

兰道尔宇宙三部曲

套 / 装 / 共 / 三 / 册



湛庐文化
Cheers Publishing
a mindstyle business 与思想有关

目 录

[叩响天堂之门](#)

[弯曲的旅行](#)

[暗物质与恐龙](#)

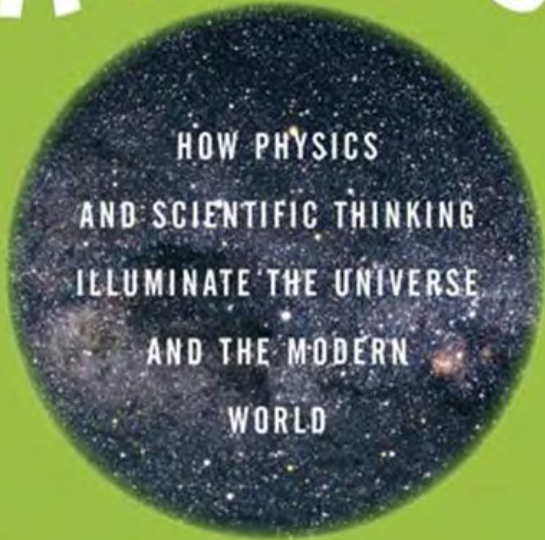
书籍免费分享微信jnztxy

Business 与思想有关

理论物理学大师
丽莎·兰道尔宇宙三部曲

KNOCKING ON HEAVEN'S DOOR

向爱因斯坦宣战的
理论物理学大师 丽莎·兰道尔
风靡世界之作



HOW PHYSICS
AND SCIENTIFIC THINKING
ILLUMINATE THE UNIVERSE
AND THE MODERN
WORLD

叩响天堂之门

宇宙探索的历程

[美] 丽莎·兰道尔 著 杨洁 符玥 译

LISA RANDALL

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社 于2016年10月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：叩响天堂之门

著者：[美]丽莎·兰道尔（Lisa Randall）

译者：杨洁，符玥

字数：390000

电子书定价：53.99美元

KNOCKING ON HEAVEN' S DOOR

Copyright 2011, Lisa Randall 仝

All rights reserved

目 录

版权信息

湛庐文化“科学素养”专家委员会 寄语

科学伴光与电前行，引领你我展翅翱翔

湛庐文化“科学素养”书系 专家委员会

重磅赞誉

推荐序

发现的激情

中文版序

宇宙的故事

..

欢迎来到科学的世界

第一部分 宇宙的故事

01 神奇的科学尺度

02 伽利略的科学求索

03 生于物质世界

04 物中之妙

第二部分 进入物质世界

05 谜般的梦幻之旅

06 “眼见”为实

07 寻找宇宙的答案

第三部分 大型强子对撞机的传奇历程

08 一个约束它们的加速环

09 加速环归来

10 有黑洞，还是没黑洞

11 这是一项风险差事

- 12 一切都是概率
- 13 进入大型强子对撞机
- 14 鉴定粒子，期待新粒子出世

第四部分 寻找“上帝粒子”

- 15 真、美与其他科学错觉
- 16 希格斯玻色子，宇宙万物为何产生
- 17 谁才是世界的下一个顶级模型
- 18 “自下而上”，还是“自上而下”

第五部分 缩放宇宙，让我们极目远眺

- 19 内外翻转
- 20 君之须弥，我之芥子
- 21 暗黑世界的来访者

第六部分 宇宙的探索仍在继续

- 22 海阔天空与脚踏实地

结语

寻路宇宙边缘

致谢

注释

译者后记

[shu籍 分.享 V信jnztxy]



你不是一个人在读书！
扫码进入湛庐“趋势与科技”读者群，
与小伙伴“同读共进”！

湛庐文化“科学素养”专家委员会

寄语

科学伴光与电前行，引领你我展翅翱翔

欧阳自远

天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

当雷电第一次掠过富兰克林的风筝到达他的指尖；
当电流第一次流入爱迪生的钨丝电灯照亮整个房间；
当我们第一次从显微镜下观察到美丽的生命；
当我们第一次将望远镜指向苍茫闪耀的星空；
当我们第一次登上月球回望自己的蓝色星球；
当我们第一次用史上最大型的实验装置LHC对撞出“上帝粒子”；
……

回溯科学的整个历程，今时今日的我们，仍旧激情澎湃。

对科学家来说，几个世纪的求索，注定是一条充斥着寂寥、抗争、坚持与荣耀的道路：

我们走过迷茫与谬误，才踟蹰地进入欢呼雀跃的人群；

我们历经挑战与质疑，才渐渐寻获万物的部分答案；
我们失败过、落魄过，才在偶然的一瞬体会到峰回路转的惊喜。
在这泰山般的宇宙中，我们注定如愚公般地“挖山不止”。所以，

不是每一刻，我们都在获得新发现。但是，我们继续。
不是每一秒，我们都能洞悉万物的本质。
但是，我们继续。

我们日日夜夜地战斗在科学的第一线，在你们日常所不熟悉的粒子世界与茫茫大宇宙中上下求索。但是我们越来越发现，虽这一切与你们相距甚远，但却息息相关。所以，今时今日，我们愿把自己的所知、所感、所想、所为，传递给你们。

我们必须这样做。

所以，我们成立了这个“科学素养”专家委员会。我们有的来自中国科学院国家天文台，有的来自中国科学院高能物理研究所，有的来自国内物理学界知名学府清华大学、北京师范大学与中山大学，有的来自大洋彼岸的顶尖名校加州理工学院。我们汇集到一起，只愿把最前沿的科学成果传递给你们，将科学家真实的科研世界展现在你们面前。

不是每个人都能成为大人物，但是每个人都可以因为科学而成为圈子中最有趣的人。

不是每个人都能够成就恢弘伟业，但是每个人都可以成为孩子眼中最博学的父亲、母亲。

不是每个人都能身兼历史的重任，但是每个人都可以去了解自身被赋予的最伟大的天赋与奇迹。

科学是我们探求真理的向导，也是你们与下一代进步的天梯。

科学，将给予你们无限的未来。这是科学沉淀几个世纪以来，对人类最伟大的回馈。也是我们，这些科学共同体里的成员，今时今日想要告诉你们的故事。

我们期待，

每一个人都因这套书系，成为有趣而博学的人，成为明灯般指引着孩子前行的父母，成为了解自己、了解物质、生命和宇宙的智者。

同时，我们也期待，

更多的科学家加入我们的队伍，为中国的科普事业共同贡献力量。

同时，我们真诚地祝愿，

科技创新与科学普及双翼齐飞！中华必将腾飞！

刘阳自述

湛庐文化“科学素养”书系 专家委员会

主席

欧阳自远 天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

委员（按拼音排序）

陈学雷 国家杰出青年科学基金获得者，国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

陈雁北 加州理工学院物理学教授

苟利军 中国科学院国家天文台研究员，中国科学院大学教授

李淼 著名理论物理学家，中山大学教授，中山大学天文与空间科学研究所所长、物理与天文学院行政负责人

王青 清华大学物理系高能物理核物理研究所所长，中国物理学会高能物理分会常务理事

张双南 中国科学院高能物理研究所研究员和粒子天体物理中心主任，中国科学院粒子天体物理重点实验室主任，中国科学院国家天文台兼职研究员和空间科学研究部首席科学家

朱进 北京天文馆馆长，《天文爱好者》杂志主编

朱宗宏

北京师范大学天文系教授、博士生导师，教育部“长江学者”特聘教授，北京天文学会理事长

重磅赞誉

韩 涛 著名理论物理学家，美国匹兹堡大学物理天文系杰出教授
匹兹堡大学粒子物理、天体物理及宇宙学中心主任

人类真的生活在一个具有多维空间的膜宇宙之上吗？暗物质真的是毁灭“地球霸主”恐龙的“幕后黑手”？发现了“上帝粒子”希格斯玻色子的大型强子对撞机，以及未来的超级对撞机，会为这些玄妙的问题提供深刻的答案吗？听天才理论物理学家丽莎·兰道尔教授用妙趣横生的案例、通俗易懂的语言，对科学求索的真相与未来娓娓道来，让人欲罢不能。这是时下科学研究前沿最振聋发聩的声音！振奋人心，启迪心智！

张双南 中国科学院高能物理研究所和国家天文台双聘研究员
中国科学院粒子天体物理重点实验室主任

我们还没有探测到暗物质，但恐龙的灭绝竟然是暗物质造成的？兰道尔“宇宙三部曲”将告诉读者，想理解地球和人类的现在、历史与未来，我们必须搞清楚物质最深层次的结构和宇宙最大尺度的规律！唉，我真为其他想写类似主题的作家们担心，再写出这么出色的书恐怕很难了。

陈学雷 国家杰出青年科学基金奖获得者
国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直开展着最前沿的科

学研究。在这套科普书中，兰道尔教授介绍了物理学家们是如何研究、探索宇宙之谜的。

她并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家们现在正在进行的猜想和探索，使读者真切地欣赏到科学研究的丰富多彩和趣味，体验科学家们在构造假说、探索未知、获得新发现时所体验到的激情。我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。

朱进 北京天文馆馆长

在兰道尔教授的笔下，额外维度、暗物质、暗能量、对撞机，这些科学家的“烧脑伙伴”也变得平易近人起来。这套科普书系通俗易懂，与晦涩无缘，揭示了即使是门外汉都读得懂的宇宙真相。

苟利军 中国科学院国家天文台研究员，中国科学院大学教授“第十一届文津奖”获奖图书《星际穿越》译者

几千年来，人类一直在试图回答“宇宙是什么”这一古老问题。现代天文观测和研究揭示，宇宙包含了时空和普通物质以及很多神秘“角色”。作为世界知名的粒子物理学家，哈佛大学物理系教授丽莎·兰道尔在她的这套系列丛书中，以其渊博的知识、广阔的视野、通俗的语言，以及丰富有趣的事例，给我们讲述了宇宙的基本组成和包含万物的时空，非常值得一读。作者大胆推断，地球上恐龙的灭绝与银河系中的某种暗物质有关。如果这能够被证实，将颠覆我们对宇宙神秘物质的现有认识。

吴岩 科幻作家，北京师范大学教授

简明扼要、通俗易懂、内容独创。地球人非读不可！

万维钢（同人于野）

科学作家，畅销书《万万没想到》作者“得到”App《万维钢·精英日课》专栏作家

过去几十年来，理论物理学中最酷的话题已经从量子力学、相对论和黑洞变成了超弦、希格斯粒子和暗物质。如果说，黑洞让人着迷、量子力学让人困惑、相对论让人脑洞大开，那这些新概念则更难让人理解！不过一旦你理解了，就会获得更大的智力愉悦感。物理学家一直致力于在不用公式的情况下让公众理解物理学，丽莎·兰道尔正是这项事业的新晋翘楚。她用一贯的机智语言告诉我们，这一代的物理学正在发生什么。[shu籍 分.享 V信jnztxy]

郝景芳 2016年雨果奖获得者，《北京折叠》作者

在这个信息爆炸的时代，我们收到的碎片化信息太多，反而难以获得真知。碎片化文章看得再多，也不如读一本真正的好书，尤其是深入浅出、结构恢弘的好书。兰道尔“宇宙三部曲”就是难得一见的、视野辽阔的好书，每一本都选择了令人好奇的话题：宇宙结构、宇宙历史、宇宙物质，并且还与恐龙灭绝这样有趣的话题相结合，更加吸引人，让人读起来手不释卷。而最为难得的是，兰道尔的文笔简洁、优美，你在书中找不到像一般物理学科普图书那种艰深晦涩的语句。她用小说一样的文笔娓娓道来，让你理解人类对宇宙最全面的认知。

比尔·克林顿 美国前总统

丽莎·兰道尔以她诙谐、通俗的一贯风格写下了《叩响天堂之门》，她让复杂的物理理论变得迷人、易懂。她的书，将会激发你产生不一样的想法，并鼓励你对世界作出更为睿智的判断。

理查德·道金斯 著名生物学家、科普作家
畅销书《道金斯传》《自私的基因》作者

科学是一场头脑与心灵的战斗，战场有两处：反对迷信以及忽略事物的某个侧面；反对伪智力的蒙昧主义。能够拥有丽莎·兰道尔那不同寻常的科学理念、清晰易懂的内在逻辑以及她美丽的外表，于我们而言，实为幸事。

史蒂芬·平克 著名认知心理学家、科普作家
畅销书《心智探奇》《思想本质》《语言本能》作者

兰道尔对物理学前沿理论及她自己那耀眼理念的诠释，既明晰易懂，又具有高度的启发性。她对科学知识的拥护，也是对科学进步的极大贡献……今日阅读兰道尔的书，将得以理解明日之科学。

克雷格·文特尔 “人造生命之父”，基因测序领域的“科学狂人”畅销书《生命的未来》作者

丽莎·兰道尔为非物理学读者解释了现代物理学的基础科学，以及最新的物理学实验有可能揭示什么。这项“科普”工作，兰道尔做得很好……《叩响天堂之门》必须一读，它可以帮助我们领会未来将要发生什么。

丹尼尔·吉尔伯特 全球幸福研究领域最具影响力和最权威的研究者之一
畅销书《撞上幸福》作者

兰道尔通过《叩响天堂之门》一书，对科学的发展以及未知宇宙如何运作，做了非常好的解释。

劳伦斯·萨莫斯 哈佛大学名誉校长

在我眼中，丽莎·兰道尔是最罕见的珍品——一位天才理论物理学家，兼有用大众能够理解、享受的方式书写、讲授物理知识的才能。《叩响天堂之门》一书让非专业人士也可以尽可能地接近宇宙背后的机制。

《纽约时报》

兰道尔是早期职业理论物理学家之一……她对科学与艺术之美之间的密切关系给出了漂亮的分析，并把理查德·塞拉雕塑作品中的对称性破缺与标准模型核心的对称性破缺进行了比较。

《自然》

从提出概念到具体实行，兰道尔这部富有信服力的著作细致地讨论了大型强子对撞机（LHC）的建成历程、它所历经过的种种磨难，带领我们走过了一段大型强子对撞机尺度之下的宏伟的科学之旅。

《书单》

兰道尔的著作是一般读者通向科学前沿不可或缺的“护照”。

《科克斯书评》

这是对粒子物理学相关科学工作的一次性灵的审视……兰道尔为这个主题带来了洪钟大吕般的热情，而她自己却无比闲适，还时而会冷幽默一把……《叩响天堂之门》真如星光一般灿烂。

推荐序

发现的激情

陈学雷

国家杰出青年科学基金获得者

国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

拿到这套书的样章，让我想起20多年前（1993年），我作为一名物理学研究生，参加了由李政道先生创办的中国高等科学技术中心组织的一个国际物理学会议。会议日程上列出的报告中有几位大名鼎鼎的学者，他们的名字，我们在粒子物理学教科书中早已熟悉。但当时还有一个我不很熟悉的名字“Lisa Randall”，而且在日程中排在十分显著的位置。会议开始后，我见到了她：一位面容美丽、身材苗条的女子。她看上去似乎比我大不了几岁，却十分高冷。而且，我听说她酷爱攀岩。然而在会议中，无论是演讲、问答还是讨论，她都显得学识渊博、机敏睿智、充满自信，与那些年龄、资历都老得多的学者辩论时，完全不落下风，成为会议的中心人物之一。这完全打破了我那时对女性物理学家的错误刻板印象。诚然，我从小遇到过很多成绩比我更优秀的女同学，但也许是因为女孩子们的谦让、文静和不好争辩，总让我怀疑她们不过是比我更用功、更擅长作业和考试而已。对于她们是否能深刻地思索或者作出创造性的发现，我内心总有一点儿怀疑。在物理学发展史上，女性物理学家，特别是理论物理学家，也确实屈指可数。然而，站在我面前的就是一位活生生的杰出的女性物

理学家，这证明之前我错了。当然，自那之后，我有幸遇到过很多优秀的女性物理学家，其中也包括丽莎·兰道尔教授的一位中国女弟子苏淑芳博士。她们都向我证明了，女性在物理学或者其他科学研究中，完全可以取得毫不逊色于男性的成就。

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直在进行着最前沿的科学研究。她有许多卓越的成就，其中最著名的是她与桑卓姆合作提出的“额外维度”模型。在这个模型中，我们所熟知的三维空间只是高维空间中的“膜”（参见《弯曲的旅行》一书）。兰道尔教授的这套科普书系，介绍了物理学家是如何研究、探索宇宙之谜的。《叩响天堂之门》一书不仅介绍了大型强子对撞机所进行的研究的意义，也用科学的道理和事实，澄清了人们对科学的各种误解；《弯曲的旅行》一书，重点是对高维空间的探索；《暗物质与恐龙》一书则介绍了作者提出的一种特别的暗物质模型，并就此提出了一个关于恐龙灭绝的有趣假说，借此又阐述了从宇宙起源到暗物质、从太阳系演化到恐龙等多方面的知识。这三部著作的共同特点是，作者并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家现在正在进行的猜想和探索，当然也清楚地说明了哪些仍仅仅是猜想和假说。这些猜想也许未必都正确，其中许多可能也会被未来的实验和观测所否定。但是，对这些内容的介绍更可以使读者真切地欣赏到科学研究的丰富多彩和趣味，体验到科学家们在构造假说、探索未知、获得新发现时所体验的激情。

我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。我也特别希望，这些书能鼓励那些喜爱科学、希望未来从事科学研究的女孩子们。

中文版序

宇宙的故事

得知我的三本书将在中国出版，我感到十分兴奋。不论在理论物理学还是在实验物理学的舞台上，中国都在扮演着日益重要的角色。

我有一些优秀的中国学生以及博士后，而且我也发现，近年来在中国这片土地上，人们对我研究方向的兴趣正在不断增长。不仅如此，中国的实验物理学也在近期取得了一些重要成果。例如，在大亚湾中微子实验室中对最轻、最重中微子混合的振荡测量，其结果震惊了世人，而且它比人们的预期早了至少一年。现有的暗物质探测器，包括PandaX与CDEX，标定了一些重要的能量范围，并且仍在不断探索，以揭示神秘的暗物质粒子的本质。展望未来，计划在中国建造的最大型的对撞机至关重要，它将成为国际主要的粒子物理学实验装置，并能够胜任探索超越已知领域的重任。

作为一位理论物理学家，我的研究领域涉猎甚广，小到物质的内部结构，大到宇宙、空间的本质。这些研究令人兴奋，然而又很难向他人解释清楚——在没有对应语境的情况下更难说清。这三本书给了我一次机会，不仅可以向世人解释我的研究，还可以同时解释作为我研究基础的量子力学、相对论、粒子物理学与天体物理学等物理学知识。我将乐于讲述一些展现这些领域中研究前沿的宏大故事。

《叩响天堂之门》一书，解释了科学的本质，并强调了尺度的重要性，也即如何在基本粒子、原子、普通物质或是宇宙的尺度上思考科学问题。《叩响天堂之门》一书也探索了科学的发展历程、什么是“对”与“错”，以及创造力在科学发展中的意义。在这本书中，我还预测了大型强子对撞机（LHC）上的物理结果。大型强子对撞机是建造在日内瓦附近的大型加速器，能让高能质子对撞，以产生新的粒子与新的相互作用，它可以用来研究人类之前所不能及的更小尺度。这本书解释了大型强子对撞机如何运作，以及在这一实验中，科学家正在研究什么以及他们未来将要研究什么。

《弯曲的旅行》一书讲述了我对空间中可能存在的卷曲的额外维度的研究——额外维度是在我们容易观察到的三个维度（左-右、前-后、上-下）之外的某个维度。额外维度可能具有重要意义，它将解释基本粒子的质量，并为它们之间的相互作用引入新的理论可能性。这些卷曲也将容许空间具有一个无穷大的额外维度，它将与我们观测到的一切事物相容。为了讲述这个故事，我回顾了前沿研究中的量子力学、相对论、粒子物理学的基础（既有理论，也有实验），还回顾了弦理论。在这本书中，我会讲述我们是如何把所有研究领域联系在一起，我们是如何得到了这一切，以及我们已经走到了哪一步的大故事。

《暗物质与恐龙》一书，既向外审视宇宙的宏大图景，又向内一窥物质的内部结构。它解释了暗物质的本质及其在宇宙演化中扮演的角色——暗物质是宇宙中捉摸不定的物质，只与引力而不与光相互作用。《暗物质与恐龙》一书也强调了物质的基本性质与我们今日所见的地球、宇宙之间的联系。这本书的内容不仅涵盖了宇宙学，还涉及星系、太阳系和地球之间的相互作用，及其与周边环境的联系。在这一旅程中，我还将解释我对暗物质的新理念：暗物质可能包含了某个小组分，这个组分通过自身的媒介物质——光进行相互作用，而普通

物质不与之相互作用。这可能会产生激动人心的结果，包括在银河系平面上，暗物质盘将形成，其引力效应可能导致巨大的流星体撞击地球，从而最终导致恐龙的灭绝。

《叩响天堂之门》《弯曲的旅行》《暗物质与恐龙》三本书包含了粒子物理学的广阔思想领域，是对我的研究以及更宽泛的粒子物理学和宇宙学的一个简述。这三本书把许多新颖且多元的理念与科学领域结合在一起，给出了对今日科学家工作状态的一个感性认知。

在完成《暗物质与恐龙》这本书之后，我已继续投身于对暗物质的研究中。在许多已有研究的基础上，暗物质已经是一个比较成熟的研究主题了。我们在实验上有了许多直接的结果，也开始着手于更好地理解用来探索宇宙的天文望远镜、人造卫星是如何阐明“暗物质是什么”这一问题的。同时，理论也在不断发展，人们已经超越了对暗物质粒子非常狭隘的假设，并对“暗物质如何相互作用”有了更深刻的想法。我正在思考关于暗物质粒子全部带电（而不只是一小部分带电）的可能性，并自问这一假设可能导出什么结果。现有的研究忽视了某些使这个假设可行的重要结果，或许这可能只是因为：暗物质着实不像粒子物理学家之前所假设的那么无趣。

理解暗物质这种神秘物质的本质，是一个非常令人兴奋的研究主题，毕竟它占据了全部物质能量的85%。我希望中国的读者能享受这一旅程：跟随我一起探索我们是由什么组成的，宇宙中的相互作用是如何发生的，以及我们这些科学家是如何研究宇宙问题的。我确信，你们将会从中学到许多新知识，同时又能提出属于自己的问题和观点。



大型强子对撞机是什么？扫码关注“庐客汇”，回复“叩响天堂之门”，听兰道尔教授用3分钟解答大型强子对撞机的秘密。

引言

欢迎来到科学的世界

我们正沐浴在一道划时代发现的曙光之下。

世界上规模最大的粒子物理学及宇宙学实验正在缓缓拉开序幕，全球顶尖的物理学家和天文学家纷纷投身于这场科学盛宴中。科学家们未来10年将要作出的发现，终将使我们对物质的基本构成乃至空间本身的认知发生巨变。这将为我们理解物理世界的本质提供更为全面的视角。关注这些进展的人们不会仅仅把它们当成后现代时期的“附属品”。我们致力于探索一些伟大的发现，这些发现有可能引出属于21世纪的科学新范式。这种范式将有助于我们探索宇宙的潜在构造，从那些尚未被人们发现的视角，改变人们对宇宙基本构造的认识。

2008年9月10日是一个历史性的日子。就在这一天，大型强子对撞机（Large Hadron Collider, LHC）首次试运行。尽管“大型强子对撞机”这个名字直白乏味、毫无创意，然而正所谓“圣人无名”，它取得的科学成就必将令世界惊叹。在“大型强子对撞机”这个名字中，“大型”是指对撞机本身，而非其中参与对撞的强子。大型强子对撞机包含一个全长26.6公里的巨型地下环形隧道，该隧道穿过了法国与瑞士的国界线，把侏罗山和日内瓦湖连接了起来。隧道中的电场同时加速两个粒子束，每一束都包含着数十亿质子（隶属于强子，对撞机也因此得名），它们以每秒11000圈的速率在隧道中狂飙。

大型强子对撞机要负责有史以来从很多方面来看都规模最大、最引人注目的物理实验的运行。其目的是，在小到无人测量过的尺度，以及高到无人探索过的能标上揭示物质结构之谜。这些能量会形成一批基本粒子，并揭示早期宇宙演化过程中出现的相互作用——这些都是在宇宙大爆炸之后大约一万亿分之一秒的时间内发生的。

大型强子对撞机的设计穷人类之智慧，可谓巧夺天工，但它实际的建造过程却存在许多预期之外的困难。令那些急于探索自然奥秘的物理学家以及其他好奇者懊恼的是，在大型强子对撞机试运行仅仅9天之后，某个不良的焊点就引发了一场爆炸。但是，随着大型强子对撞机在2009年秋天的回归，并以超出所有人的期待良好运行的这一事实，让25年前人们的承诺终于逐渐变成了现实。

2009年春天，普朗克（Planck）卫星与赫歇尔（Herschel）卫星在法属圭亚那发射。我从一个来自加州理工学院的兴奋的天文学观测组那里获知了发射时间，即5月13日早上5:30。当时，我在帕萨迪纳（Pasadena），从远处亲眼见证了这个意义重大的事件。赫歇尔卫星将提供关于恒星形成过程的观测，而普朗克卫星将为我们提供大爆炸残余辐射的细节信息——这有助于我们了解宇宙的早期历史。发射虽然激动人心，但也令人紧张——毕竟还存在2%~5%的失败率，而一旦发射失败，科学家们数年来的辛勤工作，就会随着坠毁于地球卫星中的定制科学仪器一同付诸东流。所幸，这次意义重大的发射非常顺利，它在日间发回的宝贵信息宣示着它所取得的巨大成功。即便如此，我们也还要等待许多年，以待这些卫星中关于恒星和宇宙的最有价值的信息被科学家们揭示出来。

科学之美，拓宽知识的边界

对于宇宙在大尺度上以及在高能标下的行为，物理学家已经建立了基础牢靠的理论模型。现有的理论和实验，已经让科学家们对宇宙的基本元素和结构有了深刻的理解，这种理解涵盖了极为广阔的尺度范围。这些年来，我们已经逐渐推演出了一个综合而细节明晰的架构，它可以把之前我们已经获得的理论碎片拼合在一起。这些理论成功地描述了宇宙是如何演化而来的：极小的成分形成了原子，然后聚合为恒星，继而形成星系以及更大的、遍布宇宙的结构；还描述了某些恒星如何爆炸、产生重元素，这些重元素又如何进入银河系、太阳系，并最终成为生命诞生的基石。通过大型强子对撞机以及上面提到的卫星探测的实验结果，如今的物理学家们希望能建立一个基础更加坚实、涵盖领域更广的理论，该理论可以加深人们在更小尺度与更高能标上对宇宙的理解，使理论达到之前从未达到的精确度。这将是人类迈出的一大步，而你我对此都怀有凌云壮志。

你应该已经了解过一些清晰明确的科学定义，它们比宗教信仰中的定义要明了得多。然而，这些定义形成的真实过程却极为复杂。即使我们倾向于认为，它们是外在真实世界以及物理世界运作规则的真实反映（至少我在科学之路上启程时是这么想的），然而实际的研究几乎不可避免地都是在一种不确定的状态下进行的。在这种状态下，我们期望取得进展，但不确定方向是否正确。科学家面临的挑战是：坚持那些有前途的理念，并不断地质疑它们，以确认其真实性及其蕴涵的真正意义。科学研究不可避免地涉及对那些精致、优美理念的权衡。这些理念往往处于困难、竞争与矛盾的边缘，也因此令人备加兴奋。总之，我们的目标是尽力拓展已有知识的边界。当有人开始刻意篡改数据、概念以及方程式时，即使是那些原本正确的理念，也变得不确定起来。

我的主要研究对象是基本粒子——对已知尺度最小的事物的研究，领域涉及弦理论以及宇宙学——对已知尺度最大的事物的研究。

我与同事们致力于理解以下这些问题：

- 物质的核心是什么？
- 宇宙中正在发生什么？
- 那些由实验发现的基本物理量和性质，在本质上是如何联系在一起？

像我这样的理论物理学家并不会亲自去判定什么理论适用于真实世界的实验，而是试图对可能的实验结果进行预测，以及帮助设计测试某些理论有效性的新实验方法。在可预见的未来，我们试图回答的问题对于改善人们晚餐食物这种事情毫无裨益，但这些研究将会回答以下终极问题：我们是谁？我们从哪里来？

《叩响天堂之门》讲述的是我们的研究内容，以及我们面对的最重要的科学问题。粒子物理学与宇宙学的新发展有望修正我们理解世界的理论，譬如世界的诞生与演化，以及驱动其运作的基本作用力。本书将要描述在大型强子对撞机上进行的、期望能够发现新事物的实验和理论研究；也会描述宇宙学的研究成果，即我们如何试图推断出宇宙的本质，尤其是宇宙中无所不在的神秘暗物质。

本书的视角不止如此，它还将探索适用于所有科学研究的普遍原则。虽然我们要描述当今科学研究的前沿，但是本书的核心目的在于辨析科学的本质。它将描述如下过程：

- 我们如何决定提出什么样的问题是恰当的；
- 为何科学家们对这些问题众说纷纭；
- 正确的科学观念如何最终得以成为主流观点。

本书将要探索科学发展的真实方法，并将之与人类对其他探寻真理方法的尝试相对比，给出一些科学的哲学基础以及描述一些“中间阶段”。在这些阶段，我们尚不确定孰对孰错，也不清楚前路终于何方。但是这些思考的重要之处在于，它们向我们展示了科学理念、科学方法是如何在科学之外的领域应用的，进而如何促进了那些领域中更加合理的决策的形成。

漫漫科学旅程，看尽将实而仍虚的奇迹

《叩响天堂之门》面向的是非科学专业的读者，他们可以通过本书对现代理论物理学和实验物理学有更深入的了解，对科学思想的基本原理以及现代科学的本质产生更高的审美情趣。人们往往并不真正懂得“科学是什么”“我们可以期待科学告诉我们什么”。本书是我意图纠正人们一些误解的尝试，或许还包含一些沮丧心情的发泄——这些发泄针对的是科学在当代被理解和应用的方式。

近年的工作使我拥有了一些不平凡的经历以及社交经验，从这些经历与交流中我获益良多。我打算以分享这些经历为切入点，开始我们对一些重要理念的探索。即使在本书涉及的领域中我并非全能专家，而且出于篇幅所限，我不可能在描写中做到面面俱到，然而我依旧希望，通过《叩响天堂之门》，能够带领读者走向更加丰富的思考方向，并尽可能地解释一些在这些方向上最新的、最令人兴奋的进展。本书也会帮助那些有志于在未来探寻更深层次答案的读者，帮他们确定那些最值得信赖的科学信息和误导信息的来源。本书深入浅出地表达了一些理念，这些理念有助于我们更彻底地理解隐于现代科学之下的一些思辨，从而为科学研究以及当代社会面临的重要问题铺就一条光明大道。

在这个电影前传盛行的时代，你可以认为本书是《弯曲的旅行》（*Warped Passages*）^[0]的前传，以及对“我们已经走到了哪一步”“我们正在预期什么”这些内容的更新。它补充了之前没有讲到的内容、回顾了新观点及新发现之后的科学基础，进而解释了为何我们当下静候于进展的边缘，等待新的数据以便迈步前行。

本书将在以下两点之间交替叙述：一为今日科学已经取得进展的细节，二为隐含在科学之下的那些必要话题以及概念——提出这些，对理解更为广阔的世界也很有用。本书第一部分、第二部分的第11章与第12章、第三部分的第15章与第18章以及最后一章，偏重讨论科学思想，其他章节则偏重物理学本身，包括在物理学上我们已经走到了哪里，以及对漫漫来路的回首。从某种意义上讲，《叩响天堂之门》和《弯曲的旅行》这两本书是一体的，建议读者一同阅读。现代物理学宛在天上，它晦涩难懂，与日常生活毫不相干。然而，指导我们思考的哲学和方法论基础，应当能够明晰科学及其相关的思想，正如本书给出的许多例子一样。相反，一个人只有具备一些真正的科学基础知识才能完全领会科学思想中的一些基本元素，并落实这些想法。有着更高要求的读者也许会略读或干脆跳过其中一方面的内容，然而两者的适度结合才能在我们这道科学盛宴中烹饪出绝佳美味。

全书将反复提到一个关键词“尺度”（scale）。物理定律为已经确定的理论与其对自然的描述如何结合为一个有联系的整体，确立了一致的框架：小到大型强子对撞机上进行的实验尺度，大到整个宇宙的尺度。^[1]关于尺度的这些说明，以及在这条探索的道路上我们将要邂逅的那些既定事实与理念，对我们的思考有着决定性意义。已经确定的科学理论适用于目前我们可以理解的尺度。然而，在那些从先前从未探索过的或大或小的尺度上，我们正在获得的新知识将为我们引入更精确、更基础的理论。第1章将着眼于介绍“尺度”的概念，并解

释科学家为何把万物按照其尺度分类，对于已有的物理学与建立在其上的新科学进展是必要的。

本书第一部分也会介绍并对比科学家们在攀登科学高峰时，所选择的不同道路。当你询问人们在思考科学问题时的感受时，得到的答案很可能因人而异。某些人可能会坚持他们对物理世界严格而刻板的印象；一些人也许会把它当作一个不断被修正的原理的集合；另一些人或许会声称：科学不过是另一个信仰系统，与宗教或哲学并没有什么本质区别。这些观点也许都值得商榷。

为何有如此多的争论（哪怕是在科学内部），这是科学演化本质的核心问题。第一部分会介绍一点历史知识，它揭示了今日科学如何扎根于17世纪人类智慧的土壤，以及一些有关科学观念与宗教观念的正面交锋。这一部分也将着眼于唯物主义者对物质的观点、科学-宗教问题的棘手应用，以及“谁来回答基本问题”“如何回答它们”等问题。

第二部分将转而讨论物质世界的构造。它将为我们的科学之旅导航，引领我们从熟悉的尺度一直探索到已知的最小尺度，并按照尺度来划分物质世界。沿着这条路走下去，我们会从日常所及的领域一直走到亚微观尺度，这种尺度上事物的内部结构只能通过大型粒子加速器来探索。第二部分将以对今日主要物理学实验的总结结束，这些实验主要来源于大型强子对撞机以及一些对早期宇宙的天文学观测，它们将拓展人们的认知极限。

随着一些令人兴奋的进展的涌现，这项勇气十足且雄心勃勃的事业有望使整个科学的世界观发生翻天覆地的变化。在本书的第三部分，我们将深入了解大型强子对撞机的运行，探索这部机器是如何创造对撞质子束，以产生那些目前能探及的最小尺度粒子的。第三部分也将说明实验物理学家们如何阐释他们的发现。

欧洲核子研究中心（CERN）在发布与粒子物理学实验相关的信息上不遗余力——其程度不亚于好莱坞欢乐却有误导性的大片《天使与魔鬼》（*Angels and Demons*）。这个大型粒子加速器及其建立的目的因此而广为人知：它将具有巨大能量的质子聚集到一个极小空间中，让它们对撞，以产生人们前所未闻的物质形式。正在按部就班运行的大型强子对撞机，终将改变人们对物质的基本构成乃至空间本身的认知。而对于它将找到什么，我们现在尚一无所知。

在这次科学之旅中，我们将仔细回味科学的不确定性，以及科学测量可以告诉我们什么。科研的本质决定了它只能止于我们所知事实的边界。物理实验及计算都被设计得尽可能减少乃至消除尽可能多的不确定性，并且最终精确地给出误差范围。然而，听起来滑天下之大稽而又确然如此的是，科学在实践与一些基本原则里，已经充满了不确定性。第三部分讲述了科学家是如何处理艰难研究中存在的内在挑战，以及大众在理解这个日益复杂的世界中产生的思想时，是如何从科学思维中获益的。

第三部分也将讨论可能产生于大型强子对撞机的微型黑洞，以及大众随之而来的恐慌，并把它和我们目前所面对的一些真正危险做对比。我们将会考虑有关的重要主题，包括收支分析、风险评估，以及人们如何做才会为这些问题提供有价值的意见——不论是在实验室内，还是在实验室外。

第四部分将详述人们寻找希格斯玻色子（Higgs boson）及具体模型的研究过程，这些正是科学家们对大型强子对撞机中存在什么及其搜索目标的合理猜想。如果大型强子对撞机的实验确证了某些理论家提出的观点，甚至只是偶然发现了一些预期之外的东西，那么这些结果都将改变我们看待世界的方式。在这一部分，我们将解释基本粒子如何得以产生质量的希格斯机制（Higgs mechanism），以及昭示着我

们应该能找到更多东西的等级问题（hierarchy problem）。这一部分也将讨论处理这些问题的模型，以及它们所预言的那些奇异粒子，比如那些与超对称性（supersymmetry）和空间额外维度（extra dimension）相关的粒子。

除了介绍一些具体的猜想之外，第四部分还将解释物理学家如何处理理论模型的建立以及那些指导性原则的有效性。这些原则包括“与美相伴的真理”（truth through beauty），以及“自上而下”（top-down）、“自下而上”（bottom-up）两种研究方法。它解释了大型强子对撞机正在寻找什么，以及物理学家们希望它找到什么。第四部分还将描述科学家如何试图把大型强子对撞机中看似抽象的实验数据联系在一起，并建立一些深刻、基础的理念以供人们审视。

随着对物质本质的深入研究，在第五部分我们将极目远眺。此时，大型强子对撞机在探索物质的最小尺度，而人造卫星和望远镜在探索宇宙的最大尺度，以研究宇宙的加速膨胀率以及大爆炸残余辐射的细节。这个时代也许会成为令人震惊的宇宙学进展的见证者，而宇宙学正是研究宇宙演化过程的科学。在这一部分，我们将在更大尺度上探寻宇宙的奥秘，讨论粒子物理学与宇宙学的联系，讨论神秘的暗物质以及与之相关的实验探究。

作为全书尾声的第六部分是一个综述，着眼于创造力以及创造性思维中那些丰富而多变的元素。它回顾了我们如何从源于日常所见的平凡事物出发，试图回答一些大问题。我们将以一些终极思考作为全书的尾声，包括当今科学与科学思维如此重要的原因以及科学与技术的共生关系——众所周知，它们彻底改变了现代世界。

我时常反思，试图使非科学工作者领会对其而言非常陌生的现代科学的精要，是多么棘手。这个挑战的困难性在我遇到一个班的大学

生时彰显无遗。当时我正在做一个有关物理学与额外维度的公开演讲，当获悉听众们都急切地想要提问时，我本以为他们是对空间额外维度的概念有一些困惑，不料他们却只想知道我的年龄。大众对科学缺少兴趣并非唯一的困难——那些学生确实在八卦之后转向关心严肃的科学问题了。然而，基础科学往往过于抽象、难于理解，这是无可辩驳的事实。在一个阐述基础科学重要性的国会听证会上，我曾经因为这一点与他人产生了交流障碍。当时与我一同参会的人有：美国能源部高能物理所主任丹尼斯·科瓦（Dennis Kovar）、费米国家加速器实验室（FNAL）主任皮耶·奥登（Pier Oddone）、托马斯·杰斐逊国家加速器实验室（Jefferson Lab）主任休·蒙哥马利（Hugh Montgomery）。自我多年之前在高中时代作为西屋科学竞赛（WSC）决赛的选手，被国会议员本杰明·罗森塔尔（Benjamin Rosenthal）带到这里周游一圈以来，这还是我成年后第一次走进政府大楼。罗森塔尔当时为我提供了高于其他入围者的礼遇——他们只收到了照片。

这次故地重游，我再次有幸参观了制定国家政策的那些政府办公室。众议院科学技术委员会的办公区域位于雷伯恩众议院大厦内。议员们坐在办公区后，我们面向他们而坐以“见证”他们的办公过程。有一些写有励志名言的牌匾悬挂在议员们的头顶，其中第一块写着：

没有异象，民就放肆；唯遵守律法的，便为有福。

WHEN THERE IS NO VISION THE PEOPLE PERISH. PROVERBS 29: 18.

看来，美国政府哪怕在负责科学与技术事务的国会办公区中也需要《圣经》。虽然如此，这句引文确实表达了一种高贵而精确的情感，我们都喜欢如此引经据典。

第二个不朽名言更常被人们引用，它是伟大诗人阿尔弗雷德·丁尼生（Alfred Tennyson）的：

我曾沉入到世人的目光所不能及的未来，看到了世界的幻象以及所有将实而仍虚的奇迹。

FOR I DIPPED INTO THE FUTURE, FAR AS MY EYES COULD SEE/SAW
THE VISION OF THE WORLD AND ALL THE WONDER THAT WOULD BE.

这确实也是描写我们的研究目标时，一句应该永铭于脑海的箴言。

然而颇具讽刺意味的是，我们在一个被如此布置的房间中，作为已经广为承认这些思想的科学世界“见证者”，还要面对这些光辉名句，即便我们的日常工作中无所不在地渗透着这些理念。另一方面，那些议员并不真正理解这些名句的真正含义，因为这些意义于他们的日常体验而言，只是无源之水。国会议员利平斯基在他的开场词中讲道，科学发展鼓舞了更多问题的提出以及形而上学的追问——这说明他曾经注意过那些名人名句，然而现在已显然把它们忘记了。“很少有人会看到那个高度”，他被人提醒道，他对此点头表示感谢。

处理完这些杂事，我们投入到了本职工作中——解释是什么造就了这个属于粒子物理学和宇宙学、前所未有、令人激动的时代。尽管这些议员时而提出一些多疑而尖锐的问题，我还是感受到了他们平日里受到的阻力。这些阻力来源于他们需要向选民解释，为何缩减科学研究资金是不明智之举——即便当经济形势不好时也是如此。这些问题涉及的范围广泛而细致，从具体的科学实验的目的到有关科学所扮演角色的更广泛的主题，还有对科学将把我们导向何处的追问。

由于议员们需要定期去投票，会时而缺席，所以我们对此断断续续地给出了一些支持基础科学并获益的例子。即使是着力于基础科学的研究也经常会产生意想不到的影响其他领域的硕果。我们讨论了蒂姆·伯纳斯-李（Tim Berners-Lee）对万维网的贡献，这使得来自不同国家的物理学家可以在大型强子对撞机上进行的联合实验项目中，

更加便捷地进行合作；讨论了物理理论的医学应用，比如PET扫描，即正电子成像术，这是一种以电子的反粒子来探测人体内部结构的方法；还讨论了超导磁铁的工业化生产所扮演的角色，这些本为对撞机而设计的磁铁，现在也可以用于核磁共振成像（MRI），以及广义相对论在精确测量方面的卓越应用，包括我们已经广泛应用于汽车导航之中的全球定位系统。

当然，有重大意义的科学发现并不一定能在可预期的时间内产生短期的利益。即使它最终可以产生效益，我们也往往不能在最初发现的时候便得知这一点。当本杰明·富兰克林意识到天上的闪电即凡间之电时，他并没有意识到电力将迅速改变整个世界的面貌；而爱因斯坦在提出广义相对论时，也并没有预期它可以应用于任何实际项目。

所以，我们今日首要关注的并非科学具体的实际应用，而是至关重要的纯粹科学。即使在美国科学研究的现状摇摇欲坠，很多人还是认识到了其长期价值。大众对宇宙与时空的观念被爱因斯坦改变了，正如我在《弯曲的旅行》一书中引用的《随时光流逝》（*As Times Goes By*）^[2]的歌词所显示的那样^[3]，人类特有的语言与思维，随着人们对物质世界理解的变化以及思维方式的进步而变迁。科学家们今日正在研究的课题以及我们对它们的理解，不论是我们对世界的理解，还是对这个体力与脑力相结合的社会本身的理解，都极为重要。

随着之前一些尖端研究计划的提出，我们当下生活在一个不论于物理学而言还是宇宙学而言，都令人极其激动的时代。通过对大量科学发现的介绍，《叩响天堂之门》一书将会拓展我们理解世界的不同方式，不论是从艺术、宗教还是科学的角度出发，最终都会聚焦于现代物理学的目标及方法。最后我要说的是，我们对极小尺度上事物的探索，对于回答“我们是谁”“我们从哪里来”这些终极问题而言，非常必要。我们希望了解更多的大尺度结构，它们将会揭示关于宇宙

的一切，包括宇宙的起源及其终将面对的命运。本书将讲述我们希望找到什么以及如何找到它们。

这将是一次迷人的历险——欢迎来到科学的世界！



第一部分
宇宙的故事

**KNOCKING ON
HEAVEN'S DOOR**



在让我最终选择走上物理学之路的各种原因中，最重要的一个是，我希望能做一些可以“千古流芳”的工作。如果这些工作需要大量的时间、精力以及付出，那么我希望它们最终能成为永恒以及真理的代言。和很多人一样，我也认为科学发展经得起时间的考验。

在大学时期，我主修物理学，而我的朋友安娜·布克曼（Anna Büchmann）主修英文。具有讽刺意味的是，她选择文科的理由与我选择理科的理由是相同的。她深深喜爱一个经典故事能得以流传百年的感觉。多年之后，当我和她讨论亨利·菲尔丁（Henry Fielding）的小说《弃婴汤姆·琼斯的故事》（*Tom Jones*）时，我才得知她在读研究生时为我非常喜欢的一本小说的注释工作，贡献过力量。^[1]

《弃婴汤姆·琼斯的故事》已经出版250多年了，其深邃的主题以及智慧依旧绕梁三日，余音不绝。在我首次访问日本时，我阅读了日本古典文学名著《源氏物语》（*Tale of Genji*），也为其紧凑的行文情节而惊叹，虽然此时距紫式部（Murasaki Shikibu）写下这些文字已经悠悠千年。荷马在约两千年前写下了不朽名篇《奥德赛》，即便时过境迁，我们依然能够品味奥德赛之旅的传说及其彰显出的人性之美。

科学家很少阅读古老的科学文献，我们往往把这些任务留给历史学家与文学批评家。虽然如此，我们依旧对沿袭前人提出的理论情有独钟，不管这些理论是在17世纪由牛顿提出的，还是在16世纪由哥白尼提出的。我们弃其巨著，只取其书中所蕴重要理念的精华。

科学当然不是那些一成不变的通用法则（这些法则我们在小学时期就已经接触过了），更不是一系列信手拈来的规则。科学是一种知识的演化体。我们正在研究的很多理论，最终都将被证明是不完备的或是错误的。科学理念会在我们跨越已知领域与未知版图的边界时发生巨大变化，而在未知领域中，我们也许可以对揭示事实真相的线索有着惊鸿一瞥。

当追寻永恒不变的理论时，科学家必须面对的矛盾在于：他们此时信奉的理念，彼时会因为新的实验结果或者更好理解的出现，而不断被修正甚至被抛弃。合理而值得信任的知识核心经历过实践的检验，往往被不确定的事物所构成的边界所环绕，而这些正是现代科学研究的疆域。如果被更复杂而有说服力的实验工作所证伪，曾经激动人心的理念及想法将很快会成为明日黄花。

2008年，当竞选美国总统的共和党提名人麦克·赫卡比（Mike Huckabee）选择背弃科学，站在宗教的一方时，他并没有“完全被误导”——至少按他自己的界定是这样的。这一部分是因为“科学信仰”会发生变化，而基督徒们则把权柄授予一个永恒的、不动如山的上帝。



我们所了解的有关宇宙的科学知识正如宇宙本身一样，不断在演化。长久以来，科学家们抽丝剥茧般地剥开事实的外衣，以揭开潜藏在纷呈表象之下的

事物本质。在探寻更小尺度上奥秘的过程中，人们的眼界也在不断开阔，知识在随之增长。当逐渐接触到那些难于接近的尺度时，由于未知领域的减小，人们的知识水平得以增加。而随着知识领域的扩大，“科学信仰”也随之进化了。

即便飞速发展的技术让一系列更广泛的观察得以实现，我们也不必直接否定过去曾使用过的理论，毕竟它们曾经对我们可以理解的尺度、能量、速度以及密度，作出了成功的预言。科学理论是推陈出新的体系，它会把旧知识纳入更加综合的图景，而这个图景产生于一系列更大范围理论和实验观察。这种变化并不意味着旧理论必然是错的，而只表示当更小尺度上的新要素被揭示时，原有的规则就不再适用了。因此，知识可以在随时间扩张的同时容纳旧有的理论。学无止境，永远有新的事物等待我们去探索。正如跋山涉水可使人心荡神驰，即便你不可能遍历全球（更遑论宇宙了），格物致知也是丰富我们生命的一种追求。“吾之知也无涯”，这激励着我们继续迈出资求知的脚步。

我自己的研究领域，即粒子物理学，正在探索更小的尺度，以研究越来越小的物质组分。现在的理论与实验研究试图揭示潜藏在物质内部的存在。但是，物质并非如俄罗斯套娃一样，不断由更小尺度的相同元素嵌套而成。在不断变小的尺度上进行研究的有趣之处在于，在我们进入新领域时，原有的规则将不再适用。新的基本作用力与相互作用也许会在那些尺度上出现，而这些元素的强度在我们现在所研究的尺度上极其微弱，以至于现有的手段根本无法探测到它们的存在。

尺度让物理学家能够确定与任何特定研究相关的尺寸和能量的范围，它对理解科学的进步以及我们所处世界的诸多方面而言，都至关重要。通过把宇宙按照不同尺度分割为各个可以理解的部分，我们认识到，最佳的物理定律并非必然在一切情形下都相同。我们必须把那些在某个尺度上表现更好的概念，与那些在另一个尺度上更有用的概念联系在一起。采用这种归类方式，我们得以把已知的全部事物都纳入一个统一的图景中。在这个图景中，对不同尺度上事物的描述可以截然不同。

尺度

物理学家用来确定研究对象尺寸与能量的标准，可以根据尺度的不同对事物分类。不同尺度的事物可能适用不同的物理定律。

在本章我们将会看到，按照尺度对事物进行分类，对于明晰我们对科学以及其他事物的思考颇为有益。它也可以帮助我们厘清为何在日常生活的尺度下，构建物质基石的一些微妙性质是如此难以被我们注意到。在这个过程中，本章也会详细辨析科学中的“对”与“错”，以及为什么一些明显激进的理论进展也未必会导致我们已经熟知尺度上的巨变。

科幻界的怪力乱神

人们常常分不清“科学知识的演化”与“没有科学知识”之间的区别。他们误解了一种情况，即我们在完全没有一些可信赖规则的前提下，试图发展新的物理定律。在一次加州之行中，我与编剧斯科特·德瑞克森（Scott Derrickson）的一次交谈，使我明确了一些误解的起源。那时，德瑞克森正在忙于一些电影剧本的创作，这些剧本试图提出科学与一些可能被科学家们归入超自然现象的事物之间的潜在联系。为了避免专业性错误，德瑞克森决定让一位物理学家（也就是我）来帮忙审核他那些富有想象力的故事情节。于是，我们在一个阳

光明媚的下午，在洛杉矶的一家咖啡店共进午餐，以交流我们的想法。

由于自知编剧家经常歪曲科学概念，德瑞克森希望他剧中有关鬼魂与时间旅行的故事细节在一定程度上经得起科学的推敲。作为编剧，他需要面对的最大挑战在于，不仅要给观众传达有趣的科学现象，还要把这些现象在电影荧幕上表现出来。即使没有受过正规的科学训练，德瑞克森依旧可以快速接受新鲜事物。所以我向他详细解释了为什么他的故事在物理学上站不住脚，尽管其中某些情节是别出心裁而有娱乐精神的。

德瑞克森却认为，许多科学家今天看来荒谬的现象往往会在将来被认为是正常的：“科学家们曾经不是也拒绝承认过相对论吗？”“谁曾想到随机性会在基础物理定律里扮演重要角色呢？”^[4]尽管德瑞克森十分尊重科学，且能以史为鉴，但他还是怀疑：科学家们有时不是也会在他们理论的蕴意与适用范围上犯错误吗？

有些批评家甚至更严苛。他们断言，即使科学家已经能作出一系列伟大的预测，这些预测的可靠性也仍然值得怀疑。怀疑者们坚持，即使有着明确的科学证据支持，理论中也总是有未知的隐情或漏洞。也许死者可以复生，或者他们只是进入了一个通向中世纪或者中土世界的时空之门呢？这些怀疑论者不相信任何一件被科学断言绝对不可能发生的事情。

尽管保有开放的头脑以及认识到“吾之知也无涯”是人类的大智慧，然而一个谬论^[5]深深地隐藏在这种逻辑之中。只有当我们仔细分析上述思想的意义，特别是应用尺度的概念时，问题才变得清晰明白起来。这些问题忽略了如下事实：虽然总存在我们不能探及的更小尺度与更高能标，在那些情况下物理定律可能发生改变，但在人类日常生活的尺度上，我们已经足够好地掌握了相关物理定律，而它们也在悠长的岁月中经受住了无数次考验。

当我在惠特尼美术馆遇到编舞者伊丽莎白·斯特布（Elizabeth Streib）时，我们共同受邀在有关创造力的主题上发言，而她显然大大

低估了在人类的尺度上，科学知识的基础有多坚不可摧。斯特布提出了一个与德瑞克森问过我的相似问题：“物理学家提出的那些卷曲在一个难以想象的小尺度里的小维度，为什么没有在我们的身体运动时（比如跳舞）产生任何影响呢？”

斯特布是一位成功的编舞者，通过对舞蹈与身体动作的了解，她对一些科学基本假设的看法非常有趣。然而，我们不能确认新维度是否存在以及它们所扮演的角色。因为于我们而言，它们太小或者太过“卷曲”了，以至于我们无法探测到。通过上述说明，我想表达的是：我们迄今还未在已观测的尺度上找到任何它们的影响，即使是通过最精细的测量也不能找到。额外维度产生的物理现象不够明显，还无法对人们日常的运动产生任何可见的影响，否则额外维度早就被我们观测到了。因此可以推论出，即使我们对量子引力的理解更进了一步，编舞艺术的基础也不会因此而动摇。在人类日常可及的尺度上，这些微观物理规则的影响微不足道。

当科学家后来被证明出现错误，往往是因为那时他们还未能探索在极小与极大尺度或超高能标、速度这些极端条件下事物的行为。这并不意味着他们像卢德派（Luddites）^[6]一样对所有可能的进步永远封闭了头脑，而是意味着他们只相信那些对世界进行的最新数学描述，以及那些对可观测的事物和行为的成功预言。那些科学家们断言不可能出现的现象，在他们还没有探索或经历过的尺度与速度下是被允许出现的，事实上也的确出现过。科学家们现在当然不知道那些未来的观点和理论，它们最终将由那些小尺度和高能标下的规律所支配，并成为流行理论——而现在，科学家们对它们还不熟悉。



当科学家们断言“我们已经理解了某些东西”时，他们只是想表达“我们已经有了有一套确定的观点和理论，其预言已经在某个确定范围的能标下被很好地检验过了”。这些观点和理论并非必然是永远成立的规则，也并非必定是物

理定律中不可动摇的基石，而只是在当前的技术条件允许的参数范围内，很好地符合了已有实验结果。这意味着，这些理论有可能在未来某一天被新理论替代。牛顿定律在日常领域中正确而有效，然而在研究对象的速度接近光速时就会失效，替代它的是爱因斯坦的理论。牛顿定律是正确而不完备的，它的应用被限制在了某个特定范围内。

我们通过进步的测量手段获得更先进的知识，对揭示新的不同概念而言是一个进步。我们现在所知道的许多现象，古人不可能理解或发现，是因为他们受到所处时代观测技术的局限。所以德瑞克森所言不错，科学家们有时确实会犯错误，有些他们认为不可能的现象最终却可能会出现。但这并不意味着世界完全没有规则，鬼魂和时间旅行者不会排闥而入，外星生物也不会突然穿墙而出。空间中也许存在着额外维度，但它们可能极为微小，或者“卷曲”起来，或者暂时以某种方式隐藏了起来。总之，这些都可能是我们目前尚未发现它们存在的明显证据的理由。

也许奇异现象的确会出现，但这种现象只会存在于那些难于观察的尺度上，而在这些尺度上的现象往往不能直观地被理解，并有悖于人们的常识。如果它们总是可望而不可即，那么于科学家而言，它们就没有太大的意义；如果它们对我们的日常生活没有产生任何可见的影响，那么于小说家而言，它们也就失去了借题发挥的价值。

怪力乱神实属可能，但更让人容易理解的是，非物理学家们最为关注的总是那些我们可以观察到的现象。正如美国著名电影导演史蒂文·斯皮尔伯格在针对他正在思索的一部科幻电影的讨论会上所指出

的那样：一个无法反映在银幕上且我们永远无法感受到角色的奇异世界是无法吸引观众的（图1-1给出了一些有趣的证据），只有一个我们可见、可解的新世界才能吸引眼球。抽象理念与文学作品是不一样的，它们的目标不同，即使二者都需要想象力。虽然科学理念适用的那些领域与电影或是我们日常生活中的观察毫不相关，但是对描述物理世界而言，它们却是必要的。



图1-1 XKCD漫画，它抓住了隐藏的、小而卷曲维度的实质。

南辕北辙，科学不是魔法

即便所知有限，人们在试图理解晦涩的科学概念时总是希望“走捷径”。这常常又会导致人们对科学理论应用过分热情，而这早已不是什么新现象了。

早在18世纪，当科学家们正在实验室中忙于研究磁现象时，就有妄想家幻想出了“动物磁性说”（animal magnetism）这样的概念——他们认为生物体内存在“维持生命所需的磁流体”。1784年，路易十六为此还专门组建了法国皇家调查委员会以揭露这种伪科学理论，参与者是包括本杰明·富兰克林在内的一批科学家。



如今，这种误导性的应用更多见于量子力学，人们试图把量子力学应用于宏观尺度，而在宏观尺度上，量子力学的某些特定结果会被概率抹平，从而不会留下任何可测量的特征。^[7]令人担忧的是，许多人相信正能量可以带来财富、健康和快乐，如同朗达·拜恩（Rhonda Byrne）在其畅销书《秘密》（*The Secret*）中所描述的那般。拜恩声称：“虽然在学校时我从未学习过科学或者物理学，但是如今只要我想学，就可以很好地理解那些艰深晦涩的量子力学著作。学习量子力学让我在精神层面对《秘密》一书有了更深的感悟。”

然而，便即是量子力学的开拓者、诺贝尔物理学奖获得者，尼尔斯·玻尔（Niels Bohr）也会谦虚地说：“如果谁声称自己完全搞懂了量子力学，那么他一定是没有理解量子力学。”在此我也告诉大家一个“秘密”（至少与《秘密》中的秘密一样秘密）：量子力学被大众严重误解了。我们的语言和直觉都起源于“经典的逻辑推理”（classical reasoning），而这对量子力学并不适用。当然，这不意味着任何光怪陆离的现象在量子力学的框架内都有可能发生。即使对物理定律缺乏更深刻、更基础的理解，我们现在也明白应该如何正确应用量子力学以做预言。量子力学当然永远不会支持拜恩的“秘密”，比如所谓的人与遥远事物或现象之间的相互吸引原理。在那么长的距离上，量子力学的微观原则根本不适用。人们臆想出来的对很多现象的解释与量子力学毫不相干，虽然他们执意把这些解释归于量子力学。我无法靠一直盯着实验装置来改变某个实验结果，量子力学也并非不能作出可靠的预测。事实上，绝大多数测量的精度是被实际条件而非不确定性原理（uncertainty principle）所限制的。

这样的误解是我与马克·维森特（Mark Vicente）之间一次谈话的主题。维森特是电影《我们到底知道多少》（*What the Bleep Do We Know! ?*）的导演——这部电影是科学家们的梦魇，影片声称，人类自身的影响对实验非常重要。我不知这次对话终将止于何处，但是

那时我有闲暇，因为我正在达拉斯-沃斯堡国际机场候机。那次航班因为工程师们必须修补机翼上的一个凹痕而延误了。正如一位机组人员所说，一开始他们认为这个凹痕无关紧要，但是后来“经技术测量”认为它有碍安全飞行。

在候机时，我意识到：如果我想与维森特进行一番详尽的交流，那么我必须了解他制作电影的出发点——这些观点我早已从大量听我演讲的人那里领教过了，他们问了我许多稀奇古怪的问题，而他们都是看了电影之后才提出这些问题的。维森特的回答让我感到惊奇，他得转变相当大。他透露了以下想法：他最初先入为主地提出那些科学概念时，并没有仔细考虑过。而当回顾这些想法时，他觉得它们在本质上更接近于宗教而非科学。维森特最终承认，他在电影中提出的那些概念并非科学。把量子力学现象应用于人类日常经历中也许浅薄地满足了观众的猎奇心理，但它们并不正确。

即便新理论需要开创性的假定，正如量子力学，有效的科学实验与理论推导也最终会证明它们是正确的。科学不是魔法。科学方法、数据以及有效和自洽的探索经历告诉科学家如何拓展其知识面，以把即将可探索到的尺度上的直觉，延拓为适用于相同现象但基于不同观念的知识。

在下一部分中，我们将讨论有关尺度的概念是如何系统地构建不同的理论概念，以及我们又是如何据此把它们合并为一个整体的。

有效理论，忽略细枝末节

从人类可及的最小尺度到广阔无边的宇宙，也许出于随机性与巧合，人类自身的尺度差不多处于中间位置。^[8]与物质的内部结构及其更小的组分相比，我们是如此巨大；而与恒星、星系乃至宇宙相比，我们又是那么渺小。我们最易理解的尺度正是我们平日通过五感以及最原始的测量工具接触最多的尺度，并通过观测与逻辑推导来理解更深远的尺度。这些尺度范围似乎包含了越发抽象以及难于把握的性

质，这是由于我们正在逐渐远离直接可见、可接触的尺度。但技术与理论的结合允许我们在一个广阔的尺度范围上建立描述物质性质的理论。

小到大型强子对撞机所发现的微小事物，大到星系与宇宙，已知的科学理论在如此巨大的范围上得以应用。对每一个事物可能的尺度以及介于中间的尺度，物理定律都有与之对应的一个方面。物理学家需要处理运用到这些广阔领域中的海量信息。即便在小尺度下适用的、最为基本的物理定律，同样可以支配那些大尺度下的规律，然而在实际进行计算时，我们并不使用它们，以免给计算带来困难。当那些额外的结构和基础对于一个已经足够精确的结果来说毫无用处时，我们需要更加现实可行的方法来计算并有效地应用那些更简单的规则。

物理学中最为重要的特征之一是，它可以告诉我们如何鉴别特定测量或预言所对应的尺度范围——这要根据我们已经掌握的精度来进行计算。用这种方式看待世界的优点在于，我们可以集中处理那些与我们关心的现象所对应的尺度、识别出那些在这些尺度上运作的元素，进而发现应用这些组分所遵循的法则。当科学家们建立理论或者进行计算时，时常会对出现在无法分辨的小尺度上的物理学现象取平均值或者干脆忽略（有时是无意的）。我们选择有意义的事实、忽略一些细节、集中处理那些有用尺度下的现象——只有这样，我们才能勉强得到一些进展，才能处理那些原本无法处理的、定义在稠密集（dense set）^[9]上的信息。

在合适的情况下，忽略细枝末节、集中处理那些我们关心的主题是合理的，这让我们得以不被那些非本质的细节束手缚脚而寸步难行。

哈佛大学心理系教授斯蒂芬·科斯林（Stephen Kosslyn）作的一个演讲提醒了我，科学家以及其他是如何接受信息的。在一个他给听众做的认知科学实验中，他让听众浏览一些逐条罗列在屏幕上的线段。每一条线段都有方向，如“北”或“东南”，诸如此类，它们整体构成了一条折线（见图1-2）。他要求听众闭上眼睛，然后描述刚才看到了什么。

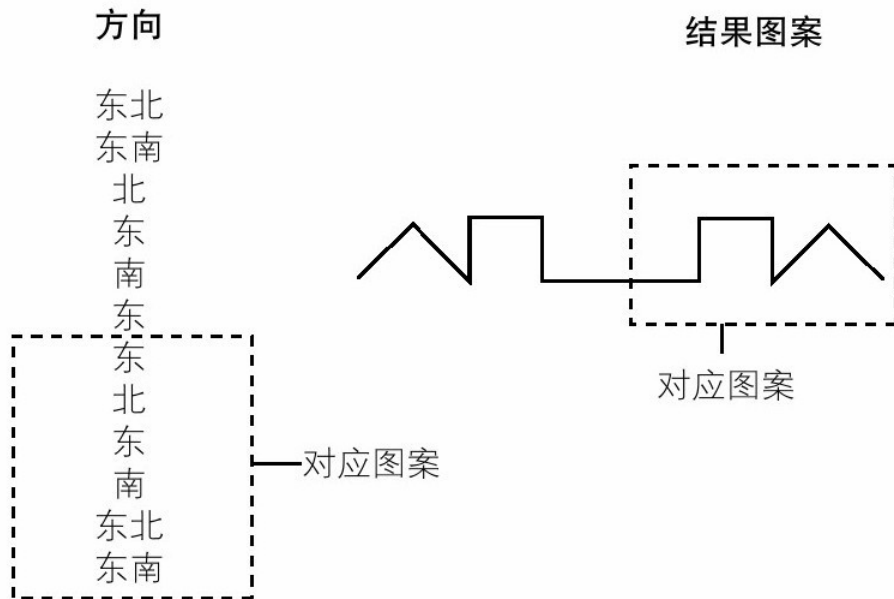


图1-2 你可以把这张图看作零散线段的集合来记忆，或是看作一个整体来记忆，比如出现两次的6条线段的组合。

我们意识到，虽然大脑在某个时刻可以记住某些线段的形态，但要长期记住它们的排列顺序，还是要把它们复原为一个整体的形状来记忆。只有当考虑那个形状的整体而非组成它的那些独立线条时，我们才能将这个图形铭记于心。

