人们认为它是一种波长很短的电磁波,索末菲(A. Sommerfeld)从理论上基本上确定它的波长就是 Å 的量级。1912年,德国的劳厄(M. V. Laue)注意到了X射线波长和晶格常数是同一量级,从而发现了晶体的X射线衍射现象,确证了晶体的周期性点阵结构。为此,劳厄获得了1914年的诺贝尔物理学奖。劳厄的这一工作使晶体几何学走出了象牙塔成为全部晶体物理学的基础,而

且方法本身也已发展成为各种凝聚态的基本观察手段。

进一步的工作是布喇格父子(W. H. Bragg 和 W. L. Bragg)在对衍射图作深入研究后做出的。他们定量地解释了衍射图像,发展了一整套 X 射线衍射晶格结构分析方法,并利用它测定了许多晶体的结构,为固体物理学的创建奠定了基础。为此,布喇格父子获得了 1915 年的诺贝尔物理学奖。

在晶态固体的周期性结构被实验证实后,电子在周期势场中的问题就自然被提出来了。研究这个问题的目的当然是要了解金属导电问题。在历史上德国的德鲁特(P. K. Drude)大概是最早对金属导电现象作出基本上合理解释的人。他在 1900 年提出金属中的电流就是汤姆逊所发现的电子的运动,并且认为金属中充满着自由电子气体,它们在电场作用下作加速运动,同时又由于与点阵原子碰撞而受到一种阻力。平衡时电子会有一个恒定的漂移速度,它与电场强度成正比。1904 年洛仑兹(Lorentz)又对德鲁特的模型作了进一步讨论,引进了不同电子速度的统计分布,但是用的是麦克斯韦尔-玻尔兹曼统计,这是一种高温下适用的经典统计。按照德鲁特-洛仑兹模型,由于电子既带电又有一定的热能,结果在给定的温度下金属的热导率与电导率成正比,这是符合实验事实的。但是,这个模型不能说明电导率随温度变化的事实,更不能解释比热的观察结果。

1928 年索末菲在量子力学的基础上重新考虑了自由电子气,并正确地运用了费米-狄拉克统计,从而消除了德鲁特-洛仑兹模型存在的困难。索末菲自由电子模型的成功在于统计用对了。他认为点阵离子所产生的周期场被公有化的许多电子所屏蔽了,因此考虑金属中的电子是自由的,这毕竟是一种简化。

同年布劳赫(Bloch)接受他导师海森堡的建议考虑金属导电问题。他从一维情况入手,依据量子力学的“波-粒二象性”,研究了电子波在周期点阵势场的晶体中的衍射,发现了以后称之为布劳赫定理的结果,即在周期势场中电子的波函数和自由电子的平面波只差一个周期性的幅度调制。这就是说,在周期势场中电子仍然能自由自在的到处跑,只是现在它在各处出现的几率不再是均等的,而是带有了周期势场的特征,具有了周期性。

布劳赫定理还揭示了重要的能带现象,一个电子在周期势中的定态

能量可以通过求解薛定谔方程得到。人们发现，与自由电子的连续能谱不同，这时电子的本征能量 E_{nk} 有 n 和 k 二个标记。 k 标记电子的动量，仍旧取连续值。 n 标记周期势的影响，它把连续能谱切割成带状。每个带内能量仍是连续变化的，而不同的带以 n 为区别。这就是电子能谱的能带现象。每个带内的能量随 k 的变化是周期性的，所以我们只需在一个最小的周期区域内表示 E_{nk} 这一最小的区域称为第一布里渊区 (FBZ)。例如一维的情况，如图 4.1 所示。类似地可以定义第二、第三等布里渊区。

典型金属，如碱金属，它们的原子最外层只有一个电子。内层电子部分地屏蔽了核的正电荷，组成原子实；外层一个电子在原子结合成晶体时很容易脱离原子，而成为在晶体中游离的电子。对这些游离的电子来说，周期势很弱，它的能量应该十分接近自由电子的能量，但仍形成能带结构。如一维情况，在周期势场中被分为图 4.2 所示的诸带。用量子力学可以证明，能带之间的能隙对应各布里渊区的边界。对此可以简单地解释如下：各布里渊区边界所对应的电子的态是一种特殊的定态，称为“驻波”，即电子在一些地方会驻定下来，一般的定态都是“行波”，电子在不停地走动。对于通常吸引的周期势，一种驻波是驻定在各点阵点处，另一种是驻定在各点阵点的间隔中间处。前一种驻波比后一种驻波更“稳定”一些，能量也低一些。这样，这二种驻波所对应的本征能量就不同了，因此就有了能隙。其实这二种驻波是一个右行和一个左行的“行波”以不同方式干涉而组合成的。这是量子力学中典型的干涉产生能隙的例子。

让我们再从原子组合成晶体的角度来看能带。当原子相距很远时，相互间的影响很小，每个原子中的电子基本上处于单个原子的本征态，可以用一套量子数来描述，它们相当于“原子轨道”。现在，把原子逐渐拉近而组成晶体。由于邻近原子的影响，一个原子中的电子有一定的几率跃迁至邻近原子上去。在邻近原子的影响不大，或二个原子态的能级间距较大时，电子基本上是跳至最邻近原子的同一原子态上；由于周期性，电子既然可以跳至最邻近原子上去，那么它也可以跳至任何一个原子处。这时电子的能量就会展成基本以所属原子态的能量为中心的能带，而带宽就决定于邻近原子的影响。这就是所谓紧束缚近似。这样，对于原子的 S 态、 P 态、……电子分别都有一条相应的能带。当然这些能带之间可能会发生重叠、也可能会有能隙；甚至会在某些 k 处有重叠，

而又会在另一些 k 处有能隙。种种复杂情况都可能发生，这取决于实际晶格的组成。

由布劳赫提出以后又经爱·威尔逊(A. H. Wilson)发展的能带理论为区分导体、半导体和绝缘体提供了一幅物理图像。按照这一幅图像，晶体中的众多电子就在一条条的能带上填充。填充是按泡利原理或一般地费米-狄拉克统计原理进行的。粗略地说来：

(a)电子恰好填满最低的一系列能带，高的各带全空。这时最高的满带称为价带，最低的空带称为导带。导带的最低处能级与价带最高处能级的差称为带隙。一般说来，带隙大的，达到 10 电子伏左右时为绝缘体；带隙小的，仅 1 电子伏左右时为半导体。

(b)电子除了填满一系列低能带外，还部分地填充了高能带，这对应于金属导体。高能带中被占据的最高能级称为费米能级，记为 E_F 。

(c)半导体仅当温度 $T=0$ 时才是非导体，由于能隙小，一旦 $T > 0$ ，就有一些电子从价带跃迁至导带，同时在价带中留下空穴(参见图 4.3)。

我们在前面曾提到碱金属的近自由电子的弱周期势近似，是一种典型的导带例子；而紧束缚近似是一种典型的价带例子。今天，根据不同的晶体组成，已发展出十分复杂、庞大的能带计算方法；能带是决定晶态固体许多物理性质，诸如电学、光学、磁学以及热学等等的重要依据，人们仍在继续探索行之有效的计算方法。

能带理论的建立为半导体理论的发展奠定了基础并从而引发了 40 年代末固体物理学的爆炸式的发明浪潮。晶体管的发展，集成电路、大规模乃至超大规模集成电路的研制成功，最终导致机、通讯技术等一系列高科技革命……。

4.2 声子——格波场的量子

佩斯(A .Pais)在他那本著名的爱因斯坦传记中援引了全书唯一的一张理论和实验对比的曲线-数据图,那是爱因斯坦 1906 年发表的关于固体比热的一篇文章中,一张将他的理论分析所得曲线与 1975 年韦伯(H .F .Weber)的金刚石比热实验数据的对比图。佩斯指出那是固体量子理论历史上发表的第一张图。

历史上,固体比热的研究可以追溯到 1819 年,那年二位年轻的法国人杜隆(P .L .Dulong)和珀替(A .T .Petit)在实验上发现一个出乎意料的结果:室温下许多金属、非金属元素的克原子比热是一个常数。最初他们相信这个结果对气体也是成立的,然而至 1830 年左右,实验证实这个结果只适用于固体。此后,陆陆续续有实验表明这个比例并不是常数,1875 年韦伯总结许多实验事实,猜测杜隆-珀替的结果在高温下是成立的,并对硼、硅、石墨和金刚石作了从-100 到 1000 的比热测量。

对杜隆-珀替结果的理论解释是半个世纪以后由玻尔兹曼(L .Boltzman)作出的。1876 年玻尔兹曼把经典能量均分定律用于三维晶格,设想每个晶格原子是一个振子,它有三个自由度,按经典能量均分定律,每一个自由度对比热的贡献是 $\frac{k_B T}{T}$,于是每个原子对比热的贡献是 $3k_B$,一个克原子有 N_0 个原子, N_0 是阿伏伽德罗常数,因此比热为 $3N_0 K_B$,这就是杜隆-珀替所发现的常数。至于低温下比热减小的反常结果,玻尔兹曼错误地认为是由于低温下近邻原子“粘着”而失去一些自由度所致。在 1906 年爱因斯坦的工作之前,一直没有人认识到金刚石比热反常是由于经典能量均分定律不适用。

固体比热,这里特别指绝缘晶体的比热问题是一个与黑体辐射十分类似的问题。它们共同的地方在于都是关于量子的统计问题,而且都是以后被称为玻色-爱因斯坦统计的那一种适用于玻色子的统计。1900 年普朗克(M .Plank)解决了黑体辐射问题,爱因斯坦(A .Einstein)在 1905 年认识到普朗克的能量子应该就是光子,并以此为依据处理了光电效应。一年以后,爱因斯坦考虑固体比热问题。他将固体(这里指绝缘晶体)的热能携带者认定为在点阵各点处作独立振动的原子,并将它们类比于光子而沿用了普朗克处理黑体辐射的方法。这就成为第一个固体的量子理论工作。但是,固体比热问题毕竟要复杂一些,热能的携带者并不是

振动着的原子！正如佩斯援引的那张曲线-数据图所表明的，在低温的一端，爱因斯坦的理论曲线比实验数据降落得快。

正确的结果，即低温时 T^3 行为，是德拜(P. Debye)在 1912 年首先获得的，紧接着玻恩和卡门(V. Karman)作了更精确的处理。德拜的考虑在物理上很有启发性。他想到，如果热能的携带者是作独立振动的原子，那么热怎么能在晶体中传播呢？如果考虑原子相互作用，热能象弹性波一样在晶体中传播，那为什么不直接考虑热能的携带者就是弹性波呢？德拜的弹性波是玻恩和卡门处理晶格振动后所得格波在晶格连续化极限下的近似。从今天的观点来看，这里热能的携带者的确是一种非常类似于光子的量子，叫声子，更确切一些，是声学声子。光子是电磁波的量子，声子是格波的量子。下面我们将以最简单的晶格，即一维单原子链为例，看下一声子是如何出现的。

原子排成点阵而形成晶体时并不是静止不动的。点阵中的几何点只是原子的平衡位置，而原子则相对其平衡位置作微幅振动。假定原子间的作用力是一种弹性力，且假设了一个适用于各原子之间的同一力常数，或称弹性系数。由虎克定律和牛顿定律写下原子的运动方程，这个方程的解是在晶格中传播的一种波，称为格波。这就是玻恩和卡门所作的“更精确的处理”。

在晶格连续化的极限下，也就是视晶格常数 a 非常小，整个晶体可以看作连续介质，这时格波的频率 ω_q 可以表示为 $\omega_q = cq$ 其中 c 是弹性波的速度， q 是波数或波长的倒数。它是德拜考虑弹性波为热能的携带者而不同于爱因斯坦的根本所在。在爱因斯坦的理论中 ω 是振子的频率，是一个常数，取正的有限值。因此在非常低的温度下，没有能量子可以激发，比热就几乎为零，这就是爱因斯坦的理论结果降落很快的原因。相反，对于德拜的考虑， ω_q 可以非常小，只要 q 很小就可以。这样，即使在非常低的温度下，也会有低频的格波模式激发，因此比热就不致于降落很快。由于整个系统的自由度有限，例如一维单原子链的情况，自由度是原子个数 N 。格波的模式个数不能超过自由度数，这个限制就给出了 q 的一个上限 q_0 ，相应的频谱的上限 ω_0 称为德拜频率，是一个非常有用的物理量。

在量子理论中晶格原子的位移和它相应的动量变成了算符，它们满足基本的量子力学对易关系。系统的哈密顿量由各原子的动能之和加上

原子之间的弹性相互作用势能(正比于相邻原子位移差的平方)之和构成。把这样一个相互作用着的晶格原子系统的运动量子化,实际上是把全部晶格原子的运动看成是象电磁波场一样的格波场,量子化是对格波场进行的。其结果类似于电磁场量子化,自然而然地获得了一个多声子系统,声子就是格波场的量子。由于格波是已经把原子之间的弹性相互作用考虑之后的结果,所以量子化所得的是独立的多声子系统。

声子是实际存在的,它有能量和动量,可以通过实验测定。声子概念的引入不仅能说明比热,在晶体的导热和热扩散问题中更是必不可少的。应该说声子是耦合的原子振动系统的等价描述所引入的物理对象,它是该系统的元激发。元激发就是基本的激发单元,我们在后面还要更多地谈到它。

这里从经典格波场到多声子的量子化方法提供了研究多体(这里是多声子)问题时引入量子场论方法的范例。回想前面我们在考虑周期场中电子问题时,最后归结为多电子在一条条的能带上填充。按照格波的范例,我们自然想到是否也能用量子场论的方法呢?回答是肯定的。我们只需把单电子的波函数的全体看作是一个经典的波场,将它量子化就能获得多电子的量子场论描述方法,这样的方法也称为二次量子化方法。与格波场量子化不同的是电子是费米子,必须用所谓的反对易关系量子化。

用了场论的方法,我们看到,电子和声子作为电子波场和格波场的量子是在一个统一的理论框架下被处理的,当然还应该包括电子和声子,电子和电子以及声子和声子的相互作用部分。这样,一个电子-声子系统的量子场论就建立起来了。这是凝聚态物理中场论方法处理问题的一个典型实例。今天,量子场论方法已成为凝聚态物理理论的基本方法之一。