对世间任何你所见、所听、所闻、所尝、所感的事物，你都有两个选择，一是近观其细节，二是远观其重要的“大图景”。不论是在赏画、品酒、阅读哲学文献时，还是在为下次出游做打算时，你都会在潜意识里按照自己的兴趣把你的想法分类——这些类别可能是大小、味道、理念、距离，以及那些当时你说不清的元素。

这种抓住核心、忽略次要问题的方法已经应用于各种场合中。

想一想，当使用谷歌地图、MapQuest或盯着iPhone那小小屏幕时，你是怎么做的吧。如果你要远行，那么你必须对目的地有一个大概的了解。当你有了大图景的时候，你会把地图放大，提高分辨率以寻求更多细节。

最开始时，你并不需要那些额外的、有关细节的信息，你只需要一些关于地点的大致信息。当你开始在地图上标出旅行细节时，就会开始关注那些更小尺度上的细节，比如你会调高分辨率以查找你需要的某条具体的街道——而这些细节在你刚开始试图确定目的地的大致位置时并不重要。

当然，你需要的精确度决定了你选择的尺度。我有一些朋友，当他们到纽约旅行的时候并不在意旅馆的具体位置。于他们而言，“街区”这个尺度等级并不重要。但对于熟悉纽约的人来说，这些细节就变得很重要，他们并不满足于“我在市中心”这种描述，而是期待更多细节。纽约本地人在意他们是在休斯敦街（曼哈顿上下城的区分界限）之下还是之上，在意他们是在华盛顿广场公园的东边还是西边，甚至还在意他们与这些地方相距两个街区还是五个街区。

尽管对具体尺度的选择因人而异，但没有人会在美国地图上寻找某家餐馆的位置，他们需要的这些细节即便是在最大的电脑屏幕上都显示不出来。另外一方面，如果仅仅是为了确定某个餐馆的位置，那你也压根不需要那些建筑平面图的细节。对于特定的问题，你会选择与之对应的尺度（见图1-3所示的例子）。



图1-3 从不同尺度上会看到不同的信息细节。

类似地，按照尺度把物理规律分类，这样我们就可以集中处理那些我们感兴趣的问题了。我们的桌面看上去是铁板一块，在很多情况下我们这么看是没有问题的。但是，实际上它是由许多原子和分子构成的。这些粒子共同作用，并以我们所熟悉的尺度表现出来——铁板一块。但是，这些原子也并非不可分。它们由原子核与电子构成，原子核由质子和中子组成，而质子和中子又由处于束缚态的更基础的组分——夸克构成。然而，在理解物质的电磁与化学性质的时候，我们并不需要对夸克有所了解（这个科学领域叫作“原子物理学”）。人们在弄清楚原子之下的物质结构之前已经研究了多年原子物理学——正如生物学研究某个细胞时并不需要了解质子中的夸克一样。



我现在还记得在上高中时，老师在讲授了几个月有关牛顿定律的内容之后，向整个班级宣布那些定律都是错误的——那一刻，我有一种被背叛的感觉。然而，我的老师并没有正确地表述他的观点。牛顿定律只适用于在牛顿时代可被观测到的尺度和速度。牛顿只考虑那些在他（以及同时代的人）当时可以测量到的精度上可应用的物理定律。

在当时的测量条件下，他给出的预测已经非常成功，而并不需要广义相对论中所描述的那些细节。即使在今天，我们在大尺度、低速度、相对低物质密度的情形下作出预测时，牛顿定律依旧适用。当物理学家与工程师们试图研究行星轨道时，他们并不需要了解太阳的详细信息。支配夸克运行的规律，对预测天体的运行并不会产生显著的影响。

试图理解物质最基础的组分对于理解大尺度上的相互作用并非行之有效的方式，因为这些小结构在大尺度上并不起什么作用。如果过分纠缠于更小的夸克的性质，那么我们在原子物理学上将寸步难行。只有当我们需要了解原子核更多细节的时候，夸克的结构才变得重要起来。在足够合理的近似条件下，我们可以放心地使用化学和分子生物学的结论，而不必考虑原子核的内部结构。伊丽莎白·斯特布的舞蹈动作不会受到量子引力尺度上规律的影响。编舞艺术只受经典物理学定律的支配。

任何包括物理学家在内的人在不需要那些高分辨率细节的时候，都期望一个可以更简洁明了地描述世界的理论。物理学家把这种直观感觉正式化，并根据物理规律适用的尺度和距离把它们分类。对特定的问题，我们使用特定的有效理论（effective theory）。“有效理论”集中处理那些在给定问题的特定尺度上“有效”的粒子和力。我们如此构建理论、方程与观测方式，使它们与我们有可能探测到的尺度相关，而非根据描述更多基础行为的不可测量的参数来描述粒子与作用力。

有效理论

物理学家们把那种简洁明了的直观感觉正式化，并根据物理规律适用的尺度和距离把它们分类。“有效理论”集中处理那些在给定问题的特定尺度上“有效”的粒子和力。

在更大尺度上应用的有效理论并不会详述隐含在其下、支配小尺度现象物理规律的细节，它只关心那些你可以看到或测量到的事物。如果某些事物对于你的研究所涉及的尺度而言太过微小，那你就需要了解它更精细的结构。这种方法并非科学上的自欺欺人，而是排除冗杂信息的一种方法，是一种获得正确答案、了解你工作系统中有什么的“有效”方式。

有效理论之所以“有效”，是因为它把那些不产生任何可测量效应的未知因素有效地忽略掉了。如果那些未知因素只在影响不可觉察的尺度、距离以及分辨率上出现，那么我们并不需要它们来进行成功的预测。当前技术所不能及的现象，并不会在那些已经考虑进理论的因素之外，对测量结果有任何影响。

这就是为什么即使缺少对相对论运动定律以及量子力学对原子与亚原子系统描述的了解，我们一样可以作出足够精确的预测。这真乃幸事，我们显然不可能考虑所有可能存在的细节。如果不忽略无关细节，那么我们将寸步难行。当集中处理那些可被实验验证的问题时，我们所限定的分辨率会使那些其他尺度上的纷乱信息变得不重要。

“不可能的”事情只会在我们没有观测到的环境下发生。那些结果与我们所知的，或至少是已经探测到的尺度无关。那些小尺度上的信息会一直隐藏起来，直到我们发明出分辨率更高、能直接观测到它们的仪器，或者能通过充分精确的测量，从那些表现在较大尺度上的小特征中来区别与鉴定出底层的理论。



科学家在作出预测时，可以合理地忽略那些太小而无法观察的事物。不仅区别那些过于微小的事物与物理过程是不可能的，而且只有当那些尺度上的物理过程的效应能够影响可测量的参数时，才能引起我们的兴趣。因此，物理学家们在有效理论中仅仅描述那些可测量尺度上的事物及其特性，并且只在这些理论所涉及的尺度上用它们处理问题。当你知道了小尺度的细节或某个理论的微观结构时，你一样可以从更加基础、精细的结构来导出那些有效描述。否则，这些细节仅仅是有待被实验验证的未知事物。这些有效理论中较大尺度上的可观测现象并不能给出更为基础的描述，但对于理论预测与实验观测而言，应用它们是非常方便的。

一个有效的描述可以对以小尺度理论再现大尺度观测的结果进行总结，然而小尺度的效应实在是太过微小而无法被观察到。这样做的好处在于，我们可以使用尽量少的参数学习、评估这些物理过程。如果不这么做，我们就会被迫考虑过多的细节。这个更小的参数集合对描述我们感兴趣的物理过程而言十分有效。此外，我们使用的这个参数集合是“普适的”——描述它们并不依赖于蕴含在这些物理过程之下更加细微的细节。为了获得它们的具体数值，我们仅仅需要在那些它们被应用的物理过程中去测量它们。

单一的有效理论可以应用于范围宽广的尺度和能标上。当那些为数不多的参数都已经被测定之后，在那个尺度上的任何结论我们都可

以通过计算导出。它给出了一个若干规则和元素的集合，可以解释大量的观测现象。任何时候我们认为是基础理论的理论，都有可能转变为有效理论——不存在无限精确的分辨率，探索永无止境。然而我们必须相信那些有效理论，因为它们在其所应用的能标和尺度范围内，可以成功地预言一系列现象。

物理学中的有效理论不仅保留了小尺度上的有效信息，还可以总结出大尺度上的效应，而其结果也许太过微小而不能被观察到。比如，爱因斯坦在发展他的引力理论时提出，我们所处的这个宇宙有可能是轻微“弯曲”着的。在大尺度上的这种弯曲涉及大尺度的空间结构。然而，这些微小的弯曲效应并不影响我们在较小尺度上进行的局域性观测和实验。只有当我们把引力纳入粒子物理学的框架中时，才需要考虑这些效应的影响，而这些影响对于许多我后文要讲的事情而言，都可以忽略不计。在这种情况下，适当的有效理论让我们得以在缺失一些需要实验测定的参数的情况下，依旧可以总结出引力的效应。

有效理论最重要的方面是，它在描述我们可以看到的尺度的同时，也把那些我们忽略掉的尺度（无论大小）做了归类。应用有效理论，我们可以在任何特定的测量中确定动力学的效应（无论已知或未知）有多显著。即使没有那些在不同尺度上的新发现，我们也可以通过数学计算来确定在我们所关心的尺度上，那些新结构对有效理论的最大影响。正如第12章所讲，只有当那些底层的物理学现象被揭示出来时，大家才会真正理解有效理论的局限性之所在。

关于有效理论的一个为人所熟知的例子是热力学，它在原子理论与量子理论出现之前就成功地解释了冰箱或引擎是如何工作的。压强、温度和体积这三个参数完美地描述了一个热力学系统的运行状态。即便我们知道这个系统又包含着由原子和分子组成的气体，而这

比前面的三个要素所能描述的结构要复杂许多。然而，为了一些特定目的，我们依然集中关注这三个参数，这样可以便捷地描述这个系统的可观测行为。

温度、压强和体积都是可以被测量的实在量。它们之间的关系背后所依附的理论已经发展得相当成熟，并且可以用以作出成功的预测。气体的有效理论并未提及底层的分子结构（见图1-4）。底层的分子元素决定了温度和压强，但不了解原子或分子并不妨碍科学家们放心地使用这些量进行计算。

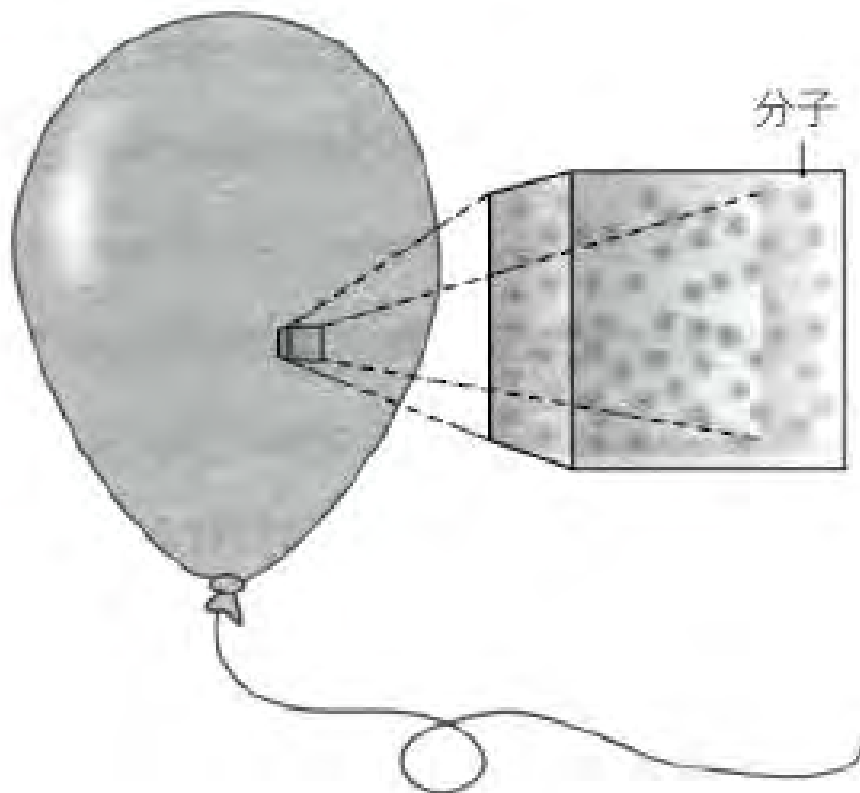


图1-4 压强和温度可以在更基础的等级上加以理解，这与单个分子的物理性质相关。

一旦理解了基础理论，我们就可以把温度和压强归为底层原子的性质，以及可以理解在什么情况下热力学的描述会失效。但在很多情

形下，我们依旧可以使用热力学来做预测。实际上，很多现象仅仅从热力学的角度来考虑，是因为我们没有足够的计算能力和存储能力，即使知道它们的存在，我们也无法追寻每一个原子的运动轨迹。有效理论是唯一的方法，它有助于理解一些有关固体或液体的凝聚态物质（condensed matter）的重要物理现象。

这个例子揭示了有效理论的另外一些重要方面。我们有时把“基础”（fundamental）当作一个相对术语。从热力学的角度来看，原子和分子的描述都是基础性的。但从粒子物理学的角度来看，它描述了原子中夸克与电子的细节，原子是组合物，由更小的元素组成。从粒子物理学的角度来看，它就是一个有效理论。

这种从已被充分掌握的知识到前沿知识对科学发展进程的彻底描述，在诸如物理学与宇宙学领域中被应用得最好，因为我们对那些功能单元及其联系有着很明晰的理解。有效理论并不一定能在新领域，譬如系统生物学中生效，因为在分子与更大层级上行为之间的联系，以及那些有关的反馈机制还有待人们去了解。

虽然如此，有效理论的理念依然在很多科学背景中得到了广泛应用。支配生物进化的那些数学方程式并不会因为新的物理结果而变化，我在与数学生物学家马丁·诺瓦克（Martin Nowak）^[10]的一次讨论中如此回答他所提出的问题。他与同事们可以在不使用更加基础描述的前提下塑造那些参数。他们也许最终与更基础的量相关——物理或者其他的量，但这并不会改变那些数学生物学家用以描述种群行为随时间演化的方程式。

对粒子物理学家来说，有效理论是必要的。我们把简单系统按照尺度分类，然后再把它们彼此联系在一起。事实上，那些尚不明了的底层结构让我们专注于可观测的尺度，并忽略更多基础效应。它们的底层相互作用隐藏得如此好，以至于我们必须付出大量的人力、物力

才能把它们搜寻出来。这些可观测尺度上更基础理论效应的微不可查，正是今日物理学家们面临的挑战。如果希望察觉到那些更加基础的物质及其相互作用所产生的效应，那么我们就需要直接探索更小的尺度，或是测量得越发精确。只有通过更高级的技术，我们才能探索那些极小或极大尺度上的事物。这是我们要设计详细实验，以期取得一些进展的原因——正如大型强子对撞机所做的一样。

光学，有效理论的典范

光学理论的历史正是一个有效理论在科学发展进程中应用的典范。在这个过程中，某些观念被抛弃，而另一些被保留，并最终作为良好的近似而应用于特定的领域之内。从古希腊人的时代起，人们就开始研究几何光学。这是有志于物理学研究的学生们在物理学GRE（即研究生入学考试）中的必考科目。这个理论假定光沿射线或直线传播，进而描述在穿过不同介质的界面时光线的行为，以及如何使用仪器去检测它们。

奇怪的是，事实上并没有人，至少在我过去曾就读以及现在所任教的哈佛大学，都没有人学习经典几何光学。也许高中课程中会涉及一点儿几何光学的内容，然而于全部课程而言，其课程量如沧海一粟。

几何光学是一个过时的学科。它于数百年前因牛顿的《光学》（*Opticks*）一书盛极一时，并延续到19世纪的前10年，那时威廉·哈密顿（William Rowan Hamilton）也许是第一个对新现象作出了数学预测。

今日，经典光学理论依然在摄影、医学、工程学、天文学等领域中应用，它依然作为制造镜子、望远镜以及显微镜的理论依据。研究

经典光学的学者和工程师们解决了各种物理现象的不同问题。然而，他们只是简单地应用已有的光学理论，而没有发展新的物理定律。

2009年，我有幸被邀请在都柏林大学做了一场哈密顿演讲——我很多令人尊敬的同事在我之前都曾受邀，这个演讲以哈密顿的名字命名。哈密顿是19世纪爱尔兰著名的数学家、物理学家。我承认“哈密顿”这个名字在物理学中实在太常见了，以至于我一开始很可笑地没有把这个名字与那个真实的爱尔兰人联系在一起。但是，我对哈密顿为数学与物理学作出的很多卓越贡献都耳熟能详，这其中就包括几何光学。

庆祝“哈密顿日”确实是一件大事。这天的活动包括沿都柏林的皇家运河而下的游行。游行止于金雀花桥，以观看社团中最年轻的成员在桥上写下一些方程式，这些方程式是当年哈密顿在途经该桥、思考自己的理论时灵光乍现而随手写下的推演。我访问了顿辛克天文台（College Observatory of Dunsink，哈密顿也曾工作、生活于此），并观看了一个300年前的拥有滑轮的木质结构望远镜。哈密顿于1827年从三一学院毕业之后就在那里任职，成为教授、爱尔兰皇家天文学家。当地人如此腹诽：尽管哈密顿有着惊人的数学天赋，而他对天文学毫无兴趣且一无所知。虽然他在理论上贡献卓著，但是他也许把爱尔兰观测天文学的发展进程拖慢了50年。

虽然如此，人们依旧设立了“哈密顿日”以表达对这位伟大的理论家卓越贡献的敬意。它们包括光学与动力学、四元数（quaternions，一种对复数的推广）的数学理论，还有其他哈密顿对数学和科学敏锐的预测与坚实的推论。四元数的发明是一件大事。它对向量微积分意义重大，是我们研究三维现象所使用数学方法的基础理论；它也在计算机图形学中得到广泛应用，这直接促进了娱乐产业以及电视游戏的兴盛。任何拥有PlayStation或是Xbox的人都应该感谢哈密顿对他们娱乐产品的贡献。

在哈密顿诸多重大贡献中，最为耀眼的是他对光学理论的发展。1832年哈密顿发现，当光线以特定角度射入到有两条独立轴的晶体上时，将会被折射并形成一中空的出射光锥。因此他预测了关于光线穿过晶体时的内锥折射与外锥折射（internal and external conical refraction）理论。这是一个重大的，也许还是第一个数学科学的伟大胜利，该预测最终由哈密顿的朋友、同事汉弗莱·劳埃德（Humphrey Lloyd）证实。想要证实一个完全由数学预测、从未见过的现象实乃大事，哈密顿也因这一发现而被封为爵士。

当我访问都柏林时，当地人骄傲地向我讲述了哈密顿对几何光学基础卓越的数学贡献。伽利略是观察科学与实验的先驱者，而弗兰西斯·培根则是归纳科学的最早一批拥趸之一。他坚信，人们做预测的出发点建立在先前经验的基础上。然而就使用数学描述一个未知现象而言，哈密顿对锥形折射的预测可谓前无古人。出于这个原因，哈密顿在科学发展上的贡献注定名留青史。

虽然几何光学在今日已不是一个研究对象，但哈密顿的发现依旧重要。所有重要现象都在很久之前被人们理解了。在哈密顿时代之后不久的19世纪60年代，苏格兰科学家詹姆斯·麦克斯韦（James Clerk Maxwell）等人发展了光学的电磁场理论。虽然几何光学是一个近似理论，但是在光的波长极小以至于其波动效应可以被忽略的情况下，把光描述为以直线传播的光线是非常好的。换言之，几何光学是一个有效理论，它的有效性被限制在一定范围内。

这并不意味着，我们要保留历史上曾提出过的一切理论，有一些观点最终被证明是错误的。欧几里得最初对光的描述复活了9世纪伊斯兰科学家艾尔-金迪（Al-Kindi）的理论，他认为光是由人的眼睛发射的。虽然古时候的一些科学家，比如波斯数学家伊本·沙（Ibn Sahl），出于错误的前提却正确地描述了一些光现象，诸如光的折射，但有着近代数理知识的欧几里得与艾尔-金迪却得出了完全错误的结论。这些理论并没有被吸纳进现代理论，而是直接被摒弃了。

牛顿没有预见到光理论的另一面。他提出了光本质的“微粒说”（“corpuscular” theory），与罗伯特·胡克（Robert Hooke）于1664年、克里斯蒂安·惠更斯（Christian Huygens）于1690年分别提出的光本质的波动说（wave theory）势不两立、水火不容，其间的争论也可谓旷日持久。19世纪，托马斯·杨（Thomas Young）与奥古斯丁·菲涅尔（Augustin-Jean Fresnel）测量了光的干涉现象，明确地证实了光的波动性质。

后来发展起来的量子理论证明，牛顿在某些方面也是正确的。现在，量子力学告诉我们，光是由叫作光子（photons）的单个粒子构成的，它们传播着电磁相互作用。近代的光子理论建立在光量子的基础上，它们是独立的粒子，共同构成光，有着特殊的性质。即使它们是构成光的单独粒子，光子的行为依旧拥有波动性。这种波动性表现为：一个自由光子在空间区域中的任何一点都可以以一个特定的概率存在（见图1-5）。

牛顿的微粒说从光学理论导出结论，但这一理论并不包含任何光的波动本质，这一点与光子对光行为的描述不同。现在我们知道，光子理论是对光行为最基础、最正确的描述，它同时涵盖了光的粒子与波这两方面的特性。量子力学给了我们当前对光的本质及其行为最为基础的描述。它在基础上是正确的，所以得以保留至今。



图1-5 几何光学与波动光学是我们对光的现代理解的两种先驱理论，而且至今依然在某些条件下被应用。

量子力学比光学更接近理论前沿。如果人们依旧从光学理论出发去思考新的科学，那他们首先要考虑那些只能通过量子力学说明的新现象。即便现代科学已经不再发展经典光学的理论了，然而它确实包含了量子光学的领域，主要研究光的量子力学性质。激光、包括光电倍增管在内的光学探测器以及把太阳能转化为电能的光电池，都遵循着量子力学的规则。

近代粒子物理学也围绕量子电动力学（QED）的理论叙述，由理查德·费曼（Richard Feynman）等人提出，它把狭义相对论与量子力学进行了结合。应用量子电动力学，我们得以研究单个粒子，包括光子，即光的粒子；还有电子以及其他带有电荷的粒子。我们可以得到这些粒子之间产生与湮灭相互作用的速率。量子电动力学是粒子物理学中常用的理论之一，它在所有科学理论中，给出过有史以来最为精确的预测。量子电动力学与几何光学几乎毫不相干，然而它们在各自适当的有效领域内都是正确的。

每一个物理学领域都使用其对应的有效理论。把旧理念整合入更基础的理论，科学就是如此演化的。旧理念依旧在很多实际场合下被应用，然而它们并不在科学的前沿领域内。虽然本章末只是集中关

注了一些若干年来对光的物理解释的特例，但是整个物理学领域都是以这种方式发展的。虽然科学在其发展的前沿存在不确定性，但是这个进程本身却是有条不紊的。特定尺度上的有效理论合理地忽略了一些我们可以确信对特定测量而言没有任何影响的效应，过去科学探索历程中去芜存菁的过程赋予了我们这种智慧和方法。然而，理论会随着我们对更广的能标和尺度范围的深入了解而不断演化，这些进步让我们得以从新的视角来研究那些表观现象之下的基础解释。

理解这个进程可以让我们更好地理解科学的本质，以及鉴赏一些物理学家（与其他领域的科学家）提出的主要问题之美。在第2章中，我们将看到，从很多方面来讲，我们今天所使用的方法论都是从17世纪开始发展起来的。

书籍免费分享微信jnztxy



今天，科学家们使用的方法基于的是过去很长时间以来各种观测研究的成果，这些研究成果可以证实或证伪不同的科学理论——后者与前者同等重要。这需要我们超越对世界的直觉理解。罗马语中用来表达“think”（思考）这一动作的词根“pensum”从拉丁语的动词“weigh”（权衡）演变而来。看来英语使用者也会“权衡”不同的观点。

很多把科学引入现代表述方法的构词法都在17世纪于意大利发展起来，而伽利略在其中扮演了关键角色。他是首先完全领会并推动间接测量（indirect measurements）发展的一批人之一。间接测量使用中间仪器进行测量，以设计以及用实验创建科学理论。此外，伽利略还构想了一系列思想实验，以帮助他建立并逐渐发展自己的理念。

2009年春天访问帕多瓦大学时，我接触了很多伽利略的深刻思想，这些思想从根本上改变了当时的科学观。我访问那里的第一个原因是为参加一场物理学会议，会议的组织者是帕多瓦大学物理学教授法比奥·茨维纳（Fabio Zwirner）；第二个原因是去接受该市“荣誉市民”的称号。我很高兴能和与会的物理学家同侪以及那些受人敬重的“市民”坐在一起，这其中包括史蒂文·温伯格（Steven Weinber

g)、史蒂芬·霍金、爱德华·威滕 (Edward Witten)。另外，作为福利，我还有一次学习科学发展史的机会。

那次旅行是在2009年，正好是伽利略进行首次天文观测的400周年。帕多瓦城的居民们对此尤为留心，因为伽利略在作出最伟大研究时恰在帕多瓦大学任教。为了纪念他，帕多瓦城（以及比萨、佛罗伦萨、威尼斯这些在伽利略科学生涯中扮演了重要角色的城市）安排了一系列展览与庆典活动，以表达对这位科学家的敬意。物理学家的会议于阿迪纳铁文化中心 (Centro Culturale Altinate, 也称圣加埃塔诺酒店) 的大厅里举行。这座建筑还为一个引人入胜的展览提供了展位，这个展览用来庆祝伽利略对科学不可磨灭的贡献，彰显他在改变以及重新定义现代科学中不可动摇的地位。

我遇到的大多数人都表达了他们对伽利略成就的敬仰以及对现代科学进展的狂热之情。帕多瓦市的市长弗拉维奥·扎诺那多 (Flavio Zanonato) 对物理学的兴趣及其深厚的物理学知识大大鼓舞了当地的物理学家。作为这座城市的领导者，他不仅出席了我在一次晚宴上作的公开演讲，并参与讨论了一些科学主题，而且在演讲过程中还以一个对大型强子对撞机中载荷流的精彩提问震惊了听众。

作为荣誉市民庆典的一部分，市长扎诺那多亲自授予了我帕多瓦市的荣誉之钥。这个钥匙十分古怪——它看起来像是只存在于电影之中的事物。它很大，通体银质，雕刻精美。我的一位同事看到它时甚至打趣说：“它是从《哈利·波特》的故事中带出来的吗？”它只是一件礼品，并没有开锁的功能，象征着“开启”——不仅意指城市，在我的想象中还指向一个丰饶而有质感的知识之门。

除了这把钥匙之外，帕多瓦大学教授马西米拉·巴尔多-秋林 (Massimilla Baldo-Ceolin) 还授予了我威尼斯纪念奖（也称奥赛拉奖 [osella]）。奖牌上面用意大利文铭刻了伽利略的名言，这句名言同样被刻在帕多瓦大学物理系的墙上：

我认为哪怕脚踏实地地找出物质的一点真相，也比高谈阔论那些于现实毫无意义的空想更有价值。

我和与会的同侪们分享了这句名言，因为这句话确实就像一个指导原则。**创造性的进展总是始于简单问题。**并非我们回答的每一个问题都具有明确的意义，然而即便看起来十分缓慢的进步，也时而会给我们的理解带来翻天覆地的变化。

本章将要介绍，现代测量及技术如何萌芽于17世纪时发展起来的一系列事物，以及那时发展起来的基础理念是如何帮助我们理解今日的理论以及所使用的实验方法的。过去的400年间，一些大问题会反复被科学家们提及，而随着理论与技术的飞速发展，一些小问题反而发生了巨大的变化与发展。

科学时代的来临

科学家格物穷理的过程正是他们对跨过已知领域与未知领域界线的一种尝试。任何时候，我们都是从一个包含了能预测目前可测量现象的规则与方程的集合出发的，但我们又总是试图进入当下实验尚不能触及的领域之中。应用数学理论与技术，我们可以系统地处理一些问题，而这些问题在过去纯粹是出于信仰或者投机而提出的主题。随着更多、更好的观测结果以及被崭新测量方式改良过的理论框架的建立，科学家们得以更加全面地理解这个世界。

当我流连于帕多瓦以及其他那些伽利略留下过足迹的邦城之间时，我更好地理解了他在发展上述思维方式中不可替代的地位。斯科罗维尼礼拜堂（Scrovegni Chapel）——就是从14世纪初期开始陈列“欧洲绘画之父”乔托（Giotto di Bondone）壁画的地方，正是这些地点中最著名的一个。这些画值得被人们注意的理由有很多，但其中最让科学家们关心的是，《博士来拜》（*Adoration of the Magi*）

一画活灵活现地描绘了1301年被人们观测到的哈雷彗星，不得不说这是个奇迹（见图2-1）。在那幅画的创作时期，哈雷彗星是肉眼可见的！



图2-1 乔托在《博士来拜》这幅画中描绘了如下情景：14世纪早期，肉眼可见的哈雷彗星划过天际。

不过，那些画面并不科学。我的导游指着拉久内宫（Palazzo della Ragione）中一幅关于星际景象的画告诉我们，她最初被告知那是银河系，而后来一位经验丰富的导游向她解释，这是由于时代局限而导致的错误理解。在这些作品的创作时期，画家们仅仅会画出他们之所见。那确实是一幅布满了星星的画面，然而与我们的星系却没有半点关系。我们今日所了解的科学在那个时代还没有诞生。

实验观测与思想实验

在伽利略之前，科学依赖于未经仔细考量的测量结果与纯粹的思辨。亚里士多德所发展的“科学”是人们试图理解世界的一种典型方式。数学可以进行逻辑推导，然而那些隐含其中的公理要么被人们无条件地相信，要么由直接的观测得来。

伽利略明确地摒弃了把研究建立在“论文的世界”（*mondo di carta*）里的想法，他希望研读并学习“自然之书”（*libro della natura*）。为此，伽利略改进了自己的观测方法，并认识到了实验的力量。他明白了如何构建及使用某些人工环境进而推导出物理定律的本质。通过实验手段，伽利略得以证实以及证伪他能（以数学手段）证明的一系列有关自然本质的假说。而证实与证伪这两个过程同等重要。

伽利略的一些实验使用了斜面，这些斜面非常重要，它们会经常出现在所有基础物理学考试之中。对伽利略来说，斜面并不只是一些教师炮制的课堂问题——正如物理学初学者所理解的一样。事实上，它是一种研究自由落体速度的方法，这种方法延长了物体运动的水平距离，所以伽利略可以精细地测量它们是如何“下落”的。他使用水時計来计量时间，还聪明地在一些特殊时间点加入小铃铛，这使他得以利用自己的音乐天赋来判断球从斜面上滚下来时的速度（见图2-2）。通过这些有关运动和引力的实验，伽利略、开普勒与笛卡儿共同为经典机械定律打下了坚实的基础，而这些定律在后世由牛顿发扬光大。

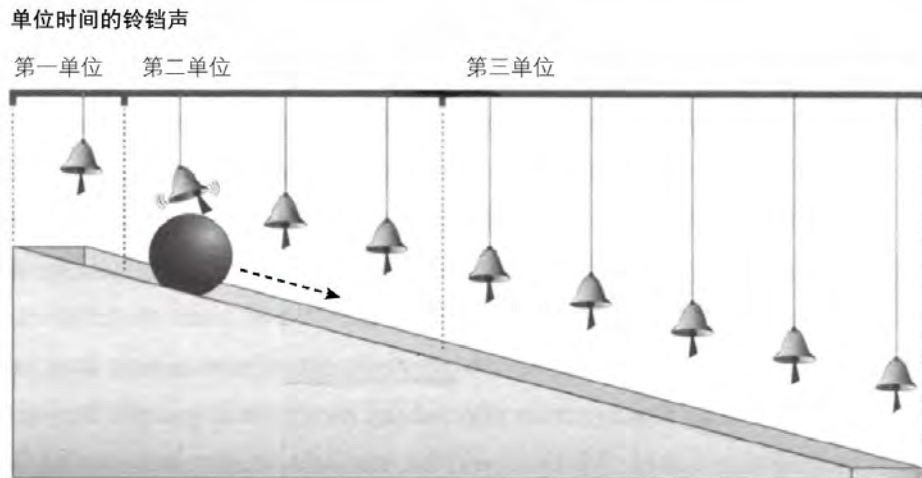


图2-2 以应用铃铛声记录球体运动过程的手段，伽利略测量了球体从斜面顶端下滚的速度。

伽利略所发展的科学也超越了他的观测水平。他发明了思想实验，即为了能预测当时还没有人可以进行的实验所做的基于所见之物的抽象理解。其中最著名的预测是：当物体在无阻力的情形里下落时，它们（在相同时间）具有相同的速度。即便伽利略不能真正在理想情况下做实验，也仍然正确地预测了这个结果。他不但正确理解了引力在物体自由下落时的作用，还明白了是空气阻力使这些物体慢了下来。好的科学理论应该能够正确理解所有因素。思想实验与实验观测共同帮助伽利略更好地理解引力的本质。



一个有趣的历史巧合是，牛顿出生于伽利略逝世的次年（如霍金在一次演讲中所言，对自己的生日正好在牛顿诞辰300年之后这一点，他感到很高兴）。牛顿作为最伟大的物理学家之一，既传承了之前的科学思想，又把它们发扬光大。设计思想实验与真实实验、理解它们

并了解其局限性的传统方法，为今日科学家们所继承，不论他们生于何日。依赖于尖端技术的现代实验更加精细，这种为了证实或证伪基于假说的预测来制造仪器的理念，明确了今日的科学及其研究方法。

将望远镜对向太空

除了实验，即用来测试其假说的人造条件之外，伽利略另外一个最终改变了科学发展命运的巨大贡献，是对技术潜力的理解与推崇。它的出现使人类对宇宙的观测手段变得越发先进。通过实验，我们可以超越纯粹的推理与思辨；而使用新仪器，我们可以对实验结果去芜存菁。

在科学发展早期，人们依靠感官观察结果。他们凭感觉与触觉目测事物，而不是使用间接的中间仪器以获得那些与直觉不同的客观事实。天文学家第谷·布拉赫（Tycho Brahe）在大量天文数据之中发现了一颗超新星，并精确测量了行星的轨道，这是伽利略登上历史舞台之前最后一个著名的天文学观测结果。第谷在测量时确实使用了一些精密的仪器，比如巨大的四分仪、六分仪以及浑天仪。第谷亲自设计、制造了比前人所制精确度更高的仪器，并得出精确的观测数据，而这些数据让他的弟子开普勒得以推断出行星的运行轨道是椭圆形的。然而，尽管第谷的观测十分细致，但是这些观测却都是通过肉眼进行的，根本没有用到透镜或者其他仪器。

值得注意的是，伽利略对艺术有着极高的洞察力——毕竟他是一位音乐家的儿子，家学渊源。但是即便如此，他也意识到了，使用技术手段作为观测的中间步骤会使他本来已经极具天赋的耳目发挥更大作用。伽利略坚信，借助观测仪器在各种尺度上完成的非直接测量取得的结果，远远超越了那些仅靠感官获得的结果。

伽利略对技术最广为人知的贡献，是他应用望远镜观测恒星的创举。他使用这种仪器改变了我们见科学、见宇宙、见自我的方式。

然而，望远镜并非由伽利略发明，而是于1608年由荷兰的汉斯·利伯希（Hans Lippershey）发明。不过这位荷兰人发明它是为了偷窥，因此这个发明在当时又被称作“偷窥镜”。伽利略是第一个意识到这个仪器潜力的人——有了这个仪器，人就可以不再依赖肉眼观测宇宙了。伽利略参照荷兰发明的“偷窥镜”原型，把它升级为了一台放大倍数为20倍的望远镜——这个发明在出现后的一年内一直被当作节日庆典上的玩具，是伽利略把它变成了一台科学仪器。

伽利略使用间接的中间仪器辅助测量的开天辟地之举是对以往测量手段的彻底抛弃。它开近代测量方法之先河，并且代表着所有近代科学必不可少的关键进程。人们最初对这种非直接测量有所怀疑。时至今日，仍然有怀疑者质疑大型质子对撞机中的现象，以及卫星或望远镜上计算机中所记录的数据的真实性。但是这些被仪器所记载的数据中，每一比特数据都与我们能直接看到的東西一样精确，从某些方面上来讲，甚至比直接所见的事物更加精确。毕竟，我们的听觉源于作用在耳鼓上的空气的振动，视觉源于作用在视网膜上的电磁波，这些作用最终被大脑加工而形成感觉。这意味着人体也是一种技术的呈现方式——而人体并不可靠，相信所有体验过光学错觉的人都会支持这一观点（例子见图2-3）。科学测量的美妙之处在于，我们可以明白无误地推断出物理实在性的方方面面，包括基本粒子的本质及其特征。这些目的要通过一些实验来达到，例如今日的物理学家们使用大型而精确的探测器所进行的实验。

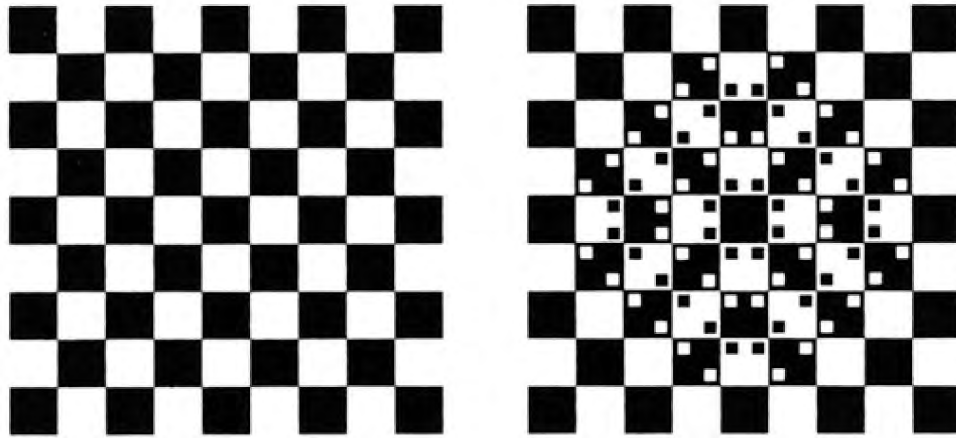


图2-3 对于确定外在的实在性，人的视觉并不总是最可靠的途径。这里的两个棋盘完全相同，然而右边棋盘上的许多小点让这些棋盘的格子看起来非常不同。

怀疑抽象结论、相信眼见为实是出于人的本能，而科学给我们上了一课，告诉我们如何超越这种人类与生俱来的倾向。使用自己设计的仪器进行测量要比肉眼观测更值得信赖，而且我们可以通过多次重复实验的方法来核实、改进它。

1611年，教会通过了关于承认间接测量有效性的决议。如汤姆·列文森（Tom Levenson）在其著作《以尺量尺》（*Measure for Measure*）^[1]中所提到的，教会的科学事务当权者们必须确定望远镜的观测结果是否值得信赖。红衣主教罗伯特·贝拉明（Robert Bellarmine）推动教会学者们来决议这个话题，1611年3月24日，教会的4名首席数学家一致认为伽利略的发现是有效的——望远镜的观测结果精确且值得信赖。

帕多瓦大学赠送给我的另一枚黄铜纪念章之上，精美地镌刻着伽利略核心思想的本质。其中一面是介绍1609年，伽利略向当时的威尼斯政府机构以及总督李奥纳多·多纳（Leonardo Dona）介绍望远镜的画面；另一面是一段致辞，大意是：这个事件“标志着近代天文望远镜的诞生”，开启了“人类向地球之外星空探索的革命性事业”，

是“跨越天文学时代的一个历史性时刻，使它成为近代科学的起点”。

伽利略对观测方式的贡献直接导致了后续一系列爆炸性的科学进展。在他不断仰望星空的过程中，伽利略发现了肉眼所不能及的一些新事物。比如，他发现了昴星团（Pleiades）以及茫茫星空中一系列人们从未见过的恒星，它们散布在我们已知的亮度更高的恒星之间。1610年，他在《星辰信使》（*Starry Messenger*）中发表了观测结果，这本书于6个星期之内匆匆写就。在印刷工人处理他手稿的时候，伽利略匆匆完成了研究工作，并且希望作为首位出版有关望远镜书籍的人，得到时为意大利托斯卡纳区（Tuscany）大公科西莫二世·德·美第奇（Cosimo II de' Medici）的关注与支持——这位大公是当时意大利最富有的家族中的一员。

伽利略那见解深刻的观测手段，促使一系列科学成就如雨后春笋般地涌现出来。他采用不同的看待问题的方式，即关注“如何做”而非“为何”。那些只有通过望远镜才能发现的观测细节，让他得出了一些激怒罗马教廷的推论。这些结果使他确信哥白尼的理论是正确的。于伽利略而言，唯一能支持他所有观测结果的世界图景必须建立在如下宇宙学假设之上：太阳而非地球才是我们所处星系的中心，而所有已知行星都围绕着太阳公转。

对木星卫星的观测在这些观测中是最重要的。伽利略观测到，这些卫星时隐时现，在其轨道上围绕木星这颗巨行星公转。在获得这个发现之前，地球固定不动似乎是唯一能解释为什么月球的公转轨道恒定不变的假说。对木星卫星观测的发现意味着，即便木星（相对地球）是运动着的，也并不妨碍它的卫星绕之公转。这便为地球可能处于运动状态，甚至以一个独立的天体为中心在公转，提供了可能性。

这些可能现象最终被牛顿的万有引力理论及其对天体之间引力的预测所解释。

为了表达对科西莫二世·德·美第奇的敬意，伽利略给木星的卫星命名为“美第奇星”，而前者在不久之后进一步阐述了他对资助的理解——这也是近代科学的一个关键方面。美第奇家族确实决定了支持伽利略的研究。然而，在伽利略被佛罗伦萨城授予终生资助对象不久之后，那些卫星被重新命名为“伽利略星”，以表达对其发现者伽利略的敬意。



伽利略还使用望远镜观测月球表面的山谷与沟壑。在他之前，人们普遍认为，宇宙作为天堂的象征，是完全不变、被绝对稳定的规律所驾驭的。在此之前，源于亚里士多德的盛行观念认为，虽然从地球到月球之间的一切都是有缺陷而易变的，但出于其神性，地球之外的天体是永恒不变的球体。彗星与流星被视作像风和云一样的气象现象，术语“meteorology”（气象学）就起源于此。伽利略详尽的观测结果暗示了，不完美性不仅仅在人类以及月球之间的领域存在着。月球并非一个完美的光滑球体，实际上它与地球的形状十分相似，这超乎人们的想象。由于对月球的这些粗略探索，关于月球到底属于人间独有还是天体的疑问顿时处在风口浪尖之上。地球不再具有独一无二的地

位，它只是一个与其他天体没什么两样的普通天体。

艺术史学家约瑟夫·柯纳（Joseph Koerner）向我解释，伽利略的艺术背景使他能够根据光线和阴影来分辨月球上的环形山。伽利略受过的有关透视画法的训练，使他得以理解他所看到的那些突起物。即使它们并非完整的三维图景，伽利略也立即意识到了这些图景背后的意义。他并非意在绘制月表地图，而是想弄清楚月球的结构；而当看到这些图案的时候，他立即就理解了。

支持哥白尼论断第三重要的结论是对金星方位的观测，如图2-4所示。这些观测结果对于断定天体是否围绕太阳运行极为重要。从这个意义而言，地球与其他行星没什么两样——金星显然并不围绕地球公转。

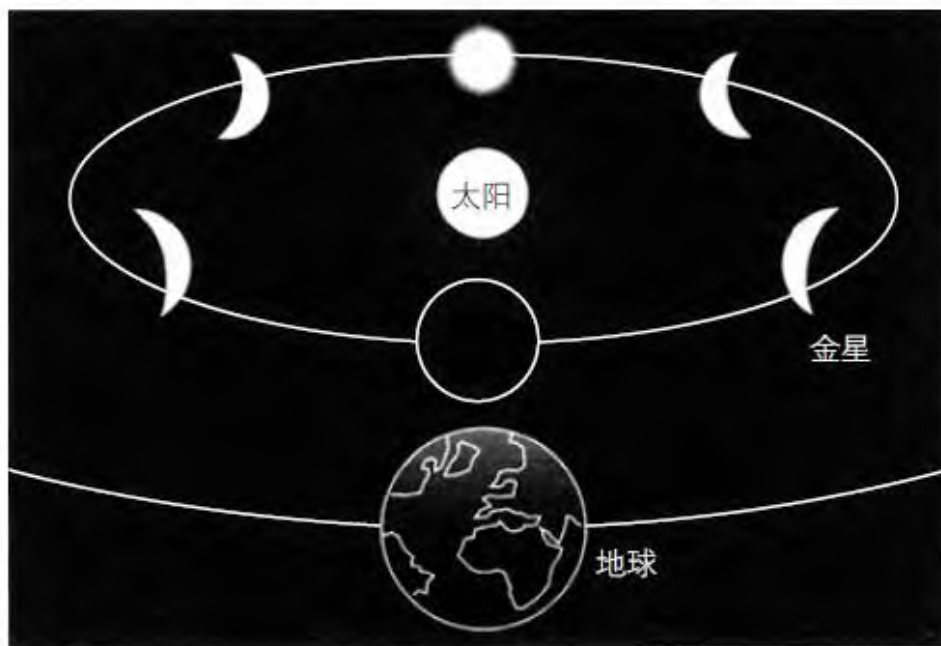


图2-4 伽利略对金星方位的观测证实了它也围绕太阳公转，而这证伪了托勒密体系（Ptolemaic system）。

从天文学的角度来看，地球并不特殊。其他行星与地球一样，像卫星围绕着它们转动一样，各自围绕太阳公转。此外，即便是地球之外的事物也并非完美无缺。地球的“缺憾”显然是人类刻意丑化的结果。伽利略发现，即便是太阳也有“污点”——太阳黑子。

应用这些观测结果，伽利略最终作出了他著名的论断。“地球并非宇宙的中心，地球围绕太阳公转，地球并非宇宙的焦点”——伽利略如此写下这些激进的观点。他用这种方式来对抗教会，即便他后来为了减轻刑罚为监视居住，而宣称自己反对哥白尼主义。



探秘微观世界

伽利略对宇宙在大尺度上的观测与推理的贡献人所共知，而这还不足以描绘他的全部，他也彻底提升了我们对小尺度上事物的认知能力。伽利略意识到，中间仪器可以如其在大尺度上的作为一样，向我们揭示小尺度上的现象，由此他在两个领域内各自拓展了其边疆。在他著名的天文学研究之外，伽利略还继续发展了其技术以探索微观世界。

我在帕多瓦城圣加埃塔诺展览的向导是一位年轻的意大利物理学家米歇尔·多罗（Michele Doro），当他肯定地告诉我是伽利略发明了显微镜时，我感到一丝惊奇。

我得说，在意大利之外，大家普遍认为显微镜由荷兰人发明，争论之处仅在于发明者到底是汉斯·利伯希还是扎卡赖亚斯·詹森（Zacharias Janssen）或其父亲。望远镜到底是不是由伽利略发明的无关

紧要（实际上几乎不可能由他发明），而显微镜确实是由他发明构造，并使用它来观测微小尺度上的事物。有了它，我们自此可以在以前所不能达到的精度上观察昆虫。

从他写给朋友与同行们的信件中，我们了解到伽利略是第一个写下有关显微镜细节及其潜力的人。这个展览展出了记载着应用伽利略的显微镜可以观察到事物的第一本出版物——这要追溯到1630年，弗朗西斯科·斯泰卢蒂（Francesco Stelluti）对蜜蜂进行的详尽研究。

这个展览还指出了，伽利略研究过骨骼，探究了在尺度发生变化的情况下，它们的结构特性会发生什么改变。显然，在他的各种其他洞见之外，伽利略还明确地知道了尺度的重要性。

总之，这个展览确信无疑地展现了伽利略对科学方法与目标的理解——定量研究、可预测性、试图描述确定事物的概念性框架，这些都以精确的规律为指导原则。当这些规则可以提供经过充分验证的对世界的预测时，它们就可以用来预测未知现象。科学家们试图找到能解释所有观测现象的最合理理论。

哥白尼革命恰当地表现了这一点。在伽利略的时代，伟大的观测天文学家第谷·布拉赫出于对太阳系本质的错误理解，作出了不同于伽利略的错误论断。他支持的是一种古怪的、托勒密体系与新体系的混杂学说，即地球是宇宙的中心，但是其他行星按照哥白尼系统的方式围绕太阳公转（图2-5是一个对比）。第谷体系的宇宙学以观测为基础，但它却不能作出最恰当的解释。比起伽利略的观念，它更符合耶稣会信仰者的图景，因为第谷理论的前提就是地球静止不动——而这正是为托勒密理论所支持、伽利略的观测结果所反对的。^[11]

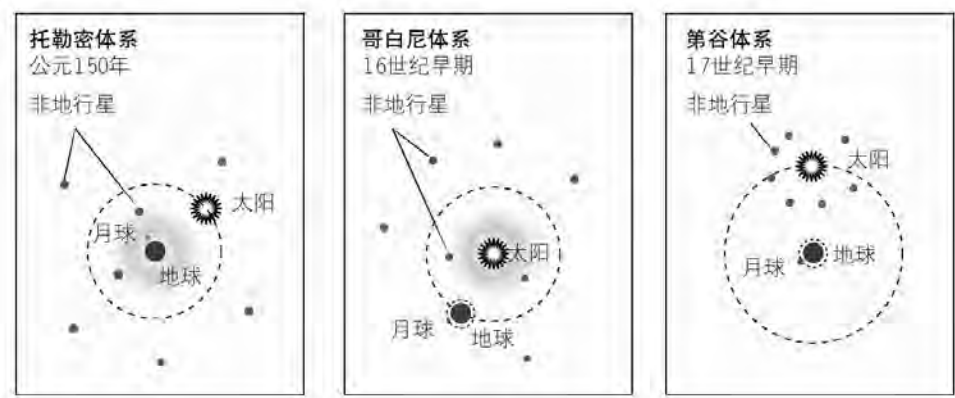


图2-5 描述宇宙的三种体系：托勒密假定，太阳与月球和其他行星一起围绕地球转动。哥白尼正确地提出，所有行星围绕太阳运动。第谷假定，地球之外的行星围绕太阳运动，而它们与太阳一起共同围绕地球运动。

伽利略意识到，第谷的解释缺乏宏观考虑，并走上了最为正确的道路。牛顿的夙敌罗伯特·胡克后来注意到，哥白尼体系与第谷体系都符合伽利略的数据，“但从追求世界之美与和谐的角度来看，我们不得不信奉哥白尼的论点”^[2]，这听起来确实更加简洁优美。伽利略追求理论美感的天性使他获得了正确的结论，当牛顿的万有引力定律解释了哥白尼设定的一致性，并预测了行星轨道的时候，伽利略的猜想也最终被证实了。第谷与托勒密的理论终于被终结了——它们是错误的理论，两人的理论都没有被后来的理论所吸纳。与有效理论的情况不同，这些非哥白尼解释并不能作为任何正确理论的近似。

第谷理论的失败之处在于，它在最初的科学解释中引入了过多的个人主观标准，而这是错误的做法。牛顿力学改变了这种思维方式。科学研究涉及对包含着观测到的结构与相互作用的底层理论及原则的探索。当我们积累了足够多的观测数据时，一个能合理容纳这些观测结果，并能成功预测基础框架的理论就会胜出。在任何时候，逻辑的作用都是有限的——一些粒子物理学家痛苦地意识到，只有我们正在

苦苦等待的那些实验数据，才能最终决定哪些描述宇宙基础本质的理论是正确的。



伽利略为今日科学家所遵循的基本原理夯实了基础。对由伽利略以及其他前辈科学家所创造理论的理解，让我们对科学的本质以及如今一些物理学家经常提到的主要问题的理解更加深刻，特别是对间接观测与实验如何帮助我们弄清正确物理描述的理解。现代科学建立在这些洞察的基础上——技术的应用、实验、理论以及数学推导，都是对试图构建符合观测结果的理论的尝试。更重要的是，伽利略意识到了构建世界物理描述的所有元素之间的相互关系。

今天，我们能够更加自由地思考。随着对宇宙的深入探索以及对可能存在的额外维度与平行宇宙（alternative universes）理论的提出，我们从哥白尼革命出发，并将它发扬光大。随着新理念的提出，人类都越发谦卑，不再以自我为中心。我们的猜想最终要由实验与观测来证实或者证伪。

伽利略引入的间接观测方法，让我们今天能在大型强子对撞机精巧的仪器中看到激动人心的结果。帕多瓦展览最后的环节描述了科学演化到现代的历史进程，甚至介绍了一些有关大型强子对撞机实验的零星片段。我们的向导承认，在这个过程中他完全被弄迷糊了，直到他意识到大型强子对撞机是今天人类发明的、能够探索最小未知尺度的显微镜。

即使今天我们已经能在极高的精确度上进行实验、建立理论，伽利略对如何设计与解释实验的真知灼见，依旧如洪钟大吕一般在我们的耳旁回响。他的遗泽，那些使用仪器以观测肉眼所不能见的对象，以及他对如何正确使用科学方法的洞见，即以实验来证实或证伪科学理念的方式，至今依旧让我们深深受惠。帕多瓦会议的与会者们在思索，近期将要作出的重大发现以及它们的意义是什么，我们似乎看到了跨越崭新的科学门槛的一道曙光。总之，“路漫漫其修远兮，吾将上下而求索”。



2008年2月，来自盐湖城犹他大学的教授、诗人凯瑟琳·科尔斯（Katherine Coles）与生物学家兼数学家弗雷德·阿德勒（Fred Adler）组织了一场主题为“一沙一世界”^[12]的跨学科会议。这场会议的主题是，尺度在不同学科中所扮演的角色，这是一个让分属不同范围、广泛组别的演讲者与参会者都能产生兴趣的话题。通过把观测结果按照尺度分类这一方法，我们可以理解、组织它们，进而把它们拼凑成一个整体。专家组（包括一位物理学家、一位建筑批评家以及一位英文教授）成员以各自有趣的方式对这一过程作出了贡献。

文学批评家、诗人琳达·格里格森（Linda Gregerson）在开幕式上以“sublime”（庄严）一词来描述宇宙。这个词精确地抓住了宇宙为何如此美妙又令人望而兴叹的本质。它们看上去超出了我们的理解和认知范畴，却又那么触手可及，引诱着我们去探究它们的奥妙。我们获取这些知识面临的最大挑战在于，要让这些关于宇宙不易理解的方面变得更加直观、更易理解，最终变得对我们来说不再陌生。人们希望学会阅读与理解自然之书，并且把这些知识引入可理解的世界中来。

为了揭示生命与世界之谜，人类引入了多种方法，向截然不同的目标激流勇进。艺术、科学与宗教也许涉及相同的创造性推动力，然

而它们在试图于人类知识的深壑之间构建高桥之时，却产生了截然不同的方法与手段。[shu籍 分.享 V信jnztxy]

所以，在回到现代物理学的世界之前，这一部分接下来将要对比这些不同思考方式之间的差异，介绍一些科学-宗教之争的历史背景，提出这些纷争至今尚未被解决的至少一个方面。回顾这些主题，我们将要探索科学的唯物主义论与机械论前提，它们都是科学试图接近真理的本质特征。十有八九，这个思想谱^[13]的两个极端永远不会改变其观点，但这些讨论却有可能让我们发掘出这些争端的根源。

未知尺度

奥地利诗人赖内·马利亚·里尔克（Rainer Maria Rilke）曾经在《杜伊诺哀歌》（*Duino Elegies*）中写下引人注目的诗句，它写出了我们在面对庄严宇宙时的矛盾心情^[14]。

美无非是可怕之物的开端，
我们尚可承受。
我们如此欣赏它，
因为它泰然自若，
不屑于毁灭我们。

在盐湖城的谈话中，琳达·格里格森以微妙、有启发性而又多少令人有些恐惧的语言，作了有关“庄严”主题的演讲。她详述了哲学家康德对“美好”与“庄严”的区分，即认为美好“让我们相信自己造化于天地，是上天的宠儿”，而“庄严”却令人恐惧。她还描述了人们“看到庄严时的不安”，这是出于一种“不匹配”的感觉，即与人类日常互动与知觉之间的不协调。

“庄严”一词再次出现在同事与我在2009年所作的一次有关音乐、艺术、科学的讨论中，这三者当时是一部以物理学为背景的戏剧

主题。其中一位导演克莱门特·鲍尔（Clement Power）认为，特殊的音乐片段偶尔展现了美丽与恐惧的共存，这已经为其他人所确认过了。对于鲍尔而言，庄严的音乐是超越了他正常理解力的山巅——这挑战着他已有的理解力。

这种庄严处于人的智力所不能及的层次，并提出了人的智力所不能解决的问题，这正是引人入胜与恐惧共存的原因。庄严的程度随时间而变化，正如我们所熟悉尺度的范围随时间增长一样。然而在任意给定时刻，我们都希望洞悉更多我们现在所能理解的、过小或者过大尺度上的行为或事件。

从很多方面来看，宇宙都是庄严的。它让人好奇，又令人畏惧甚至恐惧，这正是它的复杂本质。然而，这些元素以非凡的方式组合在一起。艺术、科学与宗教都致力于激发人们的好奇心，并以拓展已知领域边界的方式教导着我们。它们都以其不同方式承诺，可以让我们超越个人的狭隘经历，进而进入以及理解“庄严”的国土（见图3-1）。

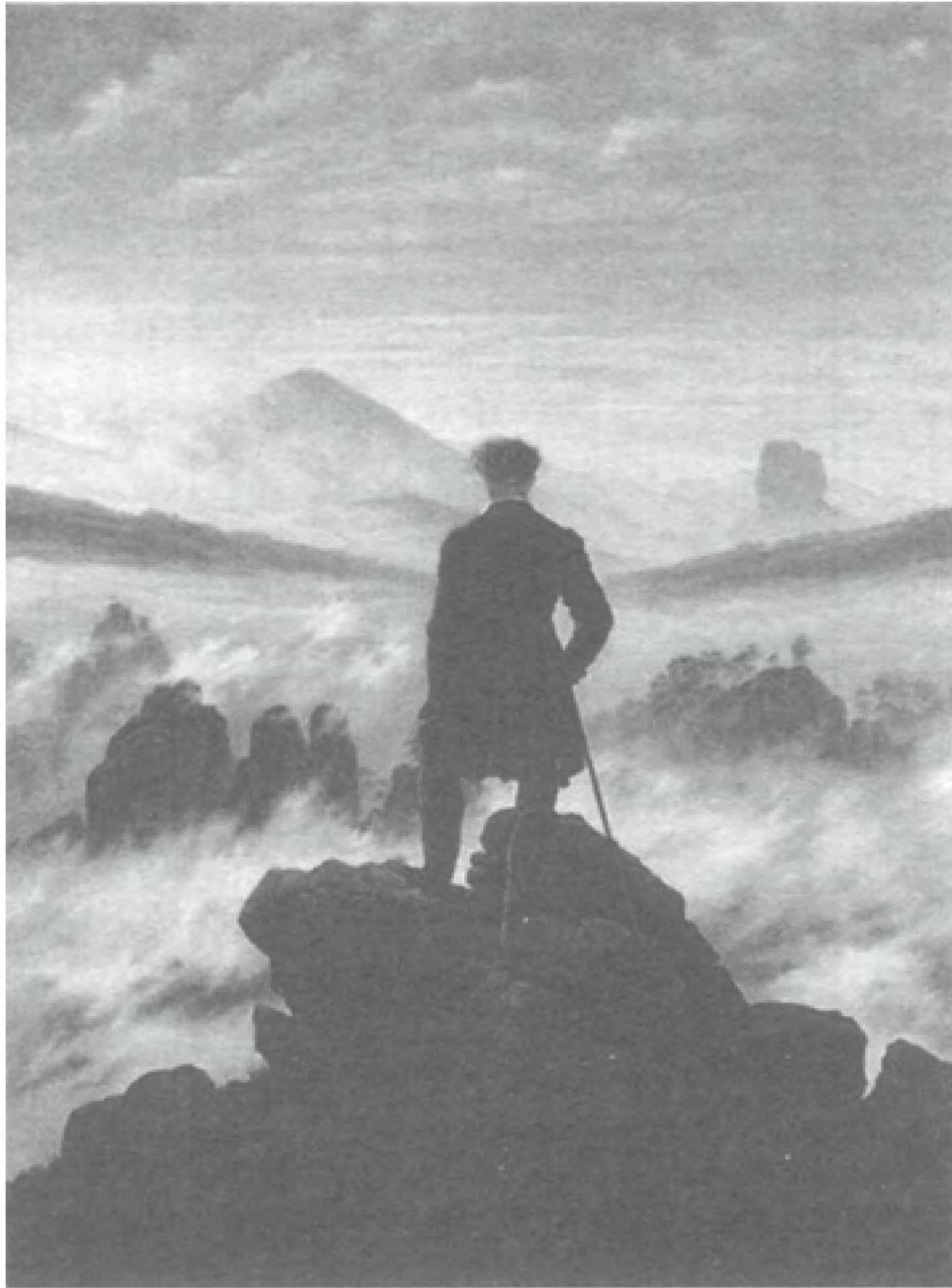


图3-1 斯帕·大卫·弗里德里希 (Caspar David Friedrich) 1818年的画作《雾海中的流浪者》(*Wanderer Above the Sea of Fog*)。这是一幅有关庄严的标志性画作。庄严是一个在艺术与音乐中会反复出现的主题。

艺术让我们可以用人类知觉与情感的滤镜来探索宇宙。它检视了人的感官如何与世界对接，以及人可以从这种互动中得到什么——它

强调人们是如何观察与“参与”这个宇宙的。于人类而言，艺术起着非常重要的作用：它让我们清楚自己的直觉，让我们知道作为人类如何理解世界。与科学不同，艺术并不追求超越人类互动之上的实在真理。艺术与我们对外在世界的物理与情感反应相关，它直接影响人类的经验、需求以及才能，而这些是科学做不到的。

从另一方面来讲，科学寻求的是关于世界客观、可证实的真理。它对组成宇宙的元素是什么以及这些元素之间如何相互作用感兴趣。即便是在形容司法调查时，福尔摩斯也以他不可模仿的风格向华生如此描述科学方法论：“侦探行业是或者应当是一门真正的科学，应该以相同的冷漠无情来对待它。你试图用浪漫主义为其着色的行为所产生的效果，与试图从欧几里得第五公设^[15]出发，导出一个爱情或者私奔故事相似……这种情况唯一值得一提的是，根据结果对原因作出的奇妙推断，让我成功地揭示了事实的真相。”^[1]

无疑，在揭示宇宙之谜上，阿瑟·柯南·道尔爵士与福尔摩斯有着相似的方法论。科学工作者们试图摒弃在明晰这一图景时，人类的局限性或偏见，这样他们就可以相信自己所提出的对现实社会的理解是不偏不倚的。他们使用逻辑推导和共同实验的方式做到这一点。科学家们试图客观地计算出事情如何发生、基础的物理框架如何解释他们所观测到的现象。

然而作为候选项，我们必须注意到，福尔摩斯使用的是归纳而非演绎逻辑，绝大多数侦探与科学家在试图把琐碎的证据拼合在一起时，都会使用这种思维方式。科学家与侦探从观测结果出发归纳，以建立一个符合所有观测现象的一致框架。一旦这种理论被建立，侦探和科学家就开始演绎工作，以预测其他现象及其关系。然而，至少于侦探而言，那一刻就意味着工作的结束。

宗教是另外一条探索的途径，它也许可以用来回应格里格森所描述的与宇宙难以接近的方面相关的挑战。17世纪的英国作家托马斯·布朗（Thomas Browne）爵士在他的《医生的宗教》（*Religio Medici*）一书中如此写道：“我喜欢在谜中迷失自我，以重新从零开始继续追寻我的理性。”^[2]对“布朗们”来说，当他们信仰独居一方的宗教时，逻辑与科学方法对追寻全部真理而言是不足的。科学与宗教之间最大的差异也许是它们选择提出问题的方式。宗教关心落在科学领域之外的问题。对那些底层作用设想的观念，宗教会问“为什么”，而科学会问“如何”。科学并不依赖于任何有关自然界底层目的的观念。这是一条判断我们是把问题留给宗教与哲学，还是完全抛弃它们的分界线。