4.3 量子液体、元激发

人类对物态变化系统研究是从 19 世纪 20 年代法拉第(M. Faraday)对气体液化的实验开始的。1823 年年轻的法拉第在一次有氯气产生的实验中,经人提醒发现密封的试管玻璃壁上有油状的东西产生。当他锉开试管要进一步研究时,突然发生的爆炸使他的双眼受伤。幸好情况并不严重,第二天他还能给人送去便条,说那油状物是液态氯。法拉第意识到这件事的重要意义,他紧接着就设计实验,在密封高压情况下产生各种气体来使它们液化。

几乎同时,另一位法国人凯纳·德拉都(Cagniard de la Tour)在 1822 年研究乙醚蒸发时,发现了以后被安德鲁斯(T. Andrews)称为临界点的现象。法拉第也看到了临界点现象与他的液化实验的关系。它意味着某些气体在常温下在任何压力下都不可能液化,需要高压和低温的组合才能成功。以后在 1835 年梯劳里厄(Thilorier)用二氧化碳液体做成了最终可达 -110 (约 160K)的低温库;1844 年法拉第构造了一个装置,可把各种气体压至 40 大气压,然后用梯劳里厄低温库冷却。这样,到 1845 年法拉第成功地将当时所知的“永久”气体除了六个以外都液化了。这剩下的六个中有三个是元素气体,它们是氧气、氮气和氢气,另三个是化合物气体,它们是一氧化碳、二氧化氮和甲烷。

在法拉第之后,许多国家的科学家们争相研究液化剩下的六个“永久”气体。由于它们临界点低,液化的努力实际上也就成了追逐低温的竞赛。人们研究出了用“连锁”降温 and 定压下气体膨胀的方法达到低温。这里,重要的是要找出“连锁”中一级级“适当”的气体,而在这一方面,1873 年范德瓦尔斯(Vander Waals)的理论工作具有极其重要的指导意义。经过许多科学家的努力,到 1896 年上述六个“永久”气体才最终都被液化了。

1895 年英国的化学家雷姆塞(Ramsay)从钷铀矿石中分离出了氦气,荷兰的翁纳斯(K. Onnes)立刻决定要液化它。为此他做了多方面的长期准备工作,仅为学会一种新的“连锁”降温的方法,他就花了 10 年的时间!翁纳斯在回忆此事时说:“我多么高兴地指给我的朋友——著名的范德瓦尔斯观看这液化了的氦。是他的理论指导我完成了这项液化气体的工作。”翁纳斯液化了最后一个“永久”气体的同时也就开创了量子

液体的研究。低温下，量子效应比较显现，液氦是一种典型的量子液体，对它的研究必须靠量子力学。这是真正的凝聚态物理意义上研究液体的开始。

翁纳斯在 1911 年又进一步发现了液氦在 $T = 2.2\text{K}$ 下开始膨胀；以后人们又发现液氦密度有一个明显的极大值和比热的 T 反常；1938 年卡比察(Kapitza)发现了毛细管中液氦的粘滞系数在从 4K 降到 T 后突然消失，这时液氦能沿着容器壁上爬，以至溢出容器，这是一种无粘滞性，也即无阻力的现象。这样，在氦实现液化 30 年以后，人们终于认识到它在 T 处发生了相变(图 4.4)。

氦有氦 3(^3He)和氦 4(^4He)二种同位素，它们的沸点是人们所知的所有物质中最低的： ^3He 在 3.19K ， ^4He 在 4.21K 。有趣的是在沸点下进一步降温，在通常的压力下，这二种液氦将保持液态直到 0K。要使它们固化，必须加压至二十多个大气压以上。有二个原因造成这种难以固化的特点。一个是氦原子的质量很小，因此它的零点能很大，零点能是量子系统特有的一个最低能量，它与质量成反比；另一个原因是氦为惰性元素，原子外的电子形成闭壳层，对称分布，只是氦原子核有很小的磁矩，因此氦原子之间的相互作用力很小。这二个因素合在一起，就使二个氦原子只能在比较大的原子间距下保持平衡，以至系统密度较低，呈现液态。唯一比氦原子质量更小的元素只有氢。但是，氢分子比氦原子容易极化，分子之间的作用力几乎是氦原子之间的 12 倍，尽管氢分子的零点能与氦原子的零点能几乎同量级，结果平衡时分子间距很小，系统稳定状态呈固态。其他元素原子或分子都比氦原子重，这就形成了氦这个独一无二的奇异特点了。

氦在 T 以下的另一相是 ^4He 的超流相，因为比热反常(图 4.4)，形状象希腊字母 λ ，故该相变点获 λ 点的名称，相应的转变温度就记为 T_λ 。对 ^4He 的这种比热反常，伦敦(F. London)在 1938 年提出了一个玻色-爱因斯坦凝聚的解释。

一个自由的玻色子系统在 0K 时，所有的玻色子都处于零动量态，这时系统的能量最低，这就是玻色-爱因斯坦凝聚，它是由玻色-爱因斯坦统计所得出的必然结论。 ^4He 是玻色子，按玻色-爱因斯坦凝聚理论，在零温时原子都处于零动量态，随着温度的上升，开始有的原子被激发到非零的动量态上去；温度越高，被激发离开零动量态的原子越多。由于

原子总数一定，因此就一定存在一个临界的温度 T_B ；此时刚好所有的玻色子都被激发了。反过来说，对于给定粒子数或粒子密度的玻色子系统，一定有一个 T_B ，当温度低于 T_B 时就会产生凝聚。按伦敦的解释，这个凝聚相就是 ^4He 的超流相。根据密度和 ^4He 原子质量计算得到 T_B 大约是 3.1K，这与 T 的 2.2K 相差不多；但是进一步的比热计算，给出图形如图 4.4 中虚线所示那样，差别很大，玻色-爱因斯坦凝聚的比热曲线是连续的有尖顶，而 ^4He 的是不连续的。

实际的液态 ^4He 系统的 ^4He 原子之间是有相互作用的，因此上述比热曲线的不同并不出乎意料。重要的一点是自由的玻色系统虽然会有凝聚现象，但不会有超流现象。考虑相互作用又怎样解释超流现象呢？从 40 年代一直到 50 年代，朗道对这一问题首先作了出色的开创性的唯象理论研究。朗道提出， ^4He 有如图 4.5 所示的能谱，这是 ^4He 的元激发谱。图中直线段部分是声子元激发，能量为 $\epsilon(p) = cp$ ，其中 p 为元激发的动量。低谷的地方是称为旋子的元激发，在那附近 $\epsilon(p) = \Delta + \frac{(p - p_0)^2}{2\mu_y}$ ，其中 Δ 是旋子元激发的能隙， μ_y 是有效质量。

这样的能谱怎么解释超流现象呢？ ^4He 液体在容器中流动有一个临界速度 V_C ，低于这个速度液体可以无阻力地流动，超过这个速度液体就受阻。所谓受阻就是容器壁和液体相互作用而激发液体，伴随着传递能量和动量给液体。临界速度当然也就是液体能够被激发的最小速度。这个最小速度自然与液体能被激发起的元激发的能谱有关。元激发是系统能够被激发的最小单元，这是朗道首先在凝聚态物理理论中引入的。他所给出的上述能谱，后来在 1958 到 1961 年被中子散射实验所证实。唯象理论中的一些参数 c 、 μ_y 、 p_0 和 A 实验上都能确定。

需特别指出的是，这里朗道所提出的元激发的概念是氦原子系统作为连续媒质的液体对平衡态的密度涨落形成的声波场的量子，即液体的声波的量子。从这一点来说，朗道可以说是最先引入声子概念的。旋子是 ^4He 所特有的元激发。

1957 年朗道又提出了适用于描述低温常规 ^3He (非超流态) 和一定条件下金属及半导体导带电子的费米液体理论。在这个理论朗道提出了系统的元激发——准粒子的能谱。它们的基态是费米面附近相互作用的费

米子系统，因此所有的动量 p 都接近费米

动量 p_F 朗道提议低温时激发态和基态的能量差 $\epsilon_{p\sigma}$ 的最低级近似 $\epsilon_{p\sigma}^0$ 为：

$$\epsilon_{p\sigma}^0 = \frac{p^2}{2m^*},$$

其中 m^* 是准粒子的有效质量。修正项依赖于系统的体积 V 和准粒子的密度涨落 n_p ，其中 σ 是准粒子的自旋。

借助于这一“简单”的谱表示，人们可以研究系统的所有热平衡性质：比热、磁化率、压缩率等等，还可以进一步在半经典的近似下研究输运性质。对于 ^3He ，在实验确定少数几个理论参量后由理论计算的物理量都与实验结果吻合得很好，理论还预言了一些新的物理现象也都被实验所证实。朗道因为他在凝聚态物质，特别是液氦方面的理论研究而获得了 1962 年的诺贝尔物理学奖。

有意思的是朗道在液氦方面的工作都是从“猜测”能谱入手的。这令人想起了玻尔。玻尔把普朗克的能量子概念引伸，“猜想”了原子的能谱，从而正确地给出了光谱线的频率法则，对量子力学的突破性的进展做出了卓越的贡献。爱因斯坦对玻尔这一工作的评论，可使我们对此有更深刻的体会。

“……大家早已知道，经典力学在物质的终极构成方面破产了，也知道原子是由带正电荷的核和围绕着核的一层结构比较松驰的粒子所组成。但是，经验上知道得相当清楚的光谱结构，同我们的旧理论所预期的东西有很大的出入，以至谁也不能对这些观察的规律找到令人信服的理论解释。而玻尔在 1913 年在量子论的路线上对最简单的光谱想出了一种解释，……作为一位科学思想家，玻尔所以会那么惊人地引起注意，在于他同时有大胆和谨慎这两种难能可贵的品质；很少有谁对隐蔽的事物具有这样一种直觉的理解力，同时又兼有这样强有力的批判能力。他不仅具有关于细节的全部知识，而且还始终坚定地注视着基本原理。他无疑是我们时代科学领域中最伟大的发现者之一。”

类似地，在液氦的研究上朗道“猜测”了 ^4He 的能谱，此谱揭示了超流现象；“猜测”了 ^3He 的能谱，此谱揭示了元激发准粒子的热平衡和输运性质。朗道与玻尔相处并不长，但他却终生以玻尔的学生自居，对朗道来说，这是少有的“谦恭”事实。在这里人们或许可以领悟到其中的缘故；同时，这一对照也许多少使人们能理解到在液氦的研究上，

朗道所表现出来的惊人的“直觉”的由来。这一从玻尔到朗道一脉相承，并有所发展的物理直觉思维方法或许是值得借鉴的。在一定的意义上可以说玻尔奠定了原子乃至量子力学的唯象基础，而朗道则奠定了凝聚态物理理论的基础。

我们知道，凝聚现象是由大量微观粒子所组成的，是一个复杂的多体系统。虽然决定凝聚体系的最基本的规律是简单的，但由于复杂的相互作用，多粒子体系的统计性质以及巨大数目的自由度，导致不同的凝聚体系或同一凝聚体系在不同条件下呈现出形貌各异的基态和元激发。凝聚体系的低激发态往往可以看成一些独立的元激发组成的稀薄气体。元激发有两类，一类是单粒子激发，如金属中的电子和空穴等；另一类是集体激发，如晶格原子振动的格波，磁系统中的自旋波等。从统计性质上元激发可以分成玻色型和费米型，分别遵从玻色统计和费米统计。当温度较低，元激发的浓度也较低，元激发之间的相互作用可以忽略；但在较高温度下，必须计入元激发之间的相互作用。

今天，元激发的概念已成为凝聚态物理中最基本的概念之一。我们不仅有声子、旋子、准粒子，还有等离子、激子、极化声子、极化子、磁子、孤粒子、相子、……还有所谓拓扑元激发如涡旋、反涡旋……。在不同的系统中寻找某种元激发已成为凝聚态物理的基本任务。

近一个世纪的科学发展的事实说明，凝聚态理论有关基态和元激发的研究，总是不断地有新的发现，又不断遇到新的挑战。以反铁磁体基态与元激发的研究为例。30年代初贝特对一维量子反铁磁海森堡模型，采用贝特方案求出了严格解，证明其基态不是简单地由相邻自旋反平行排列而成的尼尔态，而是各种可能间距的，由一对反平行自旋组成的自旋单态的叠加。这种基态总自旋为零而没有长程序，它的元激发，由于其基态具有时间反演不变性而遵守线性色散关系。50年代初安德森证明了三维量子反铁磁海森堡系统在热力学极限下的基态具有长程序，这种尼尔长程序的形成是由这样一些接近但不是完全的组态，它们在热力学极限下与自旋为零的态能量简并，从而迭加在一起而形成的。这是由于量子涨落的结果，其元激发是反铁磁自旋波。然而，二维量子反铁磁海森堡系统的基态与元激发至今仍未完全解决。在高温超导体发现后，安德森重申这是一种共振价键态(RVB)，它与一维链的基态类似。有人用计算机模拟的方法和重正化群方法，得到正方点阵上自旋 $1/2$ 的基态接近

尼尔态，但量子涨落比三维更强。此外，二维三角形点阵还存在自旋错配的问题。人们相信，二维量子反铁磁体的基态和元激发比原来设想的要丰富得多，可以肯定量子涨落起着十分重要的作用。

我们在前面曾说到将场论的方法用于凝聚态物理，能使我们在统一的框架下处理多电子、多声子系统。在那里我们所处理的电子和声子还只是自由的布劳赫电子和声子，如果考虑相互作用，例如布劳赫电子和声子的相互作用，那么这些作为元激发的电子和声子都会改变。由于有相互作用，电子的运动就会不同，它的能量就会改变，不同动量的电子分布也会改变；同样，声子的能量和不同动量的声子分布也会改变。考虑到相互作用以后的实际的元激发，称为重整化了的元激发。在本书第二篇我们曾提到在粒子物理中从裸粒子到物理粒子需经过重整化的处理，并特别强调重整化不仅仅是为了消除发散所必须，重整化的根本目的是为了得到物理粒子，因此即使没有发散也会有重整化问题。在凝聚态物理中大量遇到的就是未必有发散问题的重整化。除了自由的系统，一般系统的元激发都是重整化了的元激发。寻找重整化了的元激发必须运用场论方法，从这一点来说，在凝聚态物理理论中运用场论方法势在必行了。