我们在洛杉矶的谈话中，编剧斯科特·德瑞克森告诉我，在《地球停转之日》（*The Day the Earth Stood Still*，由他在2008年根据1951年的版本重新制作）中原本设定了这么一些台词，但是这给他带来了无数困扰，以至于他在事后数日都在深思此事。女星詹妮弗·康纳利（Jennifer Connolly）在片中的角色，是一名描述其丈夫死亡时试图解释“宇宙随机性”的女子。

德瑞克森被这些台词所深深困扰。基础物理定律确实包含随机性，然而它的整体却是对有序性的概述，这样至少宇宙中的一些方面可以被视作可预测的现象。德瑞克森告诉我，在这段台词被删掉之后，他花了几个星期的时间来寻找合适的词汇——“宇宙中立性”。而当我听到电视剧《广告狂人》（*Mad Men*）中的主角唐·德雷柏说出那些台词时，我的耳朵竖了起来，因为这些演绎方式听上去令人十分不爽。

一个对任何东西都“漠不关心”的宇宙并非一件坏事——甚至对上面提到



的事情而言，算是一件好事。科学家并不寻求宗教经常寻求的那些根本意义。客观的科学仅仅需要我们把宇宙看作中立的对象。确实，科学的中立立场有时确实移除了人类邪恶的烙印，因为它指向人类起源的物质条件而非道德条件。例如，我们现在知道了，从物理学与基因学的本源来考虑，精神病与毒瘾可以被划为“疾病”而非“道德”的范畴。

即便这样，科学也并不专注于所有的道德主题（虽然它也不否认这些）；科学也不追问宇宙行为背后的原因以及人类事务的道德准则。虽然逻辑思考确实有助于我们应付现代世界，以及今日的一些科学家确实以生理学为基础处理有关道德行为的研究，但一般来说，科学家的目的并不决定人类目前道德地位的状况。

上文所提的分界线并非绝对清晰，神学家可以提出科学问题，科学家也可以从一个启迪其内在理念与方向的世界图景下来提出它们——有时甚至是从宗教的图景下。此外，由于科学是由人类创建的，科学家在创建其理论的过程中经常会带有一些非科学的人之本性，比如对“一切问题都必定有一个答案”的信心，或者对特定信仰的情感。当然这些也反映在其他领域中，艺术家与神学家也会受科学家的观测结果与其对世界理解方式的启迪。

这种模糊的分界线并不能抹除掉分界线的两端所追求的终极目标之间的区别。科学家们追求预言性的物理学图景，它可以解释事物如何运作。从本质上来讲，科学与宗教所使用的方法与目标是不同的。科学强调物理实在性，而宗教强调人类心理学或者社会学的需求或渴望。

这种最终目的的分歧并不应该成为冲突之源——事实上，它们应该分工合作，原则上井水不犯河水。然而，宗教并不总是满足于讨论有关意义与安慰的问题。很多宗教都企图占有描述宇宙外在实在性的一席之地，这甚至反映在“宗教”这个词本身的定义中。《美国传统词典》（*The American Heritage Dictionary*）对“宗教”的定义是：“一种对神性或者超自然力量的信仰，或者一种对其创造者或统治者服从与崇拜的力量。”Dictionary.com网站对“宗教”的定义是：“一套信仰体系，它关心宇宙的起因、本质与目的，特别是当问题涉及超人力代言者的产生时；它通常包含虔诚的、仪式性的观测，以及对道德法典的构造，以统治人类事务的道德准则。”这些定义意义下的“宗教”并不仅仅关注人类与世界之间的关系——不管是从道德、情感还是精神层面，还关注世界本身，这使宗教的图景变得歪曲不堪。当科学试图侵犯宗教试图解释的知识领域时，争端就不可避免地产生了。

尽管人类生而具有求知欲，然而使用不同方法提出问题、寻求答案，以及具有不同目的的人并没有达成共识。他们为追求真理而作出的努力也没有被明确地划分界线，所以争辩时时发生。当人们运用对自然界的信仰时，就会回到那些对自然的观测结果上，而宗教必须调和这些矛盾。这在早期教会真实发生过，例如，对自由意志与上帝神力之间相容性的调和。于今日的宗教思想家而言，这依旧是个问题。

宗教与科学是否彼此相容

宗教与科学并非总是处在相互矛盾的窘境。在科学革命之前，科学与宗教是和平共处的。在中世纪，罗马天主教会对各种与《圣经》不同的解释非常大度，直到科学革命使教会感到自己的统治地位受到了威胁。伽利略为哥白尼日心说提供的证据驳斥了教会有关天堂的教义，在这种背景下，教会感到非常头痛——伽利略的著作不仅公然对

抗了教会的权威，而且明确质疑了教会对解释《圣经》不可动摇的专有地位。教士们因此非常痛恨伽利略及其理论。

更多并不久远的历史提供了许多有关科学与宗教之间冲突的例子。认为世界正逐渐向无序性演化的热力学第二定律，让相信上帝创造了理想世界的人们非常气馁。进化论也产生了类似的问题，它在最近引发了一系列与智能设计论的“争论”。“正在膨胀的宇宙”这一说法，让那些相信人类生活在一个完美宇宙中的人感到十分困扰，即便如此，作为天主教牧师的乔治·勒梅特（Georges Lemaître）却是宇宙大爆炸理论（the Big Bang theory）的提出者。

一个更有趣的有关科学家挑战自我信仰的例子，是英国博物学家菲利普·戈斯（Philip Gosse）。19世纪早期，戈斯面临一个窘境，即他意识到地球中保留了含有已灭绝动物遗体化石的岩层，这说明地球仅仅存在6000年的观念值得商榷。在他的著作《地球之脐》（*Omphalos*）中，戈斯如此解释这个疑难：地球确实是不久之前被创造的。造物主如此安排一些特殊的动物“骨骼”与“化石”，这些动物实际上从未出现过，以误导人们并使人们认为存在一些实际上从未存在过的历史。戈斯设想中的正常世界，应该能够反映出其演化轨迹，即便这些轨迹从来没有真正出现过。

这种解释听上去非常愚蠢，但是“技术地”讲，它是说得通的。然而，除了戈斯之外，似乎没有其他人严肃地对待这个解释。戈斯本人转而研究海洋生物学，以避免这些恼人的恐龙骨骼带来的对他信仰的考验。

令人开心的是，大多数正确的科学理念表面上看起来并不激进，而且随着时间的流逝，它们变得越来越容易为人所接受。最后，科学发现终于盛行起来。今日，没有人会质疑以太阳为中心的学说或是宇宙膨胀学说。然而，至少对于像戈斯一样严肃对待这些问题的忠实信仰者而言，字面的解释依然会引起一系列问题。

17世纪之前，人们很少纠缠《圣经》的字面意义，这使得这个时代之前的争端相对少很多。在午间闲聊时，学者、宗教史学家凯伦·阿姆斯特朗（Karen Armstrong）解释了为何现代科学与宗教之争在早期并不真正存在。宗教文本在很多层级上被阅读，这使得对应的解释去字面化、去教条化，因此变得不那么棱角分明。

在公元5世纪，天主教思想家奥古斯丁明确了这种观点：“非基督徒掌握了一些关于地球、天堂以及世界其他部分的事情，也许有关它们的运动以及围绕恒星的运动轨道，甚至有关它们的尺度或者距离。这些知识确实是出于理性与实验得来的。然而，于一个无信仰主义者而言，听一个基督徒对这种事情胡言乱语，而且声称他是从《圣经》中得出这些结论，这不仅失礼而且可耻。我们应该想尽办法避免这种令人尴尬情况的出现，以免无信仰者只看到基督徒的无知，从而展示他们的嘲笑与轻蔑之情。”^[3]

敏感的奥古斯丁甚至走得更远。他解释道，上帝故意在《圣经》中引入了一些谜题，以使人们获得解答它们的乐趣。^[4]这既指晦涩的字句，又指需要隐喻解释的篇章。奥古斯丁看起来对讨论合乎逻辑与不合乎逻辑这两个问题十分感兴趣，他还尝试解释一些基本的悖论问题。例如，“一个人在没有时间旅行的前提下，如何能完全地理解或者感激上帝的安排呢”。^[5]

伽利略本人坚持站在奥古斯丁的立场上。在1615年一封写给时为意大利托斯卡纳大公夫人的克莉丝汀娜·洛兰夫人（Madame Christina of Lorraine）的信中，伽利略写道：“首先，虔诚谨慎地讲，《圣经》从不妄语——无论何时它的真正含义都能被人们理解。”^[6]他甚至声称哥白尼也如此认为，说哥白尼“从未忽略《圣经》，他非常清楚，就算他的理论被证实了，也不能作为《圣经》错误的依据，直到《圣经》被人们正确地理解了”。^[7]

伽利略充满热情地如此引用奥古斯丁的话：^[8]

任何想要以《圣经》的权威来反驳显而易见的道理的人都不明白，他并非在扮演自己想象的角色。因为他对真理的反驳并非《圣经》的真正含义（事实上这些意义超出了他的理解范围），而是他自己的解释。这些解释并非出于《圣经》，而是他自己想象出来的，并且幻想存在于《圣经》之上。

奥古斯丁假设那些文本总是有其合理的意义，灵活解读了《圣经》。任何从对外在世界观测中得来的明显矛盾都是出于解读者的错误理解，即便那些解释是不明显的。奥古斯丁认为，《圣经》是人类受到神启之后的产物。

分析《圣经》中的字句，至少分析其中的一部分，并把它看作作者个人经验的体现，可以发现，奥古斯丁对《圣经》的解释在某些方面与我们对艺术的定义非常接近。教会不需要原路返回以面对奥古斯丁思维模式中的科学发现。

伽利略意识到了这一点。对于他以及一些有着相似想法的人而言，如果《圣经》中的字句被恰当地解释了，那么科学与《圣经》之间本不该有任何冲突。任何可见的矛盾并非出于科学事实，而是出于人们的理解错误。有时于人类而言，《圣经》是难以理解的，而且肤浅地看，它确实与我们的观测结果相矛盾。但是，根据奥古斯丁的解释，《圣经》从来不会出错。伽利略是虔诚的，即便出于逻辑，他也并不认为他有资格反驳《圣经》。多年之后，教皇约翰·保罗二世（Pope John Paul II）声称，伽利略是反对过他的神学家中最出色的一位。

然而，伽利略也信仰他自己发现的理论。在一堆虔诚的废话之后，他有先见之明地如此考虑：^[9]

注意，神学家们，如果你们期望通过讨论太阳和地球到底谁静止不动而建立信仰命题，那么你们正在冒的风险是：你们最终必须谴责那些声称‘地球静止不动，太阳变化位置’的人为异教徒——而我认为，最终太阳静止不动、地球运动不止将逐渐被物理与逻辑所证实。

很明显，基督教并不总是遵守这种哲学，否则伽利略就不会锒铛入狱，今日的报纸也不会用大量篇幅来报道有关智能设计论的争议了。即便很多宗教从业者有着灵活的信仰，一系列对物理现象的严格解释还是可能证明宗教理论中存在很多问题，而从字面意义来理解《圣经》是一种非常危险的尝试。随着时间的推移，由于技术让我们得以探索新的领域，科学与宗教会有更多的交叉领域，它们的潜在矛盾只会只增不减。

今天，世界宗教人口的一个重要平衡方式就是从对各自信仰更加宽容的态度出发，以避免这种冲突。它们并不依赖于对《圣经》严格的解释或是从某个特定信仰出发的对教义的解释。他们相信自己可以在接受严密科学发现的同时，保持自己精神生活的信条不变。

是“上帝的魔力”，还是实实在在的物理关联

我们要面对的本质问题是，宗教与科学之间的矛盾比任何语言与措辞所能表达的都要深刻。即便不去忧心对某个特定文本的字面解释，科学与宗教也各自依赖于彼此不相容的逻辑原则，当我们认为宗教领域通过一个外在之神的干涉作用于我们的世界与存在时更是如此。神的行为并不受科学框架的制约，不论是作用在山上还是你的道德心上。

关键性的比较介于如下两者之间：宗教是社会性或者心理性的一种经验，还是建立在一个经常通过外在干涉影响我们或我们所处世界的上帝基础之上呢？毕竟于一些人而言，宗教信仰是纯粹的私人事

情。那些在这条路上摸索的人们也许会满足于成为志趣相投宗教组织的一部分的社会联系感，或者在一个更大的世界背景之下检视自我的精神满足感。这一类人的信仰与他们的个人实践及其选择的生活方式相关。这些共有的目的成为安慰之源。

很多这样的人认为这就是信仰。宗教提升了他们存在的意义——宗教提供了背景、意义、目的以及一种群体归属感。他们并不把宗教的角色当作一种解释宇宙运作的机制。对宗教，他们只专注于个人的钦佩与怀疑，而这些也许有助于他们与他人和世界之间的互动。很多这样的人会认为，宗教与科学之间可以完美而容易地共存。

宗教不仅仅是一种生活方式或者一种哲学。大多数宗教会有一个总是以神秘方式干涉人类的神，它超越了人们与科学的描述之所能及。这种信仰方式，即使于思想开明、愿意接受科学发展的信仰者而言，也不可避免地会在他们试图调停这种思维与科学指导原则之间的矛盾时陷入窘境。即便是允许上帝或是早期作为“原动力”出现的一种有外在影响的神力存在，从科学的视角来看，上帝可以继续通过一种如春梦了无痕一般的手段，在不留下任何物质性踪迹的前提下干涉人类，这也是令人难以置信的事情。



为了理解这种冲突，更好地理解科学的本质，我们需要更充分地理解科学的唯物主义观点。这种唯物主义的特质告诉我们，科学在一个物质宇宙中应用，以及各种积极影响之间存在的物理关联。第1章中介绍过，科学的图景正是建立在这种理念之上。通过它，我们可以在每一个尺度层级上识别所有物质的组分。大尺度上的存在由小尺度上的

原料构成。即便我们没有必要通过理解所有基本物理元素的方式来解释大尺度上发生的一切，那些组分也是基本而必要的。引起我们兴趣的现象的物质构成于解释它们本身而言并不总是足够的，然而那些物理关联于解释它们的存在却很有帮助。

一些人转向宗教领域以寻求他们认为科学所不能回答的艰深问题的答案。的确，唯物主义科学观并不意味着我们可以理解一切——事实上，仅仅靠理解所有的基础组分并不能做到这一点。利用把宇宙按照尺度分类的方法，科学家们意识到，我们不可能立即回答所有的问题。即便我们所知的基本结构是本质的，它也不必立即回答我们已知的全部问题。即便掌握了量子力学，我们依旧使用牛顿定律，因为在描述一个球体如何在地球引力场中下落这一方面，从原子尺度的图景来导出结论是非常困难的。球体无疑由原子构成，但是原子尺度的图景于解释这个球体的运动轨迹并无益处，即便它一定相容于这个图景。

这种思想可以推广到日常生活中的很多现象。我们经常可以忽略底层细节或者构造，即便这些原料是本质的。

我们在驾驶车辆时并不需要掌握车辆的内部工作原理。当我们烹饪食物时，我们仅仅会关注：鱼片是否切得够薄？蛋糕有没有烤熟？燕麦片有没有泡成糊状？蛋奶酥的表面有没有膨起？除非在练习“分子烹调法”，我们是不会关心导致这些变化背后隐藏的原子尺度的结构变化的。但这并不会改变缺少这些实质的食物不可能好吃这一事实。蛋奶酥的成品与它的原料看起来截然不同（见图3-2）。食物中的分子与其他成分虽然在你看来可能不重要，但是于它们的存在而言却是必不可少的。

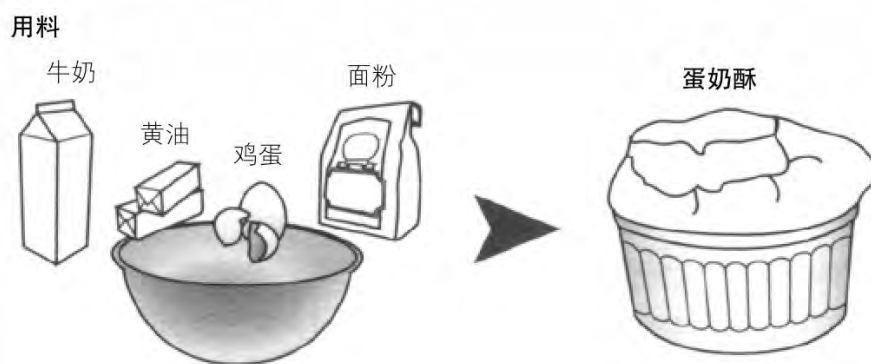


图3-2 蛋奶酥成品与组成它的原料看起来非常不同。类似地，物质之间也许有着截然不同的性质，甚至可能各自服从于天差地别的物理定律，但它们同样由更基本的物质构成。

类似地，任何人在被问到“音乐究竟是什么”的时候都会感觉非常惶恐。任何企图描述音乐现象以及我们对它的情感反应的行为，都会不可避免地涉及从原子或者中子之外的层级来审视音乐。即便我们通过耳朵倾听来自乐器的声波来欣赏音乐，音乐也不过是形成声音的空气微粒振动，或是双耳与大脑的生理反应。

这种唯物主义的图景依旧存在，而它底层的理念才是本质的。音乐产生于空气分子。如果没有耳朵对物质现象机械式的反应，那么也就不存在音乐（在真空中，没有人会听到你的大声叫喊）。只是，对音乐的某种知觉与理解是超出唯物主义描述的。如果仅仅专注于那些振动的分子，那么有关我们作为人类如何理解音乐的问题就不会被提出。对音乐的理解蕴涵着权衡和谐一致与不和谐的分量，这种方法与分子及其振动毫不相关。然而音乐需要那些振动，至少需要它们留在我们脑海中的感官印象。

同样地，理解某个动物的全部组分仅仅是我们理解生命产生过程的一小步。如果我们缺失那些组分如何结合起来，进而构成我们日常

所熟悉的那些现象的知识，那么我们就不能理解一切事物。生命是超越基础要素的一种层现现象（emergent phenomenon）。

意识也有可能被归入此类。即便我们没有关于意识的综合性理论，思维与情感最终还是扎根于大脑中电学、化学以及物理学的性质。科学家们可以观测到大脑中伴随思维与情感的辩证机械论现象，即便它们不能被一个可以解释其工作机理的统一框架所解释。这种辩证图景是基本的，而于理解世界上的全部现象而言并非必要。

我们并没有被保证可以根据最基础的单位来理解意识，却最终可以指出一些应用在更大、综合性与突发性更强的尺度上的原则。随着科学的进展，科学家就会更好地理解大脑中基础化学与电通道的原理，进而理解它的基础功能单位。最终，意识可能可以作为一种科学家们只有通过辨认与研究那些正确组分来完全理解的现象得以解释。

这意味着，不只是研究基础脑化学的神经科学家才有机会取得进展。关心婴儿的思维与成人思维之间差异的发展心理学家，^[10]或者那些对追问人类思维与狗的思维有何不同的人，也面临着取得进展的大好机会。我认为，正如音乐有很多等级与层次一样，意识也与音乐类似。通过对更高层次的质疑，我们也许会获得一些关于意识本身，以及在对其“基石”（即研究大脑的化学与物理学基础）进行研究时如何提出正确的问题的合理洞见。如同研究一块诱人的蛋奶酥，我们也必须理解自发产生的系统。虽然如此，如果无法对我们身体的物理组分施加影响，就不会产生人类的思维或者行为。

即便物理学比有关意识的理论看起来少了一些神秘感，但它却是通过对不同尺度现象的研究来发展的。物理学家在研究不同的尺度与整体时也会提出不同的问题。把宇宙飞船发射上火星时所提出的问题，与我们追问夸克之间如何相互作用时提出的问题是截然不同的。它们都是合理的问题，然而我们并不能轻而易举地通过其中一个推断

出另一个。虽然如此，那些我们送到太空中的物质也是由我们最终希望能理解的基本组分所构成的。



粒子物理学家会利用和指出人们平时不会或者不能重视的所有现象，这种唯物主义观点有时会被人们嘲笑为还原主义（reductionist）。有时这些是物理学或生物学的过程，比如飓风或大脑功能；有时它们是宗教现象——在这个领域里，我经常被人们所指的东西搞糊涂，然而我必须承认它是我们从不涉及的领域。物理理论处理从最小到最大尺度上的结构，我们可以通过实验手段来研究它们或者作出假设。随着时间的流逝，我们最终得以建立一个一致的图景，它可以描述某个层级的实在性是如何产生于邻近层级的。这些基本元素于实在性是必要的，然而好的科学家从不声称有关它们的知识可以解决一切问题。我们将继续前行，以寻求对它们的解释。

即便弦理论可以解释量子引力理论，“万物理论”（theory of everything, TOE）的名字依旧不合适。即便物理学家达到了这个几乎不可能达到的目标，即建立一个包罗万象的基础理论，我们依旧需要面对有关很多大尺度现象的问题，它们并不能简单地通过了解基本组分来回答。只有当科学家理解了在比用基本的弦所描述的更大尺度上出现的集体现象时，我们才能希冀解释超导物质、海洋中的巨浪以及生命现象。在科学研究的过程中，我们按照尺度顺序逐一去研究。如

果试图了解每一个组分，那么我们必将在比现在已经能掌握的更大的距离与尺度上，审视事物与物理过程。

即便我们聚焦于每一个实在性的层级上以处理不同的问题，唯物主义世界观也是必要的。物理学与其他科学依靠研究世界上已经存在的物质发展。科学的核心依赖于研究出于机械的原因及其效应相互作用的事物。物体由于力作用在其上运动起来，引擎消耗能量以发挥其作用。由于万有引力的作用，行星围绕太阳公转。从科学的视角来看，人类行为也最终需要化学与物理学过程，即便我们目前还不能理解它们如何运作。人的道德选择最终也必须与人的基因与进化史（至少部分地）相关。这些物理构造在人们的日常行为中扮演着重要的角色。

我们也许不会立即处理所有的重要问题，然而那些底层的基础于科学描述而言总是必要的。对一名科学家来说，辩证机械论的元素潜藏于对实在性的描述之下。这些物理关联于世界上任何现象而言都是必要的。即便它们于解释一切事物而言并非必须，却也是必要的。

唯物主义世界观对科学来说非常适合。然而，当宗教借助上帝或者其他外在实体来解释人类或者世界的行为时，它就会不可避免地带来逻辑冲突。问题在于，为了同时认可科学以及控制宇宙与人类活动的上帝（或者任何外在精神体），有关上帝具体在哪些要点上使用他的神性力量，以及他如何做到这些问题必须得到回答。于唯物主义者而言，出于科学的机械论观点，如果影响我们行为的基因是使物种得以进化的随机突变的结果，那么上帝只能通过创造这种随机突变的方式来影响我们的行为。为了引导我们今日的行为，上帝必须支配那些看上去是随机的，然而于我们的发展而言却是决定性的突变。如果他确实是这么做的，那么他是如何做到的？他应用了一种力，还是传递了能量？上帝操纵了我们大脑中的电过程吗？是他把我们推向一个

确定的方向，还是他通过为某个特定个体创造一场雷雨风暴的方式，让这个个体迷途？在更大的层级上来说，如果上帝赋予宇宙目的，那么他如何实现其意志？

问题不在于其中的某些提问看上去很愚蠢，而在于从我们理解的科学来看，这些问题看起来根本就不存在与之一致的合理答案。“上帝的魔力”怎么可能得以运作？

很明显，那些愿意相信上帝可以干涉人间的人，希冀上帝能在危难关头帮助他们或者改变世界，而这些呼吁从某种意义上来说必将引入非科学思维的祈求。即便科学并不需要告诉我们事情发生的原因，我们也知道事物确实在运动以及彼此发生相互作用。如果上帝不产生任何物理影响，那么物体就不会移动。即便是我们最终依赖大脑中电信号运动的思维也不会受到影响。

如果这样的影响在宗教中是固有的，那么逻辑与科学的思维就会指出，必须有一个使这些影响得以传播的机制。一个涉及不可见、不可知却可以影响人类活动与行为或者世界本身的力量力量的宗教，或者精神信仰制造了这么一种局面，即一个信仰者要么坚持信仰而放弃逻辑，要么完全不关心这个问题。

这种水火不相容的事实就像一个方法论与理解中逻辑的僵局，深深地打击了我。史蒂芬·杰·古尔德（Stephen Jay Gould）提出的“互不重叠的管辖区”（那些涵盖经验宇宙的科学和扩大到道德拷问的宗教）却交织在一起，而且确实也需要面对这些棘手的矛盾。尽管信仰者们也许会把科学目前还不能回答的一些有关人性深刻而基本的有趣问题降级到宗教的范畴，然而当我们谈到物质与活动的时候——不管这些问题是与大脑结构还是与天体有关，我们还是身处科学的领域之中。

这是一条鸿沟

科学与宗教之间的不相容性并非一定会困扰信仰者。在一次从波士顿到洛杉矶的航班上，这种事情就发生了。

当时我的邻座是一位年轻的演员，他曾经接受过分子生物学的训练，然而他本人却对进化论有一些不同于常人的理解。在开始当演员之前，他曾经在城市学校中协助科学教学三年。当我遇到他时，他正从奥巴马的就职典礼上返回，洋溢着热情与乐观主义，并希望通过努力把世界改造得更加美好。除了继续在演员事业上有所发展之外，他还有一个把教授科学与科学方法的学校开遍全球的大志向。

然而，话锋一转，我们的交流变得不可思议起来。他计划的课程中至少有一门涉及宗教。宗教是他生活中很重要的一部分，他认为人们应该作出自己的选择。但这并不是最大的不可思议之处。他开始解释他的信仰，即人类由亚当传承下来，而非起源于猿类。我难以理解，一个受过分子生物学训练的人怎么能不相信进化论。这种矛盾甚至比任何我上文提到的通过上帝的干涉实现对唯物主义宇宙的违背还要深刻。他告诉我，他可以领会科学并理解其中的逻辑，然而这些只是人类把事情糅合在一起的一种手段——我不理解他想表达什么。在他看来，“人类”逻辑推理的结论并不值得信赖。

这种变化加深了我对如下问题的理解：为何当试图回答有关调和科学与宗教之间矛盾的问题时，我们总是寸步难行？基于经验、由逻辑导出的科学与具有启示性本质的信仰在各自试图接近真理的尝试上，有着截然不同的方法。只有当你以逻辑为规则时，你才能导出矛盾。逻辑试图解决悖论，然而很多宗教都依靠悖论存活。如果你信仰启示性的真理，那么你就自动脱离了科学的范畴，矛盾自然也就如无源之水一样不存在了。一个信仰者可以从他自己的视角提出完全荒谬的解释世界的方式，这种方式可以与科学相容，只须引入“上帝的魔力”就可以做到这一点。或者正如我飞机上的邻座所做的，他们可以简单地决定保持这种矛盾而生活下去。

然而，上帝也许有一种避免逻辑矛盾的方法，科学却没有。宗教拥趸们希望接受某种宗教的解释，它有关世界如何运作，以及科学思想如何被迫面对一个科学发现与不可见、不可知的事物之间巨大的分歧——这是一条鸿沟，它在根本上不能被逻辑思维的方法所跨过。他们别无选择，只能暂时忽略掉有关信仰事物的逻辑（至少是字面意义）解释，或者干脆不关心这种矛盾。

不管选择哪种方式，都保留了他们成为一名学识渊博科学家的可能性，以及宗教的确可以给人的心理带来很多好处。然而任何虔诚的科学家都必须面对挑战其信仰的科学。大脑中负责宗教的那部分区域，不可能与管理科学的那部分区域同时运作——它们显然是不相容的。



我第一次听到“叩响天堂之门”这种表达是1987年在加州奥克兰的一场音乐会上，它是一首由鲍勃·迪伦（Bob Dylan）与感恩而死乐队（Grateful Dead）共同献唱的歌曲的名字。不言而喻，本书的主题与这首歌曲所指大有不同，虽然迪伦和杰里·加西亚（Jerry Garcia）的歌声依旧在我脑海中回荡。这个表达典出于《圣经》，却与《圣经》中的寓意不同，虽然本书的主题就这个理解而言，的确玩了一些文字游戏。在《马太福音》中写道：

你们祈求，就给你们；
寻找，就寻见；
叩门，就给你们开门。
因为凡祈求的，就得着；
寻找的，就寻见；
叩门的，就给他开门。

根据这些言辞，人类可以追求知识，然而最终的目标却是找到接近上帝的途径。人类对世界的好奇心以及积极的探求只不过是通向“神圣”的敲门砖——宇宙本身是次要的。也许答案即在前方，不然信仰者们也许会在追寻真理之路上加力策马奔腾。然而，如果没有上帝，那么知识就成为不可接近或者不值得追求的。人类不能自己来做这一切——我们不是最终的主宰。

这本书的主题将探讨科学中各种不同的哲理与目标。科学并非被动的理解或者信仰，有关宇宙的真理本身就是结果。科学家们满怀向往地走近知识之门，即我们已知知识领域的边缘。我们质疑、探索；事实与逻辑的力量迫使我们在必要的时候改变世界观。我们仅仅相信可以直接由实验证实的事物，或者可以由实验证实的假说推论出的事物。

科学家们已经掌握了有关宇宙的很多规律，却自知还有更多东西亟待探索，其中一些超出了现代实验手段之所能及，甚至超出了人能想象的任何实验手段。尽管人自身具有局限性，但每一次的新发现却都让人们在通向真理的阶梯上上升一步。有时，这样的一小步可以使我们看待世界的方式发生翻天覆地的变化。虽然明白雄心壮志并非总能得偿所愿，但是科学家们依然在坚持不懈地寻找更多理解世界的方式，比如技术上的进步，以把更多构成世界的要素纳入我们的视野之中。之后，我们会寻求更加全面的、能够容纳新信息的理论。

关键问题是：谁有着这个才能，或是权柄，来寻找这些答案呢？人们应当自行探究，还是应该相信更高的学术权威？在进入物理学的世界之前，这一部分将通过科学与宗教视角的对比作为尾声。

谁是“权威”

我们已经看到，17世纪，不断发展的科学思想使基督徒对待知识的态度发生了分裂，这导向了延续至今的、存在于不同概念框架之间的争端。然而，科学与宗教的第二个隔阂之源是所谓的“权威”。在教会的眼中，伽利略声称的独立思考、独立理解宇宙的能力，脱离了基督教信仰的轨道。

作为科学方法的先驱，伽利略摒弃了对权威的盲目信任，代以亲力亲为地观测和解释观测结果。根据观测结果，他会改变自己的观

点。通过这种方式，伽利略解除了手脚的束缚，开拓了一种获得关于世界知识的全新方法，这种方法将造就人类对自然更深的理解力以及影响力。然而，尽管（确切地说是因为）他的主要发现得以出版，伽利略却身羁南冠。他公开发表的那些有关太阳系的结论，即地球并非宇宙的中心，使当时的宗教政权感到恐惧，因为这动摇了他们对《圣经》的严格解释。在伽利略与其他促成了科学革命的独立思想者们的努力之下，任何对《圣经》中人类天性、起源与行为字面意义上的解释都成了被驳斥的对象。

伽利略生不逢时，他的激进思想把他推向了反宗教改革运动（Counter-Reformation，即天主教对新教分支的反击）的风口浪尖之上。宗教改革家马丁·路德（Martin Luther）提倡独立思考与直接从阅读文本来解释《圣经》，而非深信不疑地接受教会的解释方式，使天主教会感到了深深的恐惧。伽利略支持路德的观点，甚至又把它向前推进了一步，他拒绝相信权威，甚至明确地驳斥了天主教会对于宗教文本的解释。^[1]他的近代科学思想建立在对自然直接观测的基础之上，进而尝试提出最合适的假说以解释所有这些结果。尽管伽利略对天主教会有着忠诚的信仰，然而在教士的眼中，他充满好奇的理念、方法与新教徒相类。伽利略就这样无端卷入了一场宗教战争，受到池鱼之殃。

具有讽刺意味的是，反宗教改革运动也许恰巧促成了人们支持哥白尼日心说宇宙观。天主教会希望确保他们的历法是可信的，这样每年的宗教庆典就可以在正确的时间举行，宗教仪式也会得以正确地维持。哥白尼正是被教会要求修正罗马儒略历（Julian calendar）^[16]，以使它与行星和恒星的运动更相容的天文学家之一。正是这个特定的研究项目让他积累了一系列观测数据，并最终作出了激进的断言。

路德本人不同意哥白尼的理论，而且直到伽利略更进一步的观测以及之后牛顿的万有引力定律最终证实了这些之前，没有谁支持哥白

尼的理论。然而，路德确实接受了当时其他一些天文学与医学的发展，他认为这些与不加偏见地欣赏自然之美是一致的。路德并不必然是一位伟大的科学拥护者，然而他带领的革命却创造了一种思维方式、一种交流与接纳新生观念的气氛，这促进了近代科学方法的形成。也要感谢印刷术的发展，使得科学与宗教的理念得以快速传播，减弱了天主教会的权威性。

路德认为，俗世的科学与宗教研究有着同样的潜在价值。科学家们也对此抱有同感，比如伟大的天文学家开普勒。开普勒曾经致信他在图宾根大学的导师迈克尔·马斯特林（Michael Maestlin）。信中提到：“我希望成为一名神学家，长久以来我生无所息。然而现在，我在进行天文学研究，观察如何通过我的努力，使上帝在天文学中得到赞美。” [2]

在这种图景下，科学是一种承认上帝惊人的本性、它所创造的事物以及事物如何运作的解释正在被丰富与被完善的方法。科学成为一种理解被上帝创造的理性、有序宇宙的更好方法，进而可以帮助人类。尤其是，早期近代科学家们并不拒绝宗教，而是以分析他们的研究成果作为一种赞扬上帝造物伟大的形式。为了寻求科学的启示与宗教的神启，他们既阅读自然之书，又阅读上帝之书。他们研究自然的方式是一种承认与感激其造物主的形式。

在更近的年代，我们也偶尔听到这种观点。巴基斯坦物理学家阿卜杜勒·萨拉姆（Abdus Salam）由于对建立粒子物理学标准模型的贡献，荣获1979年诺贝尔物理学奖。他在获奖致辞中宣称：

伊斯兰教中神圣的穆罕默德如此强调，对科学与知识的追求是每一个穆斯林（不论男女）应尽的义务。他要求他的追随者不断求知。很显然，他的脑海中既有宗教知识，也有科学知识，更有对科学探索国际主义精神的强调。

人们为什么关心这些

尽管第3章中已经描述了一些本质的不同之处，然而某些宗教信仰者却喜欢把他们大脑中主管科学与宗教的区域分开，以继续把理解自然的图景作为一种理解上帝的方法看待。很多不积极追寻科学的人也乐于看到科学进展摆脱了镣铐。然而，科学与宗教之间的鸿沟依旧在美国以及世界其他地方存在着。它偶尔会触及某些要点，这些要点会导致暴力，或至少会干扰教育。

从宗教权威的观点来看，对宗教的挑战（比如科学）有很多值得质疑的理由，其中有一些完全无关真理或逻辑。于那些当权者而言，作为一张底牌，总是可以以祈求上帝的方式来证明他们的观点合理。任何形式的独立调查对他们而言都显然是潜在的威胁。窥探上帝的秘密会侵蚀教会的道德力量以及世俗统治者的权威。这样的质疑也会成为谦卑与共同信仰的障碍，甚至会使人忘记上帝的重要性。难怪宗教权威有时会为此忧心。

然而，为什么持有这一观念的不同个体会自发结盟为共同体呢？于我而言，真正的问题并非宗教与科学之间的区别，这些区别已经在先前的章节中被合理地勾勒过了。重要的问题是：人们为什么如此关心这些问题？为何如此多的人怀疑科学家与科学进展？为何这些围绕着权威的争端常常喷薄而出，甚至一直持续到今日？

凑巧，我是剑桥大学科学、艺术与宗教圆桌讨论的参与人之一，这个圆桌讨论是哈佛大学和麻省理工学院成员的系列讨论之一。我参加的第一场会议主题有关17世纪诗人乔治·赫伯特（George Herbert）与新无神论者，这些讨论使我对这些问题的理解获益良多。

由文学家转为法学教授的斯坦利·菲什（Stanley Fish）是最重要的演讲者，他以概述新无神论者的观点以及他们对宗教信仰的敌意开始了他的演讲。新无神论者包含了克里斯托弗·希钦斯（Christopher Hitchen

s)、理查德·道金斯 (Richard Dawkins) [17]、山姆·哈里斯 (Sam Harris)、丹尼尔·丹尼特 (Daniel Dennett) 等人。他们都在畅销书中以尖锐的批评性言辞表达了对宗教的反对之情。

在简洁地介绍了他们的观点之后，菲什开始批评他们对宗教理解的缺乏，这个视角似乎意在吸引那些易于接纳不同观点的听众。而作为一名无信仰者，我觉得我是这场讨论中的少数派。菲什认为，如果新无神论者们考虑到宗教信徒面对的那些必须依靠自己解决的挑战，那么他们会有更强的理由。

信仰需要不断的探询，许多宗教都要求这建立在严守教规的基础上。然而在同时，很多宗教，其中包括一些新教分支，却呼吁信徒拒绝，甚至主动镇压自由意志。法国著名宗教改革家约翰·加尔文 (John Calvin) 如此说：“人类在本质上倾向于欺骗性地孤芳自赏。这正是上帝的真理要求我们内审的原因，它要求一种认知，使得我们可以放下所有对自身能力的自信、剥夺我们在所有情况下的自负，并最终让我们变得谦卑。” [3]

这些特别的言辞主要应用于道德议题上，然而这种对外在导向规律的信仰却是非科学的，我们很难确定界线应当划在哪里。

这种在对知识的渴求以及对人类骄傲不信任之间的挣扎，在各种宗教作品中回响，其中包括为菲什与其他圆桌会议参会者探讨的赫伯特的诗篇。这场剑桥会话详细讨论了赫伯特与知识以及他与上帝之间联系的内在争端。于赫伯特而言，自发产生的理解是一种罪孽深重的、骄傲的标志。约翰·弥尔顿 (John Milton) 也在其作品中提出了相似的警告。即便他坚信知识探索的必要性，他也在《失乐园》 (*Paradise Lost*) 中让拉斐尔如此训诫亚当：

汝对天上星辰之运动毋须抱有好奇，因其不需汝之信仰。

惊奇的是 (至少于我而言)，出席圆桌会议的来自哈佛大学和麻省理工学院的知名教授认为，赫伯特的尝试是一种自我牺牲，并认为

压制个性、调节自身来适应这种更加伟大的力量是一件好事。（任何熟知哈佛大学与麻省理工学院教授的人也会对这种所谓的自我否定感到震惊。）

也许“人类是否能独立自主地发现真理”这个议题才是科学与宗教之争的核心问题。我们今日所闻的有关科学的消极态度，是否部分根植于那些由赫伯特和弥尔顿发表的被公认为极端的信仰？我不确认我们是不是对“世界为什么会成为这个样子”这个主题讨论了太多，包括“谁有指出事物本质的权柄”以及“我们应当信任谁的结论”这些问题。

宇宙本身是谦卑的，大自然隐藏了它大部分有趣的谜题。科学家们却傲慢地认为我们可以揭开一切问题的面纱。我们格物穷理的过程究竟是一种褻渎，还是仅仅是人类的一种自负？爱因斯坦与另一位诺贝尔物理学奖获得者戴维·格罗斯（David Jonathan Gross）把物理学家描述为与上帝对弈，以获得有关自然如何运作这个大问题的答案的人。格罗斯所指的显然不是字面意思（更明显不是出于谦逊而如此说），他意识到，人类能凭直觉来感受周围的世界是一种多么不可思议的能力。

这种不相信自己有能力指出事物本质的遗泽也在人类其他领域中继续着，我们可以在幽默故事、电影以及很多政治观点中看到它。在我们所处的这个滑稽的、时而反智的时代中，对事实的真诚和尊重在某种程度上已经不流行了。某些人否定科学的成功程度是惊人的。在一个聚会上，我曾经遇到过某个明确向我强调她不相信科学的人。所以我问她：你是否与我同样乘坐电梯到达11层？你是否使用电话？你是如何收到电子请柬的？

很多人依然认为，对事实与逻辑表现出热情是一种令人尴尬或至少令人感觉奇怪的行为。反智、反科学的一个源头也许是一种憎恨，

是对认为能够把握世界轨迹的人狂妄自大行为的憎恨。一些潜意识里认为我们没有权力肩负巨大智力挑战的人认为，那些力量属于凌驾于我们能够掌握的事物之上的领域。这种奇怪的反自我、反进步的趋向在广场和地方俱乐部中时有可闻。

于某些个体而言，认为人类可以解读世界这一观念是一种乐观主义的源泉，它会导向对更好的理解力与影响力的领会。然而于其他人而言，技艺精湛、学术渊深的学术权威与科学本身却是恐惧之源。人们根据如下标准为自我划分种类，一类是有能力从事科学工作、评价科学结论的人，另一类是在科学思想面前感到被冷落与无力，因此把这种追求真理的努力视为狂妄自大行为的人。

很多人都希望得到承认并体验到归属感。每个个体都要面对的问题是：在这个世界上，科学与宗教谁更有控制力？在哪个领域你能找到被信任、被理解以及舒适的感觉？你选择信仰，还是选择为自己找到事物答案的人们，或至少信任选择这么做的人们？人们需要的答案与引导是现在的科学所不能提供的。



科学已经告诉了我们很多有关宇宙由什么构建，以及宇宙如何运作的知识。当科学家把所有已知事物放在一起时，他们在漫漫时间长河中积累下来的那些图景竟然出人意料地吻合在一起。科学理念导向正确的预测。所以，一些人选择相信学术权威，而另一些人在悠长的生命中意识到了一些科学课程的重要性。

当探索那些我们尚不能立即进入的领域时，我们同时也在不断地超越人类直觉，然而我们还没有发现“世界以人类为中心”的科学事实。在意识到（以科学的观点来看）我们只是在一个随机运作的宇宙中，于随机尺度、随机位置上出现的事物集合中的一员时，哥白尼式的革命会周而复始地出现。

人类的好奇心与满足这种求知若渴的欲望而进行探究的能力，使人类变得极为特别。人类是一个天生会提出问题并系统性地、抽丝剥茧一般寻求答案的物种。人们质疑、互动、交流、提出假说和抽象理论；在这一切的基础上，人们最终得到对自己所处之地以及宇宙更深刻的理解。

这并不意味着科学要回答一切问题。认为科学可以解决人类的一切问题，那你就大错特错了。然而这却意味着，科学工作曾经是，将来也是一份有价值的追求。我们还不知道所有问题的答案。然而，倾心于科学的人们，不论他们是否具有宗教信仰，都试图叩响天堂之门，寻求一切未知的答案。本书第二部分将要介绍他们迄今为止已经找到了什么，以及目前科学领域的前沿是什么。



第二部分
进入物质世界

**KNOCKING ON
HEAVEN'S DOOR**



即便古希腊哲学家德谟克里特在2500年之前就发现了原子，走上了正确道路的开端，然而并没有人正确地猜测到物质真正的组分是什么。某些应用在小尺度上的物理理论太过于反直观，如果没有实验结果迫使科学家接受它们全新而令人惊惶的假设，哪怕是最有创造性、头脑最开放的人都不能想象出它们的样子。当20世纪的科学家拥有了探测原子尺度上的事物的技术时，他们屡次发现物质的内部结构与理论的预期相悖。以事实碎片拼合出假设的方式，其魔幻程度超越了我们在任何舞台剧中之所见。

对于现今粒子物理学家所研究的极小尺度上所发生的事情，任何人都很难描绘出一幅准确的视觉图像。构成我们称为物质的那些材料的基本组分，与我们直接从感官获得的经验大不相同，这些组分在我们不熟悉的物理规律支配下运作。随着尺度的减小，支配物质的性质看起来如此不同，以至于它们看似分属完全不同的宇宙的一部分。

在不熟悉不同组分的前提下，试图理解陌生的物质内部结构而产生的很多概念混淆，在各种理论得以最便利地应用的不同尺度范围之内出现。为了更全面地理解物理世界，我们需要知道存在什么，还需要对这些被各种理论所描述的尺度有一些基本的概念。稍后我们将探索与空间相关的不同尺度，这是我们的一个终极前沿。本章将首次看

向物质的内部，从我们熟悉的尺度出发，终结于物质的内景——这是另一个终极前沿。从我们日常涉及的尺度到某个原子的内部（量子力学在此必不可少），到普朗克长度（Planck scale，这个长度上引力与其他已知力的强度相似），我们将要探索已知事物以及它们如何彼此结合在一起。让我们来浏览一下物质内部结构的迷人风景吧，这些风景是物理学家与其他领域的学者在漫漫时间长河中积累下的心血结晶。

从人类到宇宙

我们的旅程始于人类尺度，即我们日常所见、所及的尺度。粗略地说，1米是衡量人体的尺度：婴儿的尺度大约是它的1/2，而成人的尺度大约是它的1.5~2倍。如果选取银河系尺度的1/100或者蚂蚁腿的长度作为我们日常测量的基本尺度，那才是怪事呢！

虽然如此，任何根据特定人类个体来定义的标准物理单位都不会是普适的，因为度量衡必须是人们共同理解、共同支持的。^[18]所以1791年，法国科学院制定了一个统一标准。1米被定义为半周期为一秒的单摆摆长，或者沿着地球子午线1/4圆长度（即赤道到北极的长度）的一千万分之一。

两种定义都与我们人类不大相关。法国人只是想找到一个客观的测量标准，使人们一致同意并且能够方便地应用。他们更倾向于采用两条定义中的后者，以排除由于地表各处引力^[19]的不同而产生的微小误差。

这种定义太过随意。它的提出者希望使1米的测量变得清晰、标准，每个人都认可它的定义，然而一千万分之一这个数字与这个理念并不相符。按照法国人的定义，度量衡应该是某种可以舒适地握在手里的东西。

大多数人的身高都可以被近似为两米，然而没有人会认可可以被近似为3米甚至10米。1米是一个人的尺度，我们对这个尺度上的事物情有独钟，至少在这个范围内我们有能力观察它们并作出反应（我们会对数米长的鳄鱼敬而远之）。我们理解这个尺度上适用的物理规则，因为我们在日常生活中不断见证着它们的发生。我们的直觉建立在一生中对物体、人类与动物的观察之上，它们的尺度可以合理地以米为单位来衡量。

有时我感觉到，我们平时处理的尺度范围非常别扭。NBA运动员乔金·诺阿（Joakim Noah）是我堂亲的朋友，我和堂亲俩人总是乐此不疲地评论他的身高。我们总会盯着门框上记录着他随着年龄的增长身高发生变化的标记，屡屡惊叹他在篮球比赛中的“盖帽”。乔金的身高令人着迷。然而事实上，他的身高只比人类平均身高高出约15%，而且他身体的运动方式与任何其他人都基本一致。精确的比例也许有所不同，这有时会带来体力上的优越性，有时则不会。然而他的肌肉与骨骼所遵循的规则与你我的大致相同。



牛顿于1687年提出的运动定律，在描述当在给定质量的物体上施加力时物体的行为依旧有效。它既适用于我们身体中的骨骼，也适用于乔金掷出的篮球。应用这些法则，我们可以计算出球从出手到落地时的轨道方程，也可以预测水星围绕太阳公转的轨道。牛顿定律告诉我们，在任何情况下物体都将保持其速度不变，除非有力作用于其上。这些作用力根据物体的质量不同而赋予其不同的加速度。任何作用力都会引起等值、反向的反作用力。

在我们能够充分理解的尺度、速度与密度范围内，牛顿定律都很好地被应用着。仅仅在被量子力学改变了其规则的极小尺度、（狭义）相对论适用的极高速度，以及广义相对论支配的极高密度（比如黑洞）这些情况下，事情才会有所不同。

任何取代牛顿定律的新理论的效应在正常的尺度、速度和密度情形下都太过微小，以至于我们观测不到。然而，我们可以应用计算与技术进入这些我们将碰到极端情况的领域。

一场小尺度的旅行

在接触到新的物理组分与物理定律之前，我们还需要沿着通往小尺度的道路走上一段路程。在一米尺度与原子尺度之间的范围上还存在很多事物。很多在日常生活与生命本身之中接触的事物都有一些重要的特征，我们只能在探索更小的系统时才能注意到它们，因为在那里不同的行为或子结构才能凸显出来（图5-1是与本章相关的一些尺度）。

当然，我们熟悉的很多事物都只是由一些单一的基本单元累次叠加而来，这些事物并没有太多我们感兴趣的细节或是内部结构。这些外延系统（extensive systems）像砖墙一样构建。我们可以用增减砖块的方式把墙变得更高或者更矮，然而那些基本单元却是相同的。从很多方面来看，一堵高墙与一堵矮墙其实并没有什么区别。这种缩放的观点在很多由重复的基本组分叠加而成的大型系统上都会用到。这能应用在很多大型组织体上，比如计算机内存条是由大量完全相同的晶体管构造而成的。

另一种应用在不同大型系统上的缩放是指数级增长，当关联而非基本元素决定一个系统行为时，它便会出现。即便这样的系统也由许多相似的单元叠加而来，它的行为却依赖于它们之间关联的数量，而

不仅是其基本单元的数量。这些关联并不只是像砖块一样存在于某个组分的邻域之上，而是可以作用在其他单元上从而延拓到全局系统。由很多突触联结组成的神经系统、存在很多相互作用着蛋白质的细胞、由大量联网的计算机构成的互联网，这些都是鲜活的例子。这些都是值得研究的主题，而某些物理学的分支也要研究这些自发产生的宏观行为。

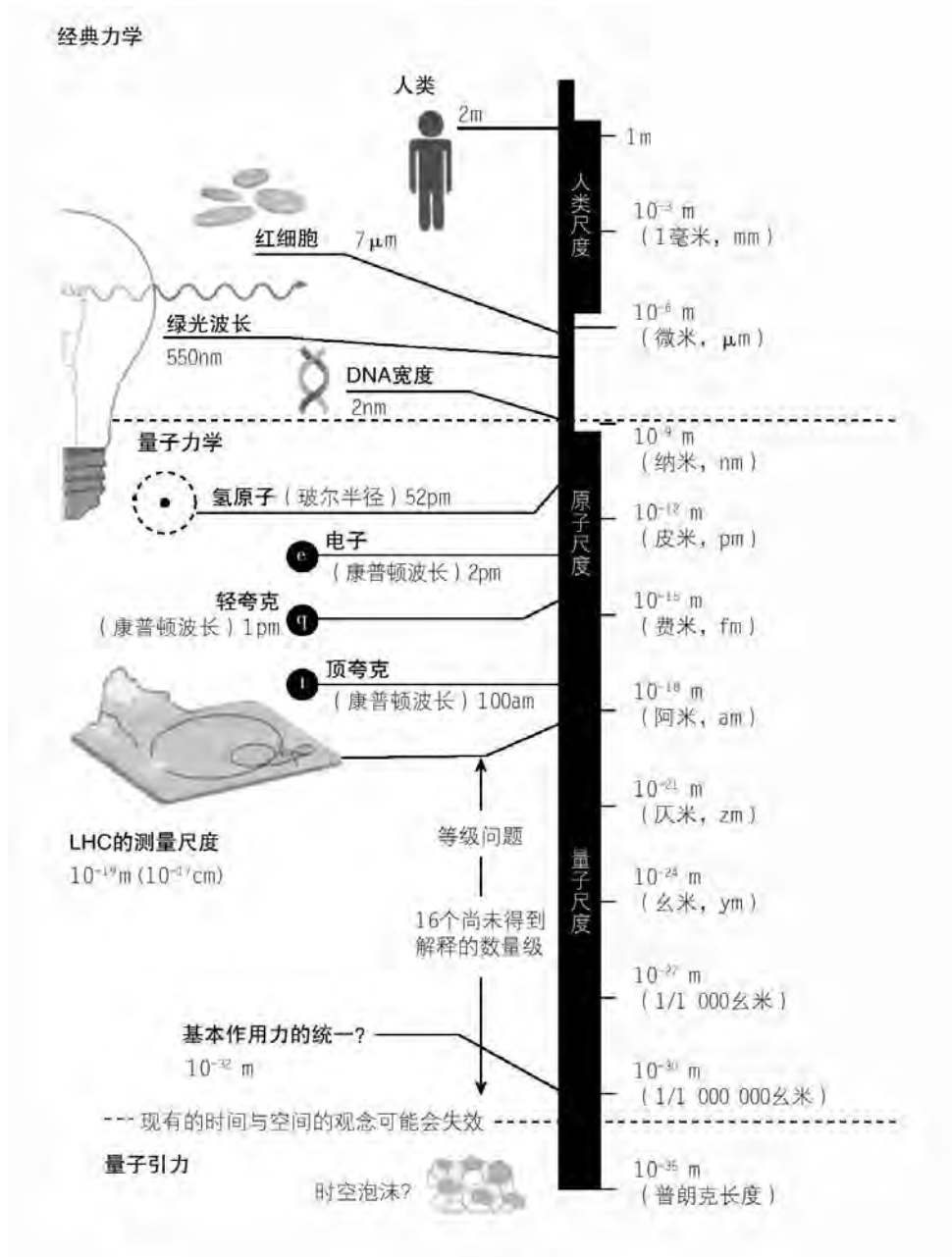


图5-1 一场小尺度上的旅行以及用以描述它们的长度单位。

基本粒子物理学并不研究这种复杂多元的系统，它致力于找出物质的基本组分以及它们所遵循的物理定律。粒子物理学是对基本物理量及其之间相互作用的研究。那些更小的组分当然与复杂的物理行为相关，后者则包含很多有趣的相互作用的组分。然而，确认那些最小的基本组分以及它们行为的方式才是我们所关注的对象。

与科学技术和生命系统一样，更大系统的个体组分也有其内部结构。例如，计算机由微处理器构成，而微处理器由晶体管制造。当医生研究人体内部时，他们会看到器官、血管以及在解剖时遇到的其他由细胞与DNA构造的组织，而DNA只能通过更先进的技术看到。那些内在元素的运作方式绝对与我们所看到的表象完全不同。在更小的尺度上，那些元素改变了，它们所遵循的规则同样随之变化了。

由于一些人类感兴趣尺度的生理学研究史与物理定律研究史在某些方式上相类似，在讨论物理学与外在世界之前，让我们用一些时间来思考一下我们自己以及我们更加熟悉的身体内部工作机制吧。

领骨（collarbone）是一个其功能只能在解剖学的角度被理解的有趣例子。它表面看上去像一个衣领，因此而得名。然而科学家在探查人体内部结构时发现，领骨上有一个像钥匙模样的部位，因此它又被命名为一个我们常用的名字：锁骨（clavicle）。

17世纪早期，生理学家威廉·哈维（William Harvey）做了一系列志在探寻人类和动物心脏、血管系统细节的细致实验之前，没有人理解血液循环或者连接动脉与静脉的毛细血管系统。虽然哈维是英国人，但他的药理学知识是在帕多瓦大学习得的。在那里，哈维从导师西罗尼姆斯·法布里休斯（Hieronymus Fabricius）那里获益良多。法布里休斯也对有关血液流动的课题感兴趣，却误解了静脉及其瓣膜所扮演的角色。

哈维不仅改变了我们对实际事物图景的认识，即从动脉到静脉再到毛细血管的枝状网络，血液在越来越小的尺度上循环；他还发现了一个重要

的过程。在人们真正看到血液在细胞之间流入流出时，没有人能明白它的形式。哈维的发现并不只是一种分类法：他发现了一个全新的系统。

然而，哈维当时没有能看到这些毛细血管的技术工具，这个成就由马尔切洛·马尔比基（Marcello Malpighi）在1661年首次完成。哈维的意见包括了一些建立在理论探讨之上的假说，而这些假说不久之后就被实验证实了。即便哈维对此进行了详尽的说明，他对自己理论的坚信程度也不如后世显微镜使用者们如列文虎克（Leeuwenhoek）那般。

循环系统中有血红细胞的参与，这些内部元素仅有7微米长，它大致是一米的十万分之一。它是一张信用卡厚度的1/100，与雾滴的尺度相当，是我们肉眼能见最小尺度（略细于一根头发丝）的1/10。

血流与血液循环并非医生随着岁月的推移而可以解读的唯一事物，对人类内部结构的探索也没有终结在微米尺度上。从那之后，全新的元素与系统在越来越小的尺度上不断被发现，不管是人体系统，还是无生命的物理系统。

下到1/10微米的尺度，即一米的一千万分之一，我们可以看到DNA这一编码基因信息的基本生命基石。这个尺度依旧是原子尺度的1000倍，此时分子物理学（即化学）已经开始扮演重要的角色了。即便没有完全被理解，与DNA共同出现的分子层级进程也蕴含于遍布地球的丰富生命广谱之下。DNA分子包含数以百万计的核苷酸，因此无疑它们遵守量子力学所支配的原子规则。

DNA本身可以被归入很多尺度中。它们具有扭曲的螺旋分子结构，因此人类DNA的总长度可以以米来计量。然而DNA的宽度仅有一微米的2%，即2纳米。这略小于微处理器上最小的晶体管门电路，后者的大小约为30纳米。单独的核苷酸长度仅有0.33纳米，与水分子尺度的量级差不多。基因的长度是核苷酸的1000~100000倍，对基因最有用的描述牵涉到问题的类型之广远多于单个的核苷酸。因此，DNA在不同的

尺度上以不同的方式运作着。关于DNA，科学家们提出了许多问题，并且在不同的尺度上建立了不同的理论。

在描述小单元共同构成可见的大尺度结构时，生物学与物理学有一些相似之处，然而生物学牵涉的内容远比理解生命系统的个体组分要广泛。生物学的目标更加雄心勃勃。即便我们坚信隐藏在人体工作机理之下的是物理定律，机能性的生物学系统却曲折复杂，它往往有着难以预料的结果。理解这些基本单位与错综复杂的反馈机制非常困难，而当它们以遗传密码的形式组合在一起时，问题就更复杂了。即便我们已经掌握了有关基本单位的知识，解决一些更加复杂的突现科学问题依旧是一件艰深的任务，尤其是那些要为生命负责的问题。

物理学家也是这样，理解了个体子单位的结构并不意味着能理解更大尺度上的过程，然而大多数物理系统就这些方面而言，比生物学系统简单。即便组合结构可能极为复杂，而且有着与小单位截然不同的特性，反馈机制和演化而来的结构往往并不影响什么。于物理学家而言，找到最简单、最基本的组分才是重要的目的。

在人体与原子之间

当我们离开生命系统的机理，“跋山涉水”以寻求能够理解基本物理元素本身的尺度时，我们将立即停在原子尺度，即100皮米（1皮米= 10^{-12} 米）上，它大约是一米的一百亿分之一。原子的精确尺度难于确定，因为它包含着永不停息地围绕原子核旋转的电子。但习惯上，我们把电子到原子核的距离称为原子的尺度。

人们想象出各种解释那些小尺度上物理过程的图景，然而它们必须以类比为基础。我们别无选择，唯有应用我们所熟悉的、日常生活中大尺度上的语言，以描述那些呈现出奇异、反直觉行为的完全不同的结构。

想以我们最惯常使用的生理学基础，即感官与人类尺度上的手工灵敏度，忠实地画出原子的内景是不可能的。例如，视觉需要借助电磁波组成的光才能使现象变得可见。那些光谱中的光波波长分布在380~750纳米的范围之内，这比尺度仅有1/10纳米的原子要大多了（见图5-2）。

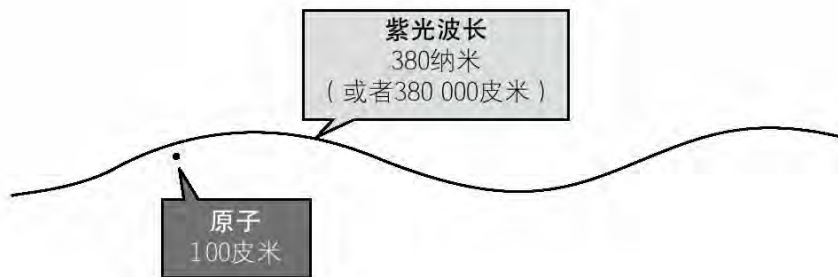


图5-2 即便与最小的可见光波长相比，一个单独的原子也只是一个微粒而已。

这意味着，试图以观测可见光的方式直接通过肉眼了解原子的行为，犹如试图戴着拳击手套穿针引线。这些问题牵涉到的波长迫使我们使用那些过度延展的、分辨率永远不够的光波，以将更小尺度上的信息模糊掉。所以从本质上来讲，我们是不可能从字面意义上“看到”夸克甚至质子的——我们没有准确看到那里存在什么的能力。

然而，混淆我们描绘现象的能力以及我们对其实实在性的信心，是科学家们所不能承受的错误。不能看到或者不能在脑海中想象，并不意味着我们不能推断出物理元素或者在那些尺度上正在发生的物理过程。

从我们假设的原子尺度的优越地位来看，世界会变得无比奇妙，因为这个尺度上的物理定律与以我们熟悉的尺度上适用的物理定律截

然不同。原子的世界与人眼可见的物质世界是全然不同的（见图5-3）。

原子的组成部分

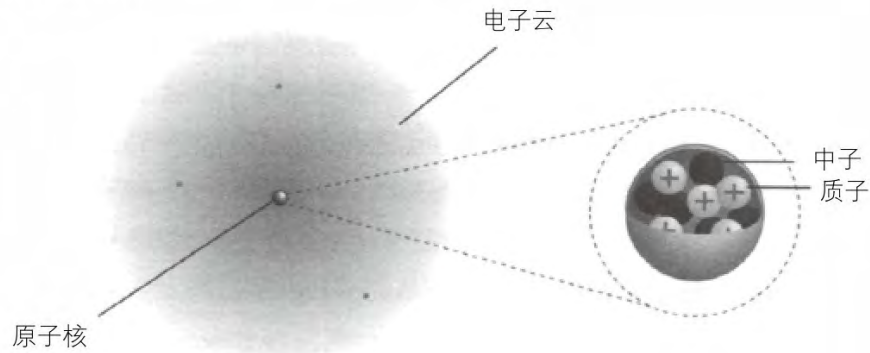


图5-3 原子由原子核与围绕原子核旋转的电子构成；原子核由带一个单位正电荷的质子与不带电的中子构成。



人们已经发现的最震撼的事实是，原子内部空无一物（当然，其间存在电磁场，然而实际上并不存在真正的物质）。原子核，即原子的中心，其半径是电子绕其旋转的轨道的万分之一。原子核的平均大小约是 10^{-14} 米，即10费米。氢原子核的大小是这个数值的1/10。与原子半径相比时，原子核的大小可以类比为与太阳系相比时，太阳的大小。原子的内部几乎是“空的”。原子核的体积只有原子体积的一万亿分之一。

这与我们把拳头打在门上或者用吸管饮用冷饮时所见到或者触及到事物时的感觉完全不同。感官让我们相信物质是连续、致密的。然

而在原子尺度上，我们发现物质的内部却空空如也。这仅仅是因为我们的感官将小尺度取了平均值，因此物质都看似是连续、致密的。然而，在原子尺度上却并非这样。

“几乎空无一物”并非原子尺度上事物唯一令人惊奇的特质。曾经在物理学界兴风作浪，今日依旧困扰着物理学家与非物理学家们的事情是：即使是最基本的牛顿物理学的前提假设也在那些小尺度上失效了。物质的波动性与不确定性原理，即量子力学的关键元素之一，对理解原子中的电子是十分关键的。它们并非如我们经常刻画的那样，沿着确定路径的简单曲线运动。根据量子力学，不可能以任意的精确度同时确定一个粒子的位置和动量，这是能够建立物体随时间运动轨迹规律的一个必要前提。由维尔纳·海森堡（Werner Heisenberg）于1926年提出的海森堡不确定性原理告诉我们，已知的位置精确度限制了人们能测量到的动量最大精确度。^[20]如果电子拥有经典意义上的轨道，我们就能同时精确地知道它的速率和运动方向，进而得知它在任意时刻的位置，而这与海森堡不确定性原理相悖。

量子力学告诉我们，在原子中，电子并不占据任何我们在经典图景中所主张的确定位置。取而代之的是概率分布的描述，它告诉了我们电子在原子空间中任意一点出现的概率，而这是我们能得知的全部信息。我们可以求出电子的坐标随时间变化函数的平均值，然而任何特定的测量都受不确定性原理的支配。

请牢记，这些概率分布并不是随意的。电子不可能具有任意的能标或者概率分布。没有适用于描述电子轨道的经典方法，它只能用概率论的术语来描述，然而这些概率分布却是精确的函数。应用量子力学，我们可以写出电子波函数的解，它可以告诉我们在空间中任意一点电子出现的概率。

另一个在经典牛顿物理学视角看来不可思议的原子性质是，原子中的电子只能占据固定的、量子化的能标。电子轨道仅仅被它的能量所决定^[21]，那些特定的能标与相伴的概率必须符合量子力学的规则。

电子的量子化层级对理解原子而言很重要。20世纪早期，一条促使经典规则发生根本性变化的线索是：在经典图景下，电子围绕着原子核旋转的结构是不稳定的。电子会在旋转中不断辐射能量，并很快掉入原子核中。这不仅与实际原子的情况不符，更不可能允许稳定原子的存在以产生我们所知的物质结构。

量子力学开拓者尼尔斯·玻尔于1912年面临着一个充满挑战的选择：是放弃经典物理学，还是放弃他相信的观测事实？玻尔睿智地选择了前者，并且假设经典定律在电子这一小尺度上不再适用。这是量子物理发展中的一个关键性洞见。



当玻尔在这个限定的领域放弃了牛顿定律之后，他提出了电子只能占据特定能标的假设，这是根据他提出的一个叫作轨道角动量（orbital angular momentum）物理量的量子化条件而作出的。玻尔认为，他的量子化规则只适用于原子尺度。这些规则与我们在宏观尺度（例如计算地球围绕太阳公转时）所用的大不相同。

从技术上来讲，量子力学同样适用于那些更大的系统。然而其效应过小，以至于我们无法测量或注意到。当你观测地球公转轨道或者任何有关的宏观事物时，都大可以忽略量子力学效应。这些效应在所有这样的测量中都大致相同，所以根据它们作出的预测与经典力学的结果完全相符。正如第1章中讨论过的，对宏观尺度上的测量而言，经典预测通常是很好的近似，好到你甚至不能辨别出事实上量子力学才

是更深的底层结构。经典预测非常像高分辨率电脑屏幕上的文字与图像。隐藏于它们之下的是像素点，这对应着量子力学支配的原子子结构。然而，那些文字或者图像本身才是我们通常想要看到的对象。

量子力学催生了一种范式的改变，而它只在原子尺度上才明显。尽管玻尔的假说比较激进，但他并不需要放弃所有已知知识。玻尔并没有认为牛顿定律一无是处，只是简单地认为经典物理定律不适用于原子中的电子。虽然原子本身的量子效应不可忽略，然而由大量原子构成的宏观物质却服从牛顿定律，至少在任何人都能测量其预测的成功性方面是这样的。牛顿定律并无错误，我们并没有在其所适用的领域抛弃它。然而在原子领域，牛顿定律并不适用，在导向量子力学新规则发展的可观察的惊人领域中，它同样不适用。

进入原子核内部

我们的“旅行”将来到原子核本身的尺度上，将继续看到不同描述、不同基本组分，甚至不同物理规律的突现。然而，基本量子力学的范式却岿然不动。

在原子内部，我们将在10费米的尺度上探索内部结构，这是原子核的尺度，即一纳米的十万分之一。以我们目前测量所能及的观点来看，电子是最基本的——这意味着，似乎没有构成它的更小组分。然而，原子核并不是最基本的事物，它由被称为核子的更小组分构成。核子包括质子和中子，质子带有正电荷，而中子是电中性的，既不带正电荷也不带负电荷。

一种理解质子与中子本质的方式是，认识到它们也不是最基本的。伟大的核物理学家、科普作家乔治·伽莫夫（George Gamow）对质子与中子的发现感到非常兴奋，他认为这是最后的“边界”，并认为不会存在更小的子结构。他如此说：^[1]

取代经典物理学中一大堆“不可分的”原子概念的是，我们仅仅保留三种本质上不同的实体：质子、电子与中子……因此，我们似乎已经触及到了有关构成物质基本元素研究的底线。

这种言辞多少有点目光短浅，不过更确切地说，它也没目光短浅到极致。确实存在更深层的子结构，即质子和中子更基本的组分，然而这些基本元素却很不容易被找到。想要找到它们，我们就必须研究比质子和中子更小的尺度，而这需要比伽莫夫作出他那不准确预言的时代已有技术更高的能量或者更小的探测器。

如果我们想要进入原子核内部，在一费米的尺度上看质子与核子——这比原子核本身约小10倍，那么我们就要遇到默里·盖尔曼（Murray Gell-Mann）与乔治·茨威格（George Zweig）假设存在于核子内部的事物了。盖尔曼创造性地把这些单位命名为“夸克”——以他自己的说法，灵感来源于詹姆斯·乔伊斯（James Joyce）《芬尼根守灵夜》（*Finnegans Wake*）中的语句（“向马克老大三呼夸克”）。核子中的上夸克与下夸克是更小尺度上更基本的事物（图5-4中包含了其中的两个上夸克与一个下夸克），一种被称为强相互作用力（又称强核力，strong nuclear force）的力把它们结合在一起，形成了质子和中子。尽管名称泛泛，然而强相互作用力却是自然界中一种特别的作用力，它与已知的电磁相互作用力、引力相互作用力，以及我们稍后将要讨论的弱相互作用力（又称弱核力，weak nuclear force）在同一层次上。

强相互作用力

又称强核力，是作用于强子之间的力，是质子、中子结合成原子核的作用力。后来进一步得知，强相互作用力是夸克之间的相互作用力。

强相互作用力被如此命名是因为它的作用确实很强——我的一位物理学家同僚如此说过。即便这听上去有些幼稚，然而这的确是事实。这就是夸克总以例如质子与中子这种结合体的形式出现的原因，

因为其他直接影响都被这种强相互作用力抹平了。这种作用力太过强大，以至于如果缺少其他影响，那些强烈地相互作用着的组分就不可能被发现。

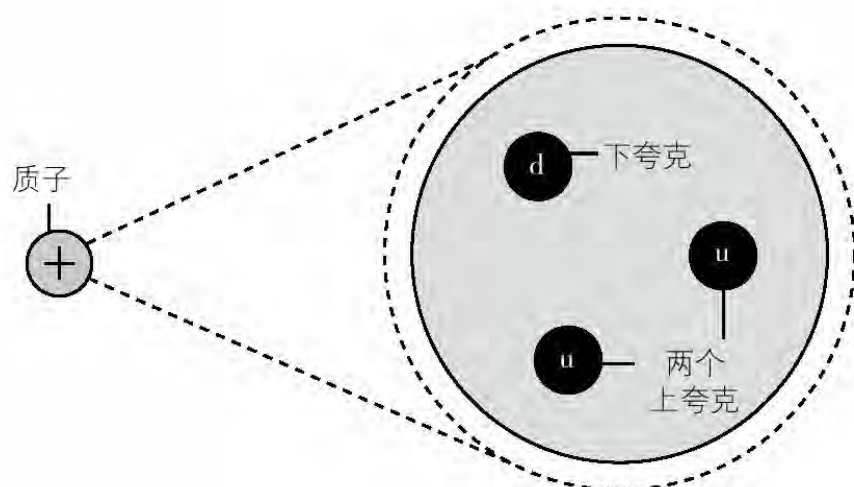


图5-4 质子的电荷由三个价夸克（valence quarks，即两个上夸克与一个下夸克）携带。

我们永远不可能分离出单独的夸克。似乎所有的夸克都携带着一种胶水，它们在长程作用时会变得黏稠无比——传递强相互作用力的粒子出于这个原因被称为胶子（gluon）。你可以想象一条橡皮圈，只有当你拉伸它的时候才有弹性力的出现。在质子或中子的内部，夸克可以自由运动，但当你试图把一个夸克移出很长的距离时，就需要额外的能量。

即便这种解释是完全正确、合理的，在使用这种解释时也要极为小心。人们不由得把夸克想象为一些被束缚在一个口袋里的东西，存在一些真实有形的障碍导致它们不能从中脱离。事实上，确实有一个描述原子和系统的模型在本质上把质子和中子精确地处理为这个样子。然而，这个模型与我们稍后将要提到的其他模型不同，它并不是针对真实发生的事情提出的假说。它的目的仅仅是为了在某个范围的

尺度和能标上进行计算，那些范围内的力太过强大，我们熟悉的方法都不适用于它。

质子与中子都不是香肠，质子中的夸克并没有包围着它的“人造肠衣”。质子是在强相互作用力的作用下，由三个夸克稳定结合而成的。在强相互作用力的作用下，三个小夸克运动一致，看上去像一个事物，即质子或者中子。

另一个有关强相互作用力与量子力学的重要结果，是质子或中子中创造出的额外虚粒子（virtual particle）。量子力学允许这种粒子的存在，它们并不能长久地存在，然而在特定的时间都可以提供能量。按照爱因斯坦著名的质能方程 $E=mc^2$ ，质量等同于质子与中子中的能量，并非仅仅由夸克本身携带，还涉及那些把它们结合在一起的束缚。强相互作用力就像把两个球捆绑在一起的橡皮圈，它本身也携带能量。“解放”出这些内蕴的能量就允许了新粒子的产生。

虚粒子

由质子或中子的相互作用所创造。虚粒子并不能长久地存在，但在特定的时间点上可以提供能量。

只要新粒子的净电荷为零，质子中的能量可以产生新粒子这一事实就不违背任何已知的物理定律。例如，产生虚粒子的时候，带有正电荷的质子不可能突变为中性粒子。

这意味着，任何时候只要载有非零电荷的夸克产生，就必然有与夸克质量相同但是带有相反电荷的反夸克（antiquark）粒子产生。事实上，夸克-反夸克对都可以产生或者湮灭。例如，夸克与反夸克都可以产生光子（传递电磁相互作用力的粒子），光子又可以产生另一个粒子-反粒子对（见图5-5）。它们的总电荷为零，即便粒子对可以产生或者湮灭，质子内部的总电荷却永远不变。

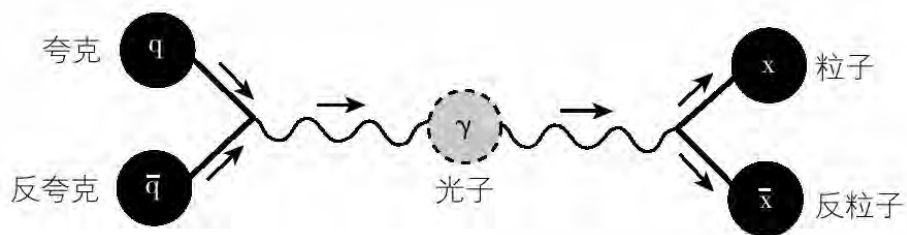


图5-5 能量足够高的夸克与反夸克可以湮灭而转变成能量，进而产生其他带电粒子及其反粒子。

除了夸克与反夸克之外，包含着被创造出虚粒子的质子海（proton sea）也包含胶子。胶子是传递强相互作用力的粒子，这个过程类似于带电粒子之间通过交换光子来传递电磁相互作用力。胶子（总共有8个不同的类别）在传递强相互作用力时也有类似的表现。它们在强相互作用力作用的载荷粒子之间交换，而这种交换的过程使夸克之间相结合或者相排斥。

然而，与光子不同，胶子本身受强相互作用力的支配，而光子本身不带电荷，所以并不直接受到电磁相互作用力的作用。因此，光子可以传递长程作用力，所以我们才可以在数里之外接收到电视信号；而胶子如夸克一样，在相互作用的过程中并不能行进太远。胶子只能在小如质子的尺度上结合事物。

如果我们以教科书式的严谨目光来审视质子且仅仅关注带有质子荷的元素，就会认为质子主要由三个夸克组成。然而，质子所包含的东西远比三个对其电荷有贡献的价夸克（两个上夸克与一个单独的下夸克）要多。在对质子荷有贡献的三个夸克之外，质子的内部是一片虚粒子的海洋，它们包括夸克-反夸克对与胶子。越接近质子，就能找到越多的虚夸克-反夸克对与胶子，真实的分布取决于我们探测它时所使用的能量。今天，我们把质子对撞在一起的能标上，可以找到它们的精确能量数值，这些能量由各种不同类型的虚胶子、夸克与反夸克

携带。对决定电荷而言它并不重要，因为这些虚粒子的总电荷为零。然而正如我们稍后将要看到的，当需要得知质子中精确存在的事物以及携带能量的媒介是什么时，它们对预测质子对撞的结果来说，却是必要的（图5-6描绘了质子中更复杂的结构）。

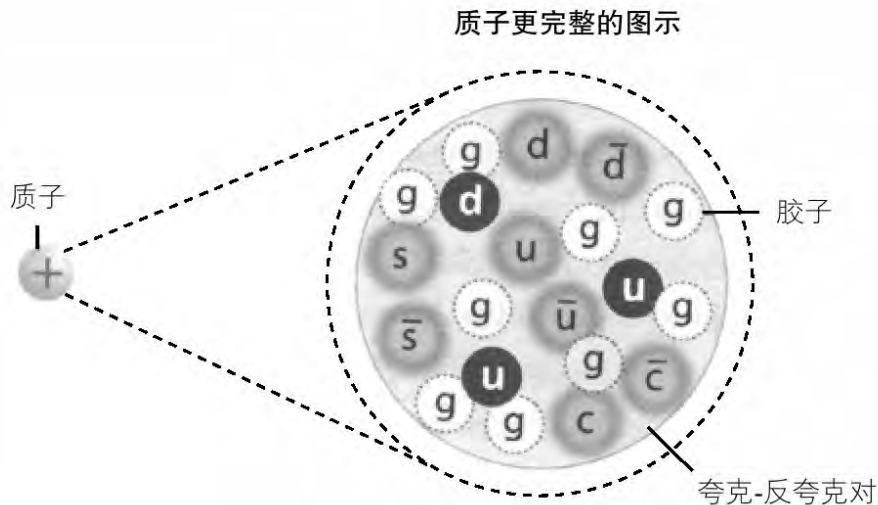


图5-6 大型强子对撞机在很高的能标上把质子对撞在一起，每个质子都包含三个夸克以及也参与对撞过程的很多虚夸克和胶子。

既然我们已经来到了由强相互作用力结合在一起的夸克尺度上，我已经可以讲述在更小的尺度上将要发生什么了。夸克之中还有其他结构吗？电子之中呢？直到现在，我们还没有证据证实这样的事情存在。至今还没有任何实验给出更深子结构存在的证据。在我们的物质内部结构之旅中，夸克与电子就是终点站了——至少迄今为止是这样的。

然而，大型强子对撞机探索使用的能标比过去高1000倍以上，因此我们可以在以往千分之一的尺度上，而这正是与质子质量相关的尺度。大型强子对撞机把两个加速到极高能标的质子束对撞在一起，这达到了前所未有的能标，堪称是里程碑式的壮举。大型强子对撞机中

的质子束包含数千束以千亿计的严格校准的质子，它们被集中在极小的空间区域中，并被送到地下隧道中环行。1232个超导磁铁排列在加速环的两端以把质子束缚在束管中，电场则把它们加速到极高的能标。其他磁铁（精确数量是392个）精细调节粒子束的取向，以避免两个粒子束彼此绕行与相撞。

之后（这一步是所有作用将要发生的一步），磁铁引导两个质子束在加速环中以精确的轨道环行，这样它们就可以在一个比人类头发丝宽度还小的区域内对撞了。当对撞发生时，这些被加速质子的某些能量将会体现为质量——正如质能方程式 $E=mc^2$ 所告诉我们的那样。在激烈的对撞并释放出的能量之时，比以往更重的新基本粒子可能产生。

当质子对撞时，夸克与胶子也偶尔在小区域、高能标内发生对撞，这有点像藏在气球中的石子，当气球对撞时这些石子也会撞在一起。大型强子对撞机提供的能量足够让我们感兴趣的事件发生，即对撞质子的个体组分撞到一起，其中就包括了对质子荷有贡献的两个上夸克与一个下夸克。然而在大型强子对撞机的能标下，虚粒子也带有质子中相当大的一部分能量。在大型强子对撞机中，虚粒子的“海洋”也随着于质子荷有贡献的三个夸克一起对撞。

当它发生的时候，粒子数与粒子类型也会发生改变，而这是整个粒子物理学的关键。大型强子对撞机中的新结果会向我们揭示在更小尺度与距离上发生的事情。除了可能存在的子结构之外，它还将告诉我们在更小尺度上物理过程的其他方面。大型强子对撞机的能标是短距离实验的最前沿，至少在可见的未来是如此的。

技术之上

现在，我们对小尺度上应用现代甚至存在于想象中的技术的旅行的介绍已经走到了终点。然而，现代人类探索能力的极限并不足以抓住现实的性质。即便看上去我们顽强地发展了探索更小尺度事物的技术，然而我们也可以试着从理论与数学的论据出发，来推导那些小尺度距离上的结构与相互作用。

自希腊时代到现在，我们已经走过了很长的一段路。我们现在意识到了，如果没有实验证据，就不可能确定在那些我们希望理解的极小尺度上存在着什么。虽然如此，即便是缺少观测结果，理论的线索也可以为我们引路，并可以启示我们发现物质与力在更小尺度上的可能行为。我们可以提出一些可能存在的假说，来帮助解释与关联那些在可测量尺度上的现象，即便我们并不能直接接触及那些基本组分。



我们现在还不知道，现有的理论推演中到底哪一个是正确的，甚至它们有可能都不正确。然而，即便缺乏接近极小尺度的直接实验手段，我们已经观察到的尺度限制着可以长久存在的事物，因为基本理论最终必须能解释我们已经看到的事物。即便是在更大的尺度上，正是实验结果限制着可能发生的事件，并把我们将推向明确而特定的方向。

因为还没有探索过那些能标，所以我们对其并不熟稔。人们甚至猜测，在大型强子对撞机实验的能量到更高的能量、更短的距离之间的尺度与能量是一片荒漠，并没有太多我们感兴趣的内容存在。也许这种观点是想象力匮乏以及工作数据缺乏的体现，然而对大多数人而言，下一个我们感兴趣的尺度必然与统一理论相关。

一个有关更小尺度事物最迷人的推断牵涉到短程作用力的统一。这个概念同时在科学界与大众的想象力中闪耀着星星之火。根据这种推想，我们所处的世界并没有显露出基本的底层理论，这个理论可以把所有已知的力（至少是引力之外所有已知的力）简洁而漂亮地统一在一起。自我们理解自然界中存在不止一种基本作用力以来，很多物理学家都认真地试图找到这种统一理论。

一个最有趣的推断由哈沃德·乔吉（Howard Georgi）与谢尔登·格拉肖（Sheldon Glashow）^[2]在1974年提出。他们认为，虽然我们在低能标下观测到了三种引力之外强度不同的基本作用力（电磁相互作用力、弱相互作用力、强相互作用力），但是有可能在高能标下仅有一种作用力，其强度是固定的（见图5-7。注意，该图与一种比乔吉-格拉肖理论更精确的统一理论版本相契合，各条线几乎汇于一点，但仍只差一步之遥。这种有缺憾的统一理论不久之后随着对作用力强度更精确的测量而被证实）。这唯一的力被称为“统一力”，因为它包含三种已知的力。这种推想被称作大统一理论（Grand Unified Theory, GUT），因为乔吉和格拉肖认为这个名字很有趣。



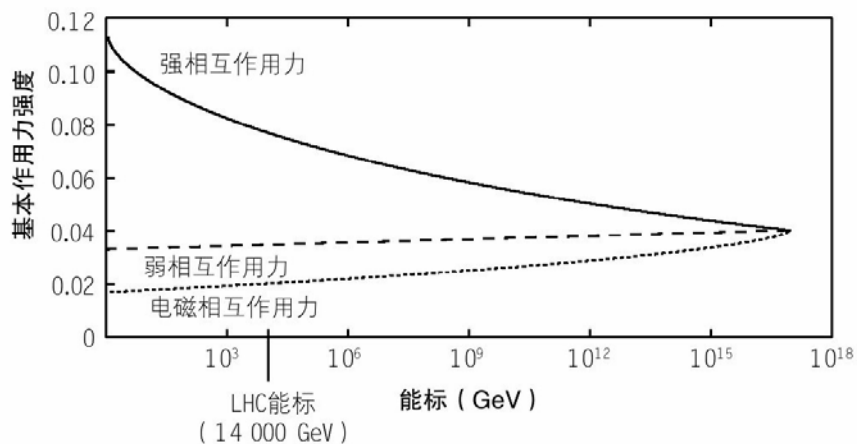


图5-7 在能标下，三种已知的非引力作用力也许有着相同的强度，因此也许可以被统一为一种力。

这种作用力强度收敛的可能性并非无意义的推断。根据量子力学与相对论所做的计算预示着情况可能的确如此。^[23]然而它可能出现的能量尺度远远高出我们通过粒子对撞机实验所能达到的。统一力出现的尺度大约是 10^{-30} 厘米。虽然这样的尺度是人们不可能直接观测到的，但是我们却可以寻找一些大统一理论的间接结果。

一种可能性是观测质子衰变。乔吉与格拉肖在夸克与轻子 (lepton) 之间引入了相互作用的新理论，并认为质子将会衰变。从两人理论的特殊性质出发，物理学家们可以计算出质子衰变的速率。迄今，我们还没有观测到任何支持大统一理论的实验证据，这排除掉了两人的特殊假定。这并不意味着大统一理论必然是错误的。有可能这个理论比他们已经提出的模型更细致、微妙。

这种对大统一理论的研究，决定了我们如何把知识延拓到超越我们能直接观测到的领域中。使用这些理论，我们可以试着推断出通过实验已经断定的、到现在为止还不能达到的能标。有时我们运气比较好，因为设计良好的实验可以让我们检验自己的推断与数据相吻合，或者检验我们是不是太过幼稚。在大统一理论的情形下，有关质子衰

变的实验让科学家们得以间接地研究尺度太小以至于不能直接观测到的相互作用。那些实验让科学家们得以检验他们的理论。一个由这些例子得到的经验是，我们偶尔可以提出对物质和力的有趣洞见，甚至通过推测那些乍看与问题无关的尺度，找到把我们对实验的解释延拓到更高能标、更广泛现象的方法。

我们理论之旅的下一站，也是最后一站，是被称为普朗克长度的尺度，即 10^{-33} 厘米。为了对这个尺度的微小给出一个直观印象，我们这样类比：这个尺度与质子大小的比例正如质子与罗德岛州大小的比例。在这个尺度上，即便是最基本的时间与空间观念都有可能不再适用。我们甚至不知道如何构造一个探测比普朗克长度更小尺度的思想实验，这是我们所能想象到的最小尺度。

对在普朗克长度上进行实验探索的匮乏，并非只是出于我们有限的想象力、技术水平甚至是资金，那些更小尺度上的不可探索性也许是出于物理规律的限制。如我们在下一章中将要看到的，量子力学告诉我们，越小尺度的探测就需要越高的能量。然而，当束缚在一个小区域中能量过大时，物质就会塌缩成为黑洞——在这一点上完全由引力统治。于是，黑洞将吸入更多的能量，这让黑洞变得更大而非更小——这与我们更熟悉的宏观情形类似，而量子力学在这个领域并不占主导地位。我们尚且不知道任何探索比普朗克长度更小尺度的方法，即便提供更高的能量也没有用。传统空间观念非常有可能在如此小的尺度上不再适用。

在我最近作的一个演讲上，在我解释了粒子物理学现在的状况以及一个有关额外维度的可能性质之后，有听众提醒我，我可能忘记了一种我提过的有关人们时空观念局限性的观点。我被问及，我如何可以调停有关额外维度的解释以及时空观念在普朗克长度不再适用的观念。

对于空间（也许还有时间）不再适用的推断仅仅发生在不可观测的小到普朗克长度的尺度上。现在我们观测所能及的最小尺度是 10^{-17} 厘米，所以在可观测的尺度上，认为几何图形是光滑平整的并不悖于任何观念。即便是空间观念本身在普朗克长度附近不再适用，这个尺度也远小于我们目前之所能及。当我们在更大、可观测的尺度上观测平均值时，光滑、可分辨的结构的出现并不会产生任何矛盾。毕竟，不同尺度的事物通常表现出不同的行为。爱因斯坦也会言及，在大尺度上，空间的几何结构是光滑的。然而他的观点在小尺度上可能并不适用——只要它们足够微小且必须服从在大尺度上可以忽略的效应，而在大尺度上，那些更基本的新要素没有我们能观测到的可分辨的效应。

不管时空观念是不是失效了，我们的方程都可以确切地告诉我们普朗克长度的一个决定性特征，即在这个尺度上，原本在我们能测量的尺度上作用于基本粒子时强度微不足道的引力会成为一个很强的力，其强度可以与我们已知其他力的强度比肩。在普朗克长度上，根据爱因斯坦的相对论提出的标准引力方程将会失效。与知道如何作出能很好地符合观测结果预测的更大尺度不同，当我们把它们同时应用在这个小区域中时，相对论与量子力学并不能很好地相容。我们甚至不知道如何试着作出预测。广义相对论建立在光滑的经典几何学基础之上。在普朗克长度之上，量子涨落（quantum fluctuation）让空间变得沸腾，这会产生大量的时空泡沫，让我们所惯用的引力方程无以为继。

量子涨落

不确定性原理允许存在全空无一物的空间（即纯粹空间）中随机地产生少许能量，前提是该能量会在很短的时间内重归消失。

为了处理普朗克长度上的物理预测，我们需要一个新的理论框架，它必须把量子力学与引力理论融合到一起，这个新的综合性理论

就是量子引力理论。统治普朗克长度的物理规律必须与在可观测的尺度上得到验证的物理规律截然不同。对这个尺度的理解必须引入令人信服的科学范式变化，正如由经典物理学过渡到量子力学时产生的变化一样。即便我们在最小的尺度上不能进行任何观察，我们也有机会通过日益发展的理论推演，得知一些有关引力、空间与时间的基本理论。

其中最流行的理论被称为弦理论。早期的弦理论是用基础弦取代基本粒子而构建的。我们现在知道，弦理论也引入了许多弦之外的基本事物（我们将在第17章详细介绍），它的名字也有时会变迁为更广泛（然而定义不那么明确）的术语，即M-理论。这个理论是现在解决量子引力问题最有希望的理论。

弦理论面临许多在观念上与数学上的挑战。现在还没有人知道如何构建能回答所有问题的弦理论，而这些问题都是我们希望量子引力理论可以回答的。此外，弦理论的尺度 10^{-33} 厘米有可能超出了我们能想象出来的任何实验之所能及。

所以，一个合理的问题是：从合理花费时间与资源的角度来看，引入弦理论是合理的吗？我经常被问及这个问题。我们为什么要研究一个实验所不能及的理论呢？某些物理学家在数学和理论物理上都已经找到了足够一致的理由。那些人认为他们可以重现爱因斯坦在发展广义相对论时的那种成功，这大部分基于纯粹的数学与理论物理推导。

另一个研究弦理论的动机（我认为它很重要）是，它可以为我们在可观测的尺度上思考新观念，提供新的方法。其中的两个观念分别是超对称理论与额外维度理论（我们将在第17章中探讨）。如果牵涉到粒子物理学中的问题，那么这些理论确实有着可供验证的实验结果。事实上，如果确定的额外维度理论被证实，并可以解释大型强子

对撞机能标上的现象，即便是弦理论的证据也有可能更低的能标下出现。超对称性或额外维度理论发现并不会被弦理论证实，然而它们却是通过缺乏直接实验结果时，仅靠思想抽象工作有效性的一个有力确证。当然它也会是研究中实验功用的确证，即便那些研究原本是由看似抽象的观念出发得来的。



只有在可以“看到”物质内部结构的工具被发明出来之后，科学家才得以解读物质的构造。“看到”这个词并非指直接观察，而是人们用以观测肉眼所不能见的尺度的间接测量技术。

这往往很复杂。尽管实验结果往往充满挑战而且不直观，其实在性却毋庸置疑。物理定律（即便是在小尺度下）可以指出可测量的结果，而这些结果最终将被更高级的研究所发现。现在有关物质及其相互作用的知识是多年来的灵感、创新和理论发展的顶峰，使我们可以始终如一地解释各种实验结果。通过由伽利略在数个世纪前提出的间接测量方法，现在的物理学家们已经推断出了物质的核心是什么。

我们接下来将要探索粒子物理学的现状以及引导我们至今日之态的理论洞见与实验现象。无疑，当我罗列出组成物质的要素及其发现方式时，这种描述看上去会像一张清单。当我们注意到不同尺度上这些不同要素行为之间的差异时，这张清单就变得更加有趣了。你所坐的椅子最终可以被还原成这些元素，但这个过程本身却复杂无比。

正如理查德·费曼俏皮地描述他的一个理论时所述^[1]：

如果你不喜欢它，那么就去别处——也许其他宇宙的规则会更加简单……我将要告诉你的理论是适用于人类的，他们为了理解这些理论付出了无比艰辛的努力。如果你不喜欢它们，那这真是糟透了。

也许你认为一些我们相信是事实的东西太过冗杂或者疯狂，以至于你不愿意接受它们。然而你的这种想法并不会改变自然界的确如此运作这一事实。

我们需要更小的波长

小尺度似乎很奇怪，因为它是人类所不熟悉的。我们需要探测器以观测小尺度上发生的事情。你正在阅读的这一页书（或者电子屏幕，如果你看的是电子版）的样子与真实存在于物质核心的事物极为不同。这是因为人类的“观看”这一特定动作需要依赖可见光。这些光由围绕着居于原子中央原子核在其轨道上旋转的电子发出。在图5-2中，可见光的波长是不足以让我们看到原子核内部样子的。

我们需要更高级的手段（或者变得更冷漠，这取决于你如何看待这件事）来探测核子在小尺度上所发生的事情。我们需要更小的波长。要相信这一点并不难。我们来想象一种虚构的光波，它的波长与宇宙等长，这种波的任何作用都不可能提供任何足以确定事物空间位置的信息。除非这种波之中有能决定宇宙结构的更小的振动，否则我们就完全无法以这种波长的波为媒介来确定特定位置上的任何事物。这有点像用一张网覆盖一堆东西，然后试图在一团杂乱中找到被藏于其下的钱包的准确位置。你不可能找到钱包，除非你在更小的尺度看向内部，在更高的分辨率下去找它。

使用光波，你需要其有着正确间隔的波峰与波谷，即我们试图决定的不论什么事物尺度的变化，以确定某个事物的位置或者它的大小与形状可能的样子。你可以认为波长就是这张网的大小。如果除了它里面有一些东西之外我们一无所知，那么我就可以确定，只有当我们试图寻找的东西的大小与这张网同样大时，我们才能确定它。如果想

要获得更多的信息，你需要一张更小的网，或者其他在更灵敏的尺度上搜寻更多变化的手段。

量子力学告诉我们，波函数描述了在任何给定位置找到某个粒子的概率。这种波函数可能与光相关，或者是量子力学告诉我们的被单个粒子秘密携带的那种波。那些波的波长告诉我们使用粒子或者辐射探测小尺度时可能获得的结果。

量子力学还告诉我们，波长越短，需要的能量就越高。这是因为它与频率相关，而这直接决定能量^[24]，所以频率最高、波长最短的波携带的能量最高。因此，量子力学把高能标与小尺度联系在一起，只有在高能标下运作的实验才能探测到物质的内部运作机理。如果我们希望探测物质的基本核心，那么这就是我们需要把粒子加速到极高能标探测器的最基本理由。

量子力学的波函数关系式告诉我们，足够高的能量可以让我们探测到小尺度上的事物，以及相应尺度上发生的相互作用。只有使用更高的能标（因此必须有更短的波长），我们才能研究那些更小的尺度。量子力学中的不确定性原理告诉我们，小尺度与大动量相联系，而狭义相对论为我们提供了能量、质量与动量之间的关系，它们一起使得这些联系变得精确起来。

在这些的基础上，爱因斯坦告诉我们，质量与能量是可以相互转化的。当粒子之间碰撞时，它们的质量可以转变为能量。所以在更高的能标上，根据 $E=mc^2$ ，更重的物质可以产生出来。这个方程意味着，更高的能量 E 允许具有更高质量 m 的更重粒子产生，而这种能量是普遍存在的，它可以产生任何在运动学上可以理解（或者说足够轻）类型的粒子。

我们目前能达到的能标越高，能达到的尺度就越小，而那些创造出来的粒子是理解在那些尺度上所适用的基本物理定律的关键。任何在小尺度上突现的新高能粒子与相互作用都保留着解码所谓粒子物理学标准模型的线索，这描述了我们目前对物质最基本元素及其相互作用的理解。我们现在要考虑一些标准模型中的关键发现，以及目前应用的一些把已有知识向前推进进一步的方法。

把电子“扯开”

每一次原子内部之旅的目标——绕核旋转的电子、质子与中子中被胶子“黏”在一起的夸克，都成功地在高能标与伴随的小尺度上被实验探测到了。我们已经知道，通过因为带有相异电性而存在的吸引力，原子中的电子被束缚在原子核附近。这种吸引力赋予这个束缚态系统（原子）的能量比那些带电的组分在孤立状态时的能量之和要低。因此，为了把电子孤立出来以供研究，我们必须提供足够的能量以把它们电离（ionize），这意味着把电子“扯开”以获得自由。一旦电子被孤立，物理学家们就可以研究它更多的性质，比如它的质量与它带有的电荷。

原子核（原子中除去电子之外的部分）的发现更加奇妙。在一个粒子物理学实验中，欧内斯特·卢瑟福（Ernest Rutherford）与他的学生通过往薄如蝉翼的金箔上发射氦原子核，而发现了原子核的存在。那些 α 粒子^[25]拥有足够的能量，以让卢瑟福辨认出原子核中的内部结构。卢瑟福与他的同事们发现，他们发射到金箔上 α 粒子的散射角有时比理论值要大很多（见图6-1）。他们本期待 α 粒子应该如被发射到一张薄纸上一样被散射，然而实际情况却是，这些薄纸的内部好像有一些坚硬无比的石头。用卢瑟福自己的话来说就是：^[2]

这是我一生中最难以置信的发现。这就像你向一张薄如蝉翼的纸上发射了一个4.5米口径的炮弹，而它反弹回来并且打到了你身上一样

不可思议。经过仔细考量，我意识到这种反弹的散射一定是单次碰撞的结果，而当我试图计算时，我发现不可能在那样巨大的数量级上产生任何有意义的结果。除非认为系统是这个样子的：大部分原子的质量都被集中在一个很小的原子核里。就是那时，我产生了‘原子具有一个载荷、微小而质量极大的原子核’的想法。

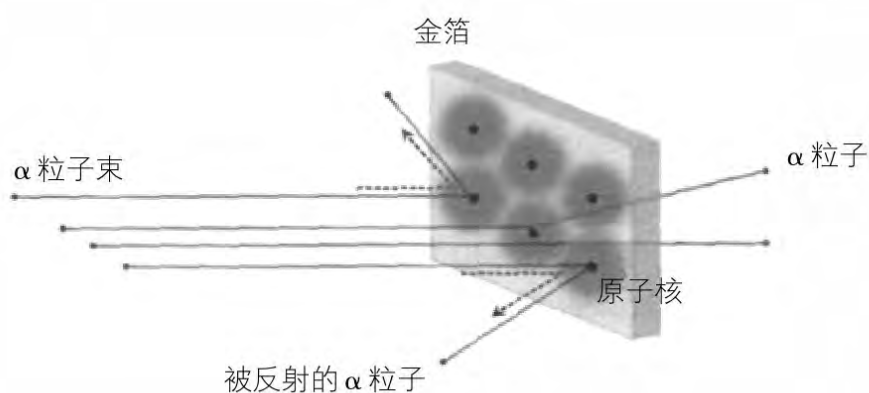


图6-1 卢瑟福实验中在金箔上散射的 α 粒子（我们现在知道它就是氦原子核）。某些 α 粒子散射角预期之外的偏差说明，在原子的中心存在着一些集中的质量——这就是原子核。

发现质子与中子中夸克的实验手段在某些方面与卢瑟福的方法非常类似，然而它需要比卢瑟福用过的 α 粒子还要高的能标。这种更高的能标需要粒子加速器，以把电子与质子加速到足够高的能标并发射出去。

第一个环形粒子加速器被命名为“回旋加速器”（cyclotron），这是因为那些粒子被加速的路径是环形的。欧内斯特·劳伦斯（Ernest Lawrence）于1932年在加州大学建造了第一台回旋加速器。这台机器直径不到0.3米，按照现代标准，它是非常无力的，它能提供的能量远远不够我们寻找夸克所需的。这个成就只有当加速器的技术发展达到一定水平时才会被人们达成（沿着这条路走下去的技术也会导致一系列的重要发现）。

在夸克与原子核的内部结构被发现之前，埃米利奥·塞格雷（Emilio Segrè）与欧文·张伯伦（Owen Chamberlain）于1955年在劳伦斯伯克利实验室（Lawrence Berkeley Laboratory）的高能质子同步稳相加速器（Bevatron）上就发现了反质子（antiproton），并因此共同获得了1959年的诺贝尔物理学奖。高能质子同步稳相加速器比回旋加速器更加精细、复杂，它可以把质子加速到6倍于其静质量的能量^[26]——这足以产生质子-反质子粒子对。高能质子同步稳相加速器中的质子束轰击靶并（根据 $E=mc^2$ ）产生奇异物质，其中包括反质子与反中子（antineutron）。

反物质

指反原子核由反质子与反中子组成的带负电荷的物质。由于物质与反物质的总荷为零，物质在与其对应的反物质相遇时会湮灭。

反物质（antimatter）在粒子物理学中扮演着很重要的角色，所以我们将用一点时间来探索这种与日常观测到的物质相对应的重要事物。由于物质与反物质的总荷为零，物质在与其对应的反物质相遇时会湮灭。例如，根据 $E=mc^2$ ，反质子（反物质的一种形态）可以与质子结合并化为纯粹的能量。

英国物理学家保罗·狄拉克（Paul Dirac）于1927年在试图寻找描述电子行为的方程时首次用数学方法“发现”了反物质。他唯一能写下的与已知的对称原理（symmetry principle）相容的方程暗示了，带有与已知粒子相同质量、相异电荷粒子的存在，这种粒子之前从未被人们发现过。

狄拉克为他的方程绞尽脑汁，最终认为这种神秘的粒子必然存在。美国物理学家卡尔·安德森（Carl Anderson）于1932年发现了正电子（positron），这证实了狄拉克的主张：“方程比我更加可信。”反质子（显然更重）在20多年之后才被人们发现。

反质子的发现不仅对确认它们的存在而言十分重要，而且对确证于解释宇宙运作规律的物理定律而言，基本的物质-反物质对称性也是必不可少的。毕竟，世界是由物质而非反物质构成的。绝大多数日常

物质的质量是由质子和中子，而非其反粒子所携带的。这种物质与反物质的不对称性对世界有着决定性的地位，然而我们还不知道它是如何出现的。

发现夸克

1967-1973年，杰尔姆·弗里德曼（Jerome Friedman）、亨利·肯德尔（Henry Kendall）与理查德·泰勒（Richard Taylor）做了一系列实验，确证了质子与中子中夸克的存在。他们在直线加速器（linear accelerator）中完成实验。这种加速器与回旋加速器和高能质子同步稳相加速器不同，它只能沿着一条直线加速电子。这个加速器中心被命名为斯坦福直线加速器中心（SLAC），它位于帕洛阿尔托（Palo Alto）。经SLAC加速的电子会辐射出光子。这些高能（因而波长很短的）光子与原子核中的夸克相互作用。弗里德曼、肯德尔与泰勒测量了逐渐增加对撞能量时相互作用的变化率。如果原子核没有内部结构，那么这个变化率应该逐渐递减；如果原子核存在内部结构，变化率依旧会减少，然而其减少的速率会变缓许多。在卢瑟福发现原子核数日之前，这些散射结果（在这种情况下是光子）就表现得非常不同，这暗示着质子并不只是一个没有内部结构的质点。

虽然如此，即便是有在必要能标下进行的实验结果的支持，想要辨认出夸克也并非一件轻而易举的任务。理论与技术都必须确保实验的特征信号可以被预期与理解。由理论物理学家詹姆斯·布约肯（James Bjorken）与理查德·费曼所做的见解深刻的实验与理论分析指出，实验确定的变化率支持原子核中存在内部结构的假说，因此确证质子与中子的内部结构——夸克就被发现了。弗里德曼、肯德尔与泰勒因此获得了1990年的诺贝尔物理学奖。

没有人能通过肉眼直接观察到夸克或是它的性质，可行的方法必然是间接的。即便如此，观测结果也确证了夸克的存在。在理论预期、测量到的性质与起初解释本质的夸克假说之间的一致，确证了它们的存在。

随着时代的发展，物理学家与工程师们发明了各种越来越先进的加速器，它们在越来越大的尺度上运行，把粒子加速到越来越高的能标上。越大、越好的加速器可以产生具有越来越高能标的粒子，它们被用来探索更小尺度上的结构。其发现确证了标准模型，因为标准模型中的每个元素都一一被发现了。

固定靶实验与粒子对撞机

发现了夸克的那类实验，即被加速的一束电子束被直接发射到固定物质上的实验，被称为固定靶实验（fixed-target experiment），其目标靶很容易被击中。

固定靶实验

指将被加速的电子束发射到固定物质上的实验。目标很容易被电子束击中。

目前最高能标的加速器与此不同。它们涉及两个粒子束的对撞，每一束都被加速到极高的能标（图6-2是一个对比）。如我们所想象的，两束粒子必须被高度集中在一个小区域里，以保证对撞能发生。这极大地减少了我们预期发生的对撞数量，因为一个粒子束更可能与一大块物质，而非另一束粒子发生相互作用。

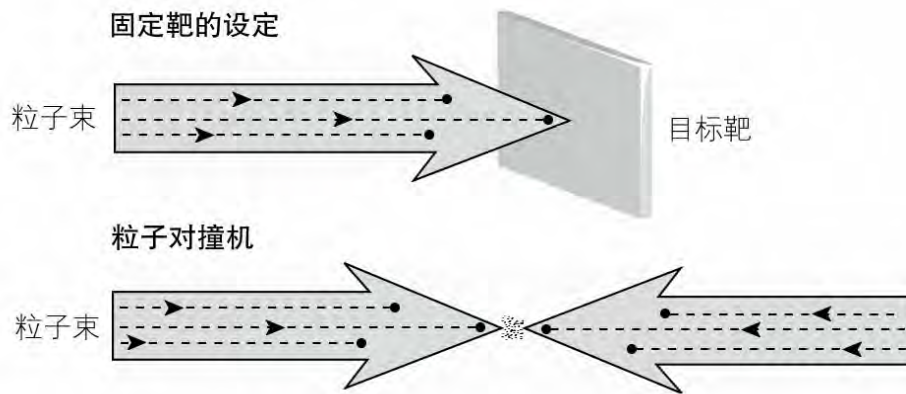


图6-2 某些粒子加速器在粒子束与固定靶之间形成相互作用，另一些则把两束粒子对撞在一起。

粒子束-粒子束对撞有一个很大的好处。这些对撞可以达到非常高的能标。

爱因斯坦曾经提到过对撞机比固定靶实验更加风靡的原因：这与系统的不变质量（invariant mass）相关。虽然爱因斯坦因相对论而成名，他却认为更应将其命名为“不变理论”。其理论的真正要点是，找到一种避免被特殊参考系误导的方法，以找到决定系统特征的不变量。

这个理念可能在一些空间性质（比如长度）上于我们而言更加熟悉。固定物体的长度取决于它在空间中如何被定位。事物具有固定的尺寸，且不随着人的观测而改变；而坐标系不同，它由事先确定的任意坐标轴与方向构成。

相似地，爱因斯坦发明了一种描述事件的方法，它不依赖于观测者本身的位置或动量。不变质量是对物体总能量的一种衡量。它告诉我们，在已知系统中能量的情况下，可能产生的物质有多重。

为了确定不变质量的数值，我们可以提出如下的问题：假使系统是静止的（完全没有速度或者动量），那么它将带有多少能量？如果一个系统没有动量，那么我们就可以应用方程 $E=mc^2$ 。因此，一个系

系统的静能量等同于知道它的不变质量。当系统不处于静止状态时，我们需要一个更加复杂版本的公式，它依赖于动量与能量的数值。

假定我们把两个具有相同能量、等大反向动量的粒子束对撞在一起，它们碰撞过程的总动量为零。这意味着整个系统已经是静止状态的。因此，全部能量（两个单独粒子束中每个粒子的能量之和）都可以转变为质量。

固定靶实验则大不相同。入射粒子束有很大的动量，而靶本身没有动量。并非所有能量都可以用来产生新粒子，因为这个由靶与入射粒子束共同构成的复合系统依旧在运动。由于这种运动的存在，并非所有对撞中的能量都可以用来产生新粒子，因为还有一部分能量以与动量相关的动能形式存在着。最终，有效能量大致只有入射粒子束与靶能量乘积的平方根的数量级。这意味着，如果我们把入射质子束的能量增大100倍，并把它与一个静止的质子碰撞，产生新粒子的有效能量只会增长10倍。

这告诉我们，在固定靶与粒子束-粒子束对撞之间存在着巨大差别。粒子束-粒子束对撞的能量远远高于粒子束-固定靶碰撞，你大概会预期前者约是后者的两倍。然而这个猜测是以牛顿思想为出发点的，并不适用于接近光速运动的相对论性粒子。固定靶碰撞与粒子束-粒子束对撞相比，能量的差异远大于之前简单的猜想，因为在近光速的情形下，相对论进入了舞台。当想要达到高能标时，我们别无选择，唯有使用粒子对撞机，把两束粒子加速到极高的能标而碰撞在一起。同时加速两束粒子可以获得更高的能量，因此可以产生更强烈的碰撞。

大型强子对撞机正是粒子对撞机的一个例子，它把被磁场约束而瞄准的两束粒子对撞在一起。决定对撞机（比如大型强子对撞机）性能的主要参数是参与对撞粒子的类型、加速之后能拥有的最高能标、

对撞机的发光度（luminosity，指结合的粒子束的强度以及出现的事件数）。

哪种对撞机才是最佳选择

一旦我们确定了同时使两束粒子参与对撞，相比固定靶实验就可以提供更高的能量（因此可以探索更小的尺度），接下来的问题就是用什么来进行对撞了。这引发了一些有趣的选择。特别地，我们必须决定要加速哪些粒子以参与对撞。

使用地球上轻而易举就能获得的物质无疑是个好主意。从原则上讲，我们可以尝试把不稳定的粒子对撞在一起，比如很快就会衰变为电子的 μ 子，或者如顶夸克这种很快就会衰变为其他较轻物质的重夸克。

在这种情况下，我们必须首先在实验室中制造这些粒子，因为它们并不是随手可得的。但是，即便我们可以制造它们，并且在它们完全衰变之前就把它加速到所需能标，我们依然必须要确保衰变过程中产生的辐射被安全地转移。这些问题并非不可逾越的天堑，尤其是 μ 子，它形成粒子束的可能性还在研究之中。然而它们确实产生了我们在处理稳定粒子时不会面对的困难。

所以，我们必须做一个直截了当的选择——直接选择那些地球上存在、不衰变的稳定粒子。这意味着我们会选择轻粒子，至少是束缚稳定态的轻粒子，比如质子。我们还希望这些粒子带有电荷，这样就可以方便地使用电场来加速它们。这使得质子和电子成为可选项，它们都是大量存在的、可以很便捷地找到的粒子。

我们应该选择哪个呢？它们各有利弊。电子的好处在于它可以完全地碰撞，而且它毕竟是基本粒子。当使用电子碰撞其他物体的时

候，电子并不会把它本身的能量分给许多其自身的子结构。正如我们目前所知，电子之下没有更多的子结构。由于电子的不可分割性，我们可以非常精确地把握它与其他物体碰撞过程中所发生的事情。

对质子而言就不是这样了。让我们回顾一下，在第5章中我们提到，质子是由被强相互作用力束缚在一起的三个夸克构成的，胶子在夸克之间被交换以把整个体系“黏”在一起。当质子在高能标下参与碰撞时，你感兴趣的能产生一些重粒子的相互作用一般只涉及质子中的单一粒子，比如单独的夸克。

夸克显然不会携带质子的全部能量。所以，即便质子可能带有极高的能量，夸克带有的能量却相对较少。它依然可以具有相当可观的一部分能量，只是这个数值依然不及假设质子可以把它的全部能量集中到那个单一夸克上所能具有的。

除了这些之外，涉及质子的对撞是非常麻烦的。这是因为质子中的其他组分依然闲置在那里，即便它们不参与我们所关心的超高能标对撞。余下的那些粒子之间依然存在着强相互作用力（这个名称非常恰当），这意味着会存在一些围绕并扰乱你感兴趣的相互作用的一团乱麻。

那么在这种情况下，为什么要使用质子进行碰撞呢？理由是：质子比电子更重。事实上，质子的质量比电子约大2000倍。这有助于我们把质子加速到极高的能标。为了获得如此巨大的能量，电场围绕着一个环存在，这样粒子就可以在一圈接一圈的旅程中逐渐被加速。具有加速度的粒子会产生辐射，而且粒子越轻，辐射越大。

这意味着，即便我们乐于使用具有超高能标的电子进行碰撞，也不会在短期之内实现这个目标。我们可以把电子加速到非常高的能标，然而高能电子在被进行环形加速的时候会把自身的很大一部分能

量辐射出去。（这就是斯坦福直线加速器中心用来加速电子的加速器是直线加速器的原因。）所以，从纯粹能量与发现粒子潜力的角度来看，都是质子胜出了。质子可以被加速到极高的能标，以至于即便是它的子组分，如夸克与胶子，其所带能标都高于被加速的电子。

事实上，物理学家们从各种类型的对撞机（如质子对撞机与电子对撞机）上发现了粒子的很多性质。使用电子束进行对撞的对撞机无法在质子加速器所达到的高能标上运作。然而对撞机上用电子束进行的实验取得了比质子对撞机更精确的测量结果，其精确度远超人们的想象。特别是20世纪90年代，斯坦福直线加速器中心与欧洲核子研究中心的大型正负电子对撞机（Large Electron-Positron collider, LEP，这个直截了当的名字不时令我莞尔）在确证粒子物理学标准模型的预测时，表现出了惊人的准确性。

这种电弱精密测量（precision electroweak measurement）实验开拓了许多使用电弱作用（electroweak interactions）知识可以预测的不同物理过程。例如，它们测量了携带弱相互作用力粒子的质量、衰变为不同粒子的速率以及探测器前段与后段的不对称性，这揭示了弱相互作用力的更多本质。

电弱精密测量明确地应用了有效理论的观念。当物理学家们做了足够充分的实验，以确定下来标准模型中为数不多的几个参数，比如每种基本作用力的相互作用强度之后，其余的事情就都变得可以被预测了。物理学家们会核实所有这些测量的一致性，以及这些数据的偏差，以确定还有什么没有找到的元素。迄今所知的所有测量数据都很好地契合于标准模型，然而依旧没有居于背后的、我们想要知道的线索。唯一可以确定的是，不管它是什么，在大型正负电子对撞机的能标下效应都必须非常微小。

这告诉我们，想要获知有关更重粒子与更高能标相互作用的信息，就需要直接考察比我们在大型正负电子对撞机与斯坦福直线加速器中心中达到过更高的能标上的相互作用过程。至少在可见的未来，电子对撞不可能达到我们所需的能标，这个能标可以确定如下问题：什么事物赋予粒子质量？为什么它们的质量是现在这个数值，而不是别的？回答这些问题需要质子对撞。

这就是物理学家们决定在一条建于20世纪80年代，用来放置大型正负电子对撞机的隧道中加速质子而非电子的原因。欧洲核子研究中心最终停止了大型正负电子对撞机的运行，以为下一个庞然大物——大型强子对撞机的准备开路。质子并不会辐射出那么多能量，大型强子对撞机可以更高效地把质子加速到更高的能标。然而，质子的对撞比只涉及电子的对撞复杂许多，充满了实验上的挑战。但是，在粒子束中加入质子，我们就有可能达到所需的足够高的能标，这个能标可以直接揭示我们数十年间孜孜以求的结果。

粒子还是反粒子

在决定使用什么进行碰撞之前，我们还有一个问题需要回答。毕竟，对撞牵涉到两束粒子。我们已经确定了，一束高能粒子由质子组成。然而另外一束粒子由什么构成呢？是质子，还是它的反粒子——反质子呢？质子与反质子质量相同，因此辐射速率也相同，所以必须以其他的标准来区分它们。

显然，自然界中的质子数目更多。我们在周围看不到反质子，是因为它们会和环境中大量存在的质子相互作用而湮灭，转变为能量或者更基本的粒子。那么，为什么要考虑使用反粒子束进行对撞呢？我们怎么获得它们呢？

答案相当多。反粒子的加速过程更加简单，因为磁场可以把粒子、反粒子导向两个不同的方向。然而最重要的理由与可能产生的粒子有关。

粒子与反粒子的质量相同，电荷相反。这意味着，入射的粒子、反粒子与纯粹的能量带有的总电荷相同——即一无所有。根据 $E=mc^2$ ，这意味着粒子与其自身的反粒子可以转变为能量，能量也可以创造其他的粒子-反粒子对，只要它们不过重，或者最初的粒子-反粒子对的相互作用过强。

原则上，这些被创造的粒子可以是新颖而奇异的，它们的电荷可能与标准模型中的粒子不同。对撞的粒子与反粒子没有净电荷，奇异粒子与它的反粒子也没有。所以，即便奇异粒子的电荷可能与标准模型中的粒子不同，原则上来说，只要粒子与其反粒子的总电荷为零，这个粒子-反粒子对就有可能产生。

让我们把这个推论应用到电子上。当把两个带有相同电荷的粒子（比如电子）对撞在一起时，我们只能制造出与参与对撞的系统总电荷相同的事物。它可能产生一个带有双倍电子电荷的事物，也有可能产生两个像电子一样的不同物体，每一个都带有相当于一个电子带有的电荷。这是一个相当强的约束。

把两个带有相同电荷的电子对撞有很强的限制。另一方面，使用粒子与反粒子对撞创造了很多新途径，而这些途径是使用同种粒子对撞所不能达到的。因为可能出现的新末态非常多，电子-正电子对撞比电子-电子对撞更有潜力。例如，涉及电子及其反粒子（即正电子）的对撞会产生大量不带电粒子，比如Z规范玻色子（gauge boson）（这正是大型正负电子对撞机的工作方式）以及许多足够轻而能够产生的粒子-反粒子对。虽然使用反粒子进行对撞会付出不菲的代价，因为它

们难以储存，但是我们的收获更大——我们发现了期望之物，即与参加对撞粒子的电荷不同的奇异粒子。

不久前，迄今为止最高能标的对撞机使用一束质子与一束反质子进行了对撞实验。这当然需要一种制造、存储反质子的方法。高效地存储质子的方法是欧洲核子研究中心已经取得的最为显赫的技术成就之一。在更早的时期，欧洲核子研究中心建造大型正负电子对撞机之前，实验室就制造了一些高能质子与反质子束。

欧洲核子研究中心进行的质子-反质子对撞实验中最重要成就，当属传播电弱作用力的电弱规范玻色子的发现，其发现者卡洛·鲁比亚（Carlo Rubbia）与西蒙·范德梅尔（Simon van der Meer）因此获得了1984年的诺贝尔物理学奖。与其他基本作用力一样，弱相互作用力也由粒子传播。传播弱相互作用力的粒子被称为弱规范玻色子，包括带正电或带负电的W矢量玻色子（vector boson）与电中性的Z矢量玻色子，这三种粒子负责传播弱相互作用力。

我依然认为两种W玻色子与Z玻色子是“天杀的矢量玻色子”（bloody vector bosons），这个说法源于一位英国物理学家。当时他酩酊大醉，步履蹒跚地来到访问学者（包括我在内）与暑期学生的集体宿舍，排闥直入，口中喃喃重复着这几个词。同时，他还表达了对美国科学统治地位的关心之情，并且前瞻了人类未来在欧洲可能作出的第一个主要发现。当欧洲核子研究中心于20世纪80年代发现了两种W矢量玻色子与Z矢量玻色子时，粒子物理学标准模型（弱相互作用力是其必不可少的一部分）就被实验证实了。

范德梅尔发现的存储反质子的方法是实验成功的关键，很显然，这是个艰巨的任务，因为反质子极易与随处可见的质子结合而湮灭。在被范德梅尔称为随机冷却（stochastic cooling）的方法中，粒子束的电信号驾驭着一个可以过滤掉动量过高粒子的装置，最终把整个

粒子束“冷却”下来，使它们不能快速移动，因此不能立即逃逸或是撞击承载它们的容器，这样即便是反质子也可以被储存。

质子-反质子对撞机的理念并不只流行于欧洲。这种类型最高能标的对撞机叫作Tevatron，它建造于伊利诺伊州的巴达维亚（Batavia）。Tevatron的能标可以达到2TeV（2000倍于质子的静能量）。^[27]质子与反质子对撞在一起，产生其他粒子，而我们可以详细研究它们。Tevatron发现的最重要的事物是顶夸克，它是标准模型中最重，也是最后一个被找到的粒子。

然而，大型强子对撞机与欧洲核子研究中心中的第一个对撞机或者Tevatron都有所不同（表6-1简述了不同对撞机之间的比较）。大型强子对撞机使用两束质子，而非质子束与反质子束进行对撞。大型强子对撞机如此选择的理由十分巧妙，值得我们在此讨论。最取巧的对撞方式是让参与对撞粒子的净电荷为零。这种类型的对撞我们已经在之前的章节中讨论过了。如果有足够的能量，那么在净电荷为零的情况下，就可以产生任何粒子与它的反粒子。如果两个电子对撞，不管产生什么，它的净电荷一定是-2，这限制了很多可能结果的出现。也许你会认为，让两个质子对撞是同等的下策。毕竟，两个质子的净电荷是2，这似乎并没有什么本质区别。

如果质子是基本粒子，这个说法确实是正确的。然而，正如我们在第5章里所讨论过的，质子是具有子结构的。质子由夸克构成，它们被胶子束缚在一起。即便三个带有电荷的价夸克（两个上夸克与一个下夸克）是存在于质子内部的全部事物，事情也没有那么美好——任意两个价夸克的净电荷都是非零。

表6-1 不同对撞机之间的比较

加速器名称 / 建造年代 / 实验室及地理位置	对撞粒子	加速器形状	能标大小
斯坦福直线加速器 (SLC) 1989 年 SLAC, 加州门洛帕克	电子 & 正电子	直线加速器 	100 GeV 3.2 km
Tevatron 1983 年 费米实验室, 伊利诺伊州 巴达维亚	质子 & 反质子	环形加速器 	1 960 GeV 6.3 km
大型电子 - 正电子对撞机 (LEP/LEP2 代)* 1989 年/2000 年 CERN, 瑞士日内瓦	电子 & 正电子	环形加速器 	90 GeV/ 209 GeV 26.6 km
大型强子对撞机 (LHC) 2008 年 CERN, 瑞士日内瓦	质子 & 质子	环形加速器 	7 000 GeV~ 14 000 GeV 26.6 km

*注: LEP被升级为了LEP2代。

质子带有的绝大部分质量并非源于夸克自身的质量。质子的质量主要取决于把质子束缚为一个整体的能量。具有极高动量的质子同样具有极高的能量。除了三个载荷的价夸克之外, 与这些能量共存的是质子中包含一片夸克、反夸克与胶子的海洋。这意味着, 如果你试图探测高能质子, 你不仅将找到三个价夸克, 还有一片净电荷为零的夸克、反夸克与胶子的海洋。

因此, 当考虑质子对撞时, 在应用属于电子对撞的那些逻辑时, 我们必须加以小心。我们感兴趣的事件是那些子结构对撞的结果。这些对撞涉及子结构, 而非质子的载荷。即便夸克、胶子的海洋对质子的净电荷并无贡献, 但它们各自确实带有电荷。当质子被对撞在一起

时，有可能质子中的某个价夸克撞上了另一个，然而这个碰撞的净电荷却非零。只要这个事件的净电荷不消失，我们感兴趣的、涉及正确数目电荷的事件就会出现，然而这个对撞并没有零净电荷的对撞所具有的广泛能力。

然而，由于虚粒子海洋的存在，很多有趣的碰撞都有可能发生。夸克可能撞上反夸克，胶子可能撞上胶子，这与零净电荷的对撞大有不同。当质子被对撞在一起时，某个质子中的夸克有可能撞上另一个质子中的反夸克——虽然这在大多数情况下不会发生。所有可能发生的过程（包括那些来自粒子海碰撞的过程），都在我们针对大型强子对撞机中发生了什么的疑问中扮演着重要的角色。事实上，质子被加速到的能标越高，粒子海中的对撞就越有可能发生。

质子的总电荷并不能决定可能产生的粒子，因为质子的其他部分避开了碰撞，一往无前。质子不参与碰撞的部分把质子其余的净电荷带走了，它们在粒子束管中消失了。这正是帕多瓦市长扎诺那多提出过的巧妙问题，即，在大型强子对撞机的质子对撞过程中，质子的电荷去了哪里。这与质子具有子结构的本质以及对撞的高能标都息息相关——后者保证了只有已知最小元素（夸克、胶子）直接进行对撞。

因为只有质子的一部分参与了对撞，以及那些部分可能是以零净电荷对撞的虚粒子，所以质子-质子对撞机与质子-反质子对撞机之争并没有那么重要。以往，在更低能标的对撞机上，使用反质子参与对撞以保证我们感兴趣的结果会出现绝对物有所值；而在大型强子对撞机的能标下，这并非是一个显而易见的选择。在大型强子对撞机将达到的高能标上，质子很重要的一部分能量是由夸克、反夸克、胶子的海洋所携带的。

服务于大型强子对撞机的物理学家与工程师们最终决定，使用两个质子束而非一束质子与一束反质子进行对撞。^[28] 这让产生很高的光

度（意味着更多对撞的发生）变得更加可能。而且，制造质子束比制造反质子束要容易许多。

所以，大型强子对撞机是质子-质子对撞机，而非质子-反质子对撞机。它有着极大的潜力，许多更加轻而易举就可以完成的质子-质子对撞将在它上面完成。



2009年12月1日早上6:00，在巴塞罗那机场附近的万豪国际酒店，我十分不情愿地起床以赶上早班机。那时，我刚刚参加了一个小西班牙歌剧的首映式，剧本由我撰写，是有关物理学与科学探索的。那个周末虽然让我很满意，但我非常疲倦，期望早点回到家中。然而，我的行程却被一个令人愉快的惊喜暂时拖延了。

那天早上，酒店放在我门口的报纸上的头条新闻赫然写着“原子加速器刷新能标新纪录”。一般情况下，头条新闻的内容都是关于灾难或者某些旋兴旋灭的猎奇事件的报道。然而这次，“大型强子对撞机前几日在创纪录的能标上运行”这件事情取而代之，成了那天最重要的新闻。在那篇文章的笔锋中，大型强子对撞机里程碑式的成就所带来的兴奋之情显而易见。

几周之后，当两束高能质子真正对撞之后，《纽约时报》在头版上以一整版的篇幅发表了一篇题为《对撞机创下新纪录，欧洲走在了美国的前面》的文章。^[1]前几日新闻报道中的能标纪录，很快就要成为近10年来大型强子对撞机将要创立的一系列里程碑中最早的一块。

大型强子对撞机正在探索我们目前研究过的最小尺度。与此同时，人造卫星与天文望远镜正在探索宇宙的最大尺度，研究宇宙膨胀的加速度，并探查作为大爆炸遗留物的宇宙微波背景辐射的一些细节。[书·免`费`分`享V.信shufoufou]

我们现在已经对“宇宙的构成”这一主题有了很多了解。然而，在绝大多数理解的过程中，随着知识的增长，会出现更多问题，其中一些问题还暴露了我们理论框架中的一些重要空白。尽管如此，在很多情况下，我们都能足够好地理解那些缺失环节的本质，以知晓需要寻找什么、如何寻找。

让我们更进一步看看什么即将来临吧——实验有了什么结果？我们期望通过它们找到什么？本章包含了本书其余部分将要探索的一些主要问题以及物理学的研究内容。

超越标准模型，粒子质量如何非零

粒子物理学标准模型告诉了我们，应当如何预测那些构成轻粒子的行为，它也可以描述具有相似相互作用的更重的粒子。通过与构成我们身体或者太阳系的粒子所参与的相同的力，那些重粒子与光、原子核相互作用。

弱相互作用力

又称弱核力，是由W及Z玻色子的交换（即发射和吸收）所引起的。这种发射最有名的就是 β 衰变。

物理学家已经对电子，以及比电子更重而载荷相同的粒子（ μ 子与 τ 子）有所了解。我们知道，那些被称作轻子的粒子与被称为中微子（neutrino）的中性粒子（不带电，因此不直接参与电磁相互作用）结对，仅仅通过弱相互作用力实现相互作用。弱相互作用力负责把中子转变为质子的放射性 β 衰变（和一般意义上的原子核 β 衰变）以及一些存在于太阳上的核反应过程。所有标准模型涉及的事物都要经历弱相互作用力。