事实上，正是为了在微观的层次上解释朗道的元激发特别是他 1957 年的费米液体的准粒子概念，人们全面地把场论方法引入统计物理并取得了一些成效；同时还用场论方法处理超导问题，效果十分引人注目，由此便开始了场论方法处理凝聚态物理问题的普及化。从 50 年代以来，量子场论的发展就不间断地对凝聚态理论起着重要的推动作用，而凝聚态理论的发展又反过来影响和促进了量子场论的发展。例如量子场论中的费曼—戴森微扰论被应用到凝聚体系，成为一种强有力的研究元激发能谱及有关性质的方法；而凝聚态理论中的超导理论的研究成果对量子场论中粒子质量的产生机理的研究起了很大的推动作用。随着粒子物理理论的深入发展，人们认识到粒子物理实质上也是多粒子系统的问题。凝聚体系的基态与元激发的概念在粒子物理中对应的是真空态与物理粒子。从这个意义上来说，凝聚态物理中的元激发概念反过来对粒子物理的一些基本概念的理解也会有很大的启示。凝聚态物理基本上是在原子的层次上处理问题，而基本粒子物理是在更深一个或几个层次上处理问题，从这一点来看，基本粒子似乎也应看作是相应的某一层次的元激发，

这或许是理论物理的这二个学科经常在概念上和方法上相互借鉴的缘故吧！

4.4 相变、序参量、对称破缺、重整化群

上一章提到，几乎与法拉第忙于气体液化的同时，1822年德·拉·都在研究乙醚蒸发时发现，在160和37个大气压左右乙醚液体变成了同一密度的气体，二种态之间没有界面。近半个世纪以后安德鲁斯对这一现象进行了深入探讨并将此状态称为临界点，他的工作导致了相变的理论研究。

平衡态可以用一组参数来描绘，这些参数有几何的、力学的、物理的或化学的等等。当然，还有反映热平衡的参数——温度 T 。在这些参数之间有一个关系式就是物态方程。最简单的均匀气体或液体的热平衡态用体积 V 和压强 P 来描绘就足够了，这时有物态方程 $f(P, T, V)=0$ 。于是，任一个态可以用三维空间 P - V - T 中的点来表示。常见的三态，固态、液态和气态分别由此空间中的不同的曲面来表示，相互间又以一定的曲面、线或点为界，组成一幅很复杂的几何图形，称为三维相图。

实际上，人们是由实验测量先知道上述三维相图在各坐标面上的投影也即二维相图，然后汇总而获得三维相图的。例如，玻义耳在1662年发现在一定温度下定量气体满足 $PV=\text{常数}$ 。以后，查理(J. Charles)发现在一定体积下定量气体有 $P=\text{常数}$ ，盖-吕萨克(J. L. Gay-Lussac)发现在一定压力下定量气体有 $V=\text{常数}$ 。它们分别给出 P - V 、 P - T 和 V - T 的二维相图，汇总后就给出了所谓理想气体的物态方程 $PV=Nk_B T$ ，其中 N 是气体的总分子数， k_B 是玻尔兹曼常数，它的三维相图如图4.6所示。由该图很易看出它的三个平面投影所对应的二维相图恰是上述三个气体定律所反映的关系。理想气体不考虑分子相互作用，是没有相变的。

图4.7是一个通常的 P - T 相图，它是对一定的体积 V 而取的。这里有三条曲线。S、L、G分别表示固、液和气相。OA称为升华线，A称为三相点，也就是说在A点所对应的 P 、 V 、 T 处，固、液和气三相同时间并存。AM线是没有终点的，因为如果有终点，那就意味着超过该终点固态和液态就没有区别了。然而这是不可能的，因为液态是不会有固态那样的晶格对称性的。当然，它或许会终止于另一种相，例如介于固态和液态的非晶态等等，这就要视各种不同物质而定了。C点就是临界点，它是汽化线的终点。从图4.7可以看到当温度高于 T_C ，也就是说在C点的温度坐标 T_C 右边，无论如何升高压力也不会使气体变成液体的；必须在 T_C 的左

边时再加压才可能图使气体液化。图 4.8 是相应的三维相图。

法拉第和安德鲁斯的研究工作上正好分别是一阶相变和二阶相变的工作。关于阶的定义我们在后面再谈，这里先看一下二种相变的特点。图 4.9 是图 4.8 三维相图的 P-V 投影，我们看到，在 T_c 以下保持温度恒定，随着压力的增大气体体积减小。到一定的值 V_G (称为饱和汽)时，进一步减小体积部分气体开始液化而压力保持恒定，在图 4.9 中就是水平线那一段；当体积被压到汽体全部液化后，再要压缩就需很大的压力，因为液体的可压缩性是很小的，反映在图上是又一段上升的曲线。水平线这一段是气体和液体共存的相，这时虽然温度固定不变，但体积发生改变(缩小)并伴随着有热量放出，称为潜热。从气态到液态是一个有序化的过程，用物理学的话说是熵减小的过程。熵是反映系统有序化高低或无序化低高的一个物理量，无序化越低或有序化越高熵就越小，反之则大。

在 P-T 投影面上，如图 4.7 所示上述水平线就成一点了；在固定的 $T(T_c$ 以下)，从气态跨过汽化线进入液态，熵会突变。而在 C 点，它在图 4.9 中无水平线段，在该点既无体积的改变也无潜热，即熵是不变的或连续的。熵在热力学中可以由一个自由能的热力学势对温度求一次导数而得。这样，在 C 点以下热力学势是连续的，但是它的一次导数有突变、不连续，这就称为一阶相变。这个定义是 1933 年厄伦弗斯给出的。一般地，如果直至 $k-1$ 阶的热力学势导数连续，而第 k 阶导数不连续，这样的相变称为是 k 阶相变。关于 C 点是二阶相变点的结论，我们在下面会看到。

1873 年荷兰的范德瓦尔斯对安德鲁斯的实验结果提出了一个唯象的理论，它是历史上第一个从物态方程出发研究相变的理论。理想气体是不计气体分子相互作用的，范德瓦尔斯对此作了修正。考虑了气体分子有一定的体积和分子间有一定的吸引力，他建立了一个比较实际的气体物态方程：

$$\left(P + \frac{aN^2}{V^2} \right) (V - Nb) = Nk_B T,$$

其中 b 是分子的体积， $a \frac{N^2}{V^2}$ 相当于气体分子间微弱吸引的“内压力”。

范德瓦尔斯方程的 P-V 图如图 4.10 所示。这里明显的有相应于图 4.9

中的C点。但在C点下的曲线与图4.9不一样，它表明范德瓦尔斯方程只是一种近似。范德瓦尔斯从分子运动论观点出发所作的修正，受到了麦克斯韦的赞赏。麦克斯韦对C点下曲线提出了修改的原则：这就是在每一条等温线的起伏处用一条水平线来代替，水平线的高低应选取在恰使水平线上右方部分的面积等于水平线下左方部的面积，这就是麦克斯韦等面积原理。按照热力学的普遍原理，一个可逆的等温循环过程对外做功为零，即可推得上述等面积原理。

范德瓦尔斯方程是V的三次方程，对任一给定P一般会有三个根，它们相应于起伏处与水平线相交有三点。但C点处只有一点，这必然是一个三重点，根据这一原则就可以求得临界点

的 V_C 、 T_C 和 P_C 。它们是： $T_C = \frac{8}{27bk_B}$ ， $P_C = \frac{a}{27b^2}$ ， $V_C = 3Nb$ 。这是一个

重要的结论。当年人们在实现气体液化以至追逐低温的“连锁”降温中，正是根据这些理论结果，结合实验来选择一级级“适当”的中介气体。

如果引入三个无量纲的量 $t = T/T_C$ 、 $p = P/P_C$ 和 $v = V/V_C$ ，则范德瓦尔斯方程变为：

$$\left(p' + \frac{3}{v'^2}\right)(3v' - 1) = 8t'$$

它是一个与气体物质无关的方程，称为对比物态方程，对一切气体都成立。二种气体，只要它们的 t 、 p 和 v 中有二个相同，另一个也必相同，这称为对应态律。

由范德瓦尔斯方程和麦克斯韦等面积原理可以求得在临界温度 T_C 附近有：当 $T < T_C$ 时

$$\frac{V - V_C}{V_C} \sim \left(1 - \frac{T}{T_C}\right)^{\frac{1}{2}}$$

而当 $T > T_C$ 时，压缩率即增加单位压力时气体体积的相对改变率 K_T 为

$$K_T \sim \left(\frac{T}{T_C} - 1\right)^{-1}$$

K_T ，正比于称为吉布斯势的热力学势 G 对 P 的二次导数，而 V 是 G 对 P 的一次导数。由上面的两个式子可见， G 对 P 的一次导数在 T_C 处连续而 G 对 P 的二次导数在 T_C 处不连续，有突变。这样，按照我们前面所提

到的关于相变阶的定义，临界点处的相变就应该是二阶相变了。

上面两个表示式中的指数，就是所谓临界指数。它曾是临界现象中长期困惑理论物理学家们的首要问题，直到一个世纪后的本世纪 70 年代，场论用于临界现象，特别是重整化群方法的引入，人们才开始对它有了真正的认识。

在热力学中描述相变，一般是先确立不同相的各自化学势，然后由化学势的相等来定分界线或称平衡线，如 P-T 图中的汽化线，它是气、液二相的平衡线。但热力学中讨论相变的这种方式没有告诉我们是否可以用一个热力学函数来描述相变，范德瓦尔斯的工作是历史上第一个启示了上述的可能性。对范德瓦尔斯方程的分子运动论基础的探索，导致了真实气态方程的统计研究以及用统计方法直接研究相变。这就是，从一个配分函数出发来研究相变。配分函数是如下定义的一个温度 T 的函数 Z(T)

$$Z(T) = \sum_n e^{-\frac{E_n}{k_B T}}$$

这里 E_n 是系统的一个可能的微观态(也就是量子力学意义上的本征态)的能量， k_B 是玻尔兹曼常数。系统的一切热力学函数都可以由 Z(T) 推导出来。

如果能用一个热力学函数来描述相变，那么怎么来区分不同的相呢？经过多年对各种相变的唯象研究，人们认识到不同的相反映了不同的序。最早用“序参量”来描述相变的，可能是 1934 年布喇格(W. Baragg)和威廉姆斯(E. Williams)关于合金有序-无序转变的理论，那是一个二阶相变问题。差不多同时，1935 年朗道在一篇关于一阶相变的论文中，声称将戴林格(Dehlinger)的序变概念定量化而引入了序参量。两年以后朗道提出了一个二阶相变的唯象理论，这个理论实际上包括了当时所知的所有的二阶相变唯象理论，例如气液临界点的范德瓦尔斯理论、铁磁体居里点的外斯(P. Weiss)理论以及合金有序-无序转换的布喇格-威廉姆斯理论等。

这里我们以铁磁体的居里点为例，看一下朗道的二阶相变理论。一个理想的铁磁体，在所谓居里温度 T_C 以下，即使没有外磁场也会有自发磁矩；在 T_C 以上，没有自发磁矩，如图 4.11 所示。在 T_C 处，磁矩大致如 $(T - T_C)$ (若 $T < T_C$)，是一个临界指数，实验值大约为 1/3。

朗道假设系统的自由能密度 f 是磁矩密度的解析函数。在临界温度附近，序参量——这里是磁矩密度 M ——应是一个很小的量，可以把 f 展开成 M 的级数，且只保留到 M 的四次方项。然后，根据系统的平衡态对应于 f 取极小，即可证明在临界点以上应有 $M=0$ ，而在 $T < T_c$ 时 M 不为 0，它满足 $M \sim (T_c - T)$ ，或者说临界指数 $\beta = \frac{1}{2}$ 。

朗道的二阶相变唯象理论看来十分简单，但它的物理含义是十分深刻的。这里我们只是先谈一下用序参量对相变分类的意

义。按照序参量的描述，一阶相变的序参量有突变，而二阶相变的序参量却是连续的。一阶相变的例子如非临界点处普通气液相变，外磁场中的超导转变等。在那里热力学势连续而比热、磁化率或压缩率等不连续。二阶相变的例子有前面提到的 ^4He 在 λ 点的超流转变，没有外磁场的超导转变点，气液相变的临界点处，以及许多磁相变的临界点处(如铁磁的居里点)等。自然界中常见的只有这二种相变。

一阶相变序参量有突变，自然能反映相变；二阶相变序参量连续无突变，那么相变在何处呢？在于对称破缺！这一概念是朗道首先提出来的。按照朗道相变理论，高对称性相中某一对称元素的突然消失，就对应于相变的发生，导致低对称性的有序相的出现。低对称性的相具有比描述系统的哈密顿量较低的对称性，这就是对称性破缺。对称从有到无或从无到有是一种突变，用一个新的序参量来描述新的对称，在高温相或对称性较高的相，这个序参量为零，而低温或对称性较低的相，它不为零，这就是朗道的思想。当然，要认识表征低温相的对称性破缺的有序化的性质常常是不容易的。

凝聚态物理中这一对称破缺的概念，在 1950 年又被南部(Y.Nambu)用到粒子物理并加以深化，提出了自发对称破缺的概念。这一概念和规范场的概念一起构成了弱电统一标准模型的二大支柱。对于自发对称破缺，本书第二和第三篇中已作了详细的介绍，这里不再赘述。只是强调指出在多体系统中这种现象相当普遍，铁磁体就是一个熟知的例子。哈密顿量具有旋转不变性，但基态的自旋在某一方向排齐了。而且，任一从基态中产生有限元激发的激发态都保持这种各向异性。此外自发对称破缺的概念又反过来使人们认识到，有些系统，特别是具有所谓宏观量子效应的系统的二阶相变理论中的对称破缺，实际上应该归于它们的基

态的自发对称破缺。

如果说场论中自发对称破缺的概念体现了凝聚态物理与粒子物理在思想上的交流，那么威尔逊(K.G.Wilson)在 60 年代中至 70 年代初将场论运用于临界现象的研究，终于在 1972 年解决了长期以来困惑理论物理学家们的临界指数问题的出色工作，则是充分体现了二者结合的重大进展。下面我们还是以铁磁的朗道二阶相变理论为例，简单地说明威尔逊的重整化群方法在临界现象的运用。

还得从朗道那简单而含义深刻的唯象理论说起。在前面介绍朗道的二阶相变唯象理论时曾用了“系统的平衡态对应于 f 取极小”，对此人们或许会用通常的热力学中平衡态对应热力学函数取极值的原理来理解，这样做也未尝不可，然而并不正确！而对此的正确理解，恰恰就是威尔逊重整化群方法用之于临界现象的最初突破口！

在铁磁问题中除了居里点的自发磁矩之外，还有居里点以上的磁关联长度问题。在原子的尺度上，磁矩是由于自旋方向没有配对的电子的磁矩引起的，在这个意义上来说磁现象的本质是量子效应。在铁磁体中，邻近原子磁矩相互平行时在能量上要低于反平行。高温时热涨落阻止了磁体中的诸原子磁矩排成有序状态，但是当温度逐渐降到接近于居里点时，一个原子磁矩的定向会引起一定距离内诸原子磁矩的偏向，这个距离称为关联长度。在居里点处，变为无限大，这样整个磁体就在一定方向排列了。实验上发现在 T_c 以上， $\xi \sim (T - T_c)^{-\nu}$ ，这里 ν 大约为 2/3。

1950 年金兹堡和朗道(Ginzburg-Landau)考虑了磁矩随空间位置不同而变化，将上述朗道二阶相变的唯象理论推广来处理关联长度问题。这

时的自由能为 $F = \int dr^3 f[M(r)]$ ，由 $f[M(r)]$ 取极小可以解得关联长度为 $\xi \sim (T - T_c)^{-\nu}$ ，但 $\nu = 1/2$ 。这表明，按朗道理论，不论是临界指数

还是 ν 都与实验不符。如此简单的理论的结果与实验不符并不出人意料，它毕竟给出了定性来说大致符合的结果。令人赞叹的是，正是威尔逊对朗道理论的正确理解，引导人们找到了对它改进的方法。

按照统计力学的观点，热力学的自由能可由系统的哈密顿量计算出来。计算时要对任一组完备可观察量的本征态求和。朗道的唯象理论相

当于在磁体中每一点处的磁矩用一个整体的平均磁矩来代替，并假设如此而得的自由能密度仍保持解析性(即为平均磁矩的解析函数)。然而这样所得到的结果并不是系统的最终自由能，还需要对各种可能的平均磁矩求和。第一步的平均，加上这第二步求和才组成了“对完备的本征态集合求和”。这样所得的 F 才是真正的系统的自由能。当取热力学极限时只有自由能密度取极小的项存在，正是 f 的极小导致了最终非解析的表达式！

在后来金兹堡-朗道讨论的非均匀的情形，实际上假设了只有原子尺度上的涨落是重要的，一旦这些涨落被平均掉，便获得了一个连续的磁矩 $M(\mathbf{r})$ ，它只在外磁场作用下涨落。同均匀的情况一样，在热力学极限下，系统的自由能对应于 $f[M(\mathbf{r})]$ 取极小，最终也得到了非解析的结果。事实上，从系统的哈密顿量直接计算自由能 F ，对于任何一个有限的系统，在 $T \neq 0$ 点结果总是解析的，而铁磁体在 $T=T_c$ 附近 ($T_c \neq 0$) 有非解析性，这只能表明此非解析性来源于热力学极限，或说系统无限大，这是在统计层次上普遍成立的结论。它的微观原因，至少在铁磁的情况，正如前面已提到的，在于此时基态的自发对称破缺。按照上述普遍成立的结论，如果只考虑有限大的系统是不可能发生非解析性的，或说这时不能取 f 的极小来定系统的自由能。正是在这个意义上，朗道二阶相变唯象理论中“系统的平衡态对应于 f 取极小”这一步，不能用通常热力学中平衡态对应热力学函数取极值的原理来理解！

有了这样的正确认识后，自然就会想到无论是将各点的磁矩用整体的平均磁矩来代替，还是用将原子尺度上的涨落平均掉后的一个连续的平均量来代替，都只是一种近似。朗道唯象理论相当于二个极端，都只考虑有一个典型尺度，前者是非常长的长波尺度，后者是原子尺度而取的近似。然而，许多现象中并无典型尺度，而是从原子尺度直到宏观波长尺度都同样起作用。上述铁磁体居里点的相变也属此类，在接近 T_c 处，关联长度变得非常大。为了简单起见，这里我们仅考虑对朗道唯象理论作适当的改进，也就是说只是考虑比原子尺度大许多的长波长的作用。

设 $M_L(r)$ 为把波长小于 L 的部分已经平均掉的磁矩，只含有大于 L 的波长部分，这里 L 是比原子尺度大许多的某一尺度。按照我们前面对朗道唯象理论的正确理解，那么对它的改进就是用 $M_L(r)$ 代替前面的 $M(r)$ 来计算自由能 F_L ，它是一个随尺度 L 变的量。这里我们考虑无外磁场的情况，另外将 3 维代之以任何 D 维。从物理上看，整个理论对 L 的依赖应该终止于关联长度 ξ 。波长长于 ξ 的涨落在物理上作用已很小可以略去。一旦直至 ξ 的所有的波长都被平均掉了，就可以运用朗道理论了。为了研究不同尺度上的涨落，我们采取逐步前进的方法，每一步仅考虑一种尺度，这就是“重整化群”方法的基本原则。

例如，我们先考虑波长在 L 到 $L + \Delta L$ 区域内的相应的涨落。设 $M_H(r)$ 是把波长小于 $L + \Delta L$ 的部分已经平均掉了的磁矩，利用 $M_H(r)$ 和 $M_L(r)$ 之间的近似关系导出 $F_{L+\Delta L}[M_H]$ 与 $F_L[M_H]$ 之间的关系，并由此建立一组微分方程，可以解出 $F_L[M_L]$ 用 M_L 展开的系数满足的一些关系式。结果表明，空间维数 D 是起很大作用的。事实上，如果 $D > 4$ ，则当 L 足够大，与朗道唯象理论的结果一致。因此，在这个意义讲，朗道理论只是对 4 维以上的空间才是正确的。而对 $D < 4$ ，我们来看一下改进后的临界指数。当 L 大到接近关联长度 ξ 时，应有

$$\xi \sim (T - T_c)^{-\nu},$$

$$\nu = \frac{1}{2} \frac{1}{1 - \epsilon/6},$$

和

$$M \sim (T - T_c)^\beta,$$

$$\beta = \frac{1}{2} - \frac{\epsilon}{2} \frac{1}{1 - \epsilon/6}$$

其中 $\epsilon = 4 - D$ 。在 $D = 3$ 时， $\nu = 0.6$ 而 $\beta = -0.1$ 。这些具体的数值结果并不重要，由于这里的计算作了许多简化，这些公式不可能是重整化群方法的精确结果。但它们至少表明 ν 和 β 的值是可以改进的。

从这个简单的例子，我们大致可以知道重整化群方法的基本思想。所谓重整化就是从一尺度到另一尺度时有关量的一种变换，所建立的方程就是所谓重整化群方程。在临界点，关联长度趋于无穷大，人们

要作无穷次的尺度变化，这就类似于一种“群”的结合运算，从而有“重整化群”的名称。

如果我们把有关量写成无量纲形式，那末就会看到在临界点， F_L 成为一个尺度变换下的不变量，或说是固定点，也即它与原子尺度的细节毫无关系！这表明一些原子尺度不同的系统，有可能在临界点处有同样的临界指数。例如气液相变点，铁磁相变点和合金有序—无序相变点等都具有同样的临界指数，这就是所谓普适性。此外，还可发现序参量 $M(r)$ 以及一些有关的量在临界点处有所谓标度性，即在尺度发生改变时，这些量只差尺度变换的某一幂因子。

威尔逊的重整化群方法并不是为处理临界指数问题而“设计”的。作为方法本身，这是他对量子场论的研究，特别是对以往的建立在微扰论基础上的重整化方法试图作出更好的理解而想到的。只是由于相对论量子场论的数学上的复杂性，使得他的方法一时很难从中抓住关键的物理东西。1970年，一次偶然的机促使他考虑把他的方法用于金兹堡-朗道的理论，由于这里数学上简单得多，物理容易显示出来。第二年他发表了将他的方法运用于临界现象的第一篇论文，在这里他给出了变换的递推公式，算出了临界指数，同时找到了固定点因而证明了标度性。以后他与费希(M.Fisher)一起发现了 $d=4-D$ 展开的有效方法，1972年他结合展开和场论算图的方法，在1971年工作的基础上不再引入任何近似，用他的重整化群方法处理了金兹堡-朗道模型。

总结上述可以说，相变、临界现象以及由此发展出的一系列新的概念和方法，如对称性破缺，序参量，标度律，普适性，标度变换和重整化群理论等是与元激发的研究并驾齐驱的凝聚态理论发展的另一条主线。

对称性破缺概念是朗道在研究连续相变中首先提出来的，这个概念在凝聚态物理中占据了核心的地位。按照朗道相变理论，液固相变是空间平移对称性破缺的结果，顺电-铁电相变对应于空间反射对称性破缺，顺磁-铁磁或顺磁-反铁磁相变对应于时间反演对称性破缺，超流、超导相变对应于规范对称性破缺，等等。为了表征相变，朗道还引入了序参量，它是某个物理量的平均值。事实上它描写了系统的一种关联或者宏观相干特性。序参量可以是标量、矢量、复数或更为复杂的量。

由于引进了重整化群思想，二阶相变点附近临界现象的理论在近 20

年来取得了重大进展，始终是凝聚态理论研究中一个十分活跃的领域。威耳逊发展的临界现象的重整化群理论，不仅为唯象的标度理论提供了微观基础，而且为计算临界指数提供了强有力的方法。特别重要的是，它可以把普适性极其自然地包含在理论之中，从而使得我们对临界现象的认识大大深入了一步。此外，威尔逊重整化群方法在临界现象运用的成功，特别是对固定点有关的物理讨论，反过来又加深了人们对以往重整化理论的理解。重整化群理论不仅可以应用于相变和临界现象，而且还可以应用于具有标度行为的其它凝聚态系统，最突出的例子是无序系统的局域化和近藤问题，值得指出的是在二类相变点附近的临界现象主要是经典涨落起作用，而对于这两者则是量子系统。

近年来，有关相变与临界现象的理论研究不断取得新的进展，开辟了一些新的研究方向，如低维相变，浸润相变，分形结构系统中的临界现象，以及动态临界现象等等。其中十分有趣的是自旋玻璃相变。自旋玻璃是一种特殊的无序磁系统，当温度降低时，其自旋将“冻结”到空间无序的各个方向上，因而没有空间长程序，但存在时间长程序。自旋玻璃的基态是高度简并的，来源于自由能有大量的局部极小。理论上特别感兴趣的是临界慢化以及基态和元激发的性质，迄今仍未解决。此外，某些解析及数值模拟研究表明转变可能不是平衡相变，而是涉及弛豫时间的多时标特征。自旋玻璃的研究还促进了其他几个领域的发展，其中之一是神经网络的自旋玻璃模型，在这个模型中，自旋方向类似于神经细胞的“开”或者“关”的状态，不同的自旋组态代表不同的记忆。神经网络的自旋玻璃模型的研究还有许多问题有待深入。量子相变和量子临界现象是近年来新开辟的又一个研究方向。我们知道，通常的相变和临界现象发生在有限温度，其中热无序与相互作用形成一对互相竞争的因素。倘若所有涨落模的特征能量远小于 kT_c 时，经典统计力学模型适用，这正是最初威耳逊重整化群理论所处理的情形。然而，如果研究零温下的相变，当某一耦合常数达到一定的阈值(临界值)时相变发生。在这种情况下，所有涨落模的能量均与热能无关，它们的统计行为将是高度非经典性的。此外，即使不在零温，只要温度足够低，使 kT_c 远小于系统的微观特征能量以至量子涨落比起热涨落来占压倒优势，也属于量子相变。赫兹最早提出这一观点，并提出量子重整化群方法，它是威耳逊重整化群理论的推广。赫兹从哈伯德模型出发，应用斯特拉脱诺维奇-哈

伯德变换，把系统的自由能转换到泛函积分的形式，并应用重整化群方法，研究了有关巡游电子铁磁性和反铁磁性的量子临界现象。近来，有人研究了二维量子反铁磁体的量子相变，得到了均匀磁化率，关联长度和核磁共振弛豫率的普适标度形式，并与镧锶铜氧的磁性质进行了比较。特别值得指出的是，由于对高温氧化物超导体正常态性质研究的推动，最近又提出来，对铜-氧平面系统，在一定的载流子浓度下，是否存在一个“量子临界”区。所以，无论从实验还是理论的角度来看，量子相变和量子临界现象是一个有意义而急需开拓的重要研究领域。

4.5 超流、超导、BCS 理论

超流和超导都是一种相变，并且都是量子效应的宏观表现。对此，1954 年伦敦曾提出一个非常有意思的观点。他说除了超流(包括 ^4He 和超导)以外，唯一的另一种无耗散的流动，就是原子和分子系统中的轨道电子流。在一个稳定的原子中，每一个电子都占有一个定态。因此他认为超流是一种量子流，应该有一个遍布整个系统的波函数。至今对超流和超导的研究都支持这一观点。

朗道对超流所提出的唯象理论是成功的。它与上述伦敦的观点有什么关系呢？朗道唯象理论中曾假设在临界速度 V_c 下整个液体有速度 V 。作为量子液体，这就隐含着系统的原子或分子由于相互作用而呈现某种“刚性”。这种“刚性”就应体现有一整体的波函数。所谓整体，就是说波函数不能分解为几个独立部分的乘积。“刚性”应显示有平移的不变性。

对于一个自由的玻色-爱因斯坦凝聚的态来说，波函数并不是一个整体的波函数，当然不会有超流。只有考虑了相互作用，系统激发态的波函数才可能成为一个整体的波函数。1955 年费曼从物理角度果然给出了这样的激发态波函数，并用变分方法证得元激发谱 ϵ_q 为 $\epsilon_q = \frac{q^2}{2mS_q}$ ，其中 S_q 是系统的静态结构因子，它是一个实验上可测的物理量。由实验上测得的 S_q 值代入上式，所得的 ϵ_q 与实验曲线定性符合。这是第一个“基本证实”朗道理论的微观理论，它同时显示了元激发谱中的声子部分和旋子部分。在量值上，旋子部分似乎偏差大了一点。后来考虑了“回流”的费曼-康恩(Feynman-Cohen)理论大大地改善了结果。沿着这个方向，今天已发展起一整套的计算方法，借助于计算机，人们已能得到与实验几乎一致的结果，然而即使如此，对 ^4He 在物理上的理解还是不完全清楚的。甚至对旋子的理解还没有一致的看法。朗道唯象理论中的 V_c 判据只是一个必要条件， ^4He 的元激发谱中除了声子外还可能有超流的湍流等等。对氦的研究，不论是 ^3He 还是 ^4He ，仍将继续是凝聚态物理中一个重要的组成部分。要知道，在整个宇宙中氦几乎占有三分之一的比例，它在天体的演化中扮演着重要的角色，是我们从微观的层次上理解液体的不可多得的“样本”。