我们还了解了在质子与中子中找到的夸克。夸克同时经历弱相互作用力、电磁相互作用力以及强相互作用力。强相互作用力是在质子

与中子之内把夸克约束在一起的力，它涉及很多复杂的计算，然而我们却理解它的基本结构。

夸克、轻子与强相互作用力、弱相互作用力、电磁相互作用力共同构成标准模型的实质（图7-1为粒子物理学标准模型简图）。

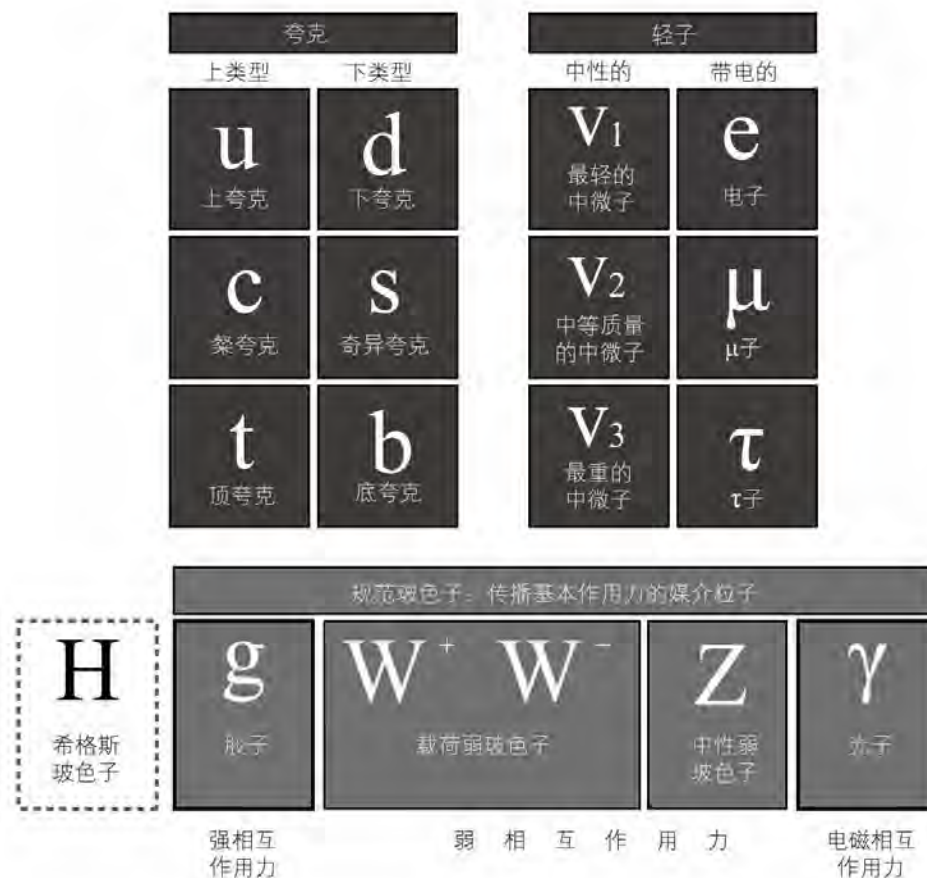


图7-1 粒子物理学标准模型中的元素描述了物质的已知基本元素及其相互作用。上夸克与下夸克参与强、弱、电磁相互作用力，载荷轻子参与弱相互作用力与电磁相互作用力，而中微子只参与弱相互作用力。胶子、弱规范玻色子与光子传递这些作用力，希格斯玻色子还没有被找到。

使用这些材料，物理学家已经成功预测了迄今为止所有粒子物理学实验的结果。我们已经很好地理解了标准模型中的粒子以及它们之间的力如何相互作用。^[29]然而，依旧有一些谜题被遗留了下来。

这些挑战的主要问题是：如何把引力纳入标准模型中？这是一个大型强子对撞机有机会去探索，然而不可能确保解决的大问题。即便从最近在地球上取得成就，以及未来将要处理的下一个谜题的视角来看，大型强子对撞机的能标已足够高；但是想要最终回答有关量子引力的问题，它还是太低。为了回答量子引力的问题，我们必须研究极小的尺度，在那些尺度上，引力效应和量子力学效应将会同时出现——而这个尺度远远超出了大型强子对撞机之所能及。如果我们足够幸运，并且引力确实在定位与质量相关的粒子问题上扮演了重要角色，那么我们将能更好地回答上面这个问题，大型强子对撞机也许能揭示有关引力和空间本身的重要信息。否则，对任何形式的量子引力理论（包括弦理论）的实验验证，都有可能还差得很远。

然而，在这一点上，引力与其他基本作用力的关联是唯一尚未得到回答的问题。我们理解中的另一个重要空缺，也是大型强子对撞机最终准备解决的问题是：基本粒子如何获得质量？这听起来是一个很奇怪的问题（除非你读过我的《弯曲的旅行》一书），因为我们倾向于认为：物体具有质量是天经地义的——这是一个粒子固有的、不可剥夺的性质。

从某种意义上来讲，这个观点是正确的。质量、电荷与相互作用分别是确定某个粒子的性质之一。粒子总是具有非零能量，然而质量却是一个内蕴性质，它可以取很多包括零值在内的可能值。爱因斯坦的一个主要洞见是：在静止时，粒子质量的数值可以决定它能量的数值。但粒子的质量并不总是取非零值。那些零质量^[30]的粒子，比如光子，从来都不会静止。

然而，基本粒子具有非零质量是它们拥有的内蕴性质，也是一个巨大的谜团。并非只有夸克、轻子与弱规范玻色子（传递弱相互作用力的粒子）具有非零质量。实验物理学家测定了它们的质量，然而即

便是最简单的物理规律也不允许它们具有质量。如果我们假定粒子确实具有那些质量，那么标准模型的预言将会成立，但我们却不知道它们最初是从哪里来的。显然，最简单的规则不适用，某些更加微妙的事情存在着。

粒子物理学家相信，非零质量仅仅因为在早期宇宙中，某些事物在希格斯机制的过程中会引人注目地出现。这个机制得名于物理学家彼得·希格斯，他率先发现了质量如何产生。至少6名物理学家提出了相似的观点，所以，也许有时你会听到“恩格勒-布绕特-希格斯-古拉尔尼克-哈根-基博尔机制”这种表述，而我坚持使用“希格斯玻色子”这一名称。^[2]这个理论是一种相变（phase transition，也许与液态水沸腾而变成蒸汽的相变类似），它的发生改变了宇宙的本质。然而在更早的时候，没有质量、乘着光速之辘呼啸的粒子在不久之后——在涉及希格斯场（Higgs field）的相变之后，就具有了质量，而且速度慢了下来。希格斯机制告诉我们，基本粒子如何从缺乏希格斯场的零质量，获得我们实验中测量到的非零质量。



如果粒子物理学家们是正确的，而且希格斯机制确实在宇宙中运行，那么大型强子对撞机将会揭露出一些泄露宇宙秘密的迹象。在它最简单的成就中，这种迹象是一个粒子——即著名的希格斯玻色子。在更详尽的、希格斯机制参与其中的物理理论中，希格斯玻色子也许会伴随着其他具有相似质量的粒子出现，也有可能完全被其他粒子替代。

希格斯机制

物理学家彼得·希格斯和其他几位物理学家共同发现的一种物理机制。这一机制解释了基本粒子获得非零质量的原因。

除了希格斯机制如何生效之外，我们还希望大型强子对撞机能找到一些有趣的事物。它也许会是希格斯玻色子，也可能是某个更奇异理论的证据，比如“技术色”（technicolor，我们将在稍后讨论它）。或者，它还可能是一些完全无法预料到的事物。如果一切都按计划进行，那么大型强子对撞机上的实验将要辨明，到底是什么让希格斯机制生效的。不管它找到了什么，这个发现都将告诉我们一些有关粒子如何获得其质量的一些有趣的事情。

粒子物理学标准模型漂亮地运行着，它描述了物质最基本的元素及其相互作用，其预言在很高的精度上被反复验证。希格斯粒子是标准模型中最后一个尚未找到的部分。^[31]我们现在假定粒子拥有质量。然而当理解了希格斯机制时，我们也会明白质量是如何产生的。我们将在第16章中进一步探讨的希格斯机制，对更令人满意地理解质量而言是很重要的。

粒子物理学中还有另外一个更大的谜题，大型强子对撞机也许有助于解开它。大型强子对撞机上的实验有可能阐明被称为粒子物理学“等级问题”的答案。希格斯机制专注于为什么基本粒子具有质量的问题，而等级问题关心为什么它们具有的质量是现在这个数值。



粒子物理学家们不仅相信质量的出现是因为弥漫在宇宙中的“希格斯场”，还相信我们知道在粒子从零质量变为非零质量的过程中出现的能量。这是因为，希格斯机制以某种可预测的方式赋予某些粒子

质量，而这种方式只取决于弱相互作用力的强度以及质量转变过程中出现的能量。

奇怪的是，从根本理论的视角来看，这种能量转变并不真的合理。如果你把量子力学与狭义相对论的结论放在一起，那么你确实能计算出对粒子质量的贡献，但是理论值远远高出实测值。建立在量子力学与狭义相对论基础上的计算告诉我们，如果没有进一步的理论，那么质量远比现在的实测值要大——事实上，要大出 10^{16} 倍。只在某些物理学家所谓“精细调节”的一派胡言中，这两个理论才不矛盾。

粒子物理学的等级问题是对物质的基本描述提出的最大挑战。我们想要知道，物质为什么表现得与我们的预期大不相同。出于量子力学的计算让我们相信，它们应该比决定其质量的弱能标要大很多。我们在表面上最简单的标准模型版本中，对弱能标理解的无能为力是通往完整理论的一块绊脚石。

可能的情况是，一个更有趣、更精妙的理论包含了这个最简单的模型——于物理学家而言，比起“精细调节”过的理论，这将是一个更加引人入胜的理论。尽管已经有了一系列雄心勃勃的问题追问到底哪个理论解决了等级问题，大型强子对撞机还是很有可能充分地理解它。量子力学与相对论不仅支配了对质量的贡献，还支配了新现象一定会在其上出现的能标。这个能标正是大型强子对撞机将要探测的。

我们预期，在大型强子对撞机上将要出现一个更有趣的理论。当新的粒子、力或者对称性出场时，有关质量之谜的理论也将呈现出来，这是我们希望大型强子对撞机实验揭开的最大秘密之一。

这个答案本身非常有趣，然而它也有可能是洞察自然其他方面本质的一把钥匙。对这个问题最有启发性的两个答案包含了时空对称性的推广，或者对空间概念本身的修正。

本书第17章将进一步告诉我们，空间有可能包含比我们熟知的上-下、左-右、前-后这三个维度更多的维度。特别是，它可能包含完全看不到的维度，而这些维度是理解粒子性质与质量的关键。如果情况确实是那样，那么大型强子对撞机就将以某种形式的粒子为这些维度提供证据，这些粒子被称为卡鲁扎-克莱因（Kaluza-Klein, KK）粒子，它们在高维时空中穿梭。

不管到底是哪个理论最终解决了等级问题，它都将在弱能标下提供实验可观测的证据。一系列理论逻辑将把在大型强子对撞机上找到的东西与最终解决这些问题的事物联系在一起。它可能是我们预料到的某些东西，也可能不是，但不管怎样，它都将非常精彩。

探秘暗物质

除了那些粒子物理学问题之外，大型强子对撞机还有助于阐明宇宙中暗物质的本质。暗物质是参与引力作用，但既不吸收光也不发射光的物质。我们看到的每一件事物——地球、你正坐着的椅子或者是你的长尾小鹦鹉宠物，都是由与光相互作用的标准模型中的粒子构成的。



我们已经理解的可见的、与光互动的物质以及它们的相互作用，只包含宇宙中占总量4%的能量密度。大约23%的能量由某种叫作“暗物质”的东西携带，它的存在还没有被确认。

暗物质的确是一种物质。也就是说，它在引力的作用下聚集成团，从而（与正常的物质一起）形成某些结构体，例如星系。然而，与我们熟悉的物质（比如构成我们人类、构成天上星星的物质）不同

的是，它既不发射光也不吸收光。因为我们通常通过被发射或者吸收的光来看事物，所以暗物质不容易被“看到”。

暗物质

暗物质是参与引力作用，但既不吸收光也不发射光的物质。它能够穿越电磁波和引力场，是宇宙的重要组成部分。

事实上，“暗物质”这名字并不恰当——所谓的“暗物质”并不真正是暗的。“暗的”物质吸收光线，当光线被吸收时，我们可以真正看到“暗的”物质。从另一方面来讲，暗物质不以任何可观测的方式与光相互作用。从技术角度说，“暗”物质实际上是透明的。不过我将继续使用人们已经习惯的称谓来描述这种奇异的物质。

我们通过引力效应得知了暗物质的存在。但是，如果不能直接看到它，那么我们就不会知道它的性质。它是由大量微小的全同粒子（*identical particle*）构成的吗？如果是，这些粒子的质量几何？如何进行相互作用？

也许我们很快就会对暗物质了解更多。引人注目的是，大型强子对撞机也许处在了正确的能标上，它可以产生可能是暗物质的粒子。暗物质的关键准则在于，宇宙应该正好包含了精确数量的暗物质，以形成我们现在所观测到的引力效应。换言之，“残留密度”（*relic density*），即宇宙模型中预言的、存留到今日的能量储存数值必须与测定值相符。而惊人的事实是，如果你有一个稳定的粒子，其质量符合大型强子对撞机将要探索的弱能标（根据 $E=mc^2$ ），其相互作用也涉及那个能标下的粒子，那么它的残留密度应该处于可能成为暗物质的大致范围之内。

因此，大型强子对撞机不但可以给我们提供一些有关粒子物理学问题的洞见，还可以给我们提供一些今日的宇宙中有什么、一切是如

何开始的线索。这些已经成为宇宙科学一部分的问题可以告诉我们宇宙是如何演化的。

就基本粒子及其相互作用而言，我们理解了惊人的宇宙历史。然而，同样就粒子物理学而言，一些大问题依旧存在。主要问题有：

- 暗物质是什么？
- 神秘的“暗能量”（dark energy）又是什么？
- 是什么导致了早期宇宙的体积在一段时期之内以“宇宙暴胀”（cosmological inflation）的形式呈指数级增长？

当下是一个极好的观测时间，它也许可以告诉我们这些问题的答案。暗物质研究是粒子物理学与宇宙学交叉区域最前沿的研究方向。暗物质与普通物质（我们可以由探测器制造的物质）的作用极度微弱，微弱到我们现在还没有找到任何引力效应之外的暗物质存在的证据。

因此，当前的研究只能期望“天降神迹”，虽然暗物质几乎不可见、几乎不与其他已知物质相互作用，但相互作用的强度还没有弱到完全无法探测。这种说法并非仅仅是我们的一厢情愿，它建立在计算的基础上。我们在上文中曾经提到过这种计算，它指出：稳定的、相互作用的、与大型强子对撞机将要探索的弱能标相关的粒子，其密度与暗物质相符。我们希望，即便现在还没有确认暗物质的存在，在不远的未来我们也会有很大的机会探测到它。

但多数宇宙学实验并不在加速器中进行。地球上专门向外部空间探索的实验主要用于确定、增强我们对宇宙学问题可能解决方案的理解。

例如，天体物理学家们把人造卫星发射到太空中，这样就可以在一个不被地表附近的灰尘以及物理、化学过程遮蔽的环境中观测宇宙。地球上的望远镜与实验在一个科学家们可以更直接控制的环境里给了我们一些额外的洞见。这些在太空与地球进行的实验已经做好准备，以阐释“宇宙为什么会是现在这个样子”这一问题。

我们希望，这些实验中某个充分强的信号（我们将在第21章中讨论）可以让我们破译暗物质之谜。这些实验可以告诉我们暗物质的本质，并且阐明它的相互作用与质量。与此同时，理论物理学家们正在努力思索所有暗物质的可能模型，以及如何使用所有已知的探测方法以研究暗物质到底是什么。

暗能量，比暗物质更神秘的存在

普通物质与暗物质还不能提供宇宙中能量的总量——它们总共只占27%。比暗物质更神秘的存在是被称为“暗能量”的东西，它构成余下73%的能量。

暗能量的发现是20世纪末期意义最深远的一记“警钟”。尽管我们对宇宙的演化尚有很多不清楚之处，但我们已经在所谓的“大爆炸理论”及其补充理论“宇宙暴胀”（宇宙体积呈指数级增长的某个时期）的基础上，对宇宙的演化过程有了非常深刻的了解。

暗能量

暗能量是宇宙学及物理学领域的一种猜想。它是一种充溢空间的、增加宇宙膨胀速度的难以察觉的能量形式。

大爆炸理论与一系列观测结果相符，包括对宇宙中微波辐射的观测，即从大爆炸时残余下来的微波背景辐射。宇宙原本是一个炎热、致密的火球。但在137.5亿年之后，它大幅地变得稀疏、冷却，留下了今日仅有2.7K的、冷却下来的辐射——仅比绝对零度^[32]高出几摄氏

度。大爆炸理论中暴胀的其他证据可以通过对早期宇宙演化过程中产生的大量原子核的研究，以及对宇宙膨胀本身的测量来找到。

那些我们用以计算出宇宙如何演化的根本方程由爱因斯坦在20世纪早期提出。这些方程告诉我们，在给定物质分布或能量分布的前提下，如何导出引力场。它们可以应用于描述太阳与地球之间的引力场，也可以用来描述整个宇宙。就一切情况而言，为了导出这些方程的结果，我们必须知道周围物质与能量的分布情况。



令人震惊的观测结果是，对宇宙特征的现有测量结果需要引入某种不被物质携带的新能量形式来解释。这些能量不由粒子或者其他组分携带，也不像通常的物质一样聚集成团。它不随宇宙的膨胀变得稀薄，而是保持恒定的密度不变。由于这些遍布宇宙的、谜一般的能量，宇宙正在缓慢加速膨胀，即便这些能量没有物质载体，一片空无。

爱因斯坦最初提出了一种被他称作“普适常数”（universal constant）的能量形式，之后被其他物理学家称作“宇宙学常数”（cosmological constant）。不久之后，爱因斯坦就认为这是一个错误概念。他试图用这个概念解释“为什么宇宙是稳态的”这一想法的确是走上了歧途。宇宙确实在膨胀，在爱因斯坦提出这个观点后不久，爱德温·哈勃（Edwin Hubble）就证实了这一点。膨胀不仅真实存在，现在看来，它的现有加速度还与爱因斯坦于1930年提出的、而又很快抛弃掉的有趣能量类型密不可分。

我们想要更好地理解谜一般的暗能量。针对这一点的观测被设计为可以确认如下问题的答案：暗能量到底只是爱因斯坦最先提出的那种背景能量，还是某种随时间改变的新能量形式；或者是某种人们完全没有预料到，以至于我们现在甚至不知道如何去推断其性质的事物？

寻找更多问题的答案

这只是一个关于我们现在正在研究什么的示例（虽然是重要的例子）。除了我已经描述过的那些，更多的宇宙学研究还在筹划之中。引力波探测器将探测从吞噬一切的黑洞中传出的引力波，以及其他令人兴奋的、涉及大质量与高能标的现象。^[33]宇宙微波实验将要告诉我们宇宙膨胀的更多细节，对宇宙射线的寻找将告诉我们有关宇宙内容的新细节，红外线探测器也许可以找到宇宙中更新奇的事物。

在某些情况下，我们将足够好地理解观测结果，以知晓对物质与物理定律的底层本质而言，它们意味着什么。对于其他情况，我们将要花费大量的时间以揭露其蕴意。不管发生了什么事情，理论与实验数据之间的相互影响将把我们带往对环绕着我们的宇宙的更高级解释，并把知识扩张到我们目前尚不能及的领域。

某些实验也许很快就会有结果，另一些可能会在多年之后才会有结果。当实验数据产生后，理论物理学家们将不得不重新考虑，有时甚至需要放弃已提出的解释，这样我们就可以改善并正确应用我们的理论。这听起来很令人沮丧，但事情并不像你想象的那么坏。当实验结果引领着研究方向、保证我们确实有所进展时，我们渴望能预期一些线索，以帮助我们回答问题——即便接受那些新结果有可能意味着我们要放弃旧观念。我们的假说最初根植于理论的一致性与优美性，

然而，正如我们接下来将看到的遍及全书的理念：是实验结果，而不是死板的信仰，最终决定了什么是正确的。



第三部分

大型强子对撞机的传奇历程

**KNOCKING ON
HEAVEN'S DOOR**



我不是一个爱说大话的人，因为我发现，伟大的成就往往会为它们自身代言。

在美国，言不由衷的溢美之词往往会给我带来麻烦。人们滥用“最……”等表述方式，就好像不用“最”的单纯赞美之词，会被误读为似褒实贬的春秋笔法。我经常被建议，在我的研究资金申请陈述中应该适当加入一些漂亮的行话或者修饰词，以避免任何形式的误解。但是，就大型强子对撞机而言，我不吝冒着巨大的风险说：毫无疑问，大型强子对撞机是一个惊人的成就。它有着神秘的权威与美丽，它是人类技术的结晶。

在本章，我们将着手探究这种令人难以置信的机器；在下一章，我们将进入到其如过山车一般的构造历程。之后几章，我们将进入记录着大型强子对撞机所创造事物的世界中。最终，我们将要专注于大型强子对撞机本身。它分离、加速高能质子，并把它们对撞在一起，以期解释物质内部的新世界。

大型强子对撞机，惊为天人的艺术品

我第一次看到大型强子对撞机时，就为它体现的一种敬畏感而惊奇——尽管我之前已经见过了许多粒子对撞机与探测器，但它的尺度

却远远不同。我们进入仪器内部，戴上安全帽，走入、穿过大型强子对撞机的隧道，在一个最终将放置ATLAS（A Toroidal LHC Apparatus，超环面仪器）探测器的深井面前停下，并最后走到了实验仪器本身面前。它还在建造之中，这意味着超环面仪器还没有如它运行时那样被覆盖起来——我们可以窥见它的全貌。

即便在一开始时，包括我在内的科学家们都对这个难以置信的精密的技术奇迹感到惶恐——它更像一件艺术品；然而我依旧难以遏制住自己的激动之情，拿出了照相机到处拍照。很难用语言来描述它的复杂性、一致性与巨大的“体格”，以及那些交错的线条与交叠的颜色。总之，唯一的印象是：它令人敬畏。

看到这件艺术品的人都会有相似的反应。当艺术收藏家弗朗西斯卡·冯·哈布斯堡（Francesca von Habsburg）观看大型强子对撞机时，她带来了一名职业摄影师。这名摄影师的摄影作品美轮美奂，曾刊载于《名利场》杂志上。当出身于书香门第的电影制作人杰西·迪伦（Jesse Dylan）首次看到大型强子对撞机时，他将其描绘为一件卓越的艺术品——一个他希望与他人分享其壮美的“终极成就”。迪伦制作了一期视频，以表达他对大型强子对撞机以及在其上进行的实验之壮丽的感叹之情。

演员、科学爱好者艾伦·艾尔达（Alan Alda）在主持一场有关大型强子对撞机的座谈会时，把它与古代世界的某座奇观联系了起来。物理学家戴维·格罗斯把它与金字塔相比。PayPal、特斯拉汽车公司的联合创始人以及太空探索科技公司SpaceX（制造把机械与货物运输到国际空间站中的火箭）的创办者、工程师、企业家埃隆·马斯克（Elon Musk）如此评论大型强子对撞机：“这无疑是人类最伟大的成就之一。”

我经常听到各行各业的人们都如此评论大型强子对撞机。网络、高速汽车、绿色能源、太空旅行都是今日应用科学中最为活跃且令人振奋的领域。然而，着手理解宇宙的基本定律却是一个独立的范畴，

它予人深刻的印象，又令人震惊。艺术爱好者与科学家们都想要理解世界、解释它的起源。你也许会对“人类最伟大成就”的本质有所争议，但我想，任何人都不会质疑，我们能做的最卓越的事情是：深思并研究隐藏于简单事物之下的复杂本质。只有人类承担着这样的责任。

我们将在大型强子对撞机中研究的对撞过程类似于大爆炸后第一万亿分之一毫秒内的过程。这个过程将会告诉我们在宇宙形成的早期，有关物质与相互作用力的本质以及超小尺度的一些事情。你可以把大型强子对撞机看作一个超级显微镜，它可以让我们在极小的尺度上看到粒子与相互作用力——这个数量级是一毫米的一亿亿分之一。

大型强子对撞机通过创造地球上能标之高前所未有的高能粒子对撞，来进行这些小尺度研究——这个能量的上限是位于伊利诺伊州巴达维亚的对撞机Tevatron的7倍。正如在第6章解释过的那样，量子力学与其在波动理论上的应用告诉我们，这些能量对研究如此小的尺度而言是非常重要的。而且，随着能量的增长，光度^[34]也会比Tevatron上的实验高50倍，这使那些能揭示自然内部运作机理的小概率事件更加可能被我们发现。

尽管我不喜欢滥用溢美之词，然而我不得不说，大型强子对撞机属于一个只能用一系列“最”来描述的世界。它不能仅称“巨大”：大型强子对撞机是人类建造过的最大的机器；它不能仅称为“寒冷”：大型强子对撞机中的超导磁铁需要在1.9K（即比绝对零度高出1.9℃）的温度下运行，这是我们在宇宙中已知的最冷的大范围区域——它甚至比太空还要寒冷；它的磁场不能仅称为“强大”：大型强子对撞机中的超导二极磁铁产生的磁场比地磁场强度大出10万多倍，这是人类工业制造过的最强磁铁。

大型强子对撞机还有数宗“最”。容纳质子的管道中的真空，其气压为大气压的十万亿分之一，是在大范围区域内制造过的最完全的真空。对撞的能标是地球上使用过的最高能标，它让我们得以向回追溯最远的时间，以研究早期宇宙中出现的相互作用。

大型强子对撞机拥有着巨大的能量。它的磁场本身就蕴含着相当于数吨TNT炸药的能量，而其中参与对撞的粒子束蕴含着前者1/10的能量。这些能量蕴含在十亿分之一克的物质中，而后者只是日常环境中亚微观结构的一粒小灰尘物质而已。当机器中的粒子束整装待发时，高度集中的能量都被倾入一个石墨罐中，这个罐长8米、直径1米，由1000吨混凝土包裹起来。

这些大型强子对撞机之最把技术推到了极致。这些技术花费不菲，每一项“最”都意味着一笔大额开销。大型强子对撞机90亿美元的身价也使它成为人类史上最烧钱的机器。欧洲核子研究中心承担了这笔费用的2/3，从德国的20%到保加利亚的0.2%，根据自身的经济实力，欧洲核子研究中心的20个会员国都各自作出了不等的贡献。余下的费用被非会员国承担，包括美国、日本与加拿大。欧洲核子研究中心为实验本身付费20%，这是由国际合作组织资助的。截至2008年，当机器最终制造完毕时，美国已经有1000多名科学家为CMS（Compact Muon Solenoid，紧凑 μ 子线圈）与ATLAS工作，并且为大型强子对撞机这项事业贡献了5.31亿美元。

25年，从设想到现实

大型强子对撞机的所在地欧洲核子研究中心是一家科研机构，它同时运作着许多项目。然而，欧洲核子研究中心的资源都集中在一些核心项目上。20世纪80年代的核心项目是SpbarpS对撞机^[35]，它最终找到了于粒子物理学标准模型十分重要的相互作用力传播粒子。1983

年所进行的主要实验找到了弱规范玻色子（两种载荷W玻色子与中性Z玻色子），它们是传播弱相互作用力的粒子。它们是当时的标准模型中缺失的关键元素，该项目的领导人最终由于该发现而获得了诺贝尔奖。

即便如此，当SpbarpS运行时，科学家与工程师们已经在计划大型正负电子对撞机的运行了。后者将把电子与它的反粒子（正电子）对撞在一起，以研究关于弱相互作用力与标准模型更加精密的细节。这个梦想实现于20世纪90年代。那时，通过一系列精确的测量，大型正负电子对撞机研究了数百万个弱规范玻色子，它们让物理学家们掌握了许多有关标准模型中物理相互作用的知识。

大型正负电子对撞机是一个周长为27公里的环形加速器。电子与正电子绕着这个加速环的轨道运行，并且不断地被加速。正如我们在第6章中讨论过的一样，环形加速器在加速如电子这样的轻粒子时效率会比较低，因为这种粒子在环形轨道上被加速时会辐射出能量。大型正负电子对撞机能标（约100GeV）下的电子束每环行一圈就失去大约3%的能量。这个损失量并不是很大，然而，如果我们希望把电子加速到更高的能标，那么在每一圈中损失的能量就会变成“破坏者”。能量每放大10倍，能量损失就会放大10000倍，而这将让加速器变得太过低效而没有实际意义。

出于这个原因，当大型正负电子对撞机还在筹划之中时，人们就已经开始思考欧洲核子研究中心的下一个核心项目——据推测，它将在更高的能标上运行。因为使用电子会造成不可接受的能量损失，如果欧洲核子研究中心想要建造一个以更高能标运行的机器，那么它将使用质子束。质子束的质量更大，因而辐射更少。建造大型正负电子对撞机的物理学家与工程师们非常清楚这种更美好的前景，所以他们把大型正负电子对撞机的隧道建得足够宽，以在这个电子-正电子对撞机被拆除后，能够容纳未来可能建造的质子对撞机。

最终，25年之后，质子束在本来为大型正负电子对撞机挖掘的隧道中穿梭了（见图8-1）。

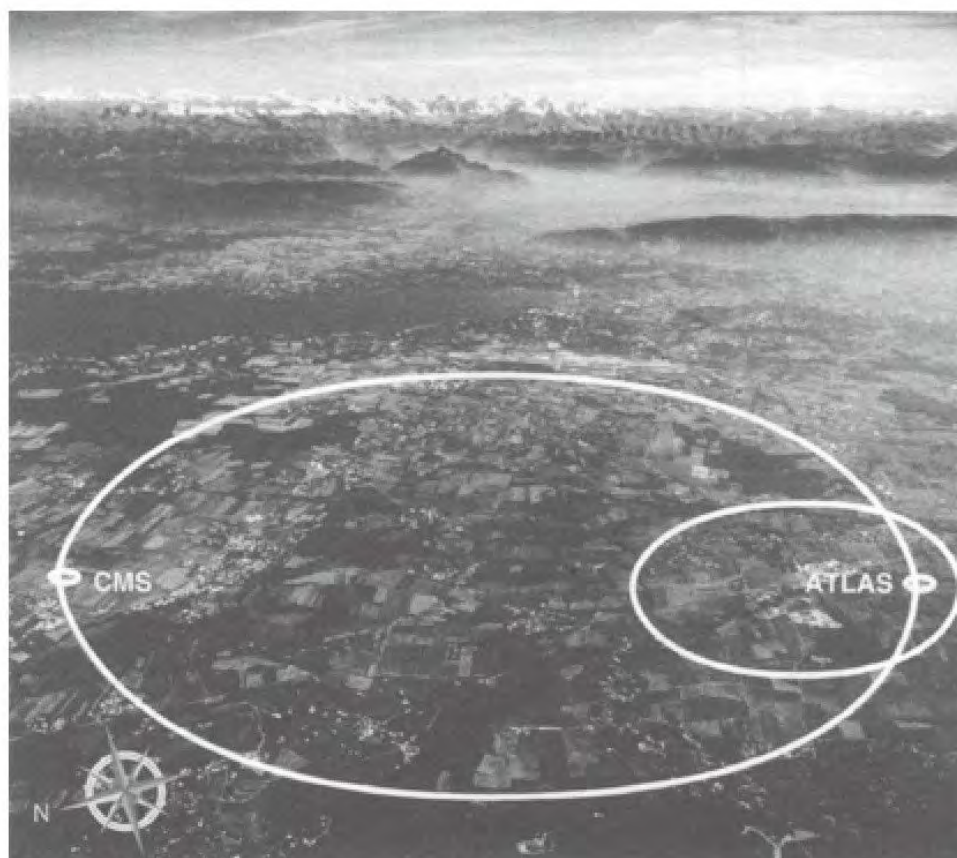


图8-1 大型强子对撞机的所在地，白色圆圈内是它的地下隧道，背景是日内瓦湖与群山。（感谢CERN友情提供图片）

大型强子对撞机比预期拖延了数年建成，其实际开销超出了预算金额约20%。这令人沮丧，但是考虑到大型强子对撞机是人类建造过的最大、最贵、能标最高、抱负最宏大、国际合作最多的实验仪器，这种超预算也是合理的。正如编剧、导演詹姆斯·布鲁克斯（James L. Brooks）在听到大型强子对撞机一系列的好事多磨时戏谑地说：“人们在挑选墙纸时耗费了几乎与此相当的时间。不过理解宇宙说不定也许有更好的效力，那里同样有着某些非常不错的墙纸。”

大型强子对撞机的几宗“最”

质子构成我们，它遍布我们周围。然而，它们通常都被束缚在原子中被电子环绕着的原子核里。它们不能孤立于电子，也不能在粒子束内部被校准（对齐成一系列）。大型强子对撞机最初的任务是分离、加速质子，再把它们导向最终的目标。这个过程就要用到大型强子对撞机的几宗“最”。

准备质子束的第一步是加热氢原子，这可以去掉电子^[36]，分离出孤立的质子（即氢原子核）。磁场控制着这些质子的方向，引导它们成为粒子束。接下来，大型强子对撞机在不同区域中对这些粒子束进行多级加速。在这个过程中，质子从一个加速器转移到另一个加速器，每一次都会增加一些能量。最终，它们形成两个平行粒子束，进而进行对撞。

最初加速的位置在欧洲核子研究中心的直线加速器中，这是一个绵延的隧道，无线电波^[37]沿着它加速质子。当无线电波的强度达到峰值时，其中的电场就会加速质子。接着，质子就会离开场区，因而当场变弱时不会减速^[38]。质子在场强再度达到峰值时回归电场，这样它们就能在峰与峰之间一直被加速。本质上，无线电波跳跃加速质子的方式类似我们推动秋千上的小孩的方法。因此，这些电磁波加速质子、提高它们的能量，不过在第一级加速中，增加的数值非常小。

在下一级加速过程中，质子被磁铁导入一系列加速环，最终在那里被加速。每个加速器的功能都类似于上面描述过的直线加速器。然而，由于下一级加速器是环形加速器，在质子环行数千圈的过程中，加速器可以不断地提高质子的能量。因此，这些环形加速器赋予质子很多能量。

在质子进入大型强子对撞机环之前加速质子的“加速环成员”包括：把质子能标加速到1.4GeV的质子同步推进器（proton synchrotron booster, PSB）、把它进一步提升到26GeV的质子同步加速器（proton synchrotron, PS），以及把它提升到所谓“注入能量”（450GeV）的超级质子同步加速器（super proton synchrotron, SPS）（请看图8-2以了解质子的旅程）。注入能量就是质子在约27公里长的巨大隧道中进入最后一级加速时所携带的能量。

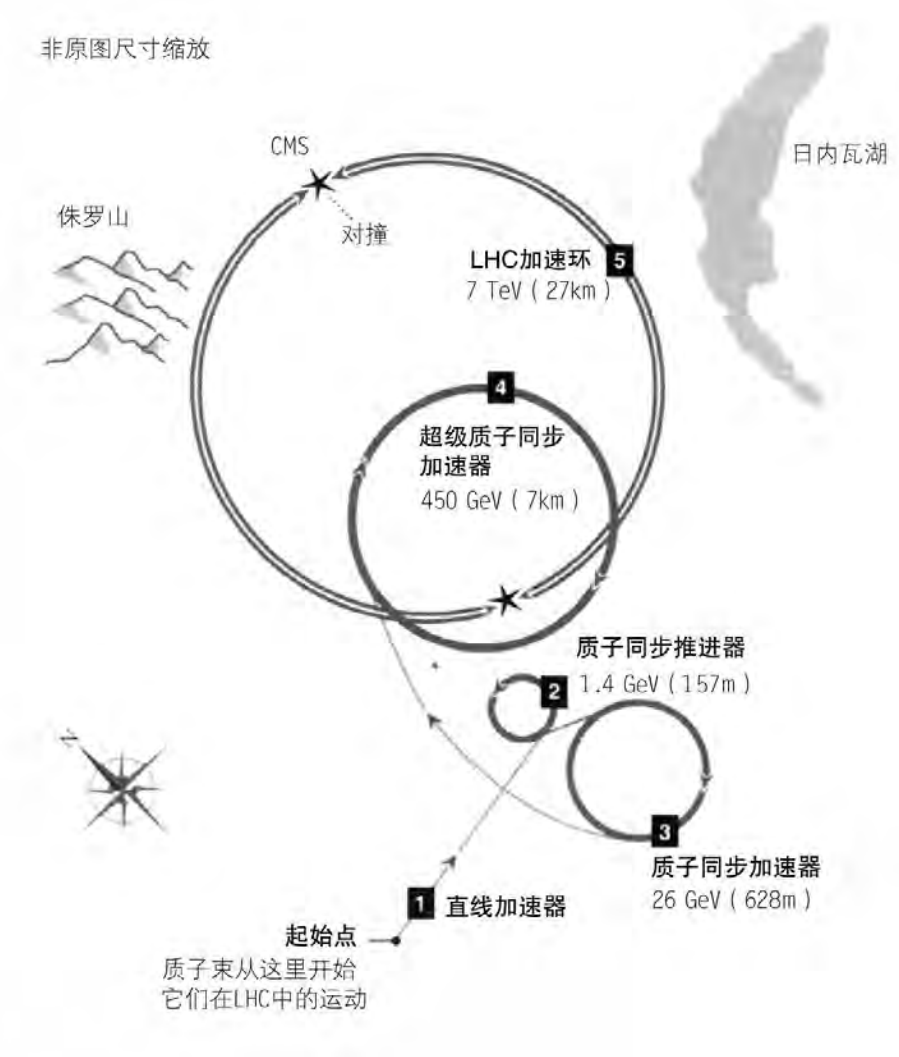


图8-2 被大型强子对撞机加速时，质子的行进路径。

其中一些加速环是欧洲核子研究中心之前的一些项目遗留下来的。最老的质子同步加速器于2009年11月迎来了它的建成50周年庆典，而质子同步推进器对欧洲核子研究中心在20世纪80年代的最后一个核心项目（即大型正负电子对撞机）而言是很重要的。

在离开超级质子同步加速器之后，质子为期20分钟的注入阶段就开始了。此时，从超级质子同步加速器中导出的、450GeV能标的质子在大型强子对撞机隧道的内部被加速到最高的能标。隧道中的两束质子沿着两根分离的、直径0.91米的狭窄管道反向行进，这些管道延伸到27公里长的大型强子对撞机地下加速环中。

这个建造于20世纪80年代、3.8米宽、现在负责质子束最后一级加速的隧道，照明、通风条件良好，空间大到能在其中舒适地环行——我有幸在大型强子对撞机尚在建造之时环行其中。在大型强子对撞机之行中，我仅仅漫步了很短的一段时光，然而我穿越这段隧道所消耗的时间依旧比被加速的高能质子长很多。后者以光速的99.9999991%运行，只需要 $1/890000000$ 秒的时间就可以环绕隧道一圈。

地下隧道的平均深度约为100米，其准确深度在50~175米之间变化。这可以保护其表面不受辐射，而且让欧洲核子研究中心不必买下（并摧毁）隧道所在位置邻域上的全部农田。然而，回溯到20世纪80年代，知识产权纠纷的确延后了这条本来为大型正负电子对撞机建造的隧道的挖掘工程。问题在于，在法国，土地拥有者拥有从地表到地心全部区域的所有权，而不仅仅是他们开垦的土地上。只有法国当局签署下“公用事业宣言”，让土地下的岩石，原则上还有更下面的岩浆都属于公共财产以“祝福”此项工程时，这条隧道才能开始挖掘。

物理学家们经常争论，隧道深度的变化到底是出于地质学原因，还是为了减小电磁辐射而故意为之。事实上，两方面原因兼备。这种参差不齐的地势实际上是对隧道深度与位置的一种有趣的约束。欧洲

核子研究中心的地下区域主要由一种被称为“磨拉石”（molasse）的致密岩石构成。然而，再下面却是一些河流与海洋沉积物，比如砂砾、碎石、含有地下水的黏土，这种地方并不适合建造隧道。地势的起伏让隧道始终处在适合建造隧道的岩层中。这也意味着，位于侏罗山脚下、居于欧洲核子研究中心一隅的隧道某个部分可以建造得更浅一点儿，这样在这个位置的垂直方向上运进、运出建筑材料就会更加方便（也更便宜）。

隧道中最终的加速电场并非精确地按照环形方式安置。大型强子对撞机有着8个大弧，大弧彼此之间又由8个700米长的直线部分相间。8个部分中的每一个部分都可以单独地加热或者冷却，这对于维护和使用而言都很重要。进入隧道之后，质子在每一个短的直线部分被无线电波加速，正如在先前把它们加速到注入能量的加速阶段时一样。这种加速在400MHz的射频谐振腔（radio-frequency cavities）中出现，这个频率与你遥控锁上汽车所用电磁波的频率相同。这个场仅能把进入谐振腔质子束的能量加速到1TeV的485/1000000000。这听起来并不大，然而质子每秒钟绕行大型强子对撞机11000圈。因此，想要把质子束从它的注入能量450GeV加速到目标能量7TeV（后者是前者的约15倍），只需要20分钟的时间。某些质子将会由于相互碰撞或者走偏而丢失，然而大多数质子将会继续环行将近半日，直到质子束中的质子几乎消耗殆尽、必须被送回地面，再被新注入的质子所取代。

在设计上，在大型强子对撞机中环行的质子并非是均匀分布的。它们被送入加速环的2808个团中，每个团中包含1150亿个质子。每个团长10厘米、宽1毫米，团之间的间距约为10米。这于加速过程很有意义，因为每个团都可以单独加速粒子。另外，把质子集成团也是保证质子团以至少25~75纳秒的时间间隔进行相互作用的手段。这段时间足以让每个团中的对撞被单独记录下来。由于在每个团中的质子都比

质子束中的质子少很多，我们可以控制同一时间对撞的次数，因为任意时间对撞的质子都在团中，而不是整体的质子束中。

至关重要的低温二极磁铁

把质子加速到高能标确实是一件显赫的成就。然而，建造大型强子对撞机过程中真正的技术绝技乃是设计、制造能产生强磁场的二极磁铁，它们可以保证质子正确地在加速环中环行。如果没有这些磁铁，那么质子就会沿直线行进。让高能质子在某个环内环行需要巨大的磁场。

由于隧道的长度是给定的，大型强子对撞机的工程师面临的主要技术困难是：如何在工业规模上制造尽可能强的磁铁——这意味着可以批量生产。强磁场是把高能质子约束在曾经放置大型正负电子对撞机的旧隧道中所必需的。让更高能标的质子环行，需要增强磁铁或是扩大隧道，以让质子的轨迹足够精确。于大型强子对撞机而言，由于隧道的尺寸已经确定了，所以目标能标仅由可达到的最大磁场决定。

假使美国的超导超级对撞机（Superconducting Supercollider, SSC）没有中途停建，那么它将要被放置在一个更大的、周长87公里的隧道中（实际上，这台超导超级对撞机是为了这个项目而特地挖掘的）。这台超级对撞机预计达到40TeV的能标，这是大型强子对撞机目标能量的3倍。这个远超大型强子对撞机的能量将可能达到，因为这部机器的建造是白手起家的，它没有已经存在的隧道尺寸的限制，也没有随之而来的对难以达到的强磁场的要求。然而，欧洲人的提议有着一些实际优势，即这条隧道与欧洲核子研究中心中有关科学、工程学与后勤的基础设施已经存在了。

当我访问欧洲核子研究中心时，给我印象最深的是柱形二极磁铁的雏形（图8-3为横截面图）。相同的磁铁总共有1232个，每个长15

米、重30吨。这些长度不仅受限于物理实验方面的考虑，还受限于放置大型强子对撞机的隧道那相对狭窄的宽度，以及在欧洲的道路上运输这些磁铁的实际需求。每一个磁铁价值70万欧元，这使大型强子对撞机中单单磁铁这一项的支出就超过10亿美元。

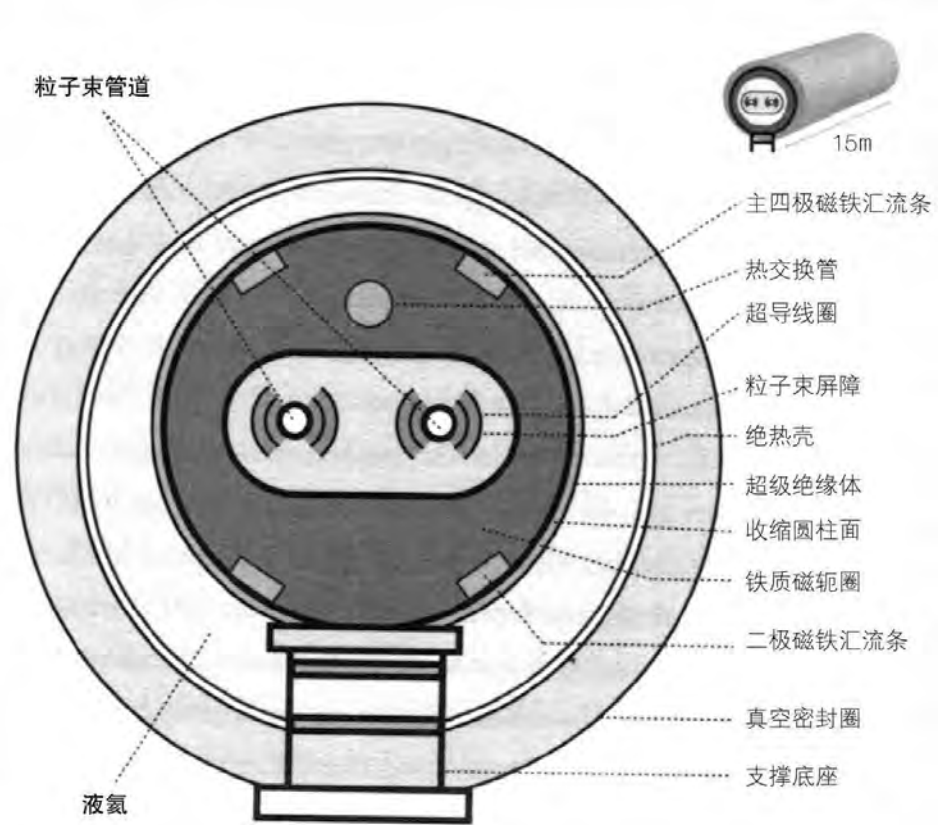


图8-3 低温二极磁铁的原理图。质子被1232个这种超导磁铁控制在大型强子对撞机的加速环中环行。

引导质子束的狭窄管道穿过磁铁中间，磁铁首尾相连，这样它们就可以贯穿大型强子对撞机隧道的内部。它们产生强度约为8.3特斯拉的磁场，这是冰箱磁铁磁场强度的约1000倍。由于质子束能量由450GeV增长到了7TeV，磁场也从0.54特斯拉增长到了8.3特斯拉，以继续引导能量又高了一级的质子环行。

这些磁铁产生的磁场太大，以至于如果没有约束，那么磁铁本身都将在场的作用下移动。这种力被线圈的几何形状减轻，然而磁铁最终却是被由4厘米厚的钢板特制的辊环约束在原来位置上的。

大型强子对撞机的高强磁铁需要超导技术。其工程师们受益于为了如下加速器的制造而研发的超导技术：超导超级对撞机，位于美国芝加哥附近的费米国家加速器实验室的Tevatron加速器，位于德国汉堡附近的德国电子同步加速器研究所（DESY）的电子-正电子加速器。

一般的电缆，比如家用铜电缆，都有电阻。这意味着，能量在电流传递的过程中会有所损失。然而超导电缆却不消耗能量，电流可以畅通无阻地从中流过。超导线圈可以带有巨大的磁场，一旦它们就位，磁场将会得以保持。

大型强子对撞机的每块磁铁都包含着许多铌-钛材质的超导线圈，每个线圈又包含着厚仅6微米的线状纤维，这比人类的一根头发丝还要细。大型强子对撞机包含1200吨这种不凡的纤维。如果你把它们拆开，其长度总和将能环绕火星公转轨道一圈。

大型强子对撞机运行时，磁铁必须极端寒冷，因为它们只能在足够低的温度下运行。超导电缆被维持在比绝对零度高1.9℃的温度上，这比水的结冰相点要低271℃。这个温度甚至比太空中的宇宙微波背景辐射温度2.7K还要低。大型强子对撞机的隧道中存在着宇宙中最冷的大区域——至少是我们已知的大区域。考虑到它们特殊的冰冷本质，这些磁铁被称作“低温二极磁铁”（cryodipoles）。

低温二极磁铁

大型强子对撞机运行时，磁铁必须极端寒冷，因为它们只能在足够低的温度下运行。超导电缆被维持在比绝对零度高1.9℃的温度上，考虑到它们特殊的冰冷本质，这些磁铁被称作“低温二极磁铁”。

除了应用于磁铁上令人瞩目的纤维技术之外，冷冻（低温）系统也是一件应该获得“之最”评价的壮观成就。实际上，这个系统是世界上最类系统中最大的。液氦把这个系统保持在极端低温状态。约97吨液氦在磁铁的周围流动，以冷却电缆。它们不是普通的氦气，而是能在足够压强下保持超流相的氦。超流氦并不像平常物质一样具有黏性，所以它可以高效地带走磁铁系统中产生的热量。10000吨液氮先被冷却，然后它们再冷却在磁铁中环行的130吨氦。

并非大型强子对撞机中的每一个部分都处在地下。地表建筑中有仪器、电子设备与冷却装置。传统的制冷装置可以把氦冷却到4.5K，之后当压强减小时，才会进入最终的冷却步骤。这个过程（以及加热）大约历时一个月，这意味着每当这台机器开、关或者试图进行修理时，都需要花费很长时间以待冷却。

如果出了某些问题，比如一小部分热量升高了温度，那么系统就会骤冷，这意味着超导性会被破坏。如果能量没有及时消散，那么这种骤冷将是灾难性的，因为储藏在磁铁中的能量将被突然释放出来。因此，有一个特殊的系统以检测骤冷现象，并把释放出的能量快速分散。这个系统寻找因丧失超导性导致的电压差异。一旦检测到这样的情况，能量就会在一秒钟之内被分散到整个系统中，所以磁铁就不再具有超导性。

即便应用了超导技术，我们也需要极大的电流，以达到所需的8.3特斯拉的磁场。电流最高能达到12000安，这大约是你书桌上台灯灯泡中流过电流的40000倍。

由于巨大电流与冷却系统的存在，大型强子对撞机的运行需要花费巨额电量，这个数量几乎等于一座小型城市的供电量，比如附近的日内瓦城。为了避免过度的能量消耗，加速器仅在瑞士的寒冬到来之前运行，那时电费将会上涨（这是2009年开机时的期望）。这个政

策给大型强子对撞机的科学家与工程师们带来了福利，让他们可以有一个美好的圣诞长假。

穿过真空而对撞

大型强子对撞机的最后一宗“最”是有关质子环行于其中的管道中的真空。这个系统必须摒除尽可能多的物质，以保持氢的温度足够低，因为任何误入歧途的分子都有可能带走热量与能量。最重要的是，质子束所在的区域必须尽可能不存在气体。如果存在气体，那么质子就有可能与它发生对撞，并破坏质子束的良好环行过程。因此，质子束内部的压强极端微小，小到是大气压的十万亿分之一——这是地表100万米的高处、极端稀薄的大气所具有的压强。在大型强子对撞机中，9000立方米的空气都被抽空，以迎接质子束的到来。

即便是在这种低到极点的压强下，管道区域中每立方厘米依旧存在大约300万个气体分子，所以质子确实经常碰撞到气体并偏离方向。如果足够多数量的质子碰撞了超导磁铁，那么它们将导致超导磁铁骤冷，并破坏超导性。碳准直器把大型强子对撞机中的粒子束校直，以除掉位于某个3毫米的孔隙之外的杂散束粒子。这个孔隙足够让宽度在毫米数量级的粒子束通过了。

然而，把质子聚合为毫米级的粒子团依旧是件棘手的工作。它由叫作“四极磁铁”（quadrupole magnet）的磁铁完成，这种磁铁可以高效地集中、压紧粒子束。大型强子对撞机中包含392个这样的磁铁。四极磁铁也让质子束从它们各自的路径上散开，这样它们才可以真正实现对撞。

这些质子束并非精确或者完全地迎面对撞，而是以一个极微小的角度对撞，这个数值大约为 $1/1000$ 弧度。这是为了确保在同一时间只

有来源于质子束的一个团发生对撞，这样数据就少一些混乱，质子束也能保持原封不动。

当来源于两个质子束的团对撞时，1000亿个质子将与另外1000亿个质子激烈碰撞。四极磁铁也用于完成极端艰巨的任务，即把沿着发生对撞、记录事件的实验所在粒子束区域的粒子束集中起来。在这些位置，磁铁把粒子束压缩到16微米的尺度上。粒子束必须极小、致密，以使在相遇时，一个团中的1000亿个质子更加可能找到另一个团中的1000亿个质子之一。

一个团中的大多数质子都无法找到另一个团中的质子，即便它们已经被引导向彼此的方向发生对撞。单独的质子直径仅有百万分之一纳米。这意味着，即使所有的质子都被约束在16微米的团内，每当团与团之间交错而过时，也仅有约20个质子迎面相撞。

事实上，这是一桩好事。如果同时出现太多的对撞，那么数据将会变得非常混乱，我们将不能准确分辨究竟哪个粒子从哪次对撞中出现了。当然，完全没有对撞出现是一桩坏事。通过把特定数量的质子集中到特定的尺度上，大型强子对撞机确保了每次两个团交错而过时发生对撞的次数是最优的。

单独质子的对撞几乎是瞬时出现的，其时间大约比1秒低25个数量级。这意味着，质子对撞的时间间隔完全取决于团之间交错的频率，其下限约为25纳秒。团之间每秒钟交错多于1000万次。通过这种频率的对撞，大型强子对撞机得以产生大量数据——每秒钟大约10亿次碰撞。幸运的是，团之间交错的时间间隔足够长，足以让计算机追踪到感兴趣的单一碰撞，而不会被不同团中发生的碰撞所扰乱。

所以，大型强子对撞机的数宗“最”对保证最高可能能标的对撞与最大对撞数量而言都是必要的。大多数能量都滞留在加速环中，只

有少数质子碰撞值得我们注意。尽管质子束中存在大量能量，每次单独的团之间对撞的能量都比几只飞行中蚊子的动能还要大一点儿。然而这是质子对撞，并非汽车或是足球运动员之间的对撞。大型强子对撞机之“最”把能量集中在一个极小的区域中，也把能量集中在实验者能捕捉到的基本粒子对撞中。我们很快就将开始思考一些研究者们也许能找到的隐藏组分，以及物理学家们希望大型强子对撞机的发现能为我们提供的有关物质与空间本质的洞见。



我从1983年开始攻读物理学研究生，而大型强子对撞机的建造计划于1984年正式提出。所以，从某种意义上来说，我在自己学术生涯的前25年，一直在等待大型强子对撞机。现在，在经历了无数等待与波折之后，我与同事们终于看到了大型强子对撞机的实验数据，并期待着实验将要很快揭示的有关质量、能量与物质的洞见。

大型强子对撞机是目前最重要的粒子物理学实验仪器。因此，当它开始运行时，我的许多物理学家同僚都非常急切、兴奋。所有的研究室中，人们都在四处打听情况：对撞将要达到什么能标？一个质子束中包含多少质子？理论物理学家们想要理解某些细节，这些细节曾经是由我们之中从事计算与概念，而非机器或实验设计工作的人抽象出来的，而实验物理学家们也是这样。在听到我们最新的猜想，并了解更多他们有可能寻找、发现到的事物之后，实验物理学家们表现出了我所见过的最大的热切之情。

即便是在2009年12月的一个计划讨论暗物质的会议上，与会者依旧热切地讨论着大型强子对撞机。那时，大型强子对撞机刚刚令人难以置信地成功完成了它的首次加速与对撞。在一年多前的那次沮丧^[39]之后，每个人都开始欣喜若狂。当实验物理学家们得到可以更好地理解探测器的实验数据时，他们感到放心多了，而理论物理学家们也为他们很快就能得到一

些答案而高兴。一切都好到难以置信：粒子束看上去很好，对撞已经进行过了，实验设备正在记录对撞事件[书·免`费`分`享V.信shufoufou]。

达到这个里程碑的过程是一个相当了不起的故事，而这一章就要讲述这段传奇。因此，请系好你的安全带，我们将要走上一段颠簸之路。

筑梦科学

欧洲核子研究中心的历史早于大型强子对撞机数十年。第二次世界大战结束后不久，在欧洲建造一家加速器中心，以主要开展研究基本粒子的实验的想法就首次提上了日程。那时，许多欧洲物理学家（其中一些人已经移居美国，另一些人还在法国、意大利与丹麦）都希望让科学的前沿回到自己的故乡。美国人与欧洲人一致同意，如果欧洲人联合起来参与这项事业，并在欧洲展开研究以弥补在战争结束不久之后遗留下来的荒芜与猜忌，那么这对科学与科学家们而言，都是最好的选择。

联合国教科文组织1950年在佛罗伦萨召开的一次会议上，美国物理学家伊西多·拉比（Isidor Rabi）提议兴建一个新的实验室，以在欧洲重建科学界的强联系。1952年，欧洲核子研究理事会（法文名为Conseil Européen pour la Recherche Nucléaire，因此其缩写为CERN）成立。1953年7月1日，来自12个欧洲国家的代表聚集在一起创建了“欧洲核子研究组织”（European Organization for Nuclear Research），其创立协议于次年得到了认可。CERN的缩写显然有点不合适了，而且我们现在研究的是亚核物理学和粒子物理学，但出于管理机构的原因，CERN的缩写依然被保留了下来。

欧洲核子研究中心的设备故意建造在欧洲的中心，一个位于日内瓦城附近、在瑞士-法国交界的地点。如果你喜欢户外运动，那么到此一游再好不过。这个极好的位置坐拥良田万顷，紧邻侏罗山，而不远处就是阿尔卑

斯山脉。欧洲核子研究中心的实验物理学家们都颇有运动天赋：滑雪、爬山、骑行，无所不及。欧洲核子研究中心占地极为庞大，足以为那些爱好运动的研究员们提供足够大的场地，让他们奔跑得精疲力竭以保持体形。那里的街道都以著名物理学家的名字命名，比如“居里路”“泡利路”“爱因斯坦路”，等等。然而，欧洲核子研究中心的建筑风格囿于它的建造时代。20世纪50年代，国际建筑风格都是平淡无奇而低矮的，所以欧洲核子研究中心的楼都不高，它的走廊蜿蜒，办公室风格乏味。“科学设施”并没有给这些建筑带来生气——看看大多数学校中的科学建筑吧，你会发现，它们往往是校园所有建筑中最丑陋的。给这个地方（与风景）带来亮色的是那些工作于此、孜孜不倦地追寻着各自学科、工程学目标与成就的人。

国际合作有助于推动欧洲核子研究中心的发展，也许这是人类创建过的最成功的国际事业。即便是在第二次世界大战之后不久，各国之间还处于冲突状态之际，来自12个不同国家的科学家们依旧坐在一起，为这项共同事业出力。

如果存在竞争，那也是为了美国人及其蒸蒸日上的科学事业的。在欧洲核子研究中心找到W、Z规范玻色子之前，几乎所有的粒子物理学发现都来自美国的加速器。我在1982年作为费米实验室的暑期学生时，那个跌跌撞撞地走入公共区域，呢喃着他们为什么“必须找到那天杀的矢量玻色子”并打破美国统治地位的醉汉物理学家，也许正表达了那时许多欧洲物理学家的想法——虽然这种措辞显然有些贫瘠和苍白无力。

欧洲核子研究中心的科学家们确实找到了那些玻色子。现在，随着大型强子对撞机的建成，欧洲核子研究中心是实验粒子物理学当之无愧的科学中心。然而，在大型强子对撞机计划首次被确认之前，这绝对没有被预先确定下来。由里根总统批准的美国超导超级对撞机的能标将是大型强子对撞机的三倍——假如美国国会一直资助这个项目，而没有让它中道崩殒的话。虽然克林顿政府原本并不支持这项由共和党前总统发起的项目，然而当他意识到某些利害关系之后，就改

变了想法。1993年6月，克林顿给美国众议院拨款委员会主席威廉·纳彻（William Natcher）写了一封信，试图继续支持这个计划。

我希望你了解我对超导超级对撞机的持续支持……在此时放弃它将是美国对其基础科学领袖地位作出妥协的一个信号，而这个领袖地位曾持续数代。现在经济形势非常紧张，政府部门支持这个项目，是因为它将在科学与技术的领域之内带来广泛的收益……我希望你继续支持这项重要而具有挑战性的努力和尝试。

当我于2005年见到克林顿时，他还提出了有关超导超级对撞机的话题，并且问我：“放弃这个项目之后我们失去了什么？”克林顿很快就承认，他也认为人类放弃了一次宝贵的机会。

大约在美国国会否决了超导超级对撞机提案的时候，纳税人们为储贷危机担负了大约1500亿美元，这远超过超导超级对撞机的预算100亿美元。作为比较，美国的年度财政赤字均摊到每个美国人身上有600美元之巨，而伊拉克战争让每位国民为此负担2000多美元。假设超导超级对撞机建成了，我们现在就能得到高能实验结果，也能达到比大型强子对撞机还高的能标。送走了储贷危机之后，美国又迎来了2008年的金融危机，而这让纳税人负担了更多的救市资金。

大型强子对撞机90亿美元的建造经费与超导超级对撞机的预计经费几乎差不多。每个欧洲人将为此承担15美元——或者如我在欧洲核子研究中心的同僚路易斯·高密（Luis Álvarez-Gaumé）所说，在大型强子对撞机的建造期间，每年每个欧洲人仅为此付出一杯啤酒的钱。评估像大型强子对撞机这种基础科学研究的价值往往是棘手的，然而基础科学研究的确推进了电力、半导体、万维网，并极大影响了我们生活中各种技术的腾飞；它们启迪了技术与科学思想，扩展到了我们经济中的方方面面。大型强子对撞机的实际效果也许难以估计，但它的科学潜力却可以估量。我觉得，欧洲人的花费物有所值。

长期项目需要信念、责任心与奉献精神。在美国，越来越难作出这样的承诺。在我们过去的观念中，提到美国就意味着巨大的科学、技术进展。然而，这种重要的长期计划正在变得越来越少。你必须把这种计划提供给欧洲共同体，他们才有能力始终如一地完成这种方案。建造大型强子对撞机的提案于25年之前首次设想，并于1994年核准。它是一个雄心勃勃的项目，然而，直到现在它的成果才“小荷才露尖尖角”。

此外，欧洲核子研究中心刚刚在国际上拓展了它的吸引力，除了它的20个会员国之外，还有53个国家都参与了大型强子对撞机的设计、建造与仪器试运行过程。现在，来自85个国家的科学家们参与其中。美国不是正式的会员国，然而，美国是所有国家里为主要实验提供最多工作人数的国家。

总共有大约10000名科学家参与了工作——这也许是全球粒子物理学家总数的一半。其中20%就居住在附近，在此全职工作。随着大型强子对撞机的运行，主餐厅变得无比拥挤，物理学家们摩肩接踵地点餐——这个情况在一座新餐厅建成之后，才有所和缓。

由于欧洲核子研究中心的人来自五湖四海，美国人会被在餐厅、办公室与走廊中回响的各种语言与口音所迷惑。美国人还会注意到香烟、雪茄、葡萄酒与啤酒，这都提醒着他们：他们不在故土。有些人会赞扬餐厅食品的高品质，然而，欧洲人很挑剔，他们对这些赞美的评论持保留意见。

欧洲核子研究中心的雇员与访客涵盖甚广：从工程师到管理员，到真正做实验的实验物理学家，再到理论物理学家。欧洲核子研究中心实行等级管理，其中有主管官员，有负责一切行政事务（包括重大决策）的委员会。其首脑被称为总干事，这也许与某些出自《吉尔伯特与沙利文》（*Gilbert and Sullivan*）的典故有关，虽然真正的原因是总干事之下有很多管理职位决定了这个名字。欧洲核子研究中心委员会是负责如计划、安排项目这样重大决策的管理体系。它尤为关

注科学政策委员会，后者是主要的咨询委员会，帮助评价提案及其科学价值。

这项有数千人参与的大型实验合作项目有其自己的体系。工作根据探测器的组分，或者分析的类型而分配。某个来自指定大学的小组可能负责仪器中某个特定的部分，或者某个特定类型潜在理论的解释。在挑选自己感兴趣的工作的权利上，欧洲核子研究中心的理论物理学家们比实验物理学家们拥有更大的自由。有时，他们的工作与欧洲核子研究中心的实验相关，但也有些人致力于构建更加抽象的、无法在短期内被验证的理念。

虽然如此，欧洲核子研究中心以及全世界所有的粒子物理学家都对大型强子对撞机充满期待。他们知道自己未来的科学研究以及粒子物理学这个领域本身，都依赖于未来10~20年之内大型强子对撞机的成功运行与发现。他们理解这项事业的挑战性，但是他们更加从骨子里认可这项事业中的那些“最”。

撞还是不撞，大型强子对撞机是怎样炼成的

林恩·埃文斯（Lyn Evans）是大型强子对撞机的总设计师。虽然我之前曾听过他用抑扬顿挫的威尔士语作演讲，但是我第一次见到他是在2010年1月初于加利福尼亚召开的一次会议上。那次恰逢其时，因为大型强子对撞机的建造正在紧锣密鼓地运行着。作为一位处事低调的爱尔兰人，他的喜悦却溢于言表。

埃文斯作了一次精彩的演讲，详述了从开始建造大型强子对撞机以来，如过山车一般的跌宕历程。他从讲述20世纪80年代有这个想法开始，那时欧洲核子研究中心负责了第一个官方研究：调查制造高能质子-质子对撞机的可行性。接下来，他叙述了1984年的会议，大多数人都认为那次会议是这个想法的正式开端。那时的物理学家与洛桑城

中的机械制造师们会面，并告诉了他们这个在10TeV能标上对撞两束质子的想法。在最终的实施上，这个数字被降到了7TeV。在几乎10年之后的1993年12月，物理学家们向欧洲核子研究中心委员会提出了一项积极的提案，希望其重大决策委员会通过以下意见：停止除大型正负电子对撞机之外的一切实验程序，在接下来的10年之内全力建造大型强子对撞机。那时，欧洲核子研究中心委员会驳回了这项提案。

原本，一个反对建造大型强子对撞机的理由是来自己超导超级对撞机的激烈竞争。但由于超导超级对撞机计划于1993年取消，大型强子对撞机成为时下唯一的极高能加速器。物理学家越来越坚信这项事业的重要性。除此之外，大型强子对撞机的研究非常成功。在大型强子对撞机的建造阶段才开始领导欧洲核子研究中心的罗伯特·艾马（Robert Aymar），在1993年组织了一个审查小组，以评估大型强子对撞机的可行性、经济性与安全性。

大型强子对撞机的规划中一个重要的困难是如何工业化生产足够强的磁铁，以把高度加速的质子约束在加速环中。正如我们在第8章中所见，已有的隧道尺寸导致了现在最大的技术挑战，因为它的半径固定，所以磁场必须非常大。在他的讲话中，埃文斯以“瑞士手表一般的准确性”来描述第一个10米长的二极磁铁原型，后者于1994年被工程师与物理学家们成功测试。他们在首次尝试中达到了8.73特斯拉——这达到了预期目标，而且是一个非常有望的标志。

即使欧洲基金会比美国基金会更加稳定，无法预料的压力依旧为欧洲核子研究中心的财政引入了不确定性。为欧洲核子研究中心资助最多的德国，由于1990年两次统一，其预算缩减了。因此，德国减少了对欧洲核子研究中心的资助，英国也不愿再大幅增加对欧洲核子研究中心的预算金额。英国物理学家、诺贝尔奖获得者克里斯托弗·卢埃林·史密斯（Christopher Llewellyn Smith）战胜了物理学家卡洛

- 鲁比亚而接任欧洲核子研究中心主管，并如他的前任一样强烈地支持大型强子对撞机的建造。通过获取资金，瑞士与法国这两个东道主国家从大型强子对撞机在自己国土上的建设与运行中获益最大，而卢埃林·史密斯部分减轻了这个严重的预算问题。

欧洲核子研究中心委员会对大型强子对撞机的技术与预算决议都很深刻，于是他们在1994年12月16日通过了决议。此外，卢埃林·史密斯与欧洲核子研究中心还确认了非会员国家也可以加入组织并参与到实验中来：日本于1995年加入，印度于1996年加入；1997年，俄罗斯、加拿大与美国陆续加入。

由于各国的支持，大型强子对撞机可以不顾原始合约上要求分两个阶段进行建设与运营的限制条款。第一个建设阶段仅仅包含2/3的磁铁。不管是从科学角度还是从节约成本的角度考虑，将磁场削弱都是愚蠢的选择。原初的打算是让预算均摊在每年之中。1996年，当德国再次削减了对大型强子对撞机的资助时，预算的情况看起来非常糟糕。1997年，欧洲核子研究中心首次被允许以贷款的方式筹资建设，以弥补资金的缺口。

在说完预算的历史真相之后，埃文斯话锋一转，开始谈论一些令人高兴的话题。他描述了于1998年12月完成的第一个二极磁铁的试验串，即一次检验结合在一起的磁铁是否可用的测试。这次试验的顺利完成证实了大型强子对撞机的某些主要部件的可行性与协调性，这是它建造过程中一块重要的里程碑。

2000年，当大型正负电子对撞机寿终正寝的时候，它被拆开以让路于大型强子对撞机的建设。然而，即便大型强子对撞机最终被放在一个已经存在的隧道中，而且相应的设施、基建、工作人员都已经就位，但要完成从大型正负电子对撞机到大型强子对撞机的转变，还需要很多人力、物力。

大型强子对撞机建造的5个时期包括：

- 挖掘实验所需的洞穴与构筑物的土木工程。
- 一般性服务的设立，以让一切可以开始运作。
- 插入冷冻链以维持加速器的低温。
- 把所有机器的构件拼装在一起，包括二极磁铁、相关的连接点与电缆。
- 硬件的试运行，以保证一切都如预期一样开展。

欧洲核子研究中心的设计者们制定了一份详尽的时刻表，以协调不同的建造阶段。但“良愿成泡影，不管是人是鼠，结局总会出其不意”^[40]，这句话应验了。预算问题是件麻烦事。我还记得2001年粒子物理学共同体的沮丧与担忧。那时我们正在等待并寻找迅速解决严重预算问题的途径，以确保大型强子对撞机的顺利建造。欧洲核子研究中心需要处理成本超支的问题，但由于其占地面积与基础设施，成本高一些是自然的。

即便是在这些资金与预算问题都被解决了之后，大型强子对撞机的发展依旧不怎么顺利。埃文斯讲述了一系列没有预料到的事件及其如何不时地拖延了大型强子对撞机的建造进程。

无疑，不会有人在挖掘紧凑 μ 子线圈洞穴的时候预见到他们会挖掘到一座14世纪的高卢-罗马庄园。产权的界定线与延续到今日的农场边界线类似。挖掘被暂停以供考古学家们研究那些埋藏的财物，其中包括奥斯蒂亚（Ostia）、里昂与伦敦（那时的庄园里居住着奥斯蒂亚人、里昂人与伦敦人）的一些货币。显然，罗马比现代欧洲更好地应用了共同货币制度——直到现在，欧元还没有代替英镑与瑞士法郎成为交易介质。这给欧洲核子研究中心的英国物理学家带来了不少麻烦，他们甚至没有用以支付打车费用的货币。

与紧凑 μ 子线圈的一路艰辛相比，2001年进行的超环面仪器洞穴开掘工作相对而言一路平安无事。挖掘这个洞穴需要挪走总重30万吨的岩石。工程师们需要面对的唯一问题是：一旦这些材料被挪走，洞穴的表面就会有一些微小的上浮，其速率约为每年1毫米。这听起来并不算多，它却会在原则上影响探测器部件的精确测量。所以，工程师们需要设置一些精密测量仪器。它们太过精密，以至于不仅能检测到超环面仪器的运动，还能被2004年的海啸以及苏门答腊岛地震触发，之后还将会有更多的自然运动可以被显示出来。

在极深的地下建造超环面仪器的过程令人印象非常深刻。屋顶浇铸在洞穴表面上，并被绳索悬吊；下面砌起高墙，直到屋顶得以盖在墙上。2003年，在一场开掘典礼之后，挖掘开始了。值得注意的是，典礼上的阿尔卑斯号角声回响在洞穴中，这被埃文斯描述为“娱乐之源”。实验仪器由下至上一个接一个地被安装，直到最终超环面仪器被这种“送入瓶子的方法”送到挖掘出来的地下洞穴中组合起来。

另一方面，紧凑 μ 子线圈的准备过程就像是面对波涛汹涌的海洋。在挖掘过程中，它再一次陷入了困境，因为紧凑 μ 子线圈的选址不仅不幸地位于一座考古遗迹上，而且那里还有一条地下河流。在多雨的那一年，工程师与物理学家惊奇地发现，插入到地下以传送材料长达70米的圆筒下沉了30厘米。为了处理这次事故，洞穴的挖掘者们用大量的冰围在圆筒周围，以冰冻土地、稳定这片区域。他们还必须安装用以稳定洞穴周围脆弱岩石的支撑结构，包括总长度达到40米的螺钉。不出意料，紧凑 μ 子线圈的开凿工程比预期长了很多。

唯一的可取之处在于，紧凑 μ 子线圈的体积相对较小，实验物理学家与工程师们已经在考虑在地面上对其进行建造、装配。在地面建造、安装这些组件更为简单。由于有更大的空间可以让很多人共同工作，这种做法也会变得更加快捷。地面建造还有另外一些重要好处，即有关洞穴的问题不再会拖延建设工程。

把如此巨大的装置送到地下是一件令人却步的事情——我在2007年首次访问紧凑 μ 子线圈时也是如此想的。的确，把实验装置送到地下是一件艰巨的任务。装置中最大的一个部件被某种特殊吊车送到100米深的紧凑 μ 子线圈井中，其速度为每小时10米，一举一动都小心翼翼

翼。由于在仪器与墙之间只有10厘米的腾挪余地，缓慢的速度与精细的监控系统必不可少。探测器的15个大型部件在2006年11月与2008年1月之间分别被送下井中，最后被送下去的部件是一个黄铜制的计时装置，这个日期很接近大型强子对撞机的开机日期。

在紧凑 μ 子线圈的地下河问题之后，大型强子对撞机建造过程中的下一个危机是：机身在2004年卡住了，问题出在被称为QRL的氦供应线上。欧洲核子研究中心的工程师发现，承接这个建造项目的法国公司用“价值5美元的垫片”取代了原本设计应使用的材料。这种替代材料的失效造成了管道内部的热收缩。出故障的部件不止一处，所有连接点都需要检查一遍。

这时，低温线已经被部分地安装好了，很多其他部分也已经就位。为了避免堵塞供给线，造成更长延迟，欧洲核子研究中心的工程师们决定修理已经生产出来的部分，同时让工厂在将剩余部件发货之前修正错误。欧洲核子研究中心工厂的运作以及移动、重新安装机器的大型部件让大型强子对撞机延迟了整整一年开机——至少，这比埃文斯等人所担心的10年延期要好得多。

如果没有这些管道和低温系统，就不可能成功安装磁铁。所以，1000个磁铁被暂存在欧洲核子研究中心的停车场上。即便是在高档宝马车与奔驰车的辉映下，价值10亿美元的磁铁仍旧超过了停车场上所有车辆价值的总和。虽然没有人会偷窃这些磁铁，但是停车场并不适合储藏技术产品，所以把磁铁重新放置到特定位置而带来的拖延又是不可避免的了。

2005年，另一个危机出现了。这次危机与建造美国费米实验室以及日本实验室的三元组磁铁（inner triplet）有关。三元组磁铁在质子束对撞之前对它们进行了最后一次调焦。它包含三个四极磁铁，配有低温与电力设备。三元组磁铁没能通过压力测试。即便这次失败令人难堪，并造成了恼人的拖延，但是工程师们可以在隧道中修好它，所以最终它并没有带来过多的时间成本。

总而言之，2005年比之前的情况都要好。虽然没有号角声相伴，紧凑 μ 子线圈开掘典礼于当年2月开幕。另一个里程碑事件也在2月发生——第一个低温二极磁铁被送到了地面之下。磁铁的建造对大型强子对撞机这项事业来说至关重要，欧洲核子研究中心与商业化工业界之间的这次紧密合作促成了磁铁及时而节约的生产过程。磁铁虽然由大型强子对撞机设计，但却由法国、德国与意大利的公司生产。原本，欧洲核子研究中心的工程师、物理学家与技师们于2000年下单购买了30个二极磁铁，以小心检测它们的质量与控制支出。之后，他们于2002年下了最终的订单——购买1000多个磁铁。虽然如此，欧洲核子研究中心依旧保留生产主要组件与原材料的职责，以确保质量具有最优性、统一性，并使支出最小化。为了达到这个目的，欧洲核子研究中心把12万吨材料挪到了欧洲，4年间，平均每日租用10辆大卡车——而这只是为建造大型强子对撞机而付出努力的一隅。

在运输之后，所有的磁铁都被测试了，并小心地沿着垂直方向被投放到坐落在侏罗山附近——那里可以鸟瞰欧洲核子研究中心。从那里开始，一种特殊的运输工具把它们沿着隧道运输到目的地。因为这些磁铁极为巨大，且隧道的墙与大型强子对撞机部件之间只有几厘米的转动余地，所以这种工具靠某种画在地板上的光学检测线自动导航。它以约每小时1.6公里的速度前行，以尽量减小振动。这意味着，需要长达7个小时的时间才能把一个二极磁铁送到加速环的另一端。

2006年，在建造5年之后，共有1232个二极磁铁被成功运输。2007年发生了两件大事：

- 最后一块低温二极磁铁被成功运输。
- 全长3.3公里的部件首次被成功冷却到了设计温度零下27 $^{\circ}\text{C}$ ，整个系统得以首次运行。

在隧道的这一部分里，数千安培的电流在超导磁铁中环行。像往常一样，欧洲核子研究中心的人们开了香槟庆祝这个时刻。

一个低温部件于2007年11月关闭，一开始，一切似乎都在正常运行——直到出现另一次灾祸。这次，问题出在所谓的“插入模块”（plug-in module），也就是PIMs上。在美国，我们并不需要关心一切有关大型强子对撞机的新闻报道。然而这个新闻却传播甚广。欧洲核子研究中心的一位同事告诉我，不仅是这个部件出了问题，同样的问题可能在加速环中无处不在。

室温状态下运行的大型强子对撞机与冷却状态下运行的大型强子对撞机之间存在接近300℃的温差，这个温差对建筑材料产生了巨大的影响——问题就出在这里。金属零件热胀冷缩，二极磁铁本身在冷却阶段会收缩几厘米。于一个15米的庞然大物而言，这听上去并不算多。然而线圈必须在0.1毫米的精度上被准确放置，以维持运行所需的强大的匀强磁场，进而正确引导质子束。

为了适应这种变化，二极磁铁被设计具有特殊的指针，在冷却时伸直以保证电学器件不分开，并在回温过程中缩回。但由于错误的铆接点，这些指针并非缩回，而是失效了。更糟的是，每一个相互连接的点都会受到这种错误的影响，我们无法得知究竟是哪处出了问题。最大的挑战在于辨认、修好每一个错误的铆接点，而且不花费太多时间。

不得不称赞工程师们的足智多谋：他们找到了一种简单方法，利用已有的、原本沿着粒子束每53米就安装一个的电子拾波器，这样电子器件就可以随着粒子束的经过而触发。他们在一个乒乓球模样的物体中安装了一个振子，这个球可以沿着质子束的路径在隧道中运行。每一个部分长3公里，这个球可以快速经过，以在每次通过拾波器时检测电子器件。如果电子器件没有反应，球就会击打那些指针，工程师

们就可以修正问题，而不用打开沿着粒子束所经路线的每一个单独的连接部件。一位物理学家戏言：“大型强子对撞机中的第一次对撞不是质子之间的对撞，而是乒乓球与失效指针之间的碰撞。”

在最后一个问题得以解决之后，大型强子对撞机看上去步入了正轨。当所有硬件就位之时，大型强子对撞机就可以开始运行了。2008年，当大型强子对撞机的首次试运行终于到来之时，很多人都在为它祈祷。

2008年9月：第一次试开机

大型强子对撞机形成质子束，并且在一系列提高能标的过程之后，把它们送入最终的环形加速器中。接下来，质子束在隧道中回旋，并得以回到它们精确的初始位置。这让质子可以在回旋加速多次之后，再定期被转向，以高效地参与对撞。其中的每一步都需要被依次检查。

第一个重要步骤是：核实质子束是否正确地在加速环中环行——它们当然会。令人惊奇的是，在经历了长时间的磨难之后，2008年9月，欧洲核子研究中心在没出什么故障的情况下就发动了两个质子束，这个结果超出了大家的预期。那一天，两束质子首次成功地以相反方向穿过了大型隧道。这一步涉及以下工作：试运行注入设备、启动控制设备与仪器、检验磁场是否成功把质子约束在了加速环中，以及确认所有磁铁按照期望整齐划一地运行。这一系列事件于当年9月9日晚上首次完成。当次日的测试进行时，一切都在计划之中，甚至比计划还要好。

每一个参与大型强子对撞机的工作人员都称2008年9月10日为他们“永世铭记的日子”。当我于一个月之后访问欧洲核子研究中心时，听到了许多关于彼时彼景的欢乐描述。人们难以置信而又激动地

在计算机屏幕上捕捉到了两个光点的轨迹。在第一圈运行中，第一个粒子束几乎成功地回来了。在开机的第一个小时内，工作人员对实际路径按照预期做了一些微调。一开始，粒子束绕着加速环环行了几圈。接下来，每一个质子脉冲依次被微调，这样质子束就可以很快环行数百圈。不久后，第二个质子束也经历了相同的过程，整个事情走上正轨耗时一个半小时。

埃文斯和其他人一样高兴，他并不知道那时的实况录像已经从工程师们所在的控制室传播到了互联网上，整个事件都被直播给了大众。太多的人都在关注屏幕上的那两个小点，以至于他们的网站由于过载而关闭了数次。当工程师们修正质子的路径，以让它们在加速环中成功环行时，全欧洲的人们（欧洲核子研究中心新闻办公室声称有几百万）都像是被施了催眠术一样。同时，在欧洲核子研究中心，当物理学家与工程师们聚集在礼堂中观看同一过程时，惊天动地的欢呼声不时爆发出来。从这一点上来看，大型强子对撞机的前途似乎一下变得无比光明。那一天是一个美妙的成功日。

然而，仅仅9天之后就乐极生悲了。那时，两个新的重要特征需要通过试运行来检验。一开始，在第一次试运行中，质子束在大型强子对撞机内部被加速到了有史以来最高的能标，这次试运行只涉及质子束注入能量，即质子在首次进入大型强子对撞机的加速环时所具有的能量。计划的第二部分是对撞这两个质子束，这当然会是大型强子对撞机运行过程中一座重要的里程碑。

在最后一刻，2008年9月19日，尽管工程师们已经进行了周详的考虑并做好了許多预防措施，但是这次试运行还是失败了。这次失败是灾难性的。与几个很少运作的氦释放阀门组合在一起的两块磁铁之间的铜框上的某个错误焊点，导致质子首次对撞的日期被拖延了一年之久。

问题在于，当科学家们试图在第八个与最后一个部分中提高电流与能量时，沿着汇流条的两个磁铁的连接点坏掉了。汇流条是连接两个超导磁铁的超导连接物（见图9-1），而出问题的正是两块磁铁之间的连接点。错误的连接点导致周围的氦被击穿而产生电弧，并造成了6吨液氦的突然泄漏（它们本应该缓慢回温的）。当液氦回温而恢复为气态时，猝息导致了超导性的丧失。

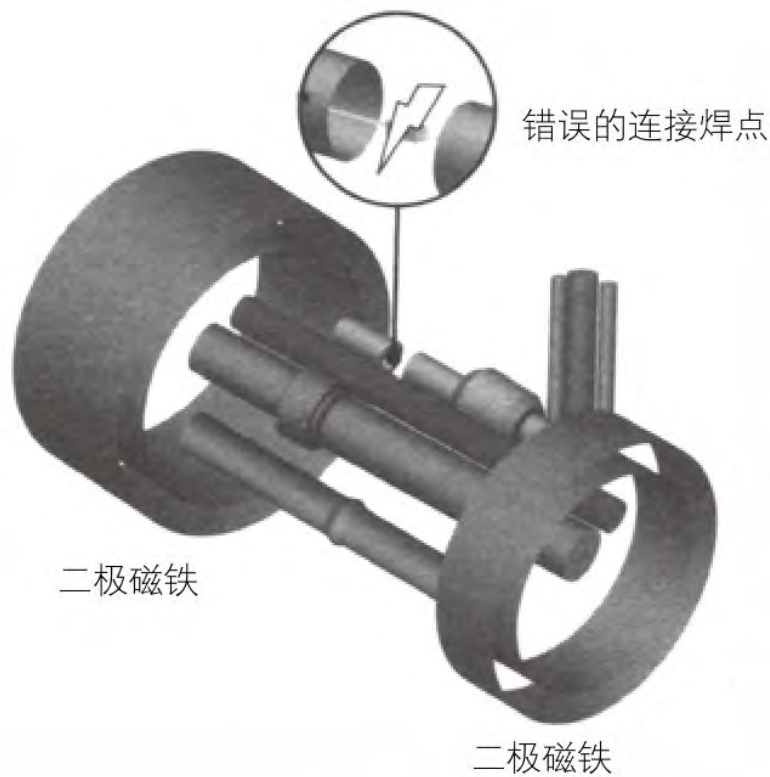


图9-1 一条把不同磁铁连在一起的汇流条。一个错误的连接焊点成为2008年那次事故的主要原因。

大量泄漏的氦气导致了巨大的气压波，进而快速地导致了一场爆炸。在30秒之内，它的能量就炸飞了一些磁铁，并且破坏了粒子束管中的真空状态、破坏了绝缘性，并污染了610米长的粒子束管，弄得到处都是脏东西。10个二极磁铁被完全摧毁，而另外29个二极磁铁损坏

程度太高，需要被替换掉。这并不是我们想要的结果。控制室中也没有人意识到这一点，直到有人注意到某台计算机控制的在隧道中的某个停止按钮被泄露的氦触发了。很快他们就意识到，粒子束丢失了。

在这次灾祸的几周后，我访问了欧洲核子研究中心，了解了更多的背景故事。记住，对撞终极目标的质心能量是14TeV，也即14万亿电子伏。欧洲核子研究中心决定，首次对撞仅需达到2TeV的能量，其目的是检测一切是否都在照常运行。之后，工程师们决定在第一次真正以收集数据为目的的运行中，把能量提高到10TeV（每个质子束5TeV）。

2008年9月12日，在某个运输机器的损坏导致的小小延迟之后，这个计划变得更加急切了。科学家们在延迟带来的时间间隔中，把隧道中8个部分的测试能量提高到了5.5TeV，并且有时间测试其中的7个部分。他们证实了这些部分可以在高能情形下正常运作，但是他们并没有时间测试第8个部分。虽然如此，他们依旧决定勇往直前，以尝试更高能标的对撞，因为一切看起来似乎都没什么问题。

一切确实都运行得很好，直到工程师们试图在最后一个未经测试的部分中提高能标。当能标从4TeV上升到5.5TeV，需要的电流值为7000~9300安培时，失稳事故发生了。这是最后一个可能出错的环节，而且确实在这里出了纰漏。

在延期的那一年中，修理费用达到了大约4000万美元。虽然修理磁铁与粒子束管需要时间，但它们并非不可能完成的任务。备用磁铁替换了39个无法修复的磁铁。总共有53个磁铁（14个四极磁铁、39个二极磁铁）在隧道中发生事故的部分被替换掉了。另外，工作人员清扫了4公里多长的真空粒子束管，为100个四极磁铁安装了一套新的约束系统，添置了900个新的氦压释放部件，并且在磁铁保护系统中增加了6500个新的探测器。

更大的问题是，磁铁之间有10000个连接点，它们有可能具有同样的问题，研究人员意识到了这些风险。但是如何能保证这个问题不在加速环的其他地方重现呢？我们需要一种检测机制，把危险扼杀在摇篮之中。工程师们升级了系统，新系统可以寻找微小的电势降落，这种电势降落代表着连接点处存在电阻，而这意味着维持系统低温的冷却系统有可能失效了。出于谨慎，我们还需要一些时间，以改进氦释放阀门系统，并进一步研究连接点与磁铁本身的铜框——这意味着距离达到大型强子对撞机的设计最高能标之前，还需要更长时间的准备。虽然如此，由于加入了多种新系统以检测、稳定大型强子对撞机的运行，埃文斯与其他工作人员非常确信，造成过损害的那种超压情况不会再出现。

在某种意义上，工程师与物理学家们能在大型强子对撞机真正运作、在机器中充满了辐射之前排除这些故障实乃幸事。那次爆炸让他们花了一年时间才能继续开始检测粒子束，并准备下一次对撞。那是一段很长的时间，但是和我们过去40年（从某些方面来讲是数千年）中对物质潜在理论的孜孜追求所耗的时间相比，这段时间并不算长。

2008年10月21日，欧洲核子研究中心行政部门确实坚持了一部分初始计划。那一天，我加入了有1500名物理学家与不少世界领袖出席的大型强子对撞机正式开幕典礼。它原本被计划得很好，但没有人能事先料到那次发生在几周之前的惨痛事件。开幕典礼当天有演讲、有音乐、有欧洲文化中不可或缺的美食。即便时辰未到，它也是令人愉快而多姿多彩的。尽管大家都还在忧心9月份的事件，然而心中却仍然充满了希望，期盼着这些实验能够阐明围绕着质量的某些谜题、引力为何如此之弱的原因、暗物质的本质，以及自然中基本作用力的性质。

即便欧洲核子研究中心的很多科学家都对那次事件感到非常不开心，然而我却认为这次庆典更倾向于是对这次国际合作凯旋而归的一种期待。那次事件并没有什么发现，却让人们意识到了大型强子对撞机与参与到其制造过程中那些国家的热情所具有的潜力。有一些演讲确实鼓舞人心。法国总理弗朗索瓦·菲永 (Francois Fillon) 讨论了基础研究的重要性，以及为何世界金融危机不应该拖延科学进展的脚步。瑞士总统帕斯卡尔·库什潘 (Pascal Couchepin) 讨论了公共服务的功绩。葡萄牙科技与高等教育部部长马里亚诺·伽戈 (Jose Mariano Gago) 教授讨论了如何在官僚作风盛行的今天重视科学，以及稳定开创重大科学项目的重要性。许多外国伙伴都是在这场庆典上首次访问欧洲核子研究中心的。在庆典中一直坐在我邻座的那个人在日内瓦为欧盟工作，但从未踏足欧洲核子研究中心一步。看到了这场盛会之后，他满怀热情地告诉我，他很快就会和同事、朋友们重返这里。

2009年11月：姗姗来迟的胜利

大型强子对撞机最终于2009年11月20日重新联机，这是一条爆炸性的新闻。不仅因为质子束在一年之后首次环行，还因为在几天之后，它们会实现对撞，创造出最终将进入实验仪器的粒子喷雾。埃文斯充满热情地描述了大型强子对撞机是如何比预期还好地运行的。从他的话语中我感受到了鼓舞，但也感觉到了一丝惊讶，毕竟他要为整个机器的成功运行负责。那时我尚未能理解的是所有环节都组合在一起的速度有多快，这都要得益于在过去的机器上所积累的经验。

紧凑 μ 子线圈的一名意大利实验员毛里西奥·皮耶里尼 (Maurizio Pierini) 向我解释了埃文斯的话意味着什么。在同一隧道进行的大型正负电子对撞机中电子与正电子束的测试，曾经在20世纪80年代花费了25天，而现在不到一个星期的时间就完成了。

质子束非常精准、稳定。它们排成一列，只有少数丢失。光学检测正常，稳定性测试正常，重组过程正常。实际的质子束与计算机程序计算出的结果精确地契合了。

事实上，当实验者们于星期日下午5点（仅在新质子束开始环行之后几天）被告知，对撞就在次日进行时，他们无比惊讶。他们本以为在第一次质子束运行结束后，还需要更长时间才能开始进行真正的对撞，而后者才是他们需要记录与测量的实验过程。这是实验者们第一次有机会应用真正的质子束测试他们的实验仪器，而不是使用他们在等待机器运行过程中一直应用的宇宙射线。这条简短的通知意味着，他们必须快速调整自己的计算机触发器，以让计算机弄清要记录哪次对撞过程。毛里西奥描述了他们所有人的焦虑感，因为他们不想错过这次机会。在Tevatron上，第一次测试就被一次粒子束与读数系统的共振毁掉了。大家都不希望这种事情再次发生。当然，除了不安感之外，在所有参与者中，一股巨大的兴奋之情也在流传开来。

11月23日，大型强子对撞机终于开始了第一次对撞实验。数以百万计的质子以注入能量900GeV发生对撞。这意味着，在多年的等待之后，实验者们终于开始收集数据，即记录下大型强子对撞机加速环中首次质子对撞的结果。来自ALICE（某个小型实验仪器）的科学家们甚至在11月28日提交了一份尚未正式发表的论文的预印本。

不久之后，在一次适当的加速过程之后，质子束达到了1.18TeV的能标，这是环行质子束具有过的最高能标。仅在大型强子对撞机首次对撞的一周之后，即11月30日，这些更高能标的质子对撞了。净质心能量为2.36TeV，达到了有史以来最高的能标，打破了费米实验室8年以来的纪录。

三个大型强子对撞机实验仪器记录了质子束对撞，在接下来的数周中，这样的对撞进行了成千上万次。这些对撞的结果不会被用来发

现新的物理理论，而是被用来检验实验仪器是否运行（它们难以置信地好用），以及被用来研究标准模型的背景——即便它不会指出任何新事物，也可能与真正的发现之间存在潜在的作用。

实验物理学家到处宣扬他们对大型强子对撞机所达到能标的满意。令人注目的是，大型强子对撞机是在千钧一发之际做了这一切——按照计划，机器应该在12月中旬到次年3月之间的冬日停机，如果12月不能进行实验，就要拖延好几个月。来自圣塔芭芭拉大学、为大型强子对撞机工作的杰夫·里奇曼（Jeff Richman）在一次我们都出席的以暗物质为主题的会议上，高兴地与我们分享了这一点，因为我们都与费米实验室的物理学家打了赌：大型强子对撞机到底能不能在2009年之前超越费米实验室Tevatron达到过的能标？他兴奋的举止宣示了这场赌局的胜利。

2009年12月18日，在大型强子对撞机首次试运行结束之后的停机阶段，兴奋的浪潮暂时停歇了一阵子。埃文斯以在2010年大幅提高能标的承诺总结了演讲。他在当时于2010年年底把能标提升到7TeV，这是前所未有的剧增。他对此自信而充满热情——当机器回到人们的视野中，并达到如此高的能标时，埃文斯的表现合情合理。在一系列跌宕起伏之后，大型强子对撞机终于按照计划运行了（图9-2是一个简略时间表）。大型强子对撞机当时计划将在2012年以7TeV或者更高的能标运行，之后再停机至少一年，以为再度提高能标、尽可能地接近大型强子对撞机14TeV的目标能标做准备。在接下来的运行中，大型强子对撞机也将试图提高质子束的光度，以增加对撞的总数量。

2009年，大型强子对撞机回归人们视野之后，实验与机器本身的运作便一直顺利地进行着。埃文斯最近一次讲话的内容还在听众的脑海中回响着：

建造大型强子对撞机的冒险历程已经到此为止了。现在，让我们走进发现物理新世界的冒险旅程吧！

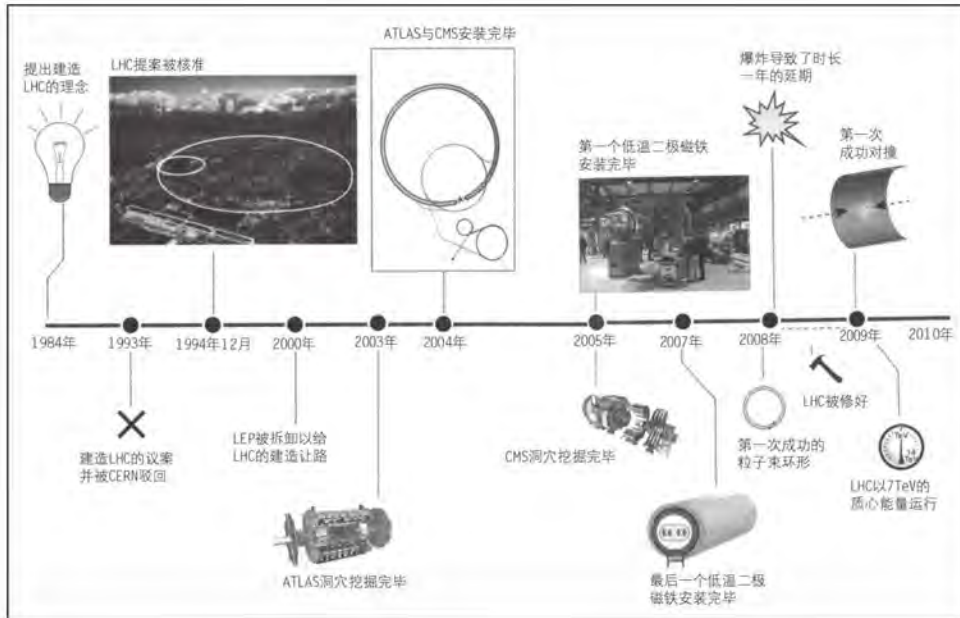


图9-2 该图简要介绍了大型强子对撞机的“发展史”。



很长一段时间以来，物理学家们都在等待着大型强子对撞机的开机运行。这些数据对科学进程而言十分重要，粒子物理学家们已经对高能情形下的实验数据求知若渴了。我们尚不知道隐藏在标准模型之下的各种假设中到底哪一个是正确的，而大型强子对撞机将提供这个答案。在这本书揭示一些有趣的可能性之前，接下来的几章中我们将绕路而行，以考虑一些有关风险与不确定性的重要问题，这些问题对理解如何解释大型强子对撞机的实验研究，以及一些有关现实世界主题的问题都是十分重要的。这场征程将从大型强子对撞机产生的黑洞，以及它们是如何收获大众杞人忧天般的关注开始。

担心世界毁灭的人们

物理学家们正在考虑大型强子对撞机最终可能发现事物的多种可能性。20世纪90年代，理论物理学家和实验物理学家们首次对某个新发现的，粒子物理学与引力本身都能在其中得以修正，并能在大型强子对撞机的能标上预测新现象的研究方向感到兴奋。这些理论中一个有趣的潜在推论激起了人们很大的兴趣，物理学领域之外的人对其兴趣尤甚，即微观低能黑洞存在的可能性。如果额外维度的观点确实正确，比如我与拉曼·桑卓姆（Raman Sundrum）提出的那套理论，这种

额外维度的微型黑洞也许真的可能产生。物理学家们乐观地预测，这种黑洞（如果被创造出来）可以为引力修正理论提供合理的证据。

并非所有人都对这种可能性抱有热情。在美国与其他地方的某些人担心，可能产生的黑洞会吞噬地球上的一切事物。我经常在许多公开演讲中被问及这种潜在的情景。在我解释了为何这种想法是杞人忧天之后，大多数的提问者都对我的回答十分满意。然而，并非每一个人都有机会聆听整个故事。

沃尔特·瓦格纳（Walter Wagner），曾是一位核安全官员，现在是律师、中学教师、夏威夷某座植物园的经理；西班牙人路易斯·桑科（Luis Sancho），一名作家、自封的时间理论研究者——这两人是这些杞人忧天的人中最激进的。他们两位激进到在夏威夷对欧洲核子研究中心、美国能源部、美国国家科学基金会与美国费米加速器中心提起了诉讼，以拖延大型强子对撞机的开机时间。如果他们的目标仅仅是拖延大型强子对撞机的运行，那还不如送出一只鸽子，让它把一片面包渣掉到大型强子对撞机的机器里更为简单呢（这种事情当真发生过，那只鸽子像是一名特立独行的特工^[41]）。但是，瓦格纳与桑科希望让大型强子对撞机永久地停止运行，所以他们依旧在奋力前进。

并不只有瓦格纳与桑科担忧黑洞可能引起的危机。由公益律师哈里·莱曼（Harry V. Lehmann）执笔的著作《量子中没有金丝雀：谁来决定大型强子对撞机是否值得以我们的地球作为赌注？》中也简明地概述了这些顾虑。一个相关主题的博客集中回顾了2008年9月的爆炸所带来的恐惧，并拷问：大型强子对撞机是否能再次安全开机？然而，最主要的顾虑并非集中于2008年9月19日灾难背后的技术失败原因，而是大型强子对撞机中可能产生的真正物理现象。

由莱曼与其他许多人对“末日审判机械”的描述所传播开的恐惧，主要集中于他们认为可能导致整个地球被吞噬的黑洞。他们担心，大型强子对撞机风险评估小组根据量子力学作出的研究缺乏可信的风险评估。

他们作出这一论断的主要依据是：

- 根据理查德·费曼与其他人的说法，“没有人真正懂得量子力学”。
- 弦理论之中尚有许多未知因素，它们有可能构成不确定性，而他们认为是有关系的。

他们的问题涉及如下要点：

- 是否不论风险多小、出于何种理由，对地球产生的威胁都不能被允许？
- 谁拥有上述问题的决议权？

虽然地球瞬间的毁灭实乃大事，然而，后一个问题往往更加适用于其他决议——比如涉及全球变暖的议题。我希望本章及下一章将能使你确信，与其担忧地球被黑洞吞噬的可能性，还不如担心美国401（K）法案^[42]的养老金被消耗光呢。尽管大型强子对撞机的计划表与预案都有风险，然而通过细致的调查、研究所修正的理论说明，黑洞并不具有任何危险。

必须明确，我这么说并不意味着我们没有必要提出这个问题。和其他人一样，科学家们也必须预估他们的行为所可能带来的危险结果。然而在黑洞的问题上，物理学家们建立在已有的科学理论和数据之上来评估风险并因此确定，并不存在令人不安的可能性。在前行到下一章讨论更一般的风险之前，本章将要探索，为什么某些人会考虑到大型强子对撞机产生黑洞的可能性，以及为什么人们对世界末日的担忧是一种完全被误导的想法。本章讨论的这些细节对下一章的一般性讨论而言并不重要，甚至对下一部分大型强子对撞机将要发现什么

的概述而言也不重要。但是，它是物理学家们如何思考、预测，并为接踵而来的风险评估打下基础的一个示例。

大型强子对撞机中的黑洞

黑洞是某种事物，它的引力之强足以把任何离它过近的物体吸引进去。在某个被称为“事件视界”（event horizon）半径之内的任何事物都会被吞噬，并被束缚在其内部。即便是如光一般轻盈的事物，也会沦陷于黑洞那巨大的引力场中。任何事物都无法从黑洞之中逃脱。一名《星际迷航》的铁杆粉丝开玩笑说，它们是“最好的博格人”。任何遭遇黑洞的事物最终都将被吞噬，因为引力的铁律宣示着“抵抗毫无意义”^[43]。

事件视界

一种时空曲隔界线。在事件视界之外的观察者无法利用任何物理方法获得事件视界之内的任何事件信息。它是从黑洞中发出的光所能达到的最远距离。



当足够多的物质被集中在一个足够小的区域中时，引力将变得不可阻挡，黑洞就产生了。形成黑洞所需区域的大小取决于黑洞的质量。更小的质量必须被集中在相应的更小区域中，而更大的质量可以分布在更大的区域中。不管怎样，当密度很大且临界物质处在所需的体积内时，引力就会变得不可阻挡，黑洞就会形成。从经典角度来看（这意味着在计算中忽略量子力学效应），当这些黑洞吸引附近的物质时会不断变大。

也是根据这种经典计算，黑洞永远不会衰亡。

在20世纪90年代之前，从没有人想过在实验室中制造黑洞，因为与典型的现代对撞机中的粒子质量或者能标相比，即便是制造最小的黑洞所需的质量也太过巨大。毕竟，黑洞有着极强的引力，然而任何我们已知单独粒子的引力都太过微小，远小于任何其他已知的相互作用力，比如电磁相互作用力。如果引力的强度与我们的预期相一致，那么在一个只有三个空间维度构成的宇宙里，在可达到的能标下，粒子的对撞远达不到所需的能量。然而，黑洞确实遍布宇宙——事实上，大多数大型星系的中心都有一个黑洞。然而，创造黑洞所需的能量至少比任何实验室能产生的能量要高出 10^{15} 倍（1后面接15个0）。

那么，为什么有人会提出大型强子对撞机中可以产生黑洞的可能性呢？理由是，物理学家们认识到，空间与引力的性质可能与我们迄今观测到的截然不同。引力也许不只在已知的三个特定维度中传播，也会在目前不可见也无法察觉的额外维度中传播。额外维度在我们已经作出的测量下没有产生任何可以辨识的效应。但当达到大型强子对撞机的能标时，额外维度的引力（如果它存在）就有可能表现出可被探测的性质。

如我们在第17章中所将见到的那样，在第7章中简洁地引入的额外维度是一个异乎寻常的观念，但是有着合理的理论基础，甚至有可能解释为什么我们已知的引力那么微弱。引力有可能在高维世界中很强，但是在我们观测到的三维世界中，它被“冲淡”而表现得很弱。或者，根据拉曼·桑卓姆和我的观点，它可能在额外维度中变化，因此在高维空间中，它可能在其他地方很强，然而在我们的位置却很弱。我们现在还不知道哪种观念是正确的。它们还远不能被确定，然

而如第17章中将要提到的，对大型强子对撞机的实验物理学家有可能发现的事物而言，他们都是最有希望的竞争者。

这样的情节意味着，当我们探索额外维度在原则上可能出现的更小尺度时，引力的一种非常不同的表现形式有可能出现。包含额外维度的理论预示着，宇宙的物理性质应该在我们很快就能探索到的高能标与小尺度下发生变化。如果额外维度的实在性确实与观测到的现象有关，那么在大型强子对撞机的能标下，引力效应将会变得比我们之前想象的更强。在这种情况下，大型强子对撞机的结果不会简单地取决于我们已知的引力，还会取决于高维宇宙中的更强引力。

如果引力真的那么强，那么质子就可以令人信服地在一个小区域中对撞，从而约束住创造高维黑洞所需的能量。如果这些黑洞能持续足够久，那么就可以吸入质量与能量；如果它们把这一过程无休止地进行下去，那么它们就会变得很危险。这正是那些担忧者们所设想的悲惨结局。

幸运的是，黑洞的经典计算（只依赖于爱因斯坦的引力理论）并不是决定这个结果的事物。霍金有很多成就，然而他的成名之作是，量子力学机制为被吸入黑洞的物质提供了一种逃逸的可能性。量子力学允许黑洞衰亡。

黑洞的表面很“热”，它的温度取决于质量。黑洞像炽热的煤块一样辐射，把能量向四面八方传播出去。黑洞依然吸引离它过近的物质，然而量子力学告诉我们，通过“霍金辐射”从黑洞表面蒸发出来的粒子带走能量，所以它慢慢地使黑洞的质量变小。这个过程甚至允许很大的黑洞最终辐射掉它所有的能量，从而消失。

由于大型强子对撞机最多也就能达到恰好可以产生黑洞的能标，唯一可能产生黑洞的形式必然是微型黑洞。如果黑洞产生时小而炽

热，正如大型强子对撞机中可能产生的那样，那么它非常有可能立即消失。根据霍金辐射而产生的衰亡会迅速让它覆灭。所以，即使更高维度的黑洞确实形成了（假设整个理论从一开始就是正确的），它们也不会存在足够长的时间而造成任何伤害。大型黑洞蒸发得十分缓慢，然而微型黑洞非常炽热，它几乎瞬间就把能量蒸发一空。从这方面来讲，黑洞是非常奇怪的。大多数物体，比如煤，会随着辐射而变冷，黑洞却随着蒸发而变热。最小的黑洞是最炽热的，因此其辐射也最为强烈。

由于我是一名科学家，我必须坚持严谨性。从技术上讲，以上论述的正确性取决于霍金辐射与黑洞衰亡确实存在。我们只理解非常巨大的黑洞，并精确地知道描述它们引力系统的方程。那些经过充分验证的引力定律给出了关于黑洞可信赖的数学描述。然而，对于极端微小的黑洞，我们并没有足够多的可信方程可以描述。于那些微型黑洞而言，量子力学将会生效——不仅是对蒸发过程，还对描述这些事物自身本质的过程。

没有人真正知道应该如何解决一个引力与量子力学共同支配系统中的问题。弦理论是物理学家们迄今为止最好的尝试，然而我们现在还不明白它的全部含义。这意味着，在原理上有可能存在漏洞。极端微小的黑洞只有通过量子引力理论来理解，它们不可能与我们应用经典引力理论导出的大型黑洞的行为相同。也许这种微型黑洞的衰亡速率压根就没有我们想象的快。

即便这样，这也并不是一个很严重的漏洞。只有一小部分人（如果存在的话）担心这类问题。只有可能长大的黑洞才可能具有危险性，微型黑洞不可能吸收足够多的物质而构成威胁。唯一潜在的风险是，微型黑洞在蒸发之前就成长到了危险的尺度。然而，即使是不能精确地知道那些黑洞是什么，我们也可以估计它们的持续时间。这些

估计显示，它们的寿命对形成一个危险的黑洞来说是如此之短，以至于即使是处于概率分布尾端最不可能发生的事情，仍然非常安全。微型黑洞的行为不会与我们熟悉的不稳定粒子有太大的不同。如这些短命的粒子一样，微型黑洞也会很快地衰亡。

某些人依旧担忧，即便符合一切已知的物理理论，霍金辐射依然有可能是错误的，而黑洞有可能完全是稳定的。毕竟，霍金辐射理论未经观察证实，因为已知黑洞中的辐射太弱，我们不可能探测到。物理学家们有理由怀疑这种说辞，因为那样他们不仅要抛弃霍金辐射理论，还要抛弃很多独立的、经过良好验证的物理理论。此外，霍金辐射理论之下的逻辑能够直接预测许多已经观察到的现象，这让我们对它的有效性有了更强的信心。

虽然如此，霍金辐射从来没有被观测到过。所以，为了确保绝对的安全，物理学家提出了如下问题：

如果霍金辐射出于某种原因并不正确，大型强子对撞机中也许能产生稳定、永不衰亡的黑洞，那么它们会是危险的吗？

幸运的是，存在更强的证据证明，那些黑洞不存在任何危险。这些讨论不需要有关黑洞衰亡的假设，也不是纯理论的探讨，而仅仅出于在宇宙中实际观测到的证据。2008年6月，两名物理学家，史蒂夫·吉丁斯（Steve Giddings）与米开朗基罗·曼加诺（Michelangelo Mangano）^[1]，以及不久之后的大型强子对撞机安全评估小组^[2]写下了明确的、以经验为基础的文章，令人信服地排除了任何黑洞灾难发生的情景。吉丁斯与曼加诺计算出了形成黑洞的可能速率，以及假使黑洞确实是稳定而不衰亡的，它们将在宇宙中产生什么样的影响。他们发现，即便我们还无法在地球上的加速器中达到足以产生黑洞的能量（即便是高维黑洞），在宇宙中这样的能量却经常可以被达到。由高能粒子组成的宇宙射线时时在空间中穿过，它们经常与其他事物碰

撞。虽然我们没有办法像我们在地球上进行的实验一样去研究这些碰撞结果的细节，但是这些碰撞的能量至少与大型强子对撞机能达到的能量相当。



如果额外维度理论是正确的，那么黑洞就有可能从天体物理学的事物中产生，甚至从地球或者太阳中产生。吉丁斯与曼加诺对某些模型（速率取决于额外维度的数量）进行了计算，结果显示，黑洞成长得太慢，无法构成任何危险——即便是在数十亿年之后，大多数黑洞依旧会保持极小的状态。在其他情况下，黑洞确实可能吸收足够的物质而长大，但是它们往往是载荷的。如果它们确实危险，那么它们很早之前就会束缚在地球与太阳的内部，并把地球与太阳的物质都吸进去。由于地球与太阳看上去一切正常，那些载荷黑洞（即便是那些迅速吸引物质的）也不会造成危险的结果。

所以，唯一的危险来自不载荷且能迅速长大的黑洞。在这种情况下，地球的引力（唯一能让它们慢下来的力）不会有效地阻止它们。这样的黑洞会穿过地球，但我们不能用地球的存在性来论证它任何的潜在危险。

然而，吉丁斯与曼加诺排除了即使在这种情况下危险性，因为大密度天体（即中子星与白矮星）的引力，能足够强到在黑洞逃逸之前阻止它们。随着极强的引力作用，超高能宇宙射线击打在这些高密度星体之上，并已经产生过大型强子对撞机中有可能产生的那种黑洞。中子星与白矮星比地球的密度还要大，大到它们的引力本身便足以把黑洞束缚在其内部。如果这样的黑洞确实被制造出来，也确实是危险的，那么它们就应该已经摧毁了我们已知存在了数十亿年的事物。宇宙中这种星体存在的数量表明，如果黑洞存在，它们就必然是安全的。即便黑洞确实形成了，它们也必须几乎立刻就消失，或者至少只留下无害而稳定的残余物。它们不会有足够的时间造成任何伤害。

除了这些之外，在吸引、毁灭物质的过程中，黑洞会释放出大量的可见光，而这种现象尚未被观测到。“我们所知的宇宙的存在”与“缺乏任何白矮星毁灭的信号”，是大型强子对撞机可能制造的黑洞并无危险性的有力证据。宇宙的这种状况告诉我们，地球不会被大型强子对撞机中的黑洞毁灭。

现在大家可以松一口气，舒缓一下心情了。然而，我还要把有关黑洞的故事继续讲下去——这一次是从我自己，一个在相关领域工作的物理学家的视角出发的。这些领域包括产生低能量黑洞所需的空间额外维度理论。

在黑洞论战吸引媒体眼球之前，我就已经对这个研究主题很感兴趣了。我的一位法国同事、朋友曾经任职于欧洲核子研究中心，现任职于一家称作奥格（Auger）的实验室。这家实验室主要研究穿过大气层来到地球上的宇宙射线。这个同事向我抱怨，大型强子对撞机分走了本可以用来研究相同能标的宇宙射线的资源。由于他的实验精确度太低，只有那些具有引人注目特征的事件才有可能被找到，比如正在衰亡的黑洞。

因此，我与一位时为哈佛大学博士后的同事帕特里克·米德（Patrick Meade）合作，计算出了他们可能观察到的这种事件的总数。在更精确的计算之后，我们发现这个数字远小于物理学家们原本乐观的预测。我在这里使用“乐观的”这一修辞，是因为我们总是期待出现新物理学的证据。我们并不关心地球上乃至宇宙中可能发生的灾难，而我相信你现在也应该认为，这并不是什么真正值得恐惧的事情。

在意识到奥格实验室不会发现微型黑洞之后，即便更高维度的粒子物理学现象解释正确，我们的计算也让我们十分好奇：为什么其他物理学家声称大型强子对撞机中可能会产生大量黑洞？我们发现，这个数量也可能被大大地高估了。虽然粗略的估计显示，在那种情况下，大型强子对撞机有可能大量地产生黑洞；然而我们更细致的计算却说明事实并非如此。

帕特里克和我并不关心危险的黑洞。我们希望知道小型、无害、很快衰亡的高维度黑洞是否能产生，以及它是否预兆着高维度引力的存在。我们的计算结果表明，就算有可能，其概率也非常小。当然，

如果真的有可能，那么这将是拉曼·桑卓姆与我提出的理论的一个良好证据。但是作为一名科学家，我必须承认计算结果。在这个结果下，我们不能抱有错误的期待。帕特里克与我（以及大多数物理学家）并不期待即便是小型黑洞的出现。

这才是科学家的工作方式。他们提出观念、粗略估算，然后再回头核实细节。事实上，细细推敲并修正原本的理念并非是愚笨的象征，而只是昭示着科学探索艰难而永无止境的标志。在我们确定理论、实验上的最佳结论之前，还要经历一些或是向前、或是向后调整的中间阶段。令人沮丧的是，帕特里克与我没有及时完成这些计算，导致我们未能阻止有关黑洞的论战深入媒体，并且引发了一场诉讼。

然而，我们的确意识到了，不论黑洞最终是否产生出来，有关其他大型强子对撞机中强烈相互作用着的粒子有趣的迹象，依旧会为引力与其他相互作用力的潜在本质提供重要的线索。我们还将看到在更高维度、更低能标下其他的信号。在看到这些奇异信号之前，我们知道并不存在制造黑洞的可能性。然而，这些其他信号本身也许最终可以阐明引力的某些性质。

这项工作证明了科学的另一个重要方面。即便科学范式有可能在不同范围的尺度下发生戏剧性的变化，但我们很少遇到数据本身发生突变的情况。已经被确认的数据有时确实会促成范式的变化，比如量子力学最终解释了已知的光谱线。在现行实验中发现的与预期的小小背离，往往会成为更加引人注目的证据的序幕。即便是科学的危险应用也需要一些时间来发展。科学家们应对核武器时代的某些方面负责，然而他们并没有出于意外就突然发现了一枚核弹。仅仅理解质量与能量的等价性是不够的。物理学家们必须更加努力，以把物质塑造成那种危险的爆炸形态。

如果黑洞有可能长大，那么它们确实有可能值得担忧，不过计算与观察都证伪了这种可能性。但是，就算它们可以长大，小型黑洞，或者至少是我们刚刚讨论过的粒子相互作用的引力效应，将会首先昭示出引力效应的变化。

最后我想说的是，黑洞不会有任何危险。但为了确保万无一失，我还是要在此承诺：我为大型强子对撞机不会制造出吞噬整个地球的黑洞负全部责任。



兰道尔和她的研究生是如何建议与核实这些事情的？扫码找答案吧！



538博客创始人内特·希尔（Nate Silver）——这位预言了2008年总统大选结果的最成功的预测家，2009年秋天为了写一本关于预测学的书来采访我。而此时，我们面临着一场经济危机、一场明显无法取胜的阿富汗战争、医疗保健费用的攀升、潜在的不可逆转的气候变化，以及其他种种迫在眉睫的威胁。带着一点互惠互利的想法——因为我也有意了解内特对概率与预测于何时、因何有效的看法，我同意了与他的会面。

尽管如此，内特为何选择采访我却令我困惑。因为我的专长是预测粒子对撞的结果，而我怀疑身处赌城拉斯维加斯的人（更不用说政府的人）可能会为此赌上一局。我猜想也许内特会问我有关大型强子对撞机的黑洞问题。但是，尽管当时被撤销的诉讼暗示着可能的威胁，鉴于前文所列举的几个更加真实的威胁，我猜想内特也会问那件事情。

事实上，内特并不关心这个话题。他问了一些更技术的问题，关于粒子物理学家如何在大型强子对撞机与其他实验中进行猜想与预测。他感兴趣的是预测学，而科学家的事情正是进行预测。他希望可以更多地了解：科学家如何选择问题，其用来猜想可能发生事情的方法（我后文会介绍），等等。

无论如何，在考虑大型强子对撞机实验与猜想可能的发现之前，本章继续讨论风险问题。有关现今的风险问题的奇怪态度，以及何时、如何应该期待它们的困惑，需要进行一番讨论。新闻曾报道了无数没有预期到的，以及未得到缓解的问题的不良结果。也许粒子物理与尺度分离的考虑可以为这些复杂问题带来一些启示。大型强子对撞机黑洞的官司当然非常有误导性，但是它与现今紧迫的事情无时无刻不在提醒我们关于风险这个专题的重要性。

进行粒子物理学预测与其他风险评估不同，而我们只能用一个章节来浅显地阐述同风险评估与缓解相关的内容。更进一步，黑洞的例子不宜推广，因为它的风险本质上说并不存在。然而，在考虑如何评估与说明风险时，它可以帮助引导我们确定某些相关的事宜。我们将看到，虽然大型强子对撞机的黑洞从来不是一个威胁，但由于预测造成的误导应用却是一个威胁。

世界上的风险

当物理学家考虑大型强子对撞机的黑洞时，我们从已经存在的科学理论中推断尚无法探索的能标上的物理学。我们有严格的理论考虑以及明确的实验证据得出结论：虽然我们还不知道将会发生什么，但未来不会发生任何灾难。在仔细考察之后，所有科学家一致认同，来自黑洞的危险风险可以被忽略，即使在宇宙的一生之中，它们也不可能构成任何问题。

这完全不同于其他潜在危险的解决方法。我仍然有些困惑：为什么几年前经济学家与金融学家未能预见即将到来的金融危机，或者甚至没有预见到一次危机被避免之后可能是为一次新的危机埋下伏笔。经济学家与金融学家在预测平稳的经济运行时并没有达成一个共识，以至于没有人进行干预，致使最后经济濒临崩溃。

2008年年末，我参与了一个跨学科会议的小组讨论。我被询问到关于黑洞威胁的问题，这并非第一次或者最后一次。坐在我右手边的高盛国际公司的副总裁同我开玩笑说：“每一个人所面临的真实黑洞危险其实是经济。”这个类比非常恰当。

黑洞通过强烈的吸引力捕获所有近邻的物质并将之转化。因为完全可以根据黑洞的质量、电荷与所谓的角动量来对它们进行分类，黑洞不记录进入的物质以及它们如何进入，所以进入黑洞的信息看似丢失了。通过走漏的辐射的微妙关联，黑洞非常缓慢地释放着信息。更进一步地，大黑洞的衰减缓慢，而小黑洞则立即衰减消失。这意味着小黑洞的寿命不会太长，而大黑洞本质上太大而不会消失。这是否使我们头脑灵光一现？信息以及债与衍生品被银行吸入而转化成不可解读的、复杂的资产。之后，信息以及其他进入的东西只能缓慢地被释放出来。

今天讲述了太多全局的现象，我们是在一个宏观的尺度上进行一个无法操控的实验。有一次在《海岸到海岸》（*Coast to Coast*）的广播秀中，我被问及，如果某个实验有可能威胁到整个世界，那么不管它有多么诱人，我是否会展开此实验。令绝大多数保守的广播听众懊恼的是，我的回答是：我们事实上已经在做这样的实验——碳排放。为什么人们不去担心它呢？

随着科学的进步，鲜有突然的变化在没有任何提前的暗示中发生。我们不知道气候将发生灾难性的变化，但是我们已经看到了来自融化的冰川与气候改变的暗示。经济也许在2008年迎来了突然的崩溃，但是许多金融学家也已知道得足够多，从而在崩塌前离开市场。新金融工具与高碳排放量是有可能酿成急剧变化的。在这些真实的世界条件下，问题不在于是否存在风险。在这些案例中，我们需要决定：

如果我们合理地考虑了可能的危险，那么应该如何决定一个可以接受的警觉程度。

计算风险是第一步

从理论上讲，第一步是计算风险。有时人们就是把概率算错了。当约翰·奥利弗（John Oliver）在《每日秀》（*The Daily Show*）中就黑洞的问题采访大型强子对撞机的一个诉讼当事人沃尔特·瓦格纳时，瓦格纳已经对大型强子对撞机的毫无信任，他说大型强子对撞机有50%的概率摧毁地球，因为这件事要么发生、要么不发生。奥利弗怀疑地回应说他“不确定概率原来是这样推算出来的”。值得庆幸的是，奥利弗是正确的，我们的确可以给出更好的（即不是那么平均的）概率估计。

情况并不总是那么简单。可以考虑不利环境变化的概率，或者中东恶性局面的概率或者经济的运行情况，这些是更加复杂的情况。不仅表述这些危机的方程难于被解决，实际上我们甚至连方程是什么都不知道。对于气候变化，我们可以进行模拟以及研究历史的记录；对于另外两个问题，我们可以试图探究历史上相似的情形，或者简化模型。但是在这三种情形中，许多不确定因素使得任何预测都显得苍白无力。

准确与可信的预测非常难以获得。甚至当人们竭尽全力把所有相关因素都考虑到模型中时，进入任何一个特定模型中的输入与假设也可能强烈地影响其结果。而如果伴随着这些底层假设的不确定性变得更多，那么一种低风险的预测就失去了意义。要使预测具有价值，不确定性的彻底性与直接性就至关重要。

在考虑其他例子之前，让我先通过回顾一个小逸事来阐述这个问题。

在我早期的物理学研究生涯中，我注意到，标准模型允许我们感兴趣的某个特定物理量拥有更广阔的数值范围。由于量子力学的贡献，该范围依赖于（当时）近期测到的、非常巨大的顶夸克的质量。当将结果在一个会议上汇报时，我被问及可否将我的新预测用顶夸克质量的函数曲线来刻画。我回答不行，因为我知道有几个不同的因素，而剩下的不确定性允许的可能性范围太宽，以至于不能给出这样一条简单曲线。然而，一个“专家”同行低估了这里面的不确定性，他给出了这样一条曲线（与现今关于现实世界所做的许多预测没什么不同），并且在一段时间内他的预测被广泛引用。最终，当测量到的物理量并没有落在他所预测的区域中时，这种不一致要追根溯源到他对不确定性过度乐观的估计上。显然，最好一开始就避免这种尴尬，无论在科学领域还是现实世界的任何情形中。我们期望预测有意义，并且如果我们对待输入的不确定性足够仔细的话，那么预测还是唯一的。

现实世界情形表述的是更为棘手的问题，要求我们更加仔细地对待不确定性与未知因素。我们必须谨慎地使用定量预测，而不能不考虑这些问题。

一个绊脚石是，如何合理地考虑系统风险，它们几乎总是很难量化。在任何大型的交互系统中，多个失效模型中的大型元件往往来自最不受注意的较小部件的互联。信息可以在转换中丢失或从一开始就没有参与进来，而且这种系统性问题可以将其他潜在风险的后果放大。

当我到一个委员会处理NASA安全问题时，第一次看到这种结构性问题。为了适应这种安抚不同选区的需要，NASA的站点遍及整个美国。即使任何独立站点会看顾各自的装置，在连接上它们却没有机构投资。接着，大型组织也援引这种做法。信息在不同次级系统的交互中很容易丢失。NASA与航空工业风险分析师乔·弗拉格拉（Joe Frago la）做了研究，他给我的一封邮件中写道：“我的经验表明，离开了项目专家之间的联合活动，系统集成团队与风险分析小组进行的风险分析注定是不全面的。特别地，所谓‘一站式’风险分析就变成了精

算练习，并且只有学术界才感兴趣。”在广度与细节之间往往有一个权衡，但是两者从长远来看都非常关键。

这种失败的一个显著案例是英国石油公司（BP）在墨西哥湾的漏油事件。

2011年2月在哈佛大学的一个演讲中，物理学家彻丽·默里（Cherry Murray）与BP公司深水地平线石油泄漏及海上钻井委员会的其中一员谈到，管理失败是BP公司石油事件的一个主要因素。理查德·西尔斯（Richard Sears），作为该委员会的资深科学家、工程顾问以及壳牌石油公司深水服务部的副总裁，认为BP公司管理层每次只解决出现的问题，而没有形成一个他所称的“超线性思维”的全局图。

虽然粒子物理学是一个特殊而艰难的行业，但是它的目标却是将简单的底层元素分离出来，以及基于我们的假设作出清楚的预测。其挑战在于达到小尺度与高能标，而不在于解决复杂的关联。即使我们不必知道哪一个基本模型是正确的，也可以作出预测（在给定模型的情形下）。例如，当大型强子对撞机的质子与质子对撞时，什么类型的事件可以发生。当小尺度被纳入到大尺度中时，适用于大尺度的有效理论可以精确地告诉我们小尺度何时进入，以及我们忽略小尺度细节所造成的误差。

然而，在绝大多数情况下，我们第1章介绍的由于尺度不同带来的清楚分隔并不适用。尽管有时可以共用方法，但是引用几位纽约银行家的说法，“金融不是物理的分支”。在气象领域与银行界，对小尺度关系的认知往往对决定大尺度的结果非常重要。

缺乏尺度分隔可能带来灾难性的后果。以巴林银行的倒闭为例。在它倒闭那一年之前，建于1762年的巴林银行是英国最老的商业银行。它曾经为拿破仑战争、路易斯安那购地案以及伊利运河提供资

金。而到了1995年，仅仅由一个小办公室的流氓交易员在新加坡作出的坏赌注就几乎将它带向了金融崩溃。

不久前，美国国际集团（AIG）的约瑟夫·卡萨诺（Joseph Cassano）的阴谋将AIG领向失败的边缘，AIG又面临着全世界主流金融行业崩溃的威胁。卡萨诺负责AIG一个相对很小（400人规模）的称为AIG金融产品的单位，即AIGFP。AIG之前一直在进行合理、稳定的投注，直到卡萨诺开始雇用信用违约互换（credit-default swaps，一种由各家银行推出的复杂投资工具）来对冲对抵押债务的投注。

在看似要恢复到传销时代的对冲计划中，卡萨诺的团队逐渐攀升到5000亿美元的信用违约互换，超过600亿美元与次级抵押贷款相联系。^[1]如果子单元可以被纳入到大系统中，就像物理中一样，那么较小部分就会以可以操控的方式在更高层次上产生信息或者活动，使得上层可以随时监管处理。但是对尺度分隔的过度违背，让卡萨诺的阴谋没有受到任何监管，并渗透到整个操作中。他的活动没有像证券、赌博、保险一样被规划。信用违约互换遍及全球，没有人研究过它的潜在意义。因此当次贷危机爆发时，AIG还没有准备好，它就崩盘了。美国纳税人最终要承担这些损失以拯救该企业。

出于对个别机构健全性的考量，安全监管机构开始在一定程度上关注传统的安全问题，但他们并没有对整个系统或与它相关联的风险进行评估。有重叠债务与义务的更为复杂的系统，要求更好地理解这些相互联系，并进行更全面的评估、比较、风险裁决，以及对可能的利益的权衡。^[2]这个挑战适用于绝大多数大系统以及被认为是相关的时间区域。