从微观层次上理解朗道唯象理论的努力中，还有另一个方向。它也是所有的努力中第一个成功的工作。这就是 1947 年博格留勃夫关于弱斥力的相互作用的玻色系统的模型。博格留勃夫的工作未能证实旋子的存在，但是证实了声子谱，并且显示了玻色-爱因斯坦凝聚在其中起了关键的作用。简单的说就是，在有相互作用的情况下，由于玻色-爱因斯坦凝

聚的存在影响了动量 $k = 0$ 态的态密度，以至在长波极限下它们呈现声子型；而这里的凝聚，即使不是全部也仍然是大量以至宏观上可观

数量的⁴He原子占有 $k = 0$ 的量子态。在以后五十年代的硬球玻色子系统模型中人们也得到了类似的结论。八十年代后，在改进对相互作用势的选择的情况下，已出现了全面证实激发谱的理论模型。

应该指出，朗道在对⁴He超流解释的总结中，拒绝了将此超流现象与玻色-爱因斯坦凝聚现象的任何类比。对此作任何评述，现在可能还为时太早。人们曾指出在朗道给出的“猜测”的理由中有这样或那样的错误，但有些并非所说的那么简单，另一些可能也不应苛求。事实上朗道本人当初也作了二次猜测。一次是 1941 年，另一次是 1947 年。前一个猜测中，声子部分和旋子部分是不连接在一起的，后一个猜测才是正确的。这说明，一个成功的唯象理论是不容易的。成功的唯象理论中的假设或猜测部分，正是未来的微观理论所要澄清的，人们不能期望猜测的理由中没有错误；事实上玻尔的原子模型中，甚至根本没有给出任何理由就假设存在一系列的定态，然而这丝毫不损害玻尔模型的成功。

应该说对超流现象的认识，极大地帮助了人们对超导现象的认识。翁纳斯是在实现了氦液化后三年，于 1911 年发现了汞(Hg)的超导现象的。为了与以后要说的高温超导区别，我们在这里有时称这最初发现的超导为低温超导。翁纳斯实现液化氦从而获得低温后，他进行的下一步研究工作就是金属的电阻率。由于汞很容易净化，因此他用汞做实验，结果发现汞的电阻率大约在温度为 4.2K 左右突然消失，这就是金属超导的发现(图 4.12)。他为什么首先想研究低温下的金属电阻？很可能他是期待有什么东西发生的。在当时，20 世纪初，人们的金属导电图像是电子象流体一样在正离子的晶格中流动。液体在冷却后会固化。因此翁纳斯也会跟其它当时的物理学家们一样，预计低温时电子流体要冻结，从而金属要变成绝缘体。这样，他期待电阻率应该变成无穷大，然而事实

却正好相反。这个“分析”如果属实的话，那末虽然事与愿违，但是低温时电子流体发生了相变，这一点却是正确的。

超导有许多实验现象，转变温度 T_c (对汞, $T_c=4.2\text{K}$) 以下电阻率为零是首要的现象。电阻率为零马上使人想到理想导体，然而 1933 年迈斯纳(W.Meissner)和奥格森弗(R.Ochsenfeld)实验发现超导体内磁场总为零，即超导体是个完全的抗磁体，这就是所谓迈斯纳效应。迈斯纳效应否定了超导体是理想导体的想法。对于一个理想导体来说，导体内电场强度为零。于是由法拉第电磁感应定律知，磁场的变化率应该为零，即磁场应为一个常数。这样一来，如果在 T_c 以上把这种导体置于磁场中，然后降温至 T_c 以下，那末它应该保持 T_c 以上就有的磁场不变。然而，实际上一旦到达 T_c 以下，进入超导态后，它的内部的磁场根本就消失了！

与许多物理现象的理解过程一样，对超导现象的理解也经历了一个从唯象理论到微观理论的过程。前面提到超导现象是一个相变的结果，要理解这一点，在唯象的层次上必须研究超导的热力学，这在物理上是头等重要的事。在这方面比热的实验结果是最有特色的(图 4.13)。由图可见超导比热曲线在 T_c 附近 $T < T_c$ 时的行为是一个指数函数，由它可以得知在 T_c 以下超导态的熵 $S(T)$ 小于正常态的熵；而在 T_c 处二者一样，这一点还可以从在磁场下超导体的热力学性质分析中得到佐证。这些事实说明：第一，超导态是比正常态更有序的态，因为它的熵比后者小；第二，在无外磁场时，正常态到超导态的转变是二阶相变，因为在 T_c 处熵连续而比热跃变；第三，在超导态电子能谱上有量级为 Δ 的能隙。此外，实验还证明在 T_c 相变前后晶格结构无变化，声子对热力学有关量的贡献也一样，这样上述第一点说明是系统中电子发生了某种有序化的变化。

基于有序的想法，1934 年高特(J.Gorter)和卡斯米厄(G.Casimir)提出了超导相的二流体唯象理论。在此理论中 T_c 以下超导态中的电子被分为二部分，一部分为超导电子，指凝聚的部分，另一部分为正常电子，指未凝聚的部分。二流体唯象理论成功地解释了超导体的热力学性质。

为解释超导的电磁现象，即零电阻和迈斯纳效应，弗·伦敦和海·伦敦(H.London)于 1935 年提出了超导体电磁学的唯象理论。在二流体唯象理论的基础上，伦敦兄弟提出超导电子的电流

密度 j_s 、电场强度 E 和磁感应强度 B 应遵循的方程组。它们与电磁

场的麦克斯韦方程组是一致的。但与正常电流的欧姆定律不同，它能解释零电阻现象。正常电流密度是与电场强度成正比的，在恒定的电场作用下所以能保持恒定的电流密度是因为有电阻。而按伦敦方程超导电流密度的变化率与电场强度成正比，于是只要电场强度不为零，超导电流就会越来越大，这表明超导电子只受电场作用而无阻力！无阻力当然就是零电阻了。

伦敦方程组中的另一个方程反映了一个超导电流的回路会有一个反向的磁通量，迈斯纳效应正是由此而产生的。简单地讲，超导电流作为总电流的一部分，按照安培定律应该有正向的磁通量，而上述反向磁通恰好抵消了这正向磁通以至有迈斯纳效应。更“严格”一些，首先考虑到我们讨论的是稳恒状态，一切量均与时间无关。按伦敦方程超导电流密度不随时间变化意味着电场强度为零，而根据欧姆定律正常电流必为零，于是超导电流密度就是总电流密度。这时我们可以利用麦克斯韦方

程组求解出磁感应强度 B ，它为 $B = B_0 e^{-\frac{x}{\lambda}}$ ，其中 λ 是与电子的电荷、质量和密度等相关的常数，量级为 10^{-6} 厘米，称为伦敦穿透深度。

当 $x \gg \lambda$ 时， $B \sim 0$ ，即对于大的超导体(厚度 $D \gg \lambda$)来说，超导体内的磁场都几乎为零，只是在表面大约 10^{-6} 厘米的薄层内有磁场存在，这就是迈斯纳效应。(图4.14)

伦敦的理论实质上还是二流体唯象理论，只不过加上了电磁场而已。超导是低温下出现的电子有序化现象，它必需靠量子力学才能正确描述。1950年金兹堡和朗道提出的量子唯象理论比伦敦理论前进了一步。我们在分析超导的比热性质时曾指出，在 T_c 处发生的正常态到超导态的转变是一种二阶相变。金兹堡-朗道理论正是推广了的朗道的二阶相变理论。他们把 T_c 附近有外磁场存在下的超导态自由能

密度用波函数 $\psi(r)$ 为序参量表达出来，取极值后得到序参量 $\psi(r)$ 所满足的方程，并导出了超导电流密度与磁场的关系。后者在波函数取常数的近似下与伦敦方程在形式上一样，只不过电子的质量 m 和电荷 e

分别换成了由序参量 $\psi(r)$ 所反映的电子的一种有序化的有效质量 m^* 和电荷 e^* 。当然，在物理上二者并不相同，例如伦敦方程中超导电流密度只是部分电流密度，而这里是在任意温度下的整个超导态的电流密度。

此外，由金兹堡-朗道方程类似可以证明在超导体内磁场为零(迈斯纳效应)，同时给出了超导体的表面层电流密度与磁场的关系。金兹堡-朗道理论还告诉我们，出现超导的根本原因在于超导态的序参量有所谓的“刚性”，它导致超导态的激发态与基态之间有能隙存在。这个理论的另一个重要预言是磁通的量子化。

考虑一个超导环(图 4.15)，在 $T > T_c$ ，即环在正常态时将其置于磁场中，然后使 T 降至 T_c ，环中的电子从正常态转为超导态。人们发现此时超导电流的走向所对应的磁场正好抵消外磁场以保证超导体内无磁场-迈斯纳效应。利用金兹堡-朗道理论给出的关系式及波函数的单值性计算穿过超导体内的回路所包围面积的磁通量，得到

$= n\phi_0$ 。其中 $\phi_0 = \frac{hc}{e}$ 。上式表明，磁通量 ϕ 是量子化的， ϕ_0 是相应的量子。

磁通量子化可以从实验上验证。对于低温超导实验证得 $e^* = 2e$ ，这是一个非常重要的结果。它说明在低温超导中，电子的有序化是两个电子的某种配对凝聚态。上述讨论对系统中任何两个电子都成立，它们都有相同的 ϕ_0 ，但是对系统本身的电子数无任何限制。也就是说对超导态来说，系统可以有多到宏观数量的电子对都具有同一“量子态”相角 ϕ_0 ——这就是一种宏观量子效应，在一定意义上类似一种玻色-爱因斯坦凝聚。

金兹堡和朗道的这一量子唯象理论竟能给出这么多的启示！但是实际上当时它并未起到这样大的作用。虽然金兹堡和朗道在 1950 年发表了他们的理论，但由于各种原因，当时这一工作并未广为人知；实验上，据说 $e^* = 2e$ 的迹象也存在，但迟至 1965 年才有人针对他们的理论而在实验上证得 $m^* = 2m_e$ 和 $e^* = 2e$ (其中 m_e 是电子的质量)，那已是 BCS 理论发表之后八年了！

金兹堡和朗道的理论对前面提到的穿透深度 $\lambda(T)$ 和相干长度 $\xi(T)$ 也给出了重要预言，对它们的物理含义作深入研究导致了第 II 类超导体的理论预言。这是阿布里科索夫(Abriko-sov)1957 年发表的重要工作。如果说金兹堡和朗道的量子唯象理论未能在人们寻求超导微观机理的最初成功的尝试中起到应有的启示作用，那末它确实首先预见第 II 类超导体的存在。第 II 类超导体是指合金超导体，通常将金属超导体称为第 II 类超导体。

类超导体，第 类超导体是有潜在应用价值的超导体，对它的预见及此后实验上的发现，意义是重大的。

在寻求超导的微观理论方面，许多人作过尝试，包括海森堡和玻恩等。然而直至人们认识到超导性是一种类似 ^4He 的宏观量子效应以后，探索才开始步入正确的方向。50 年代初金兹堡、费曼和夏夫劳斯 (Schafroth) 各自独立地证明了一个带电玻色气体在玻色-爱因斯坦凝聚后会显示迈斯纳效应。这些工作最后导致了 1954 年夏夫劳斯提出了电子配对的想法，但是他的配对想法是电子在实空间或说坐标空间中配对，是一种准束缚的局域态。有案可查，最早提出用电子对的玻色-爱因斯坦凝聚解释超导的是一位实验化学家奥格 (A. Ogg)，他是在 1946 年提出来的，他的电子对也是实空间中的对。

1950 年弗劳里希 (Frohlich) 在人们寻求超导微观机理的努力中迈出了重要的一步。他证明了电子和声子的耦合能使费米面附近的电子间产生一种有效的吸引力。简单地说来，当电子 (费米面附近的电子) 行经可以极化的晶格点阵，引起点阵形变，也就是与声子耦合了，声子追随电子运动 (如果电子的运动频率不是太高)，同时声子又影响随后的费米面附近的电子，致使电子之间产生了有效的吸引力。弗劳里希的工作没有考虑实际存在的电子间的斥力。1955 年巴丁和派因斯 (Pines) 提出了包含库仑斥力在内的完整哈密顿量，并证明了在一定条件下，吸引力可以超过排斥力。声子在低温超导中起作用的实验支持来自于 1950 年的同位素效应，在纯的同一金属的不同同位素的超导体中，测量超导转变温度 T_c ，发现 T_c 反比于同位素质量 M 的平方根。这表明电子-声子相互作用在低温超导中扮演着一定的角色。

在朗道的 ^4He 超流唯象解释中激发态的能隙起着非常重要的作用，它是“刚性”的体现。电子-声子作用引起的电子有效吸引力提供了全部问题的动力学基础，下一步就是如何在这个基础上给出能隙。库珀于 1956 年在探索超导微观机理的最后冲刺阶段，跨出了关键的一步。他考虑了这样一个问题：一个被填满了的费米海外加二个电子，电子间有吸引力，求电子系统的最小本征能量和相应的本征态。计算结果表明此时最小本征能量小于它们自由情况时最小本征能量，也即二个电子形成了某种束缚态。相对于这个束缚态来说，原先自由的情况正是一种激发态，或者说破坏这个电子对，产生了系统的元激发，二者的能量差正提供了能隙。

而这个束缚的本征态是一对动量 k 、自旋方向 s 和动量 $-k$ 、自旋方向 $-s$ 的电子。这样的一对电子，人们称之为库珀对。

这里我们看到对的形成是一个动力学的原因造成的。电子并没有“意识”到要配对成玻色子，然后发生凝聚。在这个意义上来说，超导的产生并不是玻色-爱因斯坦凝聚的结果。玻色-爱因斯坦凝聚是玻色子的一种运动学结果，在相空间(指 k 空间)上凝聚成零体积。虽然我们也说一种类似于玻色-爱因斯坦的凝聚，但只是形式上而已。

由于库珀所考虑的二体问题对 BCS 理论的重要性，我们在这里再作一些深入的讨论是完全必要的。考虑二个电子和全部已填满了的费米海。如果这两个电子没有相互作用，则它们相当于两个自由电子。设动量

量分别为 q 和 q ，考虑到费米海已经填满，其大小只可能均大于费米动量 p_F 。于是能量分别是为 $\frac{h^2 * q^2}{2m}$ 和 $\frac{h^2 * q'^2}{2m}$ ，而整个系统的总能量 $E > 2E_F$ 。

现在我们考虑前面弗劳里希所发现的费米面附近的电子因与声子相互作用而产生有效的吸引力，即相互作用势为 $-V$ ，求解相应的薛定谔方程就会发现这时系统的总能量为

$$E = 2E_F + \frac{2h\omega_D}{1 - e^{-\frac{2}{VN(0)}}} < 2E_F$$

这里 $N(0)$ 是有费米面上单位能量间隔中可能的态数。上式表明，二个电子形成了某种束缚态，本征能量小于自由时的能量。这并不出乎意料，既然它们相互吸引当然能量会降低。但是有一点很值得注意。即使 V 非常小，也就是只要有一点点的吸力，但只要 $N(0)$ 足够大，也会产生不太小的束缚能量，即总能量 E 可以明显地小于 $2E_F$ 。此外，从所得的解很

容易看到，在这个态上二个电子的动量和自旋分别为 q 、 s 和 $-q$ 、 $-s$ ，所以的确是配成对的，这就是库珀对。

在上述二体的问题中，既然二个外加电子会配对形成能量上有利的束缚态；那末费米面内费米面附近的电子也能被激发至费米面外，配对并形成能量上有利的束缚态，这一来，整个费米面就不稳定了，换句话说，应该考虑在费米面附近二边能量相当于最大声子能量，即所谓德拜能量间隔内的多电子系统。

1957年巴丁、库珀和许里弗在库珀的二体问题的基础上，考虑多电子系统的问题，终于提出了一个成功的、解释超导微观机理的理论，即著名的BCS理论。在这个理论中，系统的哈密顿量由两部分组成。一部分描写电子的自由运动，另一部分描写使电子配成库珀对的相互作用。由前面唯象理论提供的信息是“系统可以有到宏观上可观数量的电子对”，也就是说电子对数是不固定的，从这个意义上来说，超导态的电子系统是一个开放系统，它毕竟只是费米面附近一部分电子。由于电子对数不固定，或说电子数不定，很自然的“猜测”到超导基态应该是由无配对电子与有一个任意可能的库珀对的状态的叠加而成的态。

我们曾一再强调超导态是宏观量子态的平衡态。由于粒子数不定，因此应该用所谓吉布斯能 G 来描述。它可以由系统的平均能量和平均粒子数计算出来。 G 取极小对应平衡态，在温度 $T=0$ 时结果为：

$$G = G_N + \frac{2N(0)(\hbar\omega_D)^2}{1 - e^{\frac{2}{VN(0)}}} < G_N$$

其中 $N(0)$ 是粒子数的平均值， G_N 是正常态的吉布斯能。这个结果非常类似前面得到的库珀对的总能量 E ，但是由于多体的关系，这里描写束缚能的第二项正比于 $N(0)$ ，也就是说差值更大。更重要的是，从 G 我们还可以更进一步计算系统的元激发，即将一对库珀对打破，从而激发出准粒子。得到的元激发能谱如图 4.16 所示，可见这时超导的临界速度为

$$V_c = \frac{\Delta}{k_F} \neq 0$$

这就解释了超导的零电阻，这里我们只须把外电场理解为“容器壁”。

随着温度的升高，由于热激发库珀对也会被破坏。激发一个准粒子的能量大致为 Δ ，故一个对的束缚能大致为 2Δ ；由此估计使对破坏的临界温度 T_c 也就是超导态转向正常态的临界温度应该大致为 $k_B T_c \sim 2\Delta$ 。

理论计算结果为 $2\Delta = 3.5k_B T_c$ 。

超导微观理论(BCS理论)是凝聚态理论中最成功的范例之一。超导理论的成功不仅使人们第一次认识到多电子体系可以具有除了传统的费米液体(能带论)型的基态还存在其它类型的基态，而且与超流微观理论的发展相呼应，使人们认识到一个多体系统可以有一个整体位相的基态。在这样的基态上还可有拓扑型的元激发。它大大地开阔了人们对量子多

体系统的了解，也开创了一个具有重大科技应用的新领域。

BCS 理论在凝聚态物理理论发展史上是一个里程碑，它诱发了各种场论的处理方法。其中特别应该提一下的是高尔可夫(Gorkov)提出的反常格林函数计算。库珀对的形成标志着超导态的产生，它是超导相的特征，回想我们在前面提到的序参量，因此一定能有与库珀对有关的一个序参量来描绘超导相。人们发现这个序参量正对应了高尔可夫提出的反常格林函数。如此一来，不用猜测波函数，人们也可以计算序参量，从而求得超导转变温度。此外量子场论工作者对于超导理论研究的介入，推动了强耦合超导理论的形成。反过来超导理论规范不变性的研究推动了量子场论中希格斯机制理论的建立和对称性破缺理论的发展。

80 年代初实验上发现了重费米子超导电性。重费米子体系是指某些稀土或镧系元素化合物，由于很强的 f 电子关联以及 f 电子与传导电子之间的杂化，使重整化的 f 电子的有效质量可以比裸电子的质量大一二个数量级，因而称为重电子或重费米子体系。理论上有趣的问题之一是重费米子体系的超导电性。其中一些问题已得到解决，如肯定了重电子参与超导，超导能隙函数在费米面上是各向异性的。关于配对机理，倾向于是由自旋涨落产生的极化导致的，而且很可能是 p 波配对的等等。但是，仍有许多问题没有解决。例如实验发现有的重费米子体系具有超导基态，有的却是反铁磁有序基态，也有的是费米液体基态。究竟由什么因素决定，理论上尚未解决。迫切的问题是需要建立一个自洽的和具有普遍性的重费米子相干点阵理论，从而解决各种不同基态以及超导基态的机理这些尚有争议的问题。

4.6 强关联问题、高温超导

自从发现超导现象后，人们就一直在努力研究提高它的转变温度 T_c ，原因当然很简单，如果能在常温甚至更高温度下实现超导，许多电磁能的利用中的能耗可以大大地节省，其经济效益是十分可观的。然而，在长达半个多世纪的努力中，从 1911 年汞的 $T_c=4.2\text{K}$ 到 1973 年才找到铌三锗 (Nb_3Ge)，达到 $T_c=23.2\text{K}$ ，进展十分缓慢。有意思的是在人们的探索从金属转向化合物和合金时，第一步是从氧化物 (NbO) 开始的，它的 $T_c=10\text{K}$ 左右，以后却一直是在其它化合物或合金中寻找，直到 40 年以后的 70 年代才又回到氧化物中来。终于在 1986 年，贝诺斯 (Bednorz) 和缪勒 (Muller) 在铜的氧化物 La-Ba-Cu-O 化合物中发现了高达 $T_c=35\text{K}$ 的超导转变，揭开了高温超导研究的新篇章。1987 年人们又发现了在 Y-Ba-Cu-O 化合物中可达 $T_c=90\text{K}$ 左右的超导氧化物，此后高温超导材料的发现如雨后春笋层出不穷， T_c 已达 135K 以上。贝诺斯和缪勒由于这一开创性的发现而获得了 1987 年的诺贝尔物理学奖。

今天高温超导材料已多达四十多种，但都是氧化物，并且极大部分都是铜氧化物。最典型的有上述二种化合物衍生出来的 2-1-4 化合物 $\text{La}_{2-x}\text{M}_x\text{CuO}_{4-y}$ (其中 M 可以是 Ba、Br 或 Ca 等) 和 1-2-3 化合物 $\text{RBa}_2\text{Cu}_3\text{O}_7$ (其中 R 可以是 Y、Eu、Gd 等)。2-1-4 化合物如 La_2CuO_4 ，有如图 4.17 所示的结构，它没有超导态，但当掺杂后变为 $\text{La}_{2-x}\text{M}_x\text{CuO}_{4-y}$ 有图 4.18 的相图。1-2-3 化合物如 $\text{YBa}_2\text{Cu}_3\text{O}_7$ 有如图 4.19 所示的结构，它的相图如图 4.20 所示。由相图可知，在无掺杂时，即 $x=0$ 或者 $y=0$ 时，它们是一个磁性的绝缘体 ($T=0$)，掺杂至一定比例时才出现超导。这就把已有半个多世纪历史的“莫特 (Mott) 现象”又重新提了出来。

晶体的能带理论中，能带的形成关键在于相邻原子的电子波函数有重叠。由于重叠，电子能够从一个原子处逐步跳至其它原子处而形成能带。一般说来，重叠越多，能带越宽，能隙越小。可以想象，一个满带且有能隙的绝缘体，在压力作用下，原子间距离缩小，电子波函数重叠加大，最终就可能变为金属。这是在能带论即独立单电子图像中绝缘体—金属相变的观点，已得到实验证实。

1937 年玻义尔 (DeBoer) 和厄威 (Rerwey) 指出，能带理论无法解释看来应该是导体而实际上是绝缘体的 NiO 等磁性氧化物。对此，潘爱尔斯

(R.E.Peierls)指出,这是Ni的二个相同原子的d电子库仑相斥能大于因近邻原子关联而形成能带所获得的动能,以至实际上每个Ni仍维持 d^8 态的结果,虽然能带未满但却是绝缘体。此后,莫特(N.F.Mott)进一步阐述了这一观点,并考虑了因电子关联而引起的金属和绝缘体转变的普遍现象,即所谓“莫特”现象。

如果说在潘爱尔斯和莫特对“莫特现象”的解释中强调了电子关联作用的能量因素,那么以后在1961年由哈伯(J.Hubbard)提出的研究金属—绝缘体的模型已注意到所涉及到的局域化问题了。对于磁性氧化物或过渡金属化合物(如NiO),其中主要是d态或f态电子在起作用。而d态f态电子的波函数所反映的壳层半径是很小的,因此邻近电子波函数间的重叠很小,显示了电子的局域性,由此而形成了所谓窄带问题。

“莫特现象”不仅引出了哈伯模型的窄带问题,还导致了安德森(P.W.Anderson)关于局域磁矩的解释。“莫特现象”是一个电子强关联问题。而在历史上,晶体中的电子强关联问题实际上最初是出现在磁性问题中的。如铁磁的海森堡模型即包含了电子间的库仑关联又包含了泡利的自旋关联。此后,安德森的“超交换模型”,稀土金属的RKKY模型等等都是电子强关联问题。1959年,安德森借用莫特关于电子关联的想法提出了一个模型,解释了绝缘磁性材料大都是反铁磁性的问题。模型本身中并无局域磁矩出现,然而至少在哈特里—福克(Hartree—Fock)近似下,此模型显示了d电子的局域磁矩。从相变角度看,这又是一个动力学对称破缺的例子。

BCS理论中库珀对的形成是十分关键的,它是动力学对称破缺的一个典型实例。在那里,声子的传播引起了电子间有效的吸引力,最终导致库珀对的形成。在高温超导问题中,大量的实验事实表明声子可能并不起什么作用,似乎应该是别的原因导致电子系统的对称破缺或超导相。从前面二张相图可见,超导态只是其中的一部分。在目前人们还无法一下抓住这一部分物理的现实情况下,先从外围有关的态,例如磁性绝缘态、正常的金属态等着手,或许是一种明智的做法。

哈伯模型的重要,并不是它“能解释超导”——对于这一点,至今未能得到公认——只是因为它是描绘强关联的最简单的模型,而高温超导问题是强关联系统中的相变问题。因此,人们从哈伯模型入手是很自然的。但要真正解释高温超导,问题并不那么简单。

自从高温超导发现以来，人们意识到这是一个既具有巨大应用前景又具有重要理论意义的领域，从而立即在全球爆发了一场研究高温超导的热潮。这证实了金兹堡的判断：就其对社会冲击而言，高温(最终将达到室温)超导体可能是除了受控核聚变之外，物理科学中最重要的问题。经过近几年的研究，现在已经认识到高温超导体的超导相与常规超导体有许多重要的类似之处。例如，高温超导体同样存在电荷为 $2e$ 的库珀对；它属于强 d 类超导体；相干长度很小，比常规超导体要小 $1 - 3$ 数量级，因此在相干体积内只有很少的几个库珀对，从而导致明显的涨落效应等等。但是超导相中超导反铁磁涨落的竞争，包括配对的对称性，超导配对的机理，特别是层间耦合的作用等这样一些重要机理性问题，至今尚未解决。

氧化物高温超导材料另一个非同寻常的特点是其正常态的性质非常“反常”。例如，沿铜氧平面的电阻率与温度成线性关系，霍尔系数与温度的三次幂成比例，以及根据其拉曼谱，NMR 弛豫和光吸收等实验所说明的似乎存在大量的极低能扩散型激发等等。特别是，高温超导体的正常相似乎不遵从常规的费米液体行为，表现在其准粒子不能明确定义；提出了电荷与自旋自由度是否可能分离的新型元激发的概念；虽然关于费米面所包围的体积不变的定理(路庭格定理)仍然成立，但准粒子分布在费米面处的变化是连续的而非跃变的。此外似乎还存在一个由载流子浓度等诱导的金属-绝缘体相变。对此，上面已经给出了一些粗略的介绍。

由于高温超导机理极为复杂困难，目前仍未解决。某些人希望首先从解决其正常态的行为入手。近年来这方面已发展了不少理论，一类是在常规的费米液体理论上着重研究载流子和反铁磁涨落的竞争，如派恩斯等的“反铁磁费米液体理论”；另一类是瓦玛等人提出的“边缘费米液体理论”，它是一种唯象理论，但能较全面地总结大量的正常态数据。以拟合正常态的反常输运性质，理论假定了极化部分的虚部的形式，突出的特点是不存在内禀的能标。按照这一假设所计算出的波函数的重整化因子正好在费米面上趋于零；再一类是安德森的“路庭格液体理论”，试图解释为什么会有类似的费米面的性质，所引入的元激发也与常规费米液体理论十分不同。总之，目前关于高温超导体正常态的问题仍未解决，不同的理论分歧很大，它对当代的凝聚态理论提出了巨大

的挑战，在世界范围内形成了一个由各种实验物理工作者、凝聚态理论工作者和量子场论工作者共同努力的前沿热点。现在人们希望能通过进一步的实验，获得更多的有关信息，并最终为建立高温超导的整个理论铺平道路。