这让加重我们计算和处理风险的困难性的因素更多了：我们的心智、市场与政治体系对于长短期风险所采用的不同逻辑——有时是理智的，但常常是贪婪的。大多数经济学家与一些金融学家知道，市场

泡沫不会无限期地持续下去。风险不是指泡沫终究会破裂，而是指泡沫在不久的将来会破裂。驾驭或者膨胀此泡沫——甚至连你都知道是不能持续下去的泡沫，如果你随时准备好在任何时候获取盈利（或者股息）并关闭商店，那么也就不算是目光短浅。

在气候变化的情形中，我们实际上不知道如何刻画格陵兰冰层的融化量。如果我们询问它在有限时间区间（比如说在接下来的几百年）开始融化的可能性，那么它的概率就更不确定了。但是不知道这个数字，我们也无须把头埋进冰水或者格陵兰的原生态冰水中。

我们很难就以下两点达成共识：

- 气候变化到什么程度才算具有风险？
- 当环境中可能相对缓慢的后果出现时，我们如何、何时避免这些气候变化带来的风险？并且，我们也不知道如何估计行动或者不行动的代价。

如果有变化更强烈的气候事件发生，我们更可能立即采取行动。而无论我们行动多快，对环境来说都为时已晚。这意味着，即便是不会导致灾难的气候变化也同样值得人们关注。

甚至当知道某种结果的可能性时，我们倾向于将不同的标准应用于低概率且具有大灾难结果的事件，而不是应用于高概率但结果并不太具有戏剧化的事件。我们听到的关于飞机失事与恐怖袭击的事件比汽车事故的多，而实际上每年死于汽车事故的人数要多得多。人们甚至不知道概率就谈论黑洞，只是因为灾难情景的结果听起来很悲惨。另一方面，许多低（或者不是很低）概率由于在雷达下的可见度低而被忽略。甚至海上钻井一开始被认为是完全安全的，直到墨西哥湾灾难真实发生后人们才知道错了。^[3]

一个相关的问题是：有时巨大的收益或者成本来自概率分布的末端——那些最不可能的事件以及我们知道得最少的事件。^[4]从理论上讲，我们希望我们的计算结果是由已经存在的相关情形的中间或平均估计客观决定的。但是如果类似的事情从来没有发生过或者如果我们完全忽略了可能性，那么我们是不会有这样的数据的。如果成本或者收益在这些地方也足够高，假设你事先知道它们是什么，那么它们就会主导预测。无论如何，传统的统计方法不适用于这些数据太小、以至于取平均值没有意义的事件。

金融危机的发生是由于事件都出现在专家所考虑的范围之外。许多人基于可预测的方面赚钱，但是假设不可能事件决定了某些更负面的发展事态。在为金融工具的可信度建立模型之时，使用的绝大多数数据来自前几年，它们没有包含经济下滑（或急剧下滑）的可能性因素。关于是否需要调控金融工具的评估，也基于市场上升的那段时间。甚至当市场跌落的可能性也考虑进去时，跌落的假定数值也因为过低而不能准确地预言由于缺乏经济调控而造成的真实损失。本质上讲，没有人会关注可以主导危机的“不可能”事件。也许在其他情况下看起来很显而易见的风险，却从未被考虑在内。但是即使是不可能的事件，如果它们具有显著的影响，也需要被考虑进来。^[5]

由于错误的潜在假设而导致的风险评估困难，是任何风险评估的疫病之源。而没有这些估计，任何估计都存在内在偏见。除了计算中的问题与隐藏在假设中的偏见，许多实际的政策决策还包含“不知之不知”（unknown unknowns）^[44]因素，即不能或者还没有被预期到的因素。有时我们根本无法预知会造成麻烦的那些确切的不可能事件。这使得所有的预测尝试（考虑不到这些未知因素是不可避免的）变得完全没有实际意义。

缓解风险

幸运的是，我们在求知探索的过程中，百分之百地确信制造出危险的黑洞的概率微乎其微。我们不知道出现灾难结果的确切的概率数值，但我们也无须知道，因为它完全可以被忽略。如果任何事件在宇宙的一生中都不会发生一次，那么它可以被放心地忽略。

然而更一般地，量化一种可以接受的风险的水平是极其困难的。我们当然希望完全避免主要的风险——那些使生命、地球或者其他我们珍视的东西濒临绝境的风险。在我们能承担的风险之下，我们希望有方法来估计谁能受益、谁会失利，以及随之而来的可以评估与预期风险的系统。

风险分析师乔·弗拉格拉关于气候变化以及其他相伴的潜在危险的评论是：

真正的问题不是这些是否会发生，也不是它们的后果如何，而是它们发生的概率有多少、相关的不确定程度是多少。不仅基于发生的概率，也基于我们可能的举措来缓解它们的概率。有多少全球资源可供我们支配来解决这些危险？

调控往往依赖于所谓的成本-收益分析来评估风险与决定解决方案。表面上看，这个想法听起来很简单。计算支出与收益，并考察所提出的改变量是否值得。在许多情况下，这也许是最可行的方案，但是它可能产生一种危险的、但貌似数学严谨的欺骗性。而实际上，成本-收益分析是非常困难的。问题所涉及的不仅是成本与收益的度量——这本身已经是一种挑战了，而且还要事先定义什么是我们所指的成本与收益。许多假定情况涉及太多未知因素，以至于无法进行可以信赖的计算或者率先计算风险。我们当然可以尝试，但是需要将这些不确定性考虑进去，或者至少识别出来。

一种可以预期近期与将来的成本与风险的合理体系，毫无疑问是有用的。但是并非所有的交易都可以完全通过它们的成本进行评估。

如果承担风险的东西是根本不能被取代的，又当如何？^[6]要是大型强子对撞机可以在我们一生（或者甚至在百万年）中以一个相当高的概率产生可以吞没地球的黑洞，那么我们当然必须把这个实验终止。

即使我们最终会大大地得益于基础科学的研究，抛弃一个项目的经济成本也鲜少可以计算出来，因为收益很难量化。大型强子对撞机的目标包括获取基础认知，包括对质量与相互作用的更好理解，甚至可能包括对空间本质的理解。其好处也包括促使受过良好教育的大众深思与宇宙相关的问题。从更实用的角度来讲，我们一直在追随欧洲核子研究中心通过互联网创造的信息进步，网格（grid）使得全球信息处理成为可能，而电磁技术的进展产生了核磁共振等医疗设备。基础科学进一步应用的可能性可能会存在，但目前这些都是无法预期的。

成本-收益分析很难应用于基础科学之中。一个律师开玩笑地将成本-收益方法应用于大型强子对撞机并指出，与所提出的极其微小可能性的巨大风险相伴，大型强子对撞机也具有以极小可能解决世上所有问题的惊人效果。当然，尽管无数多律师都尝试过，没有哪一个结果恰好与标准的成本-收益计算相吻合。^[7]



至少科学得益于它“永恒的”真理目标。如果你发现了世界运行的方式，那么不管你多快或者多慢地发现它，它都是真实存在的。我们当然不希望科学的进展很慢。但是大型强子对撞机的延期显示了太快运行它的危险。总的来说，科学家致力于安全的前行。

成本-收益分析对于任何复杂情况来说，都充满难度，例如气候变化政策或者银行业务。虽然原则上来讲成本-收益分析是合情合理的，并且可能也不存在什么禁忌，但是你怎么应用它却可以产生巨大的差异。本质上讲，当成本-收益分析的捍卫者问我们如何可能做得更好时，他们会用成本-收益论证来证明一种方法，而且他们往往是对的。我只是倡议我们将方法应用在哪里，且更科学地应用它。我们需要清楚地知道我们所展示数据中的不确定性。在科学分析中，我们需要考虑错误、假设、偏差，并将它们公开出来。

气候变化一个很大的影响因素是：成本或者收益针对的是局部地区、国家还是全球。潜在的成本或者收益也可以越过这些范围，但我们并不总是将这些考虑周全。

美国政治家反对《京都议定书》的一个原因是：它对美国，特别是对美国的商业来说，成本将超出其收益。但这种算法并不真实，因为它没有考虑长期的全球不稳定性的成本，或者由于环境调控带来的可能繁荣的新商业收益。许多有关缓解气候变化成本的经济分析未能考虑，由于革新或者对国外依赖程度减轻，而产生的潜在附加经济收益。这里牵涉了太多未知因素。

这些例子也产生了如何评估与缓解跨国风险的问题。设想，黑洞真的造成了对地球的威胁，那么在夏威夷的某人可以合法地起诉这样一个在日内瓦进行的实验吗？根据现有的法律，答案是否定的，他只能在国内提起诉讼，干预美国财政对该实验的投入。

核扩散是另一个将全球稳定置于危险之中的问题。目前我们对于其他国家产生的危险只有有限的约束力。气候变化与核扩散看似是一国内政，但其影响力却可能会涉及其他国家或机构。当风险超越国界或法律管辖权时，如何去做就成为困难的政治问题。但它显然是一个重要的问题。

作为一个完全国际化的机构，欧洲核子研究中心的成功关键在于它的多方成员共有同一个目标。一个国家可以试图减少它自身的贡献，但是除此以外，各自的利益并不牵涉其中。所有参与的国家一起协同运作，因为科学的价值对它们来说是相同的。主办国——法国与瑞士，也许在劳动力与基础设施上占据一点经济优势，但总体而言，这不是一场零和博弈。^[45]并不存在一国受益而另一国失利的情况。

大型强子对撞机的另一个显著的特征是：如果发生任何技术或者实用问题，欧洲核子研究中心与其成员国需对其负责。2008年液氦泄露爆炸事故的维修费用需要由欧洲核子研究中心的预算支付。没有人，特别是在大型强子对撞机工作的人，可以从机械事故或者科学灾难中获利。当成本与收益不能完全匹配、受益者对风险不负全责时，成本-收益分析就无用武之地。这与科学常常试图解决的封闭体系问题所采用的论证方法非常不同。

在任何情况下，我们都希望避免道德风险，也即，当人的利益与风险不一致时，人们可能有动机承担更大的风险。我们需要合理的动机。

例如对冲基金。一般合伙人在赚钱的时候，每年会从基金获得一定比例的盈利，但是如果基金面临损失或者破产时，那么他们也不会放弃相当比例的盈利。个人获得利润，而他们的雇主（或者纳税人）分担损失。有了这些参数，对雇员来说最具有盈利性的策略将促成浮动与不稳定性。一种有效体系与有效的成本-收益分析将对风险、回馈与责任的分配负起责任。它们必须将涉及不同类型与规模的人们的因素分离开来。

银行业务在风险与收益不一致时也会面临明显的道德风险。一种“大到不能失败”的政策与弱杠杆限制的结合会产生一种情况：对损失负责的人们（纳税人）与利益的最主要代表方（银行与保险公司）不相一致。人们可以就2008年的救市是否有必要展开争论，但最初通过责任来调整风险以阻止情况恶化的举措看似是一个好方法。

而且，所有关于大型强子对撞机实验与风险的数据都是现成的。安全报告放在网上，任何人都可以查阅。当任何机构如果在失利时或者存在潜在的不稳定性时，期待获得救援，就应该提供足够的证据，以使相关的收益风险比可以被外界评估出来。准备好可供访问的可靠数据，能够帮助抵押贷款专家、监管机构或者其他机构来预测未来的金融风险或其他风险。

虽然解决方法不在它自身，但是另一个至少可以改善或者澄清分析的因素是：在利益与风险类别方面，将尺度与时限考虑在内。尺度问题可以转变成计算中牵涉的是谁的问题：它是单独个体、一个组织、一个政府还是整个世界，以及我们所关注的期限是一个月、一年还是十年？对高盛公司来说好的政策，可能最终并不能使整体经济受益或者使个人受益。这意味着即使存在完全准确的计算，也不能保证结果正确，除非它们被应用在基于问题的正确、仔细的思考上。

当制定政策或者评估成本与收益时，我们倾向于忽视可能来自全球稳定局面与帮助他人的收益——不仅出于道德考量，而且出于长期考量。一部分原因是，这些增益很难量化；还有一部分原因是，在瞬息万变的环境中做评估与制定强有力的法规所面临的挑战。不过，很显然，考虑了所有可能好处的法规，不仅仅是那些针对个人、机构或国家的法规，而且更加可靠，甚至可能创造一个更美好的世界。

时间框架也可以影响计算成本或者收益的政策决定，如同决策方所做的假设，也如我们所看到的金融危机。在其他方面，时间尺度也

有影响，因为行动过于匆忙会增加风险，而快速交易则可以提高收益（或利润）。但即使快速交易可以使定价效率更高，但快速交易并不一定能使整体经济受益。一位投资银行家向我解释可以出售股份是多么重要，但即便如此，他也无法解释为什么他们在持有仅仅几秒甚至更短的时间之后，就需要出售，而事实上，他与银行赚了很多钱。这种交易在短期内为银行与金融机构创造了更多的利润，但从长远来看，它们加剧了金融行业现有的弱势。也许一个在短期竞争中不占优势的系统却可以激发人们更多信心，在长期范围内可能更有利可图，因此是更好的选择。当然，我提到的那位银行家在一年内就为他的机构赚了20亿美元，所以他的雇主可能不会认同我的建议。但是任何最终为这笔利润埋单的人则可能会认同。

专家的角色，科学里没有金色降落伞

许多人听信了错误的教训，并得出结论认为缺乏可信的预测意味着缺乏风险。事实上恰恰相反。在我们可以确切地排除特别的假设或者模型之前，可能的结果是包含在所有可能范围之内的。尽管存在不确定性以及诸多预测危险结果的模型，但在气候或经济（或者海上钻井）中糟糕之事发生的概率是小到可以忽略不计的。也许有人会说，在一个明确的时间跨度之内，这些风险发生的机会是很小的。然而，从长远来看，除非我们有更好的信息，在很多情况下忽略这些危险将导致灾难性的结果。

只对底线感兴趣的人会联合起来反对调控，而那些对安全与预测感兴趣的人则为其辩护。人们很容易在两方阵营动摇，因为想在两者之间画一条界线本身就是一个让人望而生畏的任务。在计算风险时，不知道决策点不表示它不存在或者我们不应该定下最好的趋近目标。即使没有可以作出细致预测所必须的启示，也应当先解决结构问题。

这带给我们最后几个重要问题：

- 谁来决定？
- 专家的角色如何？
- 谁能评估风险？

鉴于大型强子对撞机所涉及的资金与管理机构以及审慎的监督，我们可以认为风险已经被充分地分析过了。而且，在它所能达到的能标上，我们并未处在粒子物理学基本出发点即将失效的新区域。物理学家很确信，大型强子对撞机是安全的，并且我们期待粒子对撞的结果。



这并不是说科学家对此不负重大责任。我们总是需要确保科学家对风险有责任而且非常谨慎。我们需要像对待所有科学机构一样，对大型强子对撞机满怀信心。如果你在创造物质或者微生物，或者其他以前不存在的东西（或者为了得到它，在地球上钻出深井或者进行其他前沿探索），你需要首先确信不会产生什么极其糟糕的后果。关键是理智地做事，而不是相信毫无理由的、从而可能阻碍进步与收益的谣言。不仅科学如此，任何有潜在风险的努力也是如此。对于想象中的未知因素，甚至对于“不知之不知”的答案就是，理性以及在需要时实施干预的自由度。在墨西

哥湾的任何人都可以证明，你需要具有在出错时将输油开关关闭的能力。

在本章前面，我总结了博客作者与怀疑论者对物理学家所采用的黑洞计算（包括基于量子力学）方法的抗议。霍金确实使用量子力学得出了黑洞衰亡的猜想。尽管费曼提出“没有人真正理解量子力学”，物理学家却明白如何应用它，即使我们对量子力学为何是正确的，无法形成什么哲学洞见。我们相信量子力学，因为它解释了数据并且解决了经典物理学无法解释的问题。

当物理学家为量子力学争辩不休时，他们争论的不是它的预测能力。它一再的成功“迫使”一代又一代的学生与研究人员接受其理论的合理性。现今关于量子力学的论辩在于它的哲学基础。是否存在其他拥有经典前提的理论，然而又可以预言出量子力学的怪异假设？即使人们在诸如此类的问题上取得了进展，它也不能对量子力学的预测产生任何影响。哲学进步可能影响我们用于表述预测的观念框架，却不能改变预测本身。

以此为鉴，我发现在这个方向上有重大进展是不可能的。量子力学可能就是一个基本理论，它比经典力学丰富。所有经典预测都是量子力学的极限形式，但是反之不然。因此很难相信我们最终可以用经典的牛顿逻辑来诠释量子力学。试图用经典基础来诠释量子力学的行为，就好像我用意大利语写这本书。所有我能用意大利语说出来的话我也可以用英语说，但是由于我的意大利语的词汇量有限，反过来则不一定了。

虽则如此，无论是否赞成哲学的出发点，所有物理学家都对如何应用量子力学具有认同感。量子力学的预测是值得信赖的，而且已经被验证过很多次了。哪怕没有它们，我们仍然有其他实验证据（以地

球、太阳、中子星与白矮星等的形式)来证明大型强子对撞机是安全的。

大型强子对撞机的危言耸听论者也反对使用弦理论的意图。的确,使用量子力学还勉强可以接受,而依赖弦理论则是不可能的。但是关于黑洞的结论却从来不需要弦理论。人们的确竭力使用弦理论来理解黑洞内部,即根据广义相对论能量密度变成无穷大的表观奇点(apparent singularity)^[46]的几何学。并且人们已经通过基于弦理论的计算给出非物理条件下的黑洞蒸发,来支持霍金的结果。但是黑洞衰亡的计算依赖于量子力学,而非完全的量子引力理论。甚至在没有弦理论的情况下,霍金也可以进行他的计算。一些博客所贴出的问题说明,缺乏足够的科学理解来权衡这些因素。

对于这种反对的一种普遍解释是:反对不是针对科学本身而是针对科学家们理论中“基于信仰”的信念。毕竟,弦理论已经超出了实验可以验证的范围。然而许多物理学家认为它是对的,并且继续研究它。但是科学界关于弦理论的各种观点,很好地说明了正好相反的观点。没有人会考虑弦理论中的安全因素。一些物理学家支持弦理论,一些则反对。然而他们也知道它既没有被证明,也没有完全被排除。在人人都认同弦理论的有效性与可靠性之前,在有风险的情况下应用弦理论将是鲁莽的行为。至于我们的安全问题,“弦理论尚未取得实验结果”,不仅是我们还不知道它是否正确的原因之一,也是它不需要我们在有生之年预测将遇到的真实现象的原因。

尽管我确信可以信赖专家,来评估来自大型强子对撞机的潜在风险,我也意识到了该策略可能存在的限制因素,而且我也不清楚如何解决它。毕竟,“专家”告诉我们,衍生品是最大限度减少风险的方法,而非创造潜在的危机。经济学“专家”告诉我们,放松管制对美国企业的竞争力是至关重要的,而不会导致美国经济发展的衰

落。“专家”还告诉我们，只有银行业的人充分理解了他们的交易，才能解决其困境。那么问题来了，我们如何确认专家的考虑是足够全面的呢？

显然，专家们也可能毫无远见，他们的利益也可能发生冲突。我们可以从科学中学到什么经验吗？

我下面的话并非出于偏见。在大型强子对撞机可以产生黑洞一事上，我们审查了可以在逻辑上设想的所有潜在风险，考虑了理论论证与实验证据。我们还想到了在宇宙中使用相同的物理条件，而没有破坏任何邻近结构的情况。

经济学家对存在的数据也作了相似的比较，如此乐观是很好的。但是卡门·莱因哈特（Carmen Reinhart）与肯尼斯·罗格夫（Kenneth Rogoff）的著作《这次不一样》（*This Time Is Different*）却不这么认为。虽然经济条件并不相同，但是在经济泡沫中，一些广泛的措施却会一再出现。

现今，许多人认为的没有人可以预期放松管制所带来的危险的论断也站不住脚。商品期货交易委员会（CFTC）的布鲁克斯丽·波恩（Brooksley Born）曾经对期货与商品期权市场作出预测。她指出了放松管制的危险，事实上她还相当合理地建议了对潜在风险的探索，但是她却被喝止了。当时并不存在关于多少警戒是合理的具体分析，但是只有一个派系认为行动缓慢对商业发展不利（正如华尔街在短期所呈现出来的那样）。

对调控与政策发表过观点的经济学家，可能在政治与金融方面拥有影响力，而这可能会干扰到事情的正确执行。理想情况下，相比于政治，科学家们更关注具有优越性的论点，甚至包括那些与风险相关的论点。大型强子对撞机的物理学家做了严肃的科学调查，以确保没有灾害发生。

虽然也许只有金融专家才了解某个特定金融工具的细节，但是任何人都可以考虑一些基本的结构性问题。大多数人甚至在没有预测、不理解可能导致崩溃发生的触发因素时，都能理解为什么一个过度杠杆调控的经济体是不稳定的。并且大多数人都可以理解，用纳税人的钱不加限制或者限制很少地提供给银行上千亿美元的资金，可能也不是什么好办法。甚至一个水龙头的安装方式也要考虑到开、关的可靠方法，或者即便是一个拖把的设计和摆放，也要考虑其对环境整洁度的影响。人们很难理解为什么同样的考虑不能应用于深海石油钻井平台的设计上。

当我们考虑专家时，心理因素介入了。《纽约时报》的经济专栏作家大卫·伦哈特（David Leonhardt）撰文道，在2010年时，将经济学家艾伦·格林斯潘（Alan Greenspan）与本·伯南克（Ben Bernanke）的错误归因于“比经济因素更重要的心理因素”。他解释说：“他们被困在传统智慧的恶性循环中。困扰‘挑战者’号航天飞机的工程师、越南与伊拉克战争的设计者，以及驾驶舱发生悲剧性错误的飞行员等人的弱点，也困扰着他们两人。两人没有充分地质疑自己的假设。这完全是人为错误。”^[8]

解决复杂问题的唯一方法是广泛听取他人意见，甚至是外行人的意见。尽管银行家有能力预测可能会陷入黑洞的经济，但孤芳自赏的银行家们却满足于现状，而不顾警告。我们都聚在一起，为正确的答案投票，在这个意义上，科学是不民主的。但是，只要有人提出一个有效的科学观点，它最终会被注意到。通常人们会关注来自更杰出科学家的发现与见解。尽管如此，一个不知名的人，只要能提出一个好的观点，最终他也会获得人们的关注。

由于一个知名科学家的聆听，一个不知名的人甚至可以马上被他人注意到。这就是为什么爱因斯坦能通过一种理论，就几乎立即动摇了科学的基础。而理解了爱因斯坦相对论见解的德国物理学家马克斯·普朗克，那时恰巧负责管理最重要的物理学杂志。

今天，各种想法都可以通过互联网快速传播，这让我们受益颇丰。任何一个物理学家都可以写一篇论文，并通过物理资料库第二天将其发送出去。当卢博胥·默托（Luboš Motl）在捷克读本科时，他解决了罗格斯大学（Rutgers University）一名杰出科学家正在研究的科学问题。汤姆·班克斯（Tom Banks）会注意任何好的想法，即使它们来自一个他从未听说过的研究所。不是每个人都那么容易接受新观点，但哪怕只有几个人关注，一个好想法最终也将进入科学的殿堂。

大型强子对撞机的物理学家与工程师为了确保安全不惜牺牲时间与金钱。他们在不丧失安全性或准确度的前提下，尽可能地节省开支。每个人的利益都是一致的。没有人可以在经不起时间考验的结果上获得任何好处。

科学注重的是声誉，科学中没有金色降落伞。^[47]

超前预测

希望大家现在都同意，我们不应该担心黑洞——虽然我们有许多其他事情值得担心。在大型强子对撞机的情况中，我们应该思考它能提供的所有好东西。它所产生的粒子将有助于我们解答有关物质基础结构的深层次、根本性问题。

回到我与内特·希尔的谈话，我意识到了我们的特殊情况。在粒子物理学中，我们可以把自己限制在足够简单的体系内，采用系统的方法在旧的基础上建立新的结果。有时，我们的预测源于基于现有证据产生的正确模型。在其他情况下，我们的预测基于那些我们有理由相信可能存在，并通过实验滤掉了一些可能性的模型。甚至在还不知道这些模型是否正确之时，我们可以预测实验证据会是什么，以及这些想法是否可以被实现。

粒子物理学家根据不同尺度各自施展才华。我们知道，小尺度的相互作用与在大尺度上的相互作用非常不同，但是它们以一种合理的方式反映在大尺度的相互作用中，给出我们已经知道的结果。

而在所有其他情况下，预测是非常不同的。对于复杂系统，我们经常需要同时解决发生在一系列尺度上的问题。这不仅对社会组织有效，例如一个不负责任的交易者会动摇美国国际集团与社会经济，而且还可以发生在其他科学领域。这些情况下的预测会包含很多变量。

例如，生物学的目标包括预测生物模式，甚至动物与人类的行为。但是，我们还没有完全理解所有的基本功能单元，或者通过基本元素产生复杂的效果来理解更高级别的组织。我们也不知道所有的反馈回路，即那些使得不同尺度下的相互作用相分离成为不可能的因素。科学家能建造模型，但是如果不能更好地理解关键性的基本元素，或者理解它们如何对相关行为负责，那么建模的人将会面临数据与各种可能性相争的谜团。

进一步的挑战是，生物模型的设计是为了与已经存在的数据相匹配，但是我们还不知道其规则是什么。因为我们还没有确认所有简单而又独立的系统，所以很难知道哪个模型是正确的。当我与研究神经科学的同事交谈时，他们谈到了同样的问题。没有新测量的质的飞跃，模型所能做到的最好的程度就是与所有现有数据保持一致。既然所有存留下来的模型必须与数据一致，所以很难确定哪一种基本假设是正确的。

与内特讨论这类他试图预测的事情不乏趣味。许多最近流行的书籍都提出一些不靠谱的假设。内特则更多地使用科学方法。他最初因对棒球运动与选举结果的准确预测而出名。他的分析建立在对过去相似情况仔细统计与评估的基础上，他会竭尽所能地纳入尽可能多的变

量——那些他可以操控、可以尽可能精确地应用到历史教训中的变量。

内特现在必须明智地选择在哪里运用他的方法。但他意识到，他所关注的各种关系，都可能是难以解释的。你可以说发动机起火导致了飞机坠毁，在失事飞机中找到一个起火的引擎并不是一件多么令人吃惊的事情。然而什么才是根本原因？同理，当你将一个突变基因与癌症相联系的时候，你也会遇到类似问题。即使它是相关因素，但它不一定是造成疾病的根本原因。

内特也意识到了其他潜在陷阱。即使存在大量数据，随机性与噪声也可以增强或抑制让人感兴趣的底层信号。因此，内特不会做与金融、地震或气候相关的预测工作。虽然他很有可能可以预测总体趋势，但是短期预测本身就是不确定的。内特现在研究的是其他方面，他的方法揭示了诸如如何最好地推广音乐与电影、如何评估NBA巨星等问题。但是他承认，只有极少数的系统可以准确地被量化。

尽管如此，内特告诉我天气预报员采用的确实是另一种预测方法。他们中的许多人都在做“超前预测”——预测人们将尝试的预测。



熟知概率与统计学对科学测量的计算大有裨益，更不用说它们在当今这个复杂世界的许多疑难事宜中所起的作用。概率逻辑之优美令我想起几年前的一件事。

我的一个朋友问我是否打算参加第二天晚上的一个活动，而我的回答是“不知道”，这着实令他沮丧了一番。然而对我而言十分幸运的是，他是一个擅用数学的赌徒，所以他没有坚持问我要一个肯定的答案，而是让我告诉他，他的胜算有多少。令我吃惊的是，这个问题要容易回答得多。即便我告诉他的只是一个粗略估计的概率，却比直接给出“是”或“否”的答案更能反映我当时的权衡和不确定性。结果是，它反而更像一个真诚的回应。

从那以后，每当我的朋友和同事认为他们不能回答某个问题时，我都尽量采用概率的方法向他们解释。我发现绝大多数人（无论是不是科学家）都有着强烈，但并非不可改变的观点——他们常常觉得概率的表达方法更合适。比如，有人很可能不知道三周后的那个周四晚上他是否想去看一场棒球赛。即便他知道自己喜欢棒球而且又想不出有没有其他工作安排，然而他还是会有所顾虑——毕竟那是在周中，所以即使他不能百分之百地确定，但也有80%的可能性会去看。虽然这只是一个估计，但是在这个点上他所给出的概率更准确地反映了他的真实期望。

在一场关于科学以及科学如何运作的谈话中，作为编剧和导演的马克·维森特（Mark Vicente）注意到，在遇到相同的事时，与绝大多数人不同，科学家不愿做明确、不合适的评论，这令他很吃惊。科学家虽然不必做到最准确，但是他们的宗旨是清楚地表达（至少是在他们作为专家的领域内）他们知道或者不知道、理解或者不理解的东西。所以他们鲜少说“是”或者“否”，因为这种答案并不能准确地反映各种可能性。因此取而代之，他们会以概率或者其他合适的方式来表达。不过具有讽刺意味的是，这种行业的差别常常导致人们误解或者低估科学家的论断。尽管科学家旨在提高精确度，非专业人士却不必知道如何权衡他们的论断，因为非专业人士一旦有足够多支持其论点的证据，他们会不加顾虑地说出更确信的东西。科学家没有百分之百的信心不代表他们缺乏认知。简单来说，这是存在于任何测量本质中的不确定性的结果，也是我们现在要展开探索的主题。**概率思维有助于澄清数据与事实，并且为更全面的决策提供可能。**在本章节，我们将思考测量能告诉我们的东西以及为什么概率表述会更准确地反映认识的程度，这些认识是在任何特定时间的对科学或其他事情的认识。

科学的不确定性

哈佛大学近来完成了一个课程评估，试图确定素质教育的核心内容。教员们考虑、讨论并打算将其纳入科学类的一个课程类别是经验推理（empirical reasoning）。这项教学提案显示了大学的宗旨应该是：教导学生如何搜集和处理经验数据，权衡证据，理解概率估计，当条件成熟时，从数据中得出推理，并在问题不能在现有证据基础上被解决时，及时作出反应。

这项教学要求的议案及后来的补充，其主旨非常好，但是它暗含了一个关于测量如何发挥作用的基本误解。科学家在解决问题时通常

会使用概率。当然在某些特定的观点和观察上，我们可以获得确证并可以通过科学来作出合理的论断。但是人们只能偶尔依据科学或者其他方法在事实的基础上绝对正确地处理问题。虽然我们可以搜集足够多的数据来确信因果关系，甚至作出令人难以置信的精确预测，但是这种事情具有相当的概率性。正如第1章所讨论的，不确定性（不管多么小）允许有趣的潜在新现象被发现。极少有事物是百分之百确定的，在任何还没有做过测试的条件下，没有哪些理论与假设完全适用。



现象只能在它们可以被测试的有效范围内，以一定的精确度呈现出来。测量往往都存在概率。许多科学测量基于一个假设：存在着我们可以通过足够清晰和准确的测量来揭示的潜在本质。我们尽量做到精确的（或者说好到足以达到我们的目标）测量以发现这些隐藏的事实。这因此存在一种情况，例如，一系列测量结果的中心区域有95%的概率包含了测量的真实值。在这种情形下，我们可以说我们有95%的把握。这样的概率告诉我们任何特定测量的可信度以及整个概率与含义的范围。如果你既不知道它伴随的不确定性也对其没有确定估计，你就不可能完全理解一个测量。

不确定性的一个来源是缺乏无限精确的实验装置。一个精确的测量可能要求测量仪器必须校准到小数点后的无穷多位，测量值因此就会精确到小数点后的无穷多位。实验物理学家不可能有这样的测量，他们只能校准他们的仪器到技术允许的可能精度——就像天文学家第

谷·布拉赫在四个世纪以前所做的那么专业。技术的不断发展促使测量仪器精度不断提高。即便如此，测量也永远达不到无限精确。一些系统的不确定性 (*systematic uncertainty*) [48]，即测量仪器本身的特质，总是会存在。

不确定性不表示科学家对所有选项或者表述都一视同仁（尽管新闻播报时常犯此错误）。这些二选一的选项仅在很少的情况下各居50%概率。但它们意味着科学家（或者任何追求完全准确的人）会作出声明，告诉人们哪些已经被测量了，哪些是以概率的形式体现的，哪怕这概率非常高。

当科学家与文人墨客都极为小心谨慎时，他们将“精确”（precision）与“准确”（accuracy）区别对待。一个装置是“精确的”意味着：当你重复测量单一数量时，你所记录的数值之间相差无几。“精确度”是描述变化程度的指标。如果重复测量的结果变化不大，那么测量就是精确的。越精确的数值所跨越的范围越窄，如果你重复测量，那么平均值也越快收敛。

精确度

当你重复测量单一数量时，你所记录的数值之间相差无几。“精确度”是描述变化程度的指标。

“准确度”告诉你的是：测量的平均值与准确结果接近的程度。换言之，它描述了测量装置是否有偏差。从技术上来讲，虽然测量装置的内禀误差不会降低它的精确度（因为你每次会犯同样的错误），但是它会毫无疑问地降低你的准确度。系统的不确定性反映了源于测量仪器本身的无法避免的准确度缺失。

准确度

“准确度”告诉我们的是测量的平均值与准确结果接近的程度。换言之，它描述了测量装置是否有偏差。

然而在许多情形下，即便你可以构造出完美测量的仪器，你仍需采取多次测量来得到正确的结果。这是因为另一个不确定性的来源^[49]是“统计性的”，也就是说测量通常需要重复很多次才能给出你所信赖的结果。所以，即便是一个准确的设备也不一定在任一特定的测量中给出一个正确的结果，但是多次测量的平均值会收敛落到正确的结果上。系统的不确定性掌控了测量的准确度，而统计的不确定性影响其精确度。一个好的科学研究在这两个方面上都要考量，因此测量要在可行的范围内、尽量多的样品上尽可能仔细地展开。理论上说，你想让你的测量既准确又精确，以至于所期的绝对误差很小，因此你可以信赖你所发现的结果。也就是说，你想让数值落到一个尽可能窄的范围内（精确），并让它们收敛到正确的数值上面（准确）。

我们可以考虑一个熟悉且重要的例子：药物疗效实验。医生通常不会讲他们可能也不了解相关的统计。当你被告知“这种药有时有效、有时没效”时，你有没有感到沮丧？不少有用资讯被这种表述所抑制，使人对这种药物的有效性充满疑惑。于是，怎么做成为一件难以抉择的事情。一个相对较为有用的表述应该告诉我们：药物或者疗程以怎样的一个比例，在年龄和胖瘦程度相似的病人身上起了什么作用。这样，即便医生们自己不懂统计学，他们也可以肯定地给出一些有用的数据和信息。

平心而论，人的差异性加之个体对药物的不同反应，使得断定一种药品是否有效成为一个复杂的问题。所以让我们先考虑一个简单一些的情形，来检测一个单独的个体。

我们用测试阿司匹林是否有助于减轻头疼的过程来作为例子。解决的方法看似相当简单：吃一颗阿司匹林看是否有效。但是实际情况要稍微复杂一点。因为即便你感觉好些，你又怎能知道一定是阿司匹林起的功效呢？为了能确切知道是不是阿司匹林的效用（即，是否你服药了头疼就会减轻，或者疼痛好得比没有服药时快），你需要有吃药与不吃药的比较。然而，你要么吃药要么不吃，单单一次测试是不足以告诉你答案的。

奏效的方法是做很多次测试。每次一旦头疼，抛硬币来决定是否需要吃一颗阿司匹林并且记录结果。当你做了足够多测试，把各种不同类型的头疼，以及变化的环境（例如，当你没那么困倦时头疼好得更快一些）取平均，并用统计来得出正确的结论。假定你的测试没有偏差，因为你是用抛硬币来决定是否吃药，并且你所采用的人口样本就只有你自己，那么你的结果会在自己所采用的足够多的测试上正确收敛。

要是总能了解药物在这样简单的程序下有效与否那该有多好。但是绝大多数药物治疗的都是比头疼严重得多的疾病，而且许多药物还有长期效应，因此哪怕你想，你也不能在一个病患身上反复地进行短期试验。

通常当生物学家或者医生测试一种药品的功效有多少时，他们不是研究单一的病患，即便从科学的角度上看他们更愿意那样做。因此他们必须接受一个事实，那就是人们对同种药物的反应不同。任何药物都产生不同结果，哪怕是在病情严重程度一样的群体上进行测试。于是在多数情况下，科学家所能做到的最好的事情是：在他们决定要给单一病患用药前，研究尽可能与其相似的群体。然而在现实中，多数医生并不自己开展研究，因此他们很难保证病人病情严重程度的相似性。

医生可能打算转而试图使用已经存在的研究——那些还没有人做过仔细设计的实验，而结果简单地来源于对已经存在的群体的观察，比如说对美国健康维护组织（HMO）成员的观察。他们就会面临如何作出正确诠释的挑战。在这些研究中，要确定相关测试导致的是因果关系而非其中的关联性，这是很困难的。比如有人可能会错误地论断说：黄色手指会导致肺癌，因为他们发现很多肺癌患者的手指都是黄色的。

这就是为什么科学家比较喜欢研究中的治疗结果或对象都是随机选取的。例如，一项吃药与否由抛硬币决定的研究会更少地依赖于人口样本，因为病患是否服药取决于抛出硬币的随机结果。相似地，随机研究原则上也能揭示吸烟与肺癌和黄色手指的关系。如果你将一个群体的成员随机地分配到吸烟组或者禁烟组，你会确定吸烟对于你所观察的病患至少是一个对黄色手指与肺癌两种现象都负有责任的潜在因素，但不能确定其是否造成了这种结果。当然这种特殊的研究是不合伦理的。

在任何可能的时候，科学家总是尽可能地简化系统，来分离出他们想研究的特定现象。选择明确的人口样本与合理的控制组，对于结果的精确度与准确度都至关重要。一些类似人体生物学中药效如何作用的复杂问题总是伴随着诸多同时发挥作用的因素。于是一个相关的问题是：这些结果需要有多少可信度？

测量的客观性

测量永远不可能完美。在科学研究（以及任何决策）中，我们需要决定不确定性的可接受程度。这使我们可以不断向前。举例来说，如果你想通过吃药来减轻头疼的症状，如果那种药可以有效帮助哪怕只有75%的一般患者（只要副作用很小），那么你可能也愿意试试。反过来，如果调整饮食可以使你从本来已经很低的心脏疾病的可能性中减低0.1个百分点的风险，比如从5%减少到4.9%，那么恐怕这也不足以令你担心到让你放弃自己最爱的波士顿奶油派的地步。

关于公众政策，决策点可能更不清晰。公众观点通常具有一个灰色区域——在改变律法或者实施法规时，人们不必赞同我们对于某些事情的了解程度有多准确。许多因素使必要的计算变得更加复杂。就

像在前文讨论的在目标与方法中存在的歧义性使得成本-收益的分析变得非常难于可靠地（即便不是完全不可能）进行。

《纽约时报》的专栏作家纪思道（Nicholas Kristof）在关于谨慎对待食品与容器中含有的潜在有害化学物质BPA^[50]的辩论中写道：“BPA的研究已经敲响警钟几十年了，证据依然复杂而且处于争论中。这就是生活：在真实世界，监管决策时常必须由有争议与有冲突的数据组成。”^[1]

所有这些并非表明我们不应该在评估政策时立足于成本与收益的定量估计。然而这表示了我们应该清楚地知道评估的意思，以及根据假设或目标，成本、收益可以改变的程度，计算中哪些被考虑了、哪些没有。成本-收益分析可以是有用的，但也可以显现出具体性、必然性与安全性的错觉，从而可能导致其在社会上被误用。

对物理学家来说幸运的是：通常我们所问的问题比公众政策的问题（至少在制定政策方面）简单得多。当我们处理纯粹的知识而非其应用时，我们会进行不同的调查。测量基本粒子至少在原则上简单得多。所有电子的本质相同。你必须考虑统计与系统误差，但不用考虑种群的差异性，一个电子的行为可以代表它们所有电子的行为。但是对相同的统计与系统误差，科学家试图在可行的范围内将这些减到最小，然而他们能达到的程度取决于他们想回答的问题。

尽管如此，即便在“简单的”物理系统中，测量也从来不是完美的，因此我们需要确定准确度目标。在应用水平上，相同的问题是，实验物理学家要重复多少次测量、测量仪器需要准确到什么程度。答案取决于物理学家自己，可以接受的不确定性范围有赖于他所问的问题。不同的目标需要不同的准确度与精确度。

例如，原子钟测量时间的稳定性达到了十万亿分之一^[51]，但是很少有测量需要对时间达到这么高的精度。测量爱因斯坦的引力理论是一个例外，它需要使用尽可能高的精确度和准确度。即便目前的所有

测试都证明这个理论正确，测量还需要不断改进。有更高的精确度，至今还未被发现的偏差所代表的新物理效应就可能显现出来，而这在以前精确度较低时是不可能的。如果有，偏差就会为我们提供对新物理现象的重要洞察力；如果没有，就可以相信爱因斯坦的理论比我们已经论证的还要更准确，就可以在更高能标、更广距离范围、更高准确度上放心地使用它。反过来，如果你要将人类送去月球，你必须充分懂得物理定律，这样才可以正确地瞄准火箭方向，但是你无须使用广义相对论——你当然也无须考虑可能存在的偏差的潜在影响。

走运的粒子物理学家

在粒子物理学中，我们会探索可以探测的统治最小与最基本物质组分的基本定律。单次实验并不测量混杂在一起的许多单次碰撞或多次复杂碰撞。我们所做的预测应用于已知粒子在给定能量上的单次碰撞。粒子进入碰撞点，相互作用，沿着探测器飞行，沿途沉积能量。物理学家刻画粒子碰撞乃是依据粒子飞出时的不同性质，即其质量、能量及电荷。

从这种意义上说，尽管我们的实验技术面对不少挑战，粒子物理学家却很走运。我们研究尽可能基本的系统，以便分离出基本的组分和定律。该想法是，让实验体系在现有资源的允许下尽量单纯。物理学家面临的挑战是获得所需的物理参数而非解开复杂体系。科学为了得到有趣的结果必须要推动认知的前沿，所以实验是很困难的。它们因此往往处在技术所能达到的能量与距离的极限处。

事实上，粒子物理学实验并不都那么简单，即使它仅仅研究基本物理量。实验物理学家在展示他们的结果时面临两种挑战。如果他们确实发现奇异的东西，那么他们还必须证明它不是寻常标准模型的事件。另一方面，如果没有发现新东西，那么他们还必须确定其准确度

足以展现一个更严格的新极限，使得事物可以出现在标准模型以外。他们必须充分了解测量设备的灵敏度以便知道哪些情况是可以排除的。



为了确定结果，实验物理学家必须将代表新物理的事件从标准模型已知物理粒子的背景事件中区别出来。这是我们需要许多碰撞来探索新发现的原因之一。许多碰撞的出现，确保了有足够多事件所代表的新物理，能有别于与之相似的标准模型的常规反应。

故此，实验需要适当的统计。由于测量本身的一些内在不确定性，需要重复进行实验。量子力学告诉我们基本事件也有不确定性。量子力学意味着不管我们如何将技术设计得何等精良，我们也只能计算相互作用发生的概率。无论我们如何测量，不确定性依然存在，即为了准确测量相互作用的强度，唯一的方法是重复测量很多次。有时这种内在不确定性比测量技术的不确定性还小，且小到可以忽视。但是有时我们仍需要将其考虑在内。

例如，量子力学的不确定性告诉我们，衰变粒子的质量是一个内禀的不确定物理量。该原理告诉我们，当测量发生在有限的时间内，所测的能量就不可能精确。因为测量的时间一定比衰变粒子的寿命短，这确定了质量测量所期的变化量。所以如果实验物理学家想通过

发现粒子衰变产物来发现新粒子，就需要反复地测量多次质量。虽然单次实验不可能精确，但多次测量的平均值可以收敛到正确的数值上。

在许多情况下，量子力学质量不确定性比测量仪器的系统不确定性（内禀误差）小。如果那样，实验物理学家就可以忽略量子力学的质量不确定性。尽管如此，由于相互作用中的概率本质，大量测量还是需要用来确保精确度。就像药品测试的例子，大量统计可以帮助我们得到正确的结论。

一个很重要的认识是：**伴随着量子力学的概率不是完全随机的。**该概率可以通过良好定义的定律来进行计算。在第14章讨论W玻色子质量时，我们将看到这一点。用一条曲线描述给定质量和寿命的粒子在一次碰撞中产生的可能性有多少，我们知道了该曲线的形状。每次能量的测量值会落在以正确值为中心的一个区域，其分布与粒子寿命和不确定性原理相容。即使单次测量不足以确定质量，多次测量却可以确定。一个明确的方案告诉我们如何从重复测量的平均值推导质量。足够多次的测量保证了实验物理学家在一定的精确度和准确度水平上决定正确的质量。

大型强子对撞机的全新可能性

应用概率来呈现科学结果与量子力学的内在概率性，两者都不代表我们什么都不知道。事实恰恰相反，我们知道很多。例如，电子磁矩（magnetic moment）是电子的一个内禀属性，我们可以用量子场论来进行极其精确的计算。量子场论结合了量子力学和狭义相对论，是一个研究基本粒子物理属性的工具。我在哈佛大学的同事、物理学家杰拉尔德·加布里埃尔斯（Gerald Gabrielse）已经测量了电子磁矩，精确到小数点后13位，并且该值与理论预测值在这个水平上是吻

合的。不确定性在万亿分之一的水平上才参与进来。这使得电子磁矩作为一个自然参数在理论与测量之间有着最精确的吻合。

物理学之外，没有谁能如此准确地预测这个世界。但是绝大多数人一旦有这样精确的数字，就会认为自己确切地知道理论与它预言的现象。科学家虽然比其他人能够给出更准确的论断，他们还是认为测量与观察的精确度不管多高，都始终为现今未知的现象与新观点留下了空间。

他们也可以给新现象的尺度设定一个确定的极限。新的假设可能改变预测，但只能发生在目前测量的不确定性或更小的水平上。有时预测的新效应是如此微小，以至于在宇宙的一生中我们都无法指望与之相逢。在这种情形下，甚至连科学家都可能断言说“这件事不会发生”。

显然，加布里埃尔斯的测量显示，量子场论在一个很高的精度上是正确的。即便如此，我们也不能断定量子场论、粒子物理学或者标准模型就是所有理论了。正如第1章所阐述，新现象的效应只能在不同能标时显现，或者当我们能做更精确测量时暗含在我们所看到的东西里。我们还没有从实验上研究该能量和尺度的范围，因此我们现在还不得而知。

大型强子对撞机实验在前所未有的高能标上进行，因此它提供了新的可能性。实验探求的新粒子与相互作用很直接，而不是那种需要用极端精确的测量来确认的间接效应。在所有可能性中，大型强子对撞机的测量不会达到足够高的能量来发现离开量子场论的偏差。但是它们可能展示其他现象，预测出与现今精度水平上的标准模型所预言的测量相偏离的结果（甚至是现在测得很准的电子磁矩）。

对于任何标准模型之外的物理模型，任何预测的小差异都是真实世界基本属性的一大线索。到目前为止还没有出现这些差异以告诉我们需要发现新事物的精度和能量水平，即便我们不知道潜在新现象的精确属性。

在开篇所介绍的有效理论是指，我们仅仅全面了解我们所学习的是什么，以及在哪一个点上它们失效的极限。有效理论整合了现有的约束——不仅在给定的尺度上归类我们的想法，而且提供了系统的方法来确定新效应在特定能量上有多大。

关于电磁相互作用力和弱相互作用力的测量与标准模型的预言在0.1%的精度上吻合。粒子碰撞速度、质量、衰变率，以及其他性质也都在这个精确度和准确度上相吻合。标准模型为新发现留下了空间，它们必须小到足以避开现今的探测。新现象或者暗含的理论效应必须小到目前还不能看见，或是因为相互作用本身很小，或是因为效应所伴随的粒子太重，以现在的反应能量还不足以产生它们。现有的测量告诉我们在多高的能量上能发现新粒子或者新的相互作用，同时在当前不确定性允许的条件下不造成对测量的更大偏差。这些测量也告诉我们新事物是多么少见。但随着测量精度变得足够高，或者实验发生在不同的物理条件下，实验物理学家就可以搜寻模型的偏差，虽然该模型迄今为止描述了所有实验粒子物理学的结果。



当今的实验基于一个通识：新想法都建立在一个成功应用于低能量的有效理论基础。这些实验的目标都是为了揭示新物质与新的相

相互作用，并牢记物理是按一个又一个能标建构知识的。通过研究大型强子对撞机的更高能标，我们希望发现并完全理解到目前为止我们所见事物的背后理论。在测量新现象以前，大型强子对撞机数据将为我们提供宝贵、严格的约束，限制标准模型之外可能存在的现象和理论。如果我们的理论考虑是正确的，那么新现象最终会在大型强子对撞机现在所研究的高能区域中出现。这样的发现迫使我们推广或者将标准模型纳入一个更复杂的理论体系中。这个更全面的模型将在一个更大的能量尺度上更准确地与实验符合。

我们不知道哪个理论会在自然中实现，也不知道何时才能给出新发现。答案依赖于有什么，并且我们还不知道它是什么或者我们应该看什么。对于任何有关何种新事物存在的特定设想，我们知道如何从实验结果中发现、如何计算以及估计它何时可能发生。在后文中，我们将探讨大型强子对撞机的实验如何工作、实验可能会发现什么。



2007年8月，西班牙物理学家和欧洲核子研究中心理论组组长路易斯·阿瓦瑞智·高密热情地鼓励我加入一次参观超环面仪器（ATLAS）实验的旅程。并且，物理学家彼得·詹尼（Peter Jenni）与法比奥拉·贾诺蒂（Fabiola Gianotti）还计划拜访诺贝尔奖获得者李政道以及其他几位物理学家。想要抗拒彼得与法比奥拉满怀感染力的热情几乎是不可能的。他们当时乃是超环面仪器实验的发言人与副发言人。他们所有的话，字里行间都充满了对实验细节的专业与熟稔。

我和同行的参观者戴上安全帽进入了大型强子对撞机的隧道。我们的第一站是一个可以向下俯视到一个深井的平台（见图13-1）。这个巨大的洞穴有竖直的管道可以将探测器从我们所站的地方输送到100米以下的地面。我们几个参观者都热切地期盼一睹其真容。

第一站以后，我们继续走到下面的地板上，那里堆放着还没有安装好的超环面仪器探测器。没完工最妙的事情就是你可以看到探测器的内部。一旦最终封闭就什么都看不见了，只有在大型强子对撞机停止实验的检修维护期间，才可能再次看见其内部结构。所以我们才有这样一个机会可以直接观看这个精密工程，它是那么色彩斑斓又庞大无比，甚至比巴黎圣母院大教堂的中殿还大。

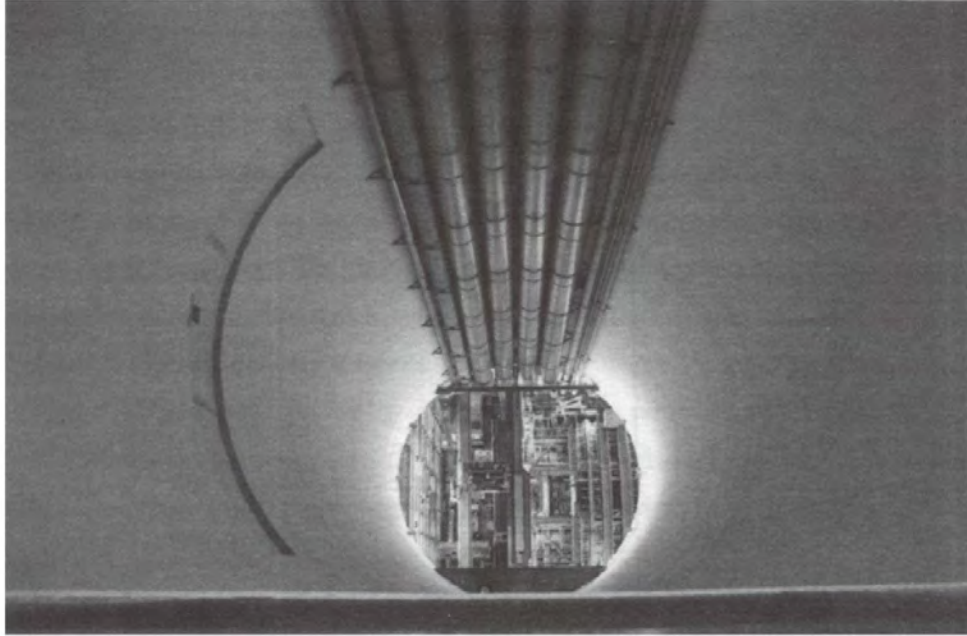


图13-1 从平台向下俯瞰超环面仪器的深井，视线中的管道可以将材料输送下去。

但是最宏伟的不仅仅是超环面仪器的规模。我们当中来自纽约或者其他大城市的人不会仅仅因为一个庞大的建筑而感到震撼。超环面仪器真正雄伟壮观的是：如此巨大的一个探测器竟然是由许许多多小的探测元件组成的。有些元件设计成可以测量到微米量级的精度。最具讽刺意味的事情正体现于此：我们需要用如此庞大的实验装置来测量如此微小的距离。当我展示一张来自报告会的探测器的照片时，我不得不强调说超环面仪器不仅巨大而且构造精密。这是它最神奇的地方。

2008年，我又重返欧洲核子研究中心并再次目睹了超环面仪器的工程进展状况。2007年时探测器的两端还开放参观，这时已经封闭。我又踏上了一次参观紧凑 μ 子线圈（CMS）的壮观旅途。紧凑 μ 子线圈是大型强子对撞机的第二个通用型探测器。这次同行的还有物理学家辛西娅·达维亚（Cinzia da Via）和我的合作者吉拉德·佩雷斯（Gilad Perez，即图13-2中的男子）。



图13-2 我的合作者吉拉德·佩雷斯站在一层层的紧凑 μ 子线圈 μ 子探测器/磁返回轭前面。

吉拉德还没有参观过大型强子对撞机的实验室，所以我借此机会通过他的兴奋回想起了我上一次的经历。我们趁着监管不严，攀上攀下，甚至俯瞰了一个粒子束管道（见图13-3）。吉拉德觉得这里可能会成为一个额外维度粒子产生的地方，进而为我以前提出的理论提供

证据。但是无论它将为我的模型还是其他模型提供证据，它都给了我们一个很好的提醒：这个管道将洞悉即将出现的新元素。

第8章介绍了大型强子对撞机，它用来加速质子并使它们对撞。本章重点讲述大型强子对撞机的两个通用型探测器——紧凑 μ 子线圈与超环面仪器，以及它们的粒子对撞产物。其他大型强子对撞机实验如ALICE、LHCb、TOTEM、ALFA以及LHCf，都设计有特殊的用途，包括对强相互作用力的理解以及底夸克的精确测量。这些实验很可能让我们可以更加细致地研究标准模型，但是它们不太可能发现超越标准模型的新能标下的物理现象，而这才是大型强子对撞机的首要目的。紧凑 μ 子线圈和超环面仪器是用来测量和揭示新现象、新物质的首要探测器。



图13-3 辛西娅·达维亚（左图）正走过那个我们可以俯瞰粒子束管道并观看其内部构造的地方（右图）。

本章会涉及大量技术细节。即便像我这样的理论物理学家也不需要知道这些因素。总的来说，如果读者仅仅对我们可能发现的新物理或者大型强子对撞机的概念感兴趣，那么可以选择跳到后面的章节。

然而，大型强子对撞机实验巧夺天工、令人难以置信，省略这些细节仍将损失一二。

一般原理

在某种意义上说，超环面仪器与紧凑 μ 子线圈探测器是几百年前伽利略和其他发明家策划改良的逻辑演变。自从显微镜发明以后，技术的不断发展让物理学家可以研究越来越小的尺度。小尺度上的研究不断地揭示物质的结构，这只能通过小的探针来观测。

大型强子对撞机的实验都设计成可以研究十亿亿分之一厘米尺度上的结构与相互作用。这大约是以往实验观测尺度的1/10。虽然大型强子对撞机探测器与此前的高能对撞实验，例如美国费米实验室的Tevatron，基于相同的原理，但新的探测器在能标与对撞速度上创下新纪录，也将面临众多新挑战，同时也迫使它们在尺度与复杂程度上开创新河。

与太空望远镜类似，这些探测器一旦建成，本身就很难再触碰。它们都被安置在很深的地下而且有大量辐射。一旦探测器开始运行，没人能再靠近它们。哪怕没有运行，要想接近任何专门的探测元件也是极其困难的。因此探测器建好之后至少可以使用10年而无须修缮。不过为了大型强子对撞机每次能持续运行两年，在两次运行之间会有一个很长的关闭时期，此时物理学家与工程师可以接触到许多探测器部件。

粒子实验与望远镜在一个重要方面上有着本质的不同。粒子探测器是不需要指向一个特定的方向的。从某种意义上讲，它们同时观看各个方向。对撞发生，粒子出现。探测器记录的任何事件都可能是有意义的。超环面仪器与紧凑 μ 子线圈是通用型探测器，它们不只记录一种粒子或事件抑或是只关注某一特殊过程。这些实验装置都设计成

可以从最广泛的相互作用与能量区域来汲取数据。实验物理学家有极强的计算能力，他们试图从实验记录的图像中明确地解读出粒子和其衰变产物的信息。

来自38个国家183个研究所的超过3000名科学家参与到了紧凑 μ 子线圈的实验中，从事建造与操纵探测器、分析数据的工作。最早作为副发言人的意大利物理学家圭多·托内利（Guido Tonelli）当时是这项合作项目的领队。

优秀的意大利女物理学家法比奥拉·贾诺蒂也从超环面仪器另一个通用型实验的副发言人转为发言人，从而打破了欧洲核子研究中心由男性物理学家主导的传统。她很适合担当此角色。她有着温和、友好、礼貌的举止，而且她在物理学和组织管理方面的贡献也很重要。然而，令我忌妒的是，贾诺蒂还是一个技艺精湛的大厨——虽然对意大利人而言，对厨艺细节的过分关注是可以理解的。

超环面仪器也有巨大的合作圈。来自38个国家174个研究所的超过3000名科学家参与了超环面仪器的实验（2009年12月）。这项合作最早形成于1992年的两个实验——精确光子、轻子与能量测量实验（Experiment for Accurate Gamma, Lepton, and Energy Measurements, EAGLE）以及超导环场仪器（Apparatus with Super Conducting Toroids, ASCOT）的合作，这两者在设计上与以前的一个提案——超导超级对撞机（SSC）探测器在某些方面有着相同的面貌。1994年最终的项目书提交了，两年后资助被批准。

两个实验在基本框架上是相似的，但是在细节的构造与应用上却不同（见图13-4）。每个实验都有不同的长处，这种互补关系使得物理学家可以交叉检验实验结果。由于粒子物理学的发现极具挑战性，两个实验有着相同的搜索目标，当它们的发现互相印证时，结果就会有更强的可信度。如果它们给出相同的结果，那么每一人都会更加确信。

两个实验的出现也引入了一种强烈的竞争元素，这是我的同行常常提醒我的。这种竞争催促他们更快、更详细地得到结果。两个实验的成员们也在互相学习。往往一个好的办法在两个实验中都有用武之地，即便它们的应用有些不同。这种竞争与合作，与两套基于不同构造与技术的独立探测系统地交融在一起，其本质是因为两个实验拥有相同的目标。

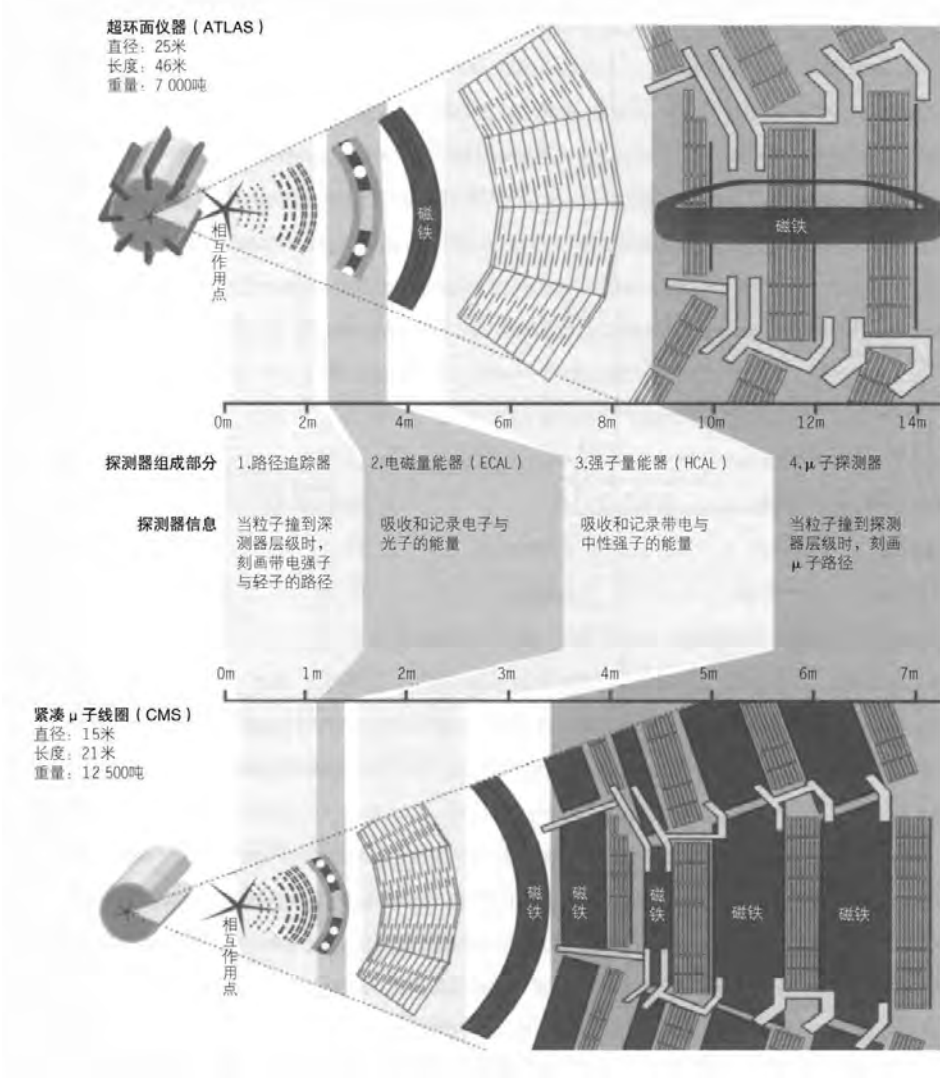


图13-4 超环面仪器和紧凑 μ 子线圈探测器的横截面（注意这个图的尺寸已经全部重新调整过）。

我经常被问到，大型强子对撞机何时将运行我的实验以检验我和合作者提出的特定模型。答案是马上，但是他们也同时在寻找其他人的方案。



理论物理学家通过引入新的搜索目标和新的策略来帮助寻找新物质。我们的研究目的是尽力确定新方法，以期在 高能标下可以显现一些新的物理元素或新的基本作用力。这样，物理学家就可以寻找、测量以及诠释这些实验结果，进而从浩如烟海的数据中洞见隐于其中的理论实质，不论它最终呈现出何种风貌。只有当这些数据被记录下来，各个分析团队的无数实验物理学家才能研究这些信息，以确定这些数据与我们的或是其他有潜力的提案相容或是相悖。

理论物理学家与实验物理学家接着会检查这些数据看它们是否遵从某些特殊假设。即使许多粒子寿命只有几分之一秒，即使我们没有直接观测到它们，实验物理学家也能使用数字数据来刻画这些组成物质核心的粒子，以及它们相互作用的“图像”。考虑到这些探测器与数据的复杂程度，实验物理学家总是会有很多令他们满意的信息。本章后文将为读者介绍具体这些是什么样的信息。

超环面仪器与紧凑 μ 子线圈探测器

到目前为止，我们跟随大型强子对撞机的质子从它们由氢原子中剥离出来，到它们约27公里的圆环上被加速到高能。两个完全平行的质子束永远不相交，两束质子也不会朝着相反方向运动。接着在沿

着环的几个位置上，二极磁铁将它们从其路径上分开，而四极磁铁则将它们汇聚，使两束质子可以在一个直径小于30微米的区域内相遇并相互作用。每一个探测器的质子-质子发生对撞的中心被称作相互作用点。

实验都设置在以这些相互作用点为中心的同心球面上，吸收和记录频繁对撞的质子所发射出的众多粒子（见图13-5）。紧凑 μ 子线圈是圆柱状的，因为尽管两个质子束以相同速度沿着相反方向运动，对撞也倾向于包含大量在两个方向上的前向运动。事实上，因为单个质子远比粒子束小得多，绝大多数质子并不发生对撞，而是沿着粒子束管道继续直直地（只稍微有点偏转）运动下去。只有极个别的单个质子对头碰撞事件能让人感兴趣。