不论将来结果如何，有一点可以肯定，那就是对电子强关联问题的研究乃至高温超导问题的最终解决，将极大地丰富我们对动力学对称破缺相变的认识。这或许会给粒子物理的希格斯(Higgs)场研究提供一些启示。理论物理的这二门学科，将会有再一次的思想交流。

4.7 强关联问题、量子霍尔效应

霍尔(Hall)效应是 1879 年霍尔发现的。其内容简单地说来如下：一块

导体二端接通电源如图4.21，于是导体中有电场 E 。如果在与 E 垂直的方向有磁场 B ，则在导体二侧P、Q点之间有所谓霍尔电势 V_H ，这种横向效应即为霍尔效应。由普通物理知识不难算出霍尔电导率 $\sigma_H = \frac{ne^2}{B}$ ，其中 n 为电荷密度。然而，一个世纪以后的1980年，冯·克立兹(Von Klitz) 、道达(Dorda)和配伯(Pepper)在低温(液氦温区)和强磁场下(10特斯拉)对MOS器件做霍尔实验，发现 $\sigma_H = \frac{Me^2}{h}$ ，其中 M 取1, 2, ...等一系列整数值， h 是普朗克常数(图4.22)。它表明 σ_H 是量子化的，以 e^2/h 为量子。由于 M 是整数，这被称为整数量子霍尔效应(IQHE)。由于这个关系非常之简明，所以人们借助于量子霍尔效应能对 e^2/h 作非常精确的测量，目前精度已达 10^{-8} 以上，这是有十分重要意义的。一方面由于 e^2/h 值的高度精确性，它可以提供另一途径得到量子电动力学(QED)中的精细结构常数 $\alpha = \frac{e^2}{hc}$ ；而另一方面，人们可以用 h/e^2 作为电阻测量的标准。

这个关系式本身是不难“证明”的。MOS 器件的结构表明，它是一个 2 维的电子系统。求解相应的薛定谔方程可知，本征能谱是一系列朗道能级，每一能级是简并的，简并数与面积成正比，单位面积有 eB/hc 个态。因此如果有 M 个朗道能级被填充，则电子的密

度为 $n = \frac{MeB}{hc}$ 。把它代回前面给出的霍尔电导率的表达式中，即得整数量子霍尔效应的结果，且易见整数 M 应该是朗道能级的填充数。

在继续讨论之前，我们在这里先来介绍一下无序引起的金属—绝缘体相变问题。朗道的费米液体理论，实质上说的是库仑相互作用仍保持动量空间中费米面的存在，或说金属—金属无相变，莫特现象反映了在窄能带的晶体场中，库仑相互作用可能导致金属—绝缘体的相变。那末在非晶体场中又怎样呢？1958 年安得森指出在一个强的随机场中电子的波函数会局域化，即有金属—绝缘体的相变。这一工作在当时并未引起

注意，因为结论似乎是在意料之中的。

从紧束缚近似提供的能带图像不难理解，在能带中心附近的态应该保持扩展态，至少对不是非常强的无序系统来说应该成立；在能带边缘的那些态是局域态。1968年莫特对这种从局域态到扩展态的过渡，提出了迁移率边缘的临界能量 E_c 的概念。莫特在1973年又进一步提出 E_c 处有一个电导率的突变，即从局域态的 $\sigma_{\min}=0$ 到扩展态 $\sigma_{\min} \neq 0$ 的突变(图4.23)。1974年沙勒斯(Thouless)等提出了局域化问题的标度描述。1979年阿布拉姆(Abrahams)等在沙勒斯等工作基础上提出了局域化标度理论，结论是电导率应该是连续变化而不是突变的。以后就有一系列用场论的重整化群方法研究无序引起的金属-绝缘体相变。结论是 $D=1, 2$ 维在无序的作用下应该是绝缘体，无相变。而在 $D=3$ 时有金属-绝缘体相变。但在有磁场存在时，上述结论并不成立。负磁阻现象($D=2$)表明，此时无序系统可能有扩展态。事实上，量子霍尔效应正是在研究电子的局域化问题时发现。量子霍尔效应表明，在有磁场时，2维无序系统应有扩展态，否则 $\sigma_{\min}=0$ 了。

再回到量子霍尔效应。怎么来解释实验中 σ_H 出现的平台呢？(见图4.22)。平台的存在说明有电子的态仅对电子密度 n 有贡献，但对 σ_H 无贡献。这就表明有局域态，为解释这一点必须考虑杂质的存在。杂质使朗道能级变宽而成了能带，并且互相重叠起来。理论计算表明大部分电子状态局域化了，即被杂质所束缚，只有那些处在能带中心的状态仍然是扩展态。改变电子浓度就改变了费米能级。当费米能级处在局域态区时霍尔电导取量子数值，而当费米能级跨过一个扩展态时，霍尔电导率就改变一个量子数(图4.24)。

然而，考虑了局域态后，又为什么霍尔电导仍是量子化了的呢？对此，普拉格(Prange)认为局域态的存在并不影响霍尔电流。当电子费米能级位于局域态时，扩展态的电子会补偿应由局域态贡献的霍尔电流。后来，劳甫林(Laughlin)又提出了规范不变的观点。所谓规范不变实质即电荷守恒。从这一点来说劳甫林的这一观点是普拉格观点的另一种更实质化、一般化的说法。然而作为物理的机理来说，哈伯林(Halperin)的“边界流”观点是十分重要的。边界流是一种拓扑元激发流。正是边界流的存在，才得以使量子化在有局域态存在时仍成立。不过，对此也有人持反对意见。应该说，即使在今天，整数量子霍尔效应(IQHE)的解

释还是不完全清楚的。

整数量子霍尔效应的发现是在 MOS 器件上作出的，这是一个 2 维有边界的现象。所谓边界，就是系统的拓扑结构。在有一定的拓扑结构下考虑无序(杂质)问题，这本身就是一个新问题。在这一方面，一些形式上拓扑问题的研究固然重要，但可能还只是问题的第一步，真正的物理还在于考虑无序后的局域态、扩展态和边界流等问题。

整数霍尔效应发现才 2 年，紧接着崔琦(Tau, 美籍华裔)、斯多麦(Stormer)和谷沙特(Gossard)又发现了分数量子霍尔效应(FQHE)，这就提出了更深层的问题。他们在 GaAs—AlGaAs 异质结上观察到，上述 σ_{xy} 的表示式中 M 为分数，而不仅是整数。实验是在高度净化，温度更低($\sim 1K$)，磁场更强(约 15 特斯拉)的条件下进行的。此后的大量实验却发现 $M=p/q$ ， p 是奇数或偶数，而对最低朗道能级， q 总是奇数。1987 年后又发现偶分母分数态 $M=5/2$ 。

如前所述整数量子霍尔效应(IQHE)可以用单粒子近似很好地描述，其物理图像已基本清楚。不过仍存在一些值得深入研究的问题。分数量子霍尔效应(FQHE)必须是高迁移率的样品在更低的温度下才能观察到。分数量子霍尔效应也是一个强磁场中的电子强关联系统，因为要解释分数量子霍尔效应必须考虑电子相互作用。这从理论上提出了一类全新的问题。

劳夫林提出奇数分母的量子霍尔态是一种不可压缩的量子流体，并构造出它的十分成功的试探基态波函数，可以很好地解释朗道能级填充因子为 $1/3$ 和 $1/5$ 的分数量子霍尔的基态和元激发的性质，在劳夫林波函数描述的不可压缩量子流体基态，可以产生单粒子激发(具有分数电荷的准电子与准空穴)，也可产生集体激发(例如磁旋子)，此外，理论上还可能产生准粒子(相反电荷的劳夫林准粒子束缚对)等。其中关于存在分数电荷的准粒子特别有兴趣。如果分数电荷的准粒子能够得到证实，那么进一步又涉及关于这类准粒子服从什么统计的问题。我们知道，自然界中迄今发现的粒子，无论是基本粒子还是复合粒子，从统计性质上只有两类，或为费米子，或为玻色子。然而上面提到的分数电荷的准粒子既非费米子，也非玻色子，于是理论上提出任意子(Anyon)的概念。目前理论上倾向与认为任意子是二维系统这一特定条件下带电粒子与磁通的复合物，分数量子霍尔态是最有希望实现分数统计的系统。

对于具有奇数分母但分子大于 1 的分数量子霍尔效应态，霍尔丹—赫尔普林提出了涡旋元激发的观念以及他们的理论描述，而简恩却基于一种“复合费米子”的观念(一个电子与偶数个磁通量子结合成一个复合费米子)提出了另一种描述。最新的实验迹象相当有利于后者，所以在分数量子霍尔效应态上是否存在“复合费米子”这种“粒子”又成为一个全新的研究方向。此外，最近又开展了填充因子具有偶数分母状态的研究。这种状态被预期为凝聚态物理中的新的物质形态。所以，强磁场下低维多电子系统又有可能成为凝聚态理论的一个新的突破口。

值得注意的是有人用在恒定磁场下相互作用电子系统，加上陈-塞门(Chern - Simmon)项的场论模型，在经典近似下获得了劳甫林态和磁旋子态，后者在分数霍尔效应的光学性质实验研究中已被证实存在。大家都知道，在拓扑中有一个欧拉示性数。它是用来描述一类拓扑上等价的几何形体特征的数，例如封闭凸多面体的欧拉示性数为 2。把这种拓扑(更确切地说代数拓扑)处理的对象对应到微分几何中的流形，便产生了实流形的欧拉示性数，而陈(Chern)示性数则是将实的流形的欧拉示性数推广到复流形上。而陈—塞门项对应于有边界的开的复流形时所增加的由边界的拓扑结构决定的“示性数”，它反映了边界效应。应该说，至今对量子霍尔效应的理解虽已有不少进展，然而仍还有不少根本问题不清楚。从物理上说，无论整数还是分数效应应该有一个统一的方法来处理。目前的进展表明从拓扑上考虑边界作用作为第一步是有希望的统一方法。然而，如何从物理上自然地引出所谓相应的陈—塞门项，可能是首先要考虑的问题。此外，还得考虑杂质的作用。

量子霍尔效应是一个很有意思的物理现象，是近 10 多年来凝聚态理论取得重大进展的低维系统中极为重要的领域。它提供了凝聚态物理中继氦之后又一个有着丰富物理内容的量子液体，这里有通常意义上的元激发，还有拓扑元激发。此外，它还紧密地联系着规范场中的角、瞬子等问题。对它的深入研究会给人们许多新的物理启示。

4.8 小量子系统

这里再简单介绍凝聚态物理的一个新兴分支—小量子系统或统称介观物理—所涉及的一些基本理论问题。介观物理是 80 年代发展起来的。由于固体器件微型化和微加工技术的高度发展，高分辨率观察手段的空前提高，已经能够用人工方法制备尺寸小到几百纳米甚至更小的结构，称为介观体系。在实验上对这些小量子系统的细致研究已经实现，由此发现和提出了许多新的物理问题。除量子霍尔效应外，新的物理问题之一就是电子的量子干涉效应。小量子系统尺度的上限就是由能保持量子相干效应的最大尺度决定的，下限可以一直到原子尺度。

实际的问题都是在温度 $T \rightarrow 0$ 的情况下发生的。这时系统内有元激发如声子，电子的相干性因与元激发的非弹性散射而被破坏，它决定了相干效应的最大尺度 L_0 。此外，系统中有杂质，特别是静态杂质。电子与杂质的散射决定了一个所谓弹道平均自由程 l 。它是电子扩散运动尺度的下限。或者说在小于 l 的尺度上电子作弹道运动。在温度足够低时 $L_0 \gg l$ ，于是我们就有了一个研究无序系统电子量子相干的尺度 L ，它满足 $l \ll L \ll L_0$ ，这就是介观尺度。例如 $T \sim 1\text{K}$ 时，典型的尺度大约为 $10^{-4} \sim 10^{-5}$ 厘米。介观物理就是研究这一尺度上的物理现象的新学科。由于这时电子波函数的位相相干长度会大于体系的线度，因而电子的行为表现出很强的量子相干性，导致许多宏观系统中所没有的新的物理现象，如阿哈洛诺夫—玻姆效应，普适电导涨落，量子点接触的电导量子化，弹道输运中的各种非定域效应，量子点共振隧穿中的库仑阻塞，以及介观正常金属环中的持续电流等。人们对介观体系的研究，除了它本身物理的兴趣外，重要的还在于它的潜在的应用前景。毫无疑问，介观物理理论的发展，将为新一代的量子器件提供概念与理论基础。这里我们只想涉及几个带基本性质的理论问题。

首先我们来介绍一下所谓的阿哈洛诺夫—玻姆(Aharonov—Bohm)效应或 AB 效应。它是与一个电子在有外磁场的环境中运动，研究外磁场的存在如何影响电子的定态波函数的问题相联系的。由于薛定谔方程有电磁规范不变性，通过适当的规范变换很容易证明，这时的电子定态波函数是在自由运动定态波函数的基础上添加了一个相位的改变，它由外磁场矢量势 A 沿一条路径的积分决定。通常人们一直认为只有磁场强

度才是可以观测的物理量， A 似乎是一个无法观察到的量。但是在下述的AB效应中，理论上预言了应该能够观察到 A 的存在。

设在平面上一个小区域(图 4.25 中阴影区)有磁场，考虑电子从 P 点沿不同路径绕过此区域在 Q 点干涉，不难证明干涉强度

$$I = 2|\psi'(Q)|^2 \left(1 + \cos 2\pi \frac{\Phi}{\Phi_0} \right)$$

其中 $\Phi = \frac{hc}{e}$ ， Φ 为穿过由 l_1 和 l_2 两条路径所包围的面积磁通量。这就是阿哈洛诺夫(Aharonov)和玻姆(Bohm)1959年提出的AB效应的基本思想。它表明 I 以 Φ_0 为周期振荡，而沿 l_1 或 l_2 事实上都没有磁场，这就证明了 \bar{A} 的存在。AB效应当然反映的是电子的量子相干性，实验在1960年对于理论讨论的真空情况基本证实了这个效应，80年代改进了的实验完全证实了此效应。然而，在金属中电子会有散射，是否仍然能观察到AB效应呢？这就对介观物理提出了问题。1983年以来人们相继在多次介观尺度的实验中观察到了AB效应。这说明杂质引起的弹性散射并不破坏电子的干涉效应。电子的相位既然未受破坏，人们可以预计在一个介观尺度范围内的环，可能会有持续电流，当然考虑到电子相互作用，问题就不那么简单了。实验上证实确有持续电流，但理论的计算在量级上还差不少，问题还有待深入。

有趣的是人们在类似的实验中不仅看到了以 Φ_0 为周期的振荡，还看到了以 $\Phi_0/2$ 为周期的振荡。它的理论上预言早在1981年就由阿苏勒(Altshuler)，阿洛诺夫(Aronov)和斯皮伐克(Spivak)在1981年提出了(又称AAS效应)，是70年代末由局域化标度理论引起的弱无序系统电导的深入研究的成果之一。

AAS效应最初出现在解释负磁阻现象，在那里AAS提出了相背散射的干涉效应。简单地说来如下：如图4.26所示，在 O 点处沿实线的回路和沿虚线的回路，即使在有杂质时电子的波函数也会相干(虚线回路是沿反向绕实线的回路，只是为了明显起见我们才另画回路了)，因为这只是以相反的方向沿同一路径走的。在有磁场时，由产生AB效应的同样道理可知，二个波函数各自多一个因磁场而引起的相因子，不难证明这二

部分的相因子差为 $2\pi \frac{\phi}{\phi_0/2}$ ，这就带来了磁电阻或磁电导的以 $\phi_0/2$ 为

周期的振荡现象。

我们再来介绍一个与介观系统有关的统计物理问题。介观系统介于微观与宏观之间，它与宏观体系的显著区别是不满足热力学极限，因而破坏了宏观体系通常条件下所遵从的自平均行为，表现在介观系统具有很强的涨落。典型的例子是普适电导涨落。这个现象是 1984 年在实验上发现的(图 4.27)。低温下金属导线的电导在磁场下显示了出乎意料的涨落。涨落呈现与时间无关的随机性，在给定温度下对一定的样品这种涨落可以重复，对不同的样品涨落型式也不同。但是涨落的幅度与样品尺寸无关甚至与样品中杂质引起的无序程序也无关(只要还保持是金属)都为 e^2/h 量级，这就是所谓普适电导涨落。此外，实验表明随着温度升高涨落逐渐平缓消失。由于样品的尺寸很小，这一事实提示我们某种与温度有关的尺度随温度升高而逐渐变得小于样品的尺度，导致涨落减小；这种涨落是一种量子干涉的效应。我们可以如下来理解这种它的由来。

在本书的第三篇我们曾经提到过费曼的路径积分量子化。按照这种观点，电子在导线中将以各种可能的轨道运动；由于杂质的存在，电子将被杂质随机地散射而穿过导线。对一个给定的样品杂质的分布是一定的。计算表明电子的每二条可能的轨道所对应的作用量之差将给电导带来一个相因子。这样，在有外磁场的情况下，二条分别能形成 AB 效应或 AAS 效应的轨道，它们对电导贡献的幅度都是 e^2/h 。这种随机轨道所产生的量子干涉就是普适电导涨落的由来。

AB 效应实质上是一种推广了的贝瑞(Berry)几何相因子。贝瑞在 1984 年发现：一个量子系统在经“绝热”演化过程回到它的初始物理态时，波函数除了获取一个通常的动力学因子外，还获取了一个由演化路径决定的几何因子。此结果以后又被阿哈洛诺夫和阿那旦(Anandan)推广到非绝热和非封闭路径的情况。前面霍尔效应理论研究中的拓扑问题都与贝瑞相因子有关，从这个意义上来看，介观尺度的霍尔效应研究是很有意义此外，已有实验证实介观电导的时间涨落，对此人们自然会想到应考虑非静态的杂质，沿此方向的研究可能有助于更深入地理解 $1/f$ 噪声及介观系统的非线性现象。在微观层次上的这些研究意义将是深远的。

理论上还发现，在扩散区电导几率分布函数虽然主体部分接近高斯

分布，但带有非高斯型的、对数正态分布的“尾巴”，而且随无序程度的增加，“尾巴”所占的比重越来越大，当趋近强无序区时，电导几率分布整体上变成对数正态分布。这时，仅仅用平均值和方差不再能描写介观系统统计涨落的行为，而必须代之以研究整个分布函数的特征。

介观体系具有很强的“个性”，这是介观系统区别于宏观系统的另一特征。例如对于介观金属样品电导随磁场的变化的涨落图样，特定的样品有特定的图样，而且同一个样品在相同条件下涨落图样可以重复出现，因而被称为样品的“指纹”。

再有，根据统计力学，在热力学极限下不同的平衡系综是等价的，可以任意选用正则系综或巨正则系综去计算宏观系统的平衡性质。但对介观系统，用不同的系综得到的结果可能不同。突出的例子是对介观正常金属环中持续电流在多环实验条件下的理论分析：如果用巨正则系综计算，持续电流接近于零，只有用正则系综才能得到非零的结果。

另一个问题涉及如何正确地描述介观体系与宏观体系的联结。通常在理论处理时作如下理想化：假定介观样品与宏观热库和粒子库由理想导线相连，并假定载流电子位相的无规化都在库中实现，这种理想化显然与实际情况有差别。如何恰当地描述宏观与介观系统的联结，需要进行更深入的理论探讨。

近年来，有人提出把与导线通过隧穿弱耦合的量子点当做一个“小型人造实验室”，来研究强电子关联效应。其中一个问题是研究近藤态与混价态之间的连续过渡(CROSSOVER)。由于可以通过调节势垒高度来控制量子点与导线之间的耦合强度，还可以通过调节栅压控制量子点内局域态能级和外部相连电极的费米能级之差，从而可以相当方便地改变有关的参数。虽然至今实验上尚未观察到这类过渡(很可能是该体系相应的近藤温度太低)，但这方面的研究显然是很有意义的。

小量子体系研究的另一个重要方向就是人造“原子”，它是介观尺度以下一直到几个原子的尺度范围。今天，在实验室中人们已能做出由几个乃至几百个原子形成的微观和细小聚集体，它们被称为团簇。典型的例子如以 C_{60} 为代表的富勒伦斯(Fullerene)系列。

1985年人们为解释星际光谱中一个未知的新宽峰，用激光气化石墨，结果导致由60个碳原子组成的稳定的 C_{60} 团簇的发现。此后又发现一系列稳定的 C_{20} 、 C_{24} 、 C_{28} 、 C_{32} 、 C_{36} 、 C_{50} 以及 C_{70} 等等。由于 C_{60} 的结

构是由 20 个六边形和 12 个五边形所构成；或者说是一个截平了的 20 面体。这很类似于建筑师富勒(B. Fuller)设计的多面体圆顶结构，故而得名为富氏团系列(Fullerene)(图 4.28)。

碳是有机物构成的基本成分之一，以往人们所知道的由碳原子组成的稳定形态有石墨和金刚石，但都不是团簇， C_{60} 不仅是团簇，并且有高度的对称性。正如苯环在有机化学中占有相当大的重要地位，据说有人认为 C_{60} 的研究可能会开辟有机化学，乃至生物学的新领域。

从物理上来看，人们已实现由团簇再组成晶体，如 C_{60} 组成的面心立方晶体。但是和通常固体物理学的晶体不同，这里并无平移对称性。因为在不同的晶格点处的 C_{60} 可以有不同的定向。实验上已发现这样的晶体经掺杂后有超导现象，如 K_3C_{60} 的超导转变温度在 19K 左右等等。

团簇的实现是人造“原子”的标志。这里出现的一个首要问题就是稳定性问题。人造“原子”是人工实现的原子的有序态。稳定性问题反映了这种有序化有临界状态。实验上团簇的实现为我们研究这种微观的临界态提供了实例，并将极大地推动这方面的理论研究。这对我们认识宏观世界的复杂性会是十分重要的。

当今的凝聚态物理的许多方面已开始涉及了一个共同课题：结构和形态。如果说过去的研究几乎无一例外地属于微观结构和宏观物性的研究，那末当今的研究已出现相当一部分的关于微观结构和宏观几何形态的研究，这无疑地是一个进入复杂性研究的重要环节。或许还是联系凝聚态物理和生命现象的一条通道。在这一方面，团簇的实现和研究也许具有带基础意义的重要性。

介观物理中还涉及许多其它基本性的理论问题，如微腔中真空涨落引起的量子电动力学效应，量子混沌在介观系统中的实现，以及量子力学的测量问题等。总之，人们对于这样一个重要特征尺度下的新系统的认识仅有十年稍多一点的历史，仅就上面所列的一些问题就可看出这个领域将可能成为下一世纪凝聚态物理的重要课题之一。

早在 1959 年费曼在一次探索新材料的科学和工程的会议上谈到：“一旦我们能掌握小尺度东西的安置，我们就能获得具有巨大范围各种可能性质的材料，对此我几乎毫不怀疑。”

这是一位理论物理学家的预言，数十年后的今天，人们果然踏上了这一新时期的征途。这个新时期无论对物理学本身还是对人类的未来都

会是意味深长的。因为从现代宇宙学的观点来看，我们所生活的世界是从一次大爆炸以后，由一些基本粒子，如电子、正电子、光子、中微子及少数重子逐步在膨胀而带来的冷却过程中凝聚产生的。最初是原子核形成，然后是原子、分子……。这是一个自然创世的过程。而今天人类在经过将近一个世纪对微观世界的探索后，开始了自己的“创世”——人类自己构成组成物质世界的基本砖块——人造“原子”。

4.9 结束语

凝聚态理论是理论物理的一个重要领域，它涉及的范围极广，问题错综复杂。凝聚态理论的最大特点之一是几乎在每一类型的系统中，透过复杂现象的种种侧面，常能发现它是由一种或几种相当基本的因素在起着主导作用，而它(们)所服从的规律又常常是与当前理论物理发展前沿密切联系的。尽管凝聚态物理与高能物理、核物理、宇宙论等其它领域的具体研究对象不同，但它们的基本概念往往是相互沟通的，这实际上反映了物质结构概念的统一性。

如前所述，自本世纪 30 年代以来，凝聚态物理取得了巨大的进展。20 年代末，著名理论物理学家狄拉克曾经断言，在非相对论量子力学建立后，固体物理已经没有什么需要物理学家研究了，“剩下的是化学了”（大意）。幸好，物理学家并没有管狄拉克的断言，他们坚持在这个领域里不断探索，取得了一个又一个的辉煌成就。事实上，物质结构和运动是十分复杂的，它可以分成若干不同的层次。物质结构的每一层次除了受到一些最基本最普遍的规律控制外，还有它自身所特有的规律，物理学的不同领域，如高能物理、核物理、凝聚态物理、宇宙论等等，也可以说就是以物质结构相应的层次的特有的规律为对象的。

由于凝聚态物理与材料科学及新技术有着紧密的联系，因此，凝聚态理论的几乎每一进展，都直接或间接地为新技术的发展提供了概念和理论基础，充分体现了理论的巨大推动作用。40 年代末巴丁等应用固体能带论研究当时新发现的半导体材料，提出了晶体管理论，导致电子学乃至电子工业革命性的变革，为开创信息时代和太空飞行创造了条件。又如激光器的发明，50 年代末在对受激辐射的量子力学本性深刻认识的基础上，提出了激光的概念并很快在实验室中造出了激光器，以后被广泛应用于高技术，乃至医学等各个领域。超导的研究虽然早期偏重基础方面，60 年代发现约瑟夫森效应以及 A_1 类超导体的研究进展，为超导在强磁场、精密测量等方面的应用提供了基础。一个有趣的历史反复是在 70 年代末期，由于密度泛函理论的成就和大型计算机的发展，国际上又出现这样一种舆论：凝聚态理论只需进一步的精确计算而理论上“不再有什么可做的”。可是 80 年代又是物理学发展上激动人心的 10 年。凝聚态物理与理论更是硕果累累，高潮不断：异质结超晶格与各种人造结

构，低维系统，量子霍尔效应，纳米结构与介观物理，重费米子与强关联电子系统，高温超导， C_{60} 及富勒烯系列，等等。展望 90 年代乃至 21 世纪，可以肯定凝聚态物理学这一领域仍将蓬勃发展，生机盎然。

至此，我们仅仅提及了凝聚态物理中重要的然而毕竟是极小的一部分。从基本相互作用的观点看，凝聚态物理基本上只涉及电磁相互作用。在这方面，我们只提及了电子，几乎未提光子，而光子正越来越引人注目起来。且不说从所周知的激光，就说光导纤维。由于掺掺杂光导纤维实现了光放大器，它促进了人们对掺掺杂非晶态特别是玻璃的研究。这必将反过来推进光学通讯的发展，它与电子通讯互为补充，相互结合，将会带来新的工业革命。此外，还有更多涉及到生物学的研究……新现象、新问题还会层出不穷。

凝聚态物理理论只是理论物理的一部分，作为理论物理，如果以本世纪初作为开端(代表人物如普朗克)，那末经过了几代人的努力，它的辉煌成就，它所体现的人类智慧、信念和勇气将继续鼓舞这一代人。这里我们引用爱因斯坦的话作为结束。

“……

为了整理和理解我们的感觉印象世界，我们试图借助物理学理论，通过观察到的事实的迷宫，找出一条出路。我们希望观察到的事实能从我们的实在概念逻辑地推论出来。要是不相信我们的理论结构能够掌握实在，要是不相信我们世界的内在和谐，那就不可能有科学。这种信念是，并且永远是，一切科学创造的根本动力。在我们的一切努力中，在每一次新旧观点之间的戏剧性的冲突中，我们都认识到寻求理解的永恒欲望，以及对于我们的世界和谐的坚定信念；这种欲望和信念都随着寻求理解的障碍的增大而不断地加强。”

参考文献

- [1] R.P. Feynman, Lectures on physics. (1, 2, 3 卷) 1963, 64, 65.
- [2] C. Kittel, Introduction to solid state physics, 1986 6ed.
- [3] K.G. Wilson Rev. Mod. Phys. 47 (1975) 773.
- [4] P.W. Anderson, Basic notions of condensed matter physics, 1984
- [5] A. Isihara, Condensed matter physics, 1991.

[6] D.Pines & P.Nozieres , The theory of quantum liquids , 1966.

[7] P.Nozieres & D.Pines , The theory of quantum liquids , 1984.

[8] 冯端等，凝聚态物理学新论，上海科学技术出版社，1992.

[9] 阎守胜甘子钊，介观物理，1995.