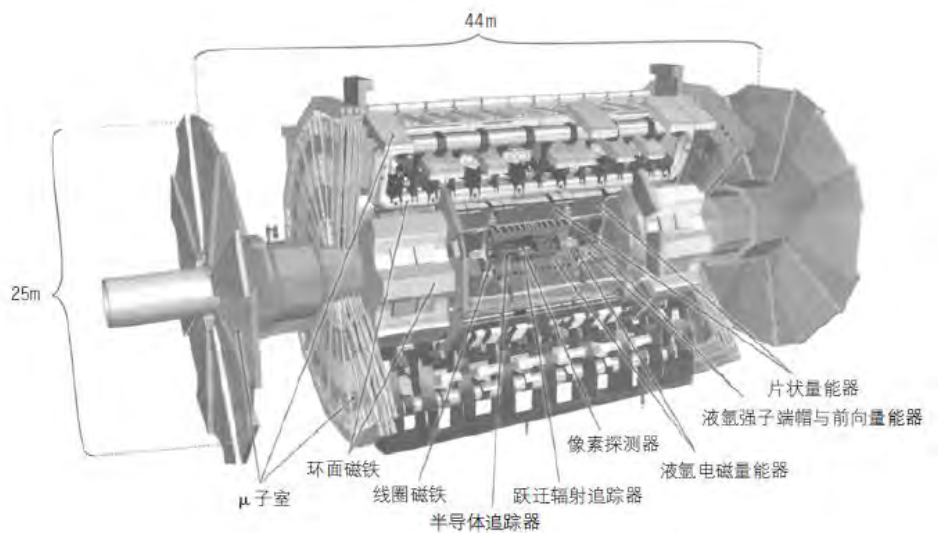


图13-5 电脑绘制的紧凑 μ 子线圈分离示意图，显示了各个探测单元。（感谢CERN与CMS友情提供图片）

这意味着，虽然大多数粒子继续沿着粒子束管道运行，但是潜在的有趣事件是一个与粒子束运动明显垂直的粒子束。考虑到沿着粒子束的方向粒子会广泛分散，柱形探测器被设计成可以尽可能多地探测这些相互作用产物。紧凑 μ 子线圈探测器被安置在法国塞西（Cessy）

地下的一个质子对撞点处，该地接近日内瓦边境；而超环面仪器相互作用区域则是在瑞士梅兰镇（Meyrin）的地下，靠近欧洲核子研究中心主体建筑的地方（见图13-6，该图模拟粒子从对撞点出来，沿着超环面仪器探测器的截面射出）。

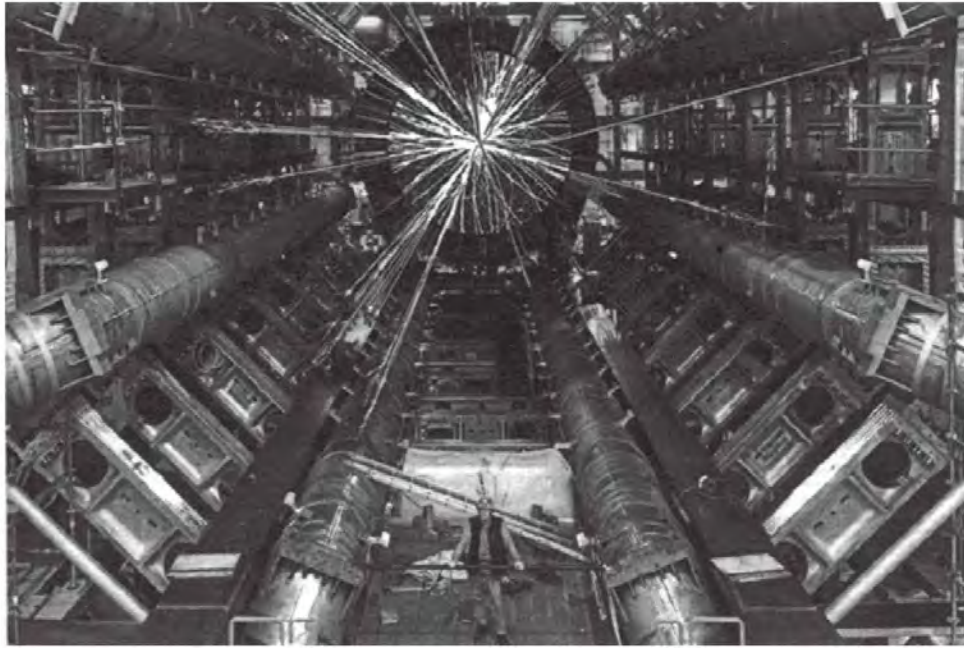


图13-6 模拟超环面仪器探测器的一个事件。该图模拟横向粒子束穿过探测器各层（注意，图中的人物是用来进行对比的，对撞发生时是没有人处在其中的）。其中，特别的环形磁铁清晰可见。（感谢CERN与ATLAS友情提供图片）

标准模型粒子都被其质量、自旋以及相互作用类型所刻画。不管最终产生出什么，两个实验都依赖已知标准模型的力和相互作用来探测它。那就是所有的可能性。不带电的粒子会不留痕迹地离开相互作用区。

当实验测量标准模型的相互作用时，它们可以确认通过的是什么，这就是探测器设计出来的目的。紧凑 μ 子线圈与超环面仪器都测量光子、电子、 μ 子、 τ 子以及强相互作用力粒子，它们被纳入沿着几乎平行方向运动的喷射流中。检测质子对撞区域所发出粒子的探测

器被设计成可以测量能量与电荷以此甄别粒子，它们还包含复杂、设计精良的计算机硬件、软件与电子器件来处理铺天盖地的数据。实验物理学家可以确认带电粒子，因为它们与其他已知的带电物质相作用。他们还发现了其他由强相互作用力联系物质。

探测器组件最终依赖电线与电子来记录所通过的物质——那些电子是由反应产生的粒子与探测器材料相互作用所产生的。有时因为产生了众多的电子和光子，出现了带电粒子雨（particle shower）^[52]；有时仅限于材料被记录的电荷所电离。但不管哪一种情形，电线记录信号并将之传送给物理学家以供他们在计算机上处理和分析。

磁铁对于两个探测器都至关重要，它们对测量电荷符号与带电粒子的动量都是不可或缺的。电磁相互作用中带电粒子在磁场中根据它们运动速度的大小发生相应的偏转。有着更大动量的粒子，路径更直，并且带不同符号电荷的粒子的偏转方向不同。大型强子对撞机的粒子能量（以及动量）如此之高，实验中需要采用非常强的磁铁，才有机会测量高能带电粒子那弯曲很小的轨道。

紧凑 μ 子线圈仪器是两大通用型探测器中尺寸较小的一个，但是它更重，重量高达12500吨。它的“紧凑”尺寸为：长21米、直径为15米，它比超环面仪器小，但其占地面积可以占一个网球场。

紧凑 μ 子线圈最显著的特征是它的强磁场达到4特斯拉，这是它名字中“线圈”部分的由来。探测器内部的线圈由一个直径为6米的柱状超导电缆线圈组成。在探测器外围一圈的磁返回轭也非常可观，占据了绝大部分的重量，它的含铁量比巴黎埃菲尔铁塔还高。

你也许对紧凑 μ 子线圈中“ μ 子”的来源感到好奇（我第一次听到这个名字时也一样）。快速确认高能电子与 μ 子（电子的较重部分，可以穿透到探测器外层）在新粒子探测中是非常重要的，因为这

些高能粒子有时产生于重物质的衰变。由于这些重物质不参与强相互作用，它们很可能是新的东西，因为质子并不是自动产生它们的。而这些轻易可以确认的粒子（电子和 μ 子）为此可以指示从对撞中产生的衰变粒子。紧凑 μ 子线圈中磁场的设计初衷就是重点关注高能 μ 子，以便能够引发它们。这表明它将记录任何与 μ 子相关的事件数据，哪怕它被迫为此要牺牲掉许多其他数据。

超环面仪器与紧凑 μ 子线圈相似，它的名字中也指示了磁铁的存在，因为大磁场也是它运作的一个关键。如前所述，超环面仪器的英文是“A Toroidal LHC ApparatuS”。“环面”（toroid）指的是磁铁，它的磁场没有紧凑 μ 子线圈那么强，但是遍布广泛。这个巨型环形磁铁使得超环面仪器成为两个通用型探测器中尺寸更大的一个，并且是迄今为止最大的实验仪器。它长46米、直径为25米，贴合地安置于一个55米长、40米高的洞穴中。超环面仪器大约重7000吨，是紧凑 μ 子线圈重量的一半多。

为了测量粒子的性质，越来越多的大型柱状探测器元件被安置在对撞发生的区域。紧凑 μ 子线圈与超环面仪器都有几个测量粒子轨道与电荷的嵌入单元。对撞产生的粒子首先遇到内层追踪器（inner tracker），它可以精确测量带电粒子接近对撞点的路径；接着遇到的是量能器（calorimeter），它可以测量由被截停粒子造成的能量沉积；最后是外围 μ 子探测器，它可以测量穿透力极高的 μ 子能量。每一个探测元件都有很多层级来增加每个实验的精度。我们现在从最里层的探测器出发到最外层，来一次实验之旅，以解释粒子束离开对撞点后如何变成可记录、可确认的信息。

三个追踪器，截获信号

仪器最里层的部分是追踪器，即用来记录带电粒子离开反应区域的位置，以便它们的路径可以得到重现并且动量可以被测量。在超环面仪器与紧凑 μ 子线圈中，追踪器由几个同心部件组成。最接近粒子束与相互作用点的层级，划分最精细，也能获得最多的数据。硅像素（silicon pixel）有着极其微小的探测器元件，位于最里层区域，从粒子束管道里面几个厘米的地方开始。它们专为精确地跟踪非常接近粒子密度最高的相互作用点附近的区域。硅被用在现代的电子器件中，因为它的精致结构使它可以被切割成更小的部分，而粒子探测器也看中了这一点。超环面仪器和紧凑 μ 子线圈的像素元件被设计成能以高分辨率探测带电粒子。通过将点与点的连接与产生粒子的相互作用点相连，在内层区域非常接近粒子束的地方，实验物理学家找到了粒子的运动轨迹。

紧凑 μ 子线圈探测器的前三层沿径向（向外11厘米）排列着100~150微米量级的像素，总计6600万个。超环面仪器的内层像素探测器也有相似的精度，它最里层的探测器可以被读取的最小单元是大小为50微米 \times 400微米的像素。超环面仪器的总像素为8200万，稍微比紧凑 μ 子线圈多一点。

有着上千万元件的像素探测器需要精细的电子读数。对于两大探测器来说，电子读数系统的广度与速度以及探测器内部承受的巨大辐射是两个主要的挑战（见图13-7）。

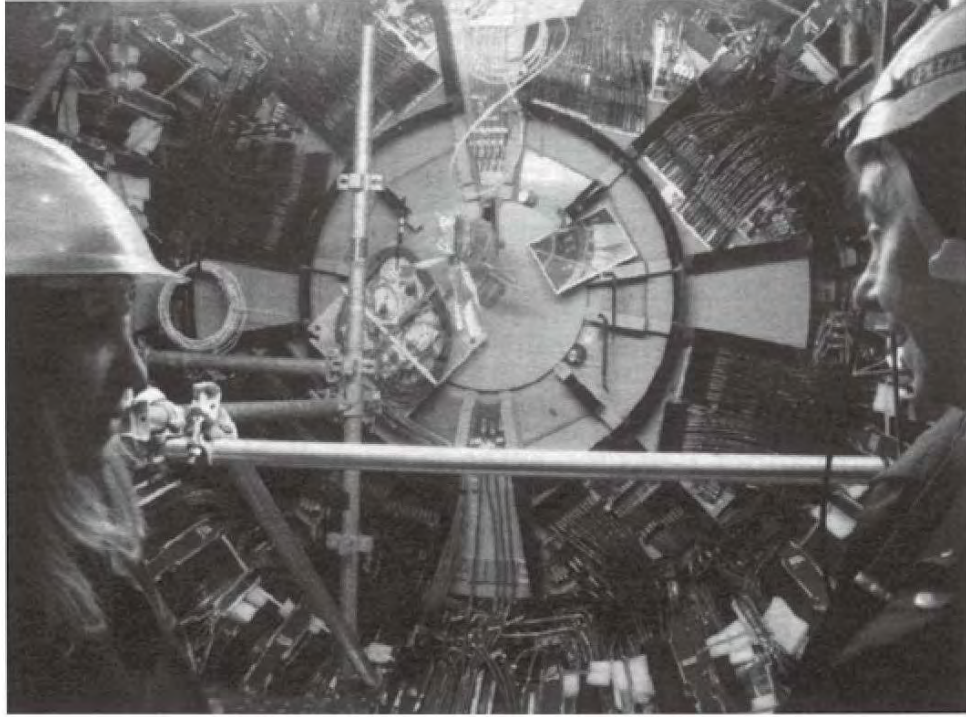


图13-7 辛西娅和工程师多梅尼科·达拓拉 (Domenico Dattola) 站在紧凑 μ 子线圈其中一个硅追踪器舱壁前面的脚手架上。

因为内部追踪器有三层，任何寿命足够长的带电粒子通过时，它们会记录三次闪光。这些轨迹一直延续到像素层以外的外层追踪器，产生一个强健的信号以便能够确认粒子。

我与合作者马修·巴克利 (Matthew Buckley) 花了很多精力在内层追踪器的几何学上。我们意识到，由于纯粹的巧合，一些猜想的新带电粒子通过弱相互作用力衰变成其中性粒子时，将会留下一条只有几厘米长的轨迹。那意味着在这些特殊情形中，轨迹可能仅仅在内层延伸，以至于可读取的信息也全部都在这里了。若实验物理学家只能依赖于探测器内部最里层的像素探测器，我们认为他们会面临更多的挑战。

然而绝大多数带电粒子可以存留到进入下一个追踪器单元，因此探测器会记录到一条更长的轨迹。所以，在内层高分辨率像素探测器

之外，在两个方向上有着不同大小的硅条，其中一个方向较为粗糙。较长的硅条与柱状实验相容并覆盖了一个更大的可用区域（半径越大、面积越大）。

紧凑 μ 子线圈的硅追踪器在它的中心区域总共有13层，前面与后面区域^[53]各有14层。在如前所述的前3个高像素层级后面有4层硅条，延展到半径55厘米处。在这里的探测器元件都是10厘米长、180微米宽的长条。其余6层在较为粗糙的方向上更不准确，由20厘米长、宽度从80微米~205微米渐变、延展到半径1.1米处的硅条组成。紧凑 μ 子线圈内层探测器总的硅条数目为960万。这些硅条对于重现那些带电最多的粒子的轨迹是至关重要的。总的来说，紧凑 μ 子线圈的硅覆盖总面积有整个网球场那么大（远比以前最大硅探测器的2平方米优越得多）。



超环面仪器的内层探测器延展到稍小一点的半径，即1米处，沿纵向有7米长。与紧凑 μ 子线圈一样，在内3层硅像素层之外的半导体追踪器（semiconductor tracker, SCT）由4层硅条组成。在超环面仪器中，其尺寸为长12.6厘米、宽80微米。半导体追踪器总的面积也极大，能覆盖61平方米。这些像素探测器对于重现接近相互作用点的精细结构非常有用，半导体追踪器是所有追踪器中最重要的一环，它的高精度与大面积覆盖区域（尽管只是在其中一个方向上），非常重要。

与紧凑 μ 子线圈不同的是，超环面仪器的外层探测器不是由硅做成的。跃迁辐射追踪器（transition radiation tracker, TRT）——内部探测器的最外层组件由充满了气体的麦管（straw tube）组成，它同时被用作追踪器与跃迁辐射探测器。当带电粒子电离麦管中的气体时，它们可以被追踪和测量。麦管长144厘米、直径4毫米，有金属丝延伸到其中心探测电离度。麦管在横向方向上精度最高。麦管探测轨迹的精度为200微米，比最内层的追踪器精度低，但覆盖面更广。根据它们产生的所谓跃迁辐射，麦管探测器还可以区分速度非常接近光速的不同粒子。因为轻粒子通常运动速度更快，这也区分了不同质量的粒子，从而有助于确认电子。

如果你觉得这些细节有点难以消化，那么你只须记住，这些信息对于绝大多数物理学家来说也有些过多了。它们只是提供给你一个关于尺度与精度的概念，但对研究特殊探测器组件的研究者来说很重要。哪怕是对于其中一个组件极其熟悉的人也不必对其他知识都一一了解。我也是在我出于研究要确认探测器图片是否正确时偶然学到的。假如你没有在第一时间掌握这些信息，你也不要觉得沮丧，很多实验物理学家也不必通晓每一个细节。

电磁量能器（ECAL），搜寻粒子流

一旦通过了上述三种追踪器，粒子在其沿径向往外的征程中，遇到的探测器的下一个环节是电磁量能器（electromagnetic calorimeter, ECAL）。它可以记录经停的带电粒子和中性粒子（主要是电子和光子）的能量沉积以及它们离开的位置。其探测机制是，搜寻入射电子或者光子与探测器材质发生碰撞时产生的粒子流。探测器的这一部分会产生对粒子精确能量与位置的追踪信息。

紧凑 μ 子线圈实验中电磁量能器的材料是一个奇迹。它是由钨酸铅晶体做成的，选材原因是由于这种晶体足够致密而又光学透明，正好可以用来截停和探测电子与光子（见图13-8）。最重要的是这种晶体的令人难以置信的清晰度，你恐怕从未见过这样的致密程度和透明程度。这种晶体之所以有用，是因为其测量电磁能量可以达到不可思议的精确度，这也是第16章要介绍的发现希格斯粒子的关键。

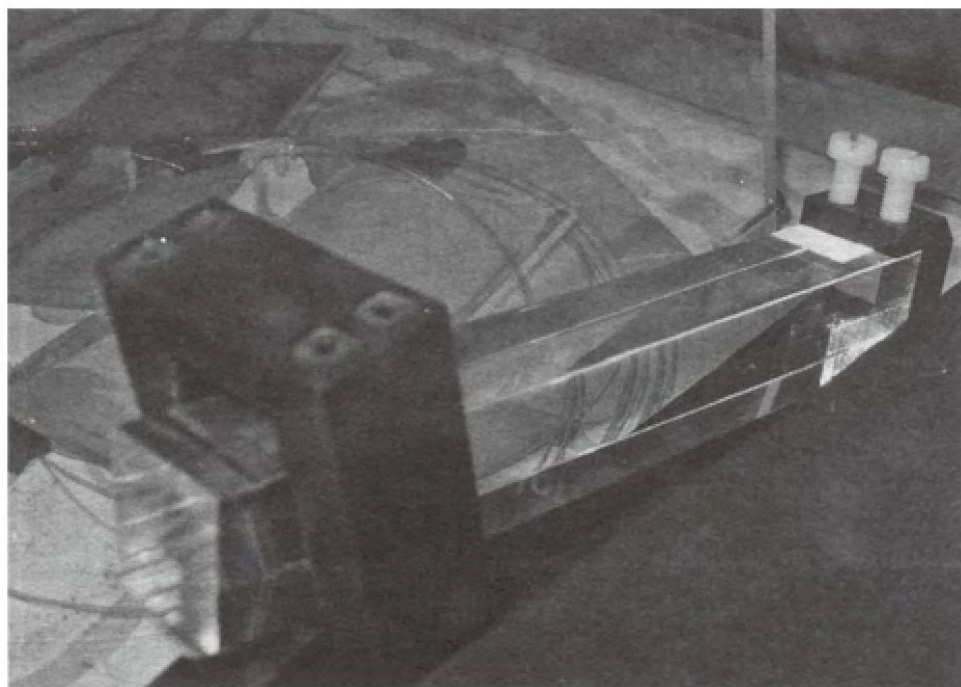


图13-8 用在紧凑 μ 子线圈电磁量能器中的钨酸铅晶体。

超环面仪器探测器用铅来截停电子和光子。在这种吸收材料中的相互作用，将能量从初始的带电轨迹转化成可以检测到能量的粒子雨。氩是一种惰性气体，与其他元素没有化学反应，并且抗辐射，液体氩被用于从粒子雨的能量中取样来导出入射粒子能量。

尽管我偏好理论，但超环面仪器的探测元件也令我着迷。法比奥拉参与了量能器的前沿几何设计与构建——沿着径向呈手风琴状层叠的铅板，层与层之间有一个薄层的液体氩与电极。她描述了这个几何形状如何能更快地读取电子，因为电子更靠近探测元件（见图13-9）。

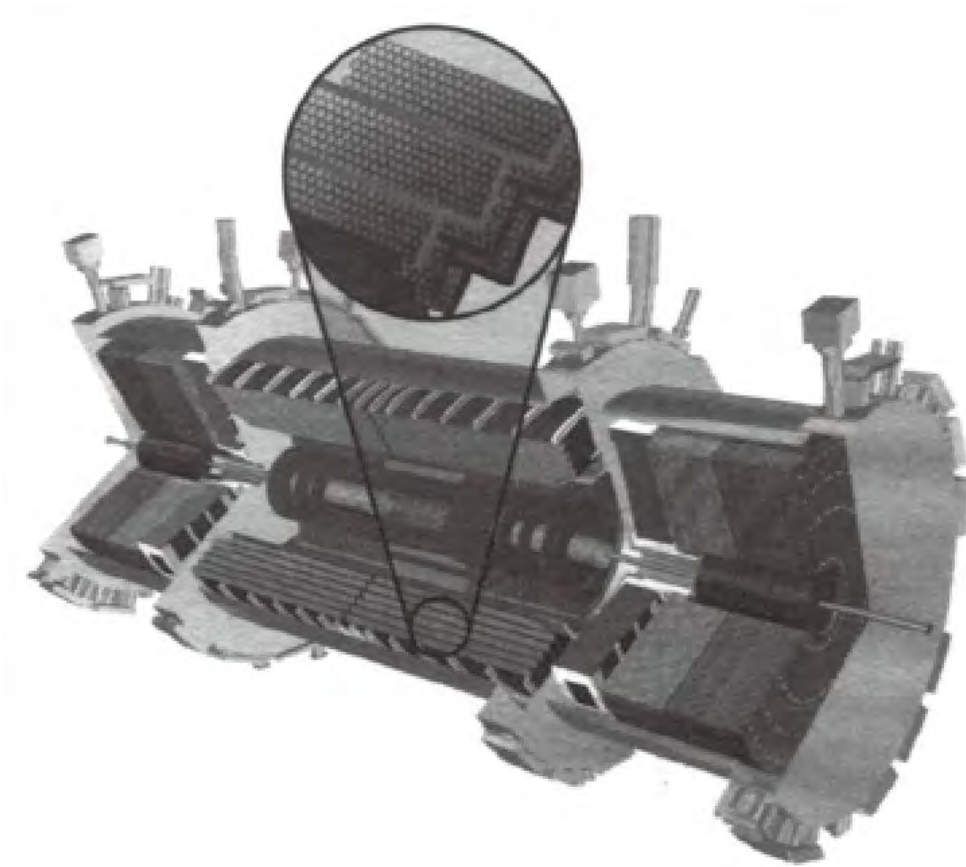


图13-9 超环面仪器电磁量能器的手风琴状结构。

定位粒子的强子量能器（HCAL）

在粒子束管中，我们沿着径向往外的旅程，下一站是强子量能器（hadronic calorimeter, HCAL）。强子量能器主要用来测量强子（参与强相互作用力的粒子）的能量与位置，然而它测量的精度比电磁量能器测量电子与光子能量的精度低得多。这是必然的，因为强子量能器是巨大的。举例来说，强子量能器直径为8米、长12米。如果将强子量能器按照跟电磁量能器一样的精度划分，那么将需要极其高昂的成本，所以追踪器的精度必须降低。最重要的是，能量的测量对于强相互作用力粒子来说更难，抛开划分不说，因为强子雨的能量涨落更大。

紧凑 μ 子线圈中的强子量能器含有致密材料层——黄铜或者钢，与塑料的闪烁片交叠安放，通过闪光的强度，记录穿过强子的能量与位置。虽然在超环面仪器中心区域的吸收材料是铁，但是其中的强子量能器的工作方式也基本相同。

μ 子探测器，发现有趣的碰撞

任何通用型探测器最外层都是 μ 子室（muon chamber）。 μ 子是类似电子的带电粒子，但比电子重200倍。它们在电磁或者强子量能器中不会被截停，相反，它们会高速穿过探测器厚厚的外层（见图13-10）。

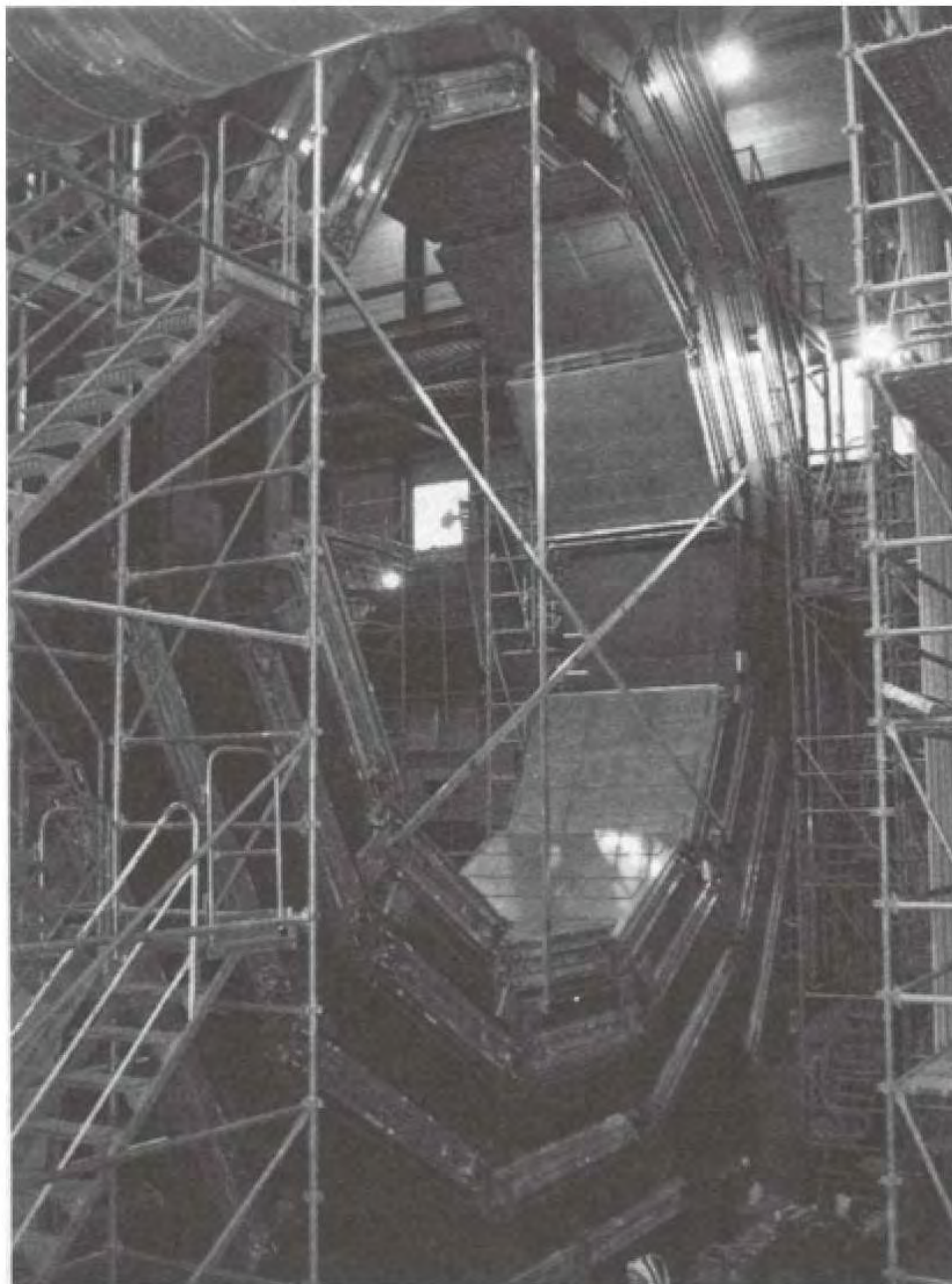


图13-10 紧凑 μ 子线圈中在建的磁回路线圈与 μ 子探测器交错排列。

高能 μ 子对于寻找新粒子非常有用，因为与强子不同，它们相当孤立，因此它们的探测与测量结果相对纯净。实验物理学家想记录所有高能 μ 子在垂直方向上的事件，因为 μ 子很可能伴随着更加有趣的

碰撞。 μ 子探测器对于验证其他能达到外层探测器的稳定的带电重粒子也是有用的。

μ 子室记录 μ 子留下的到达最外层探测器的信号。它们与内层探测器中的轨迹在某些方面是类似的——磁场使 μ 子的轨道弯曲，所以轨迹和动量能被测量到。然而在 μ 子室中，磁场是不同的，并且探测器的厚度大得多，因此允许测量一些曲率更小、动量更高的粒子（动量越高的粒子在磁场中偏转得越小）。在紧凑 μ 子线圈中 μ 子室沿径向从3米处向外延伸到7.5米处，而在超环面仪器中它们从4米处伸展到11米处。这些巨型结构能够测量50微米粒子的轨迹。

端帽，记录所有动量

最后介绍的探测元件是端帽（endcaps），即实验仪器前后两端的探测器（见图13-11）。我们不再从内向外“参观”（ μ 子室是那个方向的最后一步），我们现在沿着柱状探测器的轴向前进到两个端点处，并将端帽切开。探测器柱状部分被端帽探测器截断的地方是确保其能记录到尽可能多粒子的地方。因为端帽是探测器的最后一个安装组件，2009年^[54]我参观超环面仪器时才可以看到安置于探测器内部的众多层级。

探测器被安置在这些端点区域，以确保大型强子对撞机实验可测量到所有粒子的动量。目标是让实验装置密封性更好，即在各个方向都被覆盖而没有空洞和缝隙。密闭实验确保了即使不参与相互作用或者参与非常弱的相互作用的粒子也能被发现。如果能观测到横向动量缺失，那么必然有不可直接测量的一个或更多粒子产生。这样的粒子带有动量，它们所带走的动量使得实验物理学家知道了它们的存在。

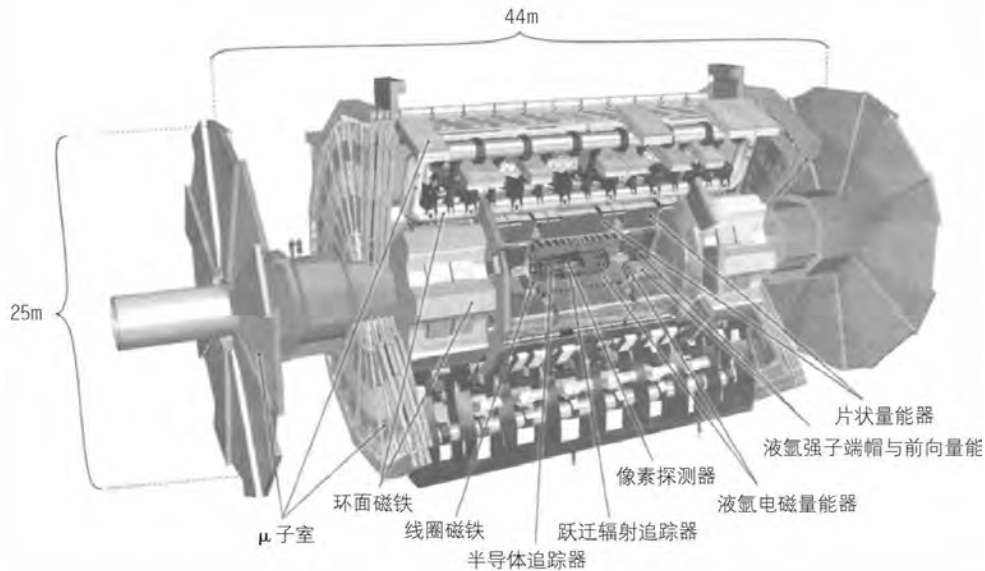


图13-11 超环面仪器的计算机三维图显示了它的诸多层级与分离的端帽。（感谢CERN与ATLAS友情提供图片）

如果你知道探测器测量了所有横向动量，而垂直于粒子束方向上的动量在对撞后不守恒，那么一定有物质没有被探测到，或消失了并且带走了动量。如我们所见，探测器测量垂直方向上的动量非常仔细。前向与后向的量能器保证了极少的垂直于粒子束的能量或动量可以不被察觉地逃脱，以此确保探测器的密闭性。

紧凑 μ 子线圈仪器在两端有钢吸收板和石英纤维，其致密性可以使其将粒子轨道分离得更好。在端帽的黄铜是一种可回收的材料，最初它被用于制造炮弹的外壳。超环面仪器在前向区域使用液体氩量能器来检测的粒子不仅有电子、光子，还有强子。

重要成员：磁铁

两个探测器还需要详细介绍的剩余部件是磁铁。磁铁并不是探测器的基本因素，因为它不能记录粒子的性状。但是磁铁在粒子探测中至关重要，因为它可以帮助确认动量与电荷，这是用于鉴定与表征粒

子轨迹的重要属性。粒子在磁场中偏转，所以它们的轨迹是弯曲的而非直线的。偏转的方向与程度依赖于粒子的能量与电荷。

紧凑 μ 子线圈巨螺线管磁铁长12.5米、直径为6米，是由冷冻的超导铌-钛线圈制成的。该磁铁是此探测器的基本特征，并且是这一类型史上最大的磁铁。螺线管导线线圈绕在一个金属芯上，当电流通过时产生磁场。磁铁中存储的能量与半吨TNT炸药的能量相当。不用说，必须考虑危险预防措施，例如磁铁万一猝息或突然失去超导性。2006年9月，螺线管4特斯拉的测试成功完成，但是当它正式运行时，磁场会略微低一些（3.8特斯拉），以确保它可以运行更长时间。

螺线管大到可以容下追踪器与量能器。位于探测器外围的 μ 子探测器在螺线管外面。 μ 子探测器的4层结构与一个巨大的环绕在引导磁场的线圈上的钢铁结构交织在一起，从而确保了均匀性和稳定性。磁返回轭长21米、直径14米，延伸到探测器半径7米处。事实上，它也是 μ 子探测器的一部分，因为 μ 子是唯一已知的带电粒子，可以穿透10000吨铁并穿进 μ 子室（在现实中，高能强子有时也会进入 μ 子室，这令实验物理学家很头疼）。在磁返回轭中的磁场让外层探测器中的 μ 子偏转。 μ 子在场中偏转的程度依赖于它们的动量，所以磁轭对测量 μ 子动量与能量至关重要。同时，巨大而又稳定的磁铁结构还起到另一个作用：它支撑整个实验装置，并且保护此装置不受来自自身磁场产生的巨大作用力的破坏。

超环面仪器中的磁场构造是完全不同的。它有两套磁铁系统：一个是2特斯拉的螺线管，它把跟踪系统包围起来；另一个是外层区域的巨大环状磁铁，它与 μ 子室交织在一起。当你看到超环面仪器的图片（或者实验本身）时，最引人注目的部件就是这8个巨大的环状结构（见图13-6），以及端帽处的两个附加的环状结构。它们产生的磁场

在粒子束轴向方向展开有26米长，从 μ 子谱仪沿着径向展开有11米长。

在访问超环面仪器实验期间，我所听过的许多有趣的故事中有一个是关于最初施工人员是如何将磁铁降到地面下的。他们从一个（侧面看上去为）椭圆形的结构出发。工程师在安装磁铁之前将重力因素分离出来，所以他们能正确地计算出磁铁自身的重量，在一段时间以后，这些磁铁构型会变成圆形。

另一个给我留下了深刻印象的故事，是关于超环面仪器的工程师如何考虑以下问题：由于开凿工程会导致洞穴内静水压力发生变化，从而导致洞顶每年会稍微上升1毫米。他们重新设计实验，使得这样一个微小的变动刚好可以在2010年将仪器摆到最优位置（最初的计划是2010年开始首次最大限度的运行）。由于大型强子对撞机的延期，结果并不是那样。但是到目前为止，实验的地基已经安置在一个停止移动的地点上，所以在整个运作中它都会保持在正确的位置上。尽管美国职业棒球大联盟前教练尤吉·贝拉（Yogi Berra）的训诫是“作出预测是困难的，尤其是关于未来的预测”，超环面仪器的工程师却做到了。

庞大的计算能力

一个缺少了描述大型强子对撞机的庞大计算能力的介绍是不完全的。除了前面考虑的追踪器、量能器、 μ 子系统和磁铁中的卓越硬件之外，世界范围内的协同计算对于处理众多对撞产生的海量数据也至关重要。不仅大型强子对撞机比此前能量最高的对撞机Tevatron的能量高出7倍，而且它的速度也快50倍。大型强子对撞机需要处理极其高分辨率的图片——记录发生速率为每秒10亿次的对撞事件。每个事件的“图片”包含1MB的数据。

这一数据对于任何计算系统来说都太巨大。所以触发系统要决定哪些数据需要保留，哪些需要丢弃。到目前为止，最频繁的对撞是在普通质子之间通过强相互作用发生的作用。没人关心这些绝大多数的对撞，因为它们代表了已知的物理过程但没有新物质。

质子对撞与两个装满豆子的袋子的碰撞，在某些方面有相似之处。因为袋子是软的，大部分时间它们皱缩而疲软，在碰撞中不会产生有趣的现象。但当袋子偶尔猛撞在一起时，单个豆子会以较大的力度互相碰撞——这种力可能大到让袋子破损。在这种情况下，发生对撞的单个豆子会快速飞出，因为单个豆子十分坚硬，其局部的能量很高。而其他没有发生对撞的豆子会沿着原路继续向前飞行。

类似地，当粒子束中的质子彼此撞击时，其单个子单元对撞产生了有趣的事件，而其余的质子成分会继续沿着粒子束管道照着相同方向飞行。

与豆子的碰撞不同，豆子仅仅碰撞然后改变运动方向。当质子猛撞对方时，内部成分（夸克、反夸克、胶子）互相撞击时，最初的粒子可以转变为能量或者其他类型的物质。在低能时，对撞只涉及三个带有质子电荷的夸克；而在高能时，量子力学的虚拟效果产生了可观数目的胶子与反夸克，正如我们在第6章所看到的。有趣的对撞来自这些质子中亚组分之间的碰撞。

当质子能量很高，它内部的夸克、反夸克、胶子的能量也相应会很高。然而它们的能量不是质子的全部能量，只是其中一部分。更多的情况是，夸克与胶子以质子的少部分能量发生碰撞，以至于不能产生比较重的粒子。正是由于这些较低的相互作用强度或者较重新粒子的可能性，使得迄今未见的粒子或者相互作用产生的概率比预料之中的标准模型碰撞的概率低了很多。

正如装满了豆子的袋子的碰撞一样，大部分粒子束对撞也是无趣的。要么是质子彼此擦肩而过，要么对撞产生了我们所知道的标准模型的事件。另一方面预测告诉我们，大型强子对撞机产生新粒子（如希格斯玻色子）的概率大约为十亿分之一。

结果是，不仅在很短而且很需要运气的一段时间之内才能产生好成果。这也是为什么我们一开始就需要那么多的对撞，但绝大多数没有发现新物质。极少事件是非常特殊且带有信息量的。

触发系统中的软件与硬件的设计使其可以确认有潜力的有趣事件并将它们搜寻出来。要理解这项极其艰巨的任务（假设你考虑了各种不同的反应道），假设你有一台15000万像素（每束交叉信息量一样）的相机，每秒可以拍摄4000万张照片。每束交叉有20~25个事件发生，这相当于每秒10亿物理事件。触发系统就相当于负责从其中选出一些有趣照片的装置。你还可以把触发系统想象成垃圾邮件过滤器。这项工作是要确定只有有趣的数据才能进入实验物理学家的计算机。

触发系统需要确定潜在的有趣对撞，并将没有新东西的数据抛弃。这些事件本身（离开相互作用点并被探测器记录下来的）必须足以能从标准模型的普通过程中区分出来。识别出特殊事件的方法即筛选出需要保留下来事件的方法，这使得能辨认出新事件的概率更低。触发系统有一个强大的任务，它们负责从每秒十亿事件中筛选出几百个可能是有趣的事件。

硬件与软件相结合的“大门”完成了这项使命。每一个连续触发级别会拒绝它收到的没有新东西的大部分事件，剩下一些容易处理得多的数据。这些数据接下来会被全球160个研究机构的计算机系统分析。

第一级触发系统修建于探测器中，是基于硬件的系统，在鉴定不同的性质时起到一个总通行证的作用。例如选择包含高能 μ 子或者量能器中大的横向能量沉积事件。在等待第一级触发系统出结果的几微秒中，来自各个束交叉的数据被储存在缓冲区内。更高级的触发系统是基于软件的。筛选算法运行在探测器旁边一个大型计算机群上。第

一级触发系统将每秒10亿事件减少到每秒10万事件，然后软件触发系统再进一步将此比率减少到千分之一，即几百事件。

每个通过触发系统的事件都携带巨大的信息量（此前我们讨论的从探测元件读取的），并超过1MB。每秒几百事件，实验物理学家需要每秒往计算机写入100MB的数据，等于1000万亿字节，也即 10^{15} 字节，相当于每年成千上万台DVD所存的信息。

蒂姆·伯纳斯-李首先开发出万维网来处理欧洲核子研究中心的数据，并让全世界的实验物理学家可以实时共享计算机的数据。大型强子对撞机计算机网络是欧洲核子研究中心的另一个主要计算优势。网络是在2008年年底（在广泛的软件开发之后）开始进行计算的，以帮助处理实验物理学家要处理的海量数据。欧洲核子研究中心的网络使用私用光缆以及高速部分的公用网络。它名字的由来是因为数据不是仅仅与单一的位置相关联，而是分布在全世界的计算机中——就像都市的电力不是只与哪家特别的电厂相关一样。

一旦触发系统审核通过的事件被存储起来，它们就通过网络遍布全球。在网络的帮助下，世界范围内的计算机网络都准备好接收冗余存储的数据。当网络共享信息时，网络却在众多参与项目的计算机中共享计算能力和数据存储。

在网络的帮助下，分层的计算中心会分别处理数据。第0层是欧洲核子研究中心的中心设备，数据被记录下来并从其原始形式预处理到一个更适于物理分析的形式。高带宽连接将数据送达12个大型国家级计算中心，组成第1层计算中心。分析组可以访问这些数据，假如他们选择这样做的话。光缆会连接第1层与50个左右的设在大学里的第2层分析中心。第2层有足够的计算能力来模拟物理过程，并进行具体的分析。最后，全球任何一所大学的群组都可以做第3层分析，绝大多数具体物理数据最终在此被提取出来。

此时，各地的实验物理学家可以通盘考虑他们的数据以找出高能质子对撞可以揭示的东西。这些可能是令人欣喜的新结果。但是为了确定能否出现这些情况，实验物理学家的首要任务是推断出已经存在的东西究竟是什么。我们将在接下来的一章继续探讨这些内容。



粒子物理学的标准模型简洁地归类了我们目前对基本粒子及其相互作用的理 解（见图14-1的总结）。

图14-1包括分别罗列了左手与右手粒子条目，这些粒子根据其手征性区分。对于零质量粒子，手征性表示沿着运动方向的自旋。质量将两种手征性混合起来，例如左手和右手的电子。图中严格的左、右手区分性质没有它们之间相互作用的差别更重要。如果粒子都是零质量的，那么将上夸克变成下夸克或者将带电轻子转变为中性轻子的弱相互作用只作用在左手粒子上。另一方面，强相互作用力与电磁相互作用力在左手与右手粒子上均发生作用，但是在强相互作用力下只有夸克带荷。

不仅包括上、下夸克与处在物质核心的电子，也包含许多其他有着同种相互作用的更重的粒子（通常不易在自然中发现的粒子），那些我们只能 在高能对撞实验中仔细研究的粒子。绝大多数标准模型的要素，例如大型强子对撞机目前正在研究的粒子，都一直被完全埋没，直到20世纪后半叶由于新理论与新实验方法的出现，它们才得以被发现。

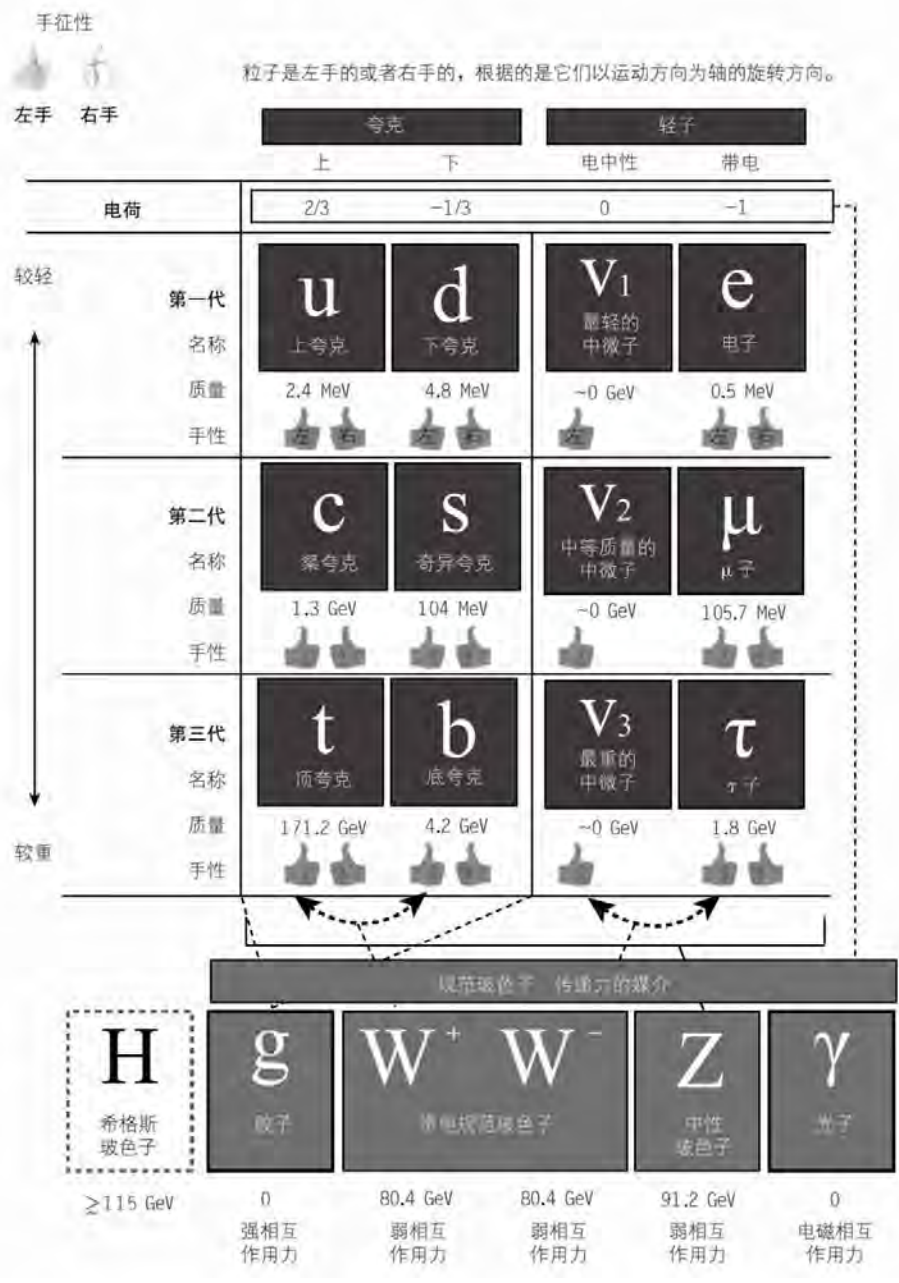


图14-1 粒子物理学标准模型的元素，质量已标注，同时标注的还有左手或右手粒子。弱相互作用力可以改变粒子类型，并只作用在左手粒子上。

在大型强子对撞机中，超环面仪器与紧凑 μ 子线圈实验都被设计成可以探测和鉴别标准模型粒子。当然，真正的目的超越了我们目前了解的知识，乃是为了发现新的粒子和相互作用来解决悬而未决的难

题。为了能这样做，物理学家需要区分标准模型背景事件，确认可能从新奇粒子衰变出来的标准模型粒子。大型强子对撞机的实验物理学家就像侦探一样分析数据、拼凑线索，确定那里有什么。他们只有排除了所有熟悉的东西才能推导出新东西的存在性。

在介绍完通用型实验之后，我们在这一章将重新“造访”它们，以便更好地理解大型强子对撞机的物理学家如何鉴别个别粒子。更熟悉粒子物理学的现状以及标准模型的粒子如何被发现，将有助于我们在后文讨论大型强子对撞机的探索能力。

发现轻子

粒子物理学将标准模型的基本粒子分成两类。一类是轻子，包括类似于电子、不参与强相互作用力的粒子。标准模型也包括两种较重版本的电子，它们与电子的电荷相同，但是质量大了很多，分别称为 μ 子与 τ 子。事实证明每一种标准模型的基本粒子都有三个版本，都有相同的电荷，但是相邻的每一代（generation）^[55]都比前一代重一些。但我们不知道为什么正好是三个版本的粒子，为什么这些粒子都具有相同电荷。诺贝尔奖获得者伊西多·拉比（Isidor Isaac Robi）听说存在 μ 子时，夸张地表达了他的困惑与感叹：“谁下的订单？”

较轻的轻子是最容易被发现的。虽然电子与光子都在电磁量能器中沉积能量，因为电子带电而光子不带电，所以电子可以很容易从光子中区分出来。电子在到达量能器进行能量沉积之前，会在内层探测器中留下一条轨迹。

μ 子的确认也相对直接。就像所有其他较重的标准模型粒子， μ 子衰变太快以至于无法在普通的物质中找到它们，因此我们很少在地球上发现它们。然而 μ 子的寿命又足够长，使得它们在衰变前可以到达探测器的外层。因此它们留下了长长的从内层探测器到外层 μ 子室

的清晰可见的轨迹。 μ 子是唯一的标准模型粒子，它可以到达这些外层探测器并且留下可见信号，所以它们容易被挑选出来。

τ 子虽然可见，却不容易被找到。 τ 子与电子和 μ 子一样是带电粒子，它重得多。与许多重粒子一样，不稳定是其共性，也就是说衰变只留下其他粒子的痕迹。 τ 子快速衰变成一个较轻的带电轻子与两个叫作中微子的粒子，或者衰变成一个中微子与一个叫作 π 子的经受强相互作用的粒子。实验物理学家会研究这些衰变产物——从初始的粒子衰变得到的粒子来推断它们是否来自一个重的衰变粒子，以及在这种情况下，它的性质是什么。即便 τ 子没有直接留下一个轨迹，所有实验记录的衰变产物的信息也可以帮助确认它与它的性质。

电子、 μ 子与更重的 τ 轻子带有电荷-1，与质子所带电荷相反。对撞也产生了伴随这些带电轻子的反粒子——正电子、反 μ 子与反 τ 子。这些反粒子带电荷+1，在探测器中留下相似的轨迹。然而由于它们的电荷相反，所以它们在磁场中会往相反的方向偏转。

除了上述三种带电轻子，标准模型中也包含中微子，它们是不带电的轻子。三种带电轻子经受电磁力与弱相互作用力，中微子是电中性的，因此不受电磁影响。20世纪90年代以前，实验结果都暗示中微子没有质量。然而在那之后的10年，一个非常有趣的发现表明，中微子有非0的但是非常小的质量，它为标准模型的结构提供了重要信息。

尽管中微子很轻，因此对撞机有足够的能量可以产生它们，大型强子对撞机却不可能直接探测它们。因为它们不带电，所以耦合很弱。它如此之弱，以至于即便每秒会有50万亿中微子从太阳而来并穿过我们，要是没人告诉我们，我们也根本无从得知。

尽管中微子不可见，物理学家沃尔夫冈·泡利（Wolfgang Pauli）却猜想它们的存在，作为解释中子衰变时能量走失问题的“救命稻草”。没有中微子带走一些能量，该过程的能量守恒就被打破，因为衰变后探测到的质子和电子的能量加起来与中子衰变前的能量不相等。即使卓越的物理学家尼尔斯·玻尔，那时也打算牺牲该守恒原理，承认能量有缺失。然而泡利更忠诚于已知的物理理论，他猜想能量的确是守恒的，只是实验物理学家不能看到电中性粒子把剩余的能量带走了。事实证明他是正确的。



泡利将这种假想的粒子命名为“中子”，该名称此后还被用于其他方面，也即，命名处在核子中的质子的电中性伙伴。故此，意大利物理学家恩里科·费米（Enrico Fermi，发展了弱相互作用理论，最有名的是设计了世界上首个核反应堆）给此粒子起了一个可爱的名字——“中微子”，在意大利语中是“小中子”的意思。当然中微子不是个头小的中子，但是它像中子一样不带电，并且中微子比中子轻了太多。

与其他标准模型的粒子类似，存在三种中微子。每一种带电轻子（电子、 μ 子和 τ 子）都有一种通过弱相互作用力来相互作用的中微子。^[56]我们已经知道如何发现电子、 μ 子与 τ 子。剩下的实验问题是实验物理学家如何发现中微子。因为中微子不带电荷而且耦合很弱，所以当它们离开探测器时，它们根本不会留下任何轨迹。大型强子对撞机的人要怎样知道它们在哪里？

动量（当粒子缓慢运动时等于速度乘以质量；粒子以接近光速运动时，更像能量在特定方向的移动）在各个方向守恒。与能量相同，我们从未发现任何证据来表明动量可以缺失。所以，如果探测器中测到的粒子动量小于进入探测器的动量，那么一个其他粒子（或者多个粒子）必定已经逃离，而且在此过程中带走了缺失的动量。这种逻辑导致泡利在第一时间（当时是在核衰变的 β 衰变中）推断出中微子的存在性，这也是现今我们如何知道看似不可见的弱相互作用粒子的存在性的逻辑。

在强子对撞中，实验物理学家会测量所有与粒子束相垂直的动量，并且计算是否有缺失的部分。他们关注于垂直方向，因为大量的动量被粒子沿着粒子束管道方向带走，所以在该方向上太难追踪。而垂直于初始质子的动量方向则容易测量和考虑得多。既然对撞前垂直于粒子束方向的总动量本质上为零，那么对撞后也应如此。因此如果测量与预期不同，实验物理学家就可以“测到”有物质缺失。剩下的问题是如何区分这些是哪种不参与相互作用的可能粒子。对于标准模型过程，我们知道中微子是不可探测的元素的其中一种。基于我们接下来要简单介绍的中微子已知的弱相互作用力，物理学家计算和预测了中微子的产生速率。而且，物理学家已经知道W玻色子的衰变应该如何，例如衰变得到的孤立电子或者 μ 子的横向动量带有相当于W玻色子一半质量的能量，这是相当独特的。所以使用动量守恒与理论输入，中微子可以被“找到”。显然，这些粒子的定义标签比我们可以直接看到那些粒子的少。只有理论的考虑与缺失能量的测量相结合，才可以告诉我们会有什么。

我们在考虑新发现时保持这样的想法很重要。相似的考虑也适用于其他不带电或者所带电量低到无法被直接检测的新粒子上。只有一个将缺失能量与理论输入综合起来的考虑可以用来推断会存在什么。这就是为什么密闭性（检测尽可能多的动量）如此重要。

发现强子

我们已经考虑了轻子（电子、 μ 子与 τ 子以及它们伴随的中微子）。标准模型中其余类别的粒子被命名为强子，即相互作用为强相互作用力的粒子。这个类别包含了所有由夸克与胶子组成的粒子，例如质子、中子以及其他被称为 π 子的粒子。强子也有内部结构——它们是夸克与胶子通过强相互作用力结合起来的束缚态。

然而标准模型没有列出许多可能的束缚态。它列出的是更基本的由夸克与胶子束缚成的强子态。除了处于质子与中子内的上夸克与下夸克，重的夸克称为粲夸克、奇异夸克、顶夸克以及底夸克。与带电和中性的轻子一样，重夸克与它们较轻的伙伴——上夸克和下夸克带有相同的电荷。重夸克在自然界中不易被发现。对撞机也需要研究它们。

强子（通过强相互作用力耦合的粒子）与轻子（不参与强相互作用力的粒子）在粒子对撞机中看起来非常不同。这主要因为是夸克和胶子有强相互作用力，它们从来不会单独出现。它们总是在可能包含初始粒子的喷射流中，但同时还包含了参与强相互作用力的其他粒子。喷射流不是只包含单个粒子，而是包含对初始粒子形成一层“防护”的强耦合粒子流，可以从图14-2中看到。

强相互作用力将在第一时间从引发喷射流的夸克与胶子之中产生许多新夸克与胶子，即使初始事件中并不包含它们。质子对撞产生大量喷射流，因为质子本身是由强相互作用力的粒子组成的。这些粒子会产生许多附加的强相互作用力的粒子流，这些粒子流会伴随它们一起运动。它们有时产生的夸克与胶子会沿着不同的方向飞离，并构成它们自己独立的喷射流。

我在《弯曲的旅行》一书中所引用的电影《西区故事》（*West Side Story*）中的《喷射机帮派之歌》（*Jet Song*）^[57]把强子喷射流描述得很好：

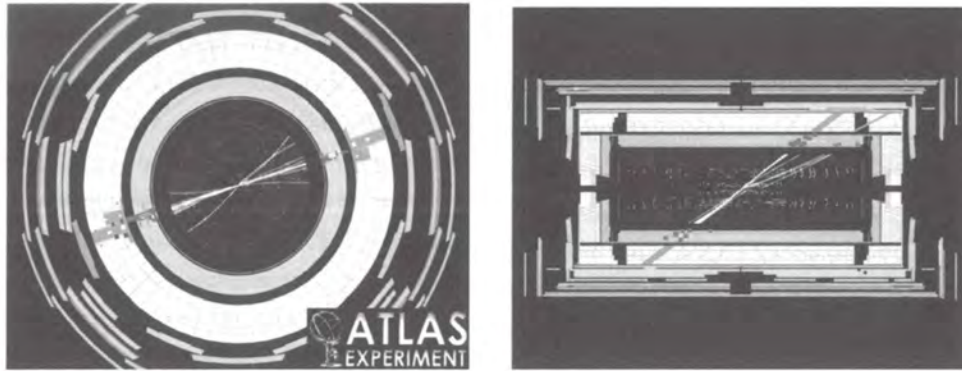


图14-2 喷射流是环绕着夸克和胶子的强相互作用力粒子流。这张图片显示了它们在追踪器和强子量能器中的检测。（感谢CERN友情提供图片）

你从不孤单，
你从不会被疏远！
即便你独自守候，
只要呼唤陪伴，
必有人到你身边誓死保护你！
You' re never alone,
You' re never disconnected!
You' re home with your own:
When company' s expected,
You' re well protected!

单独的夸克（以及许多“帮派”成员）不会被发现，不过在一群相关的强相互作用力粒子中会被发现。

喷射流通常留下可见轨迹，因为喷射流中的一些粒子是带电的。当一个喷射流到达量能器，它将能量沉积下来。细致的实验研究、分析与计算机计算帮助实验物理学家在第一时间推导出产生喷射流的强

子的性质。虽然如此，强相互作用力与喷射流使得夸克与胶子的测量非常巧妙。你不需要探测夸克与胶子本身，而是它们所处的喷射流。这使得绝大多数夸克和胶子流无法与其他流区分开。它们都沉积了很多能量并且留下许多轨迹（图14-3所示为探测器如何确认标准模型的关键粒子）。

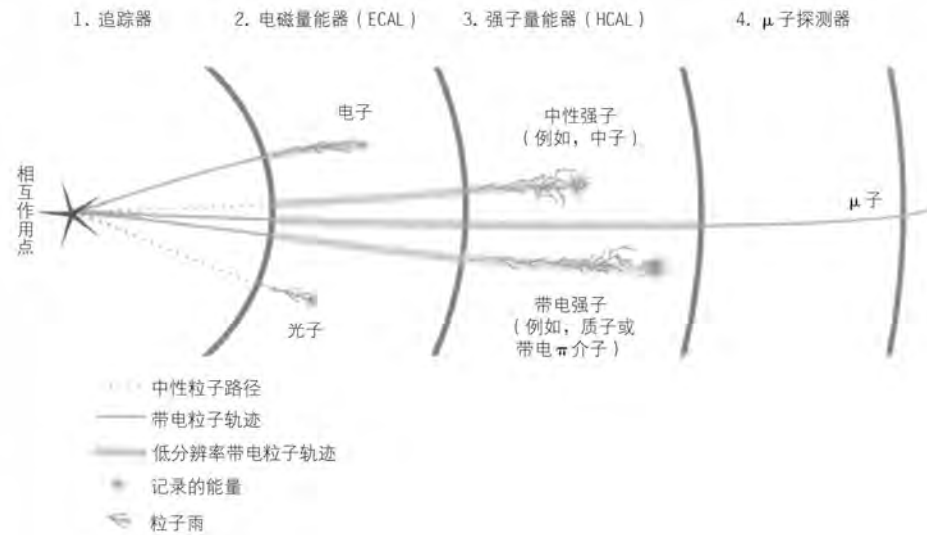


图14-3 一个关于如何将标准模型粒子从探测器中区分出来的概括。中性粒子在追踪器中没有记录。带电和中性强子可以在电磁量能器中沉积少量能量，但是会在强子量能器中沉积几乎全部的能量。 μ 子会穿透到外层探测器。

测量完喷射流的性质，要分辨出不同的夸克或胶子引发的喷射流，虽然这样并非完全不可能，但也极具挑战性。与下夸克带有相同电荷的底夸克（以及较重的奇异夸克）是这一类中最重的夸克，它是该原则下的一个特例。底夸克之所以特殊的原因是，它比其他夸克衰变得慢。其他不稳定夸克在产生后立即衰变，所以它们的衰变产物似乎在质子对撞的相互作用点就开始它们的运行轨迹了。然而底夸克存在的时间足够长——大约1.5皮秒（picosecond, 10^{-12} 秒），也即有足够的时间以光速来运行半毫米，它们会在相互作用点留下一条可以

观测到的足够长的轨迹。内层硅探测器探测这个偏离顶点，如图14-4所示。

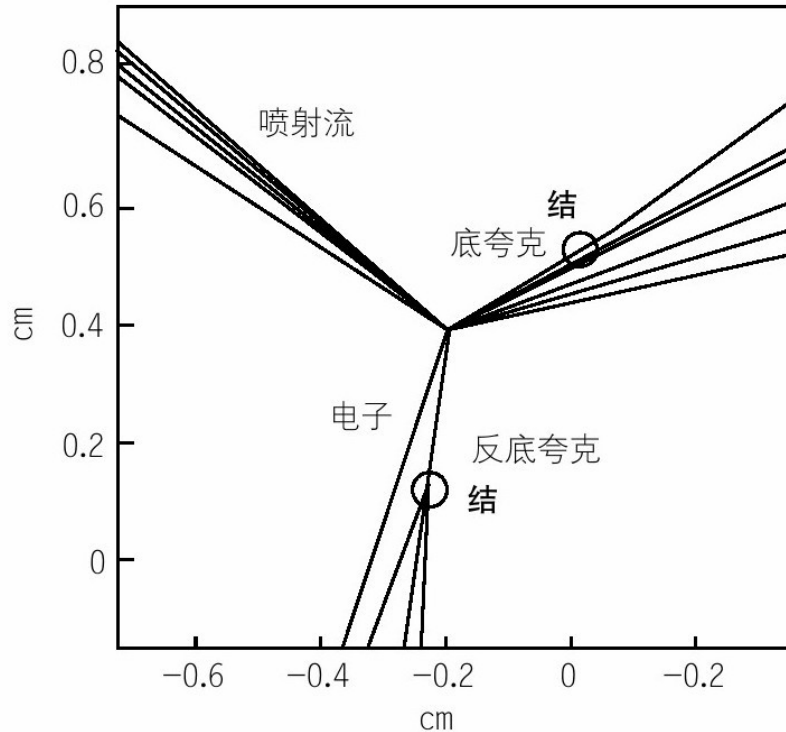


图14-4 底夸克构成的强子存在的时间足够久，在衰变成其他带电粒子前，在探测器中会留下一条可见的轨迹。在硅顶点探测器中留下一个结（kink），可以被用于鉴定底夸克。这张图来源于顶夸克衰变图。

当实验物理学家从底夸克重建轨迹时，它并不会延伸回事件中心的初始相互作用点。该轨迹反而看起来是从内部追踪器的底夸克衰变点出发的，在轨迹中留下一个结——一个连接进入的底夸克和衰变产物的结。A在硅探测器的精细划分之下，实验物理学家可以看到靠近粒子束的轨迹的细节，从而在相当一部分时间中确认底夸克。

A如果初始的b介子（meson）是电中性的，你只能看到从衰变点出发的轨迹，而看不到从中性态出发的标识轨迹。

另一种实验上有优势的夸克是顶夸克，它很特殊，这是因为它的质量非常大。与上夸克所带电荷相同的三种夸克（还有一种是粲夸克）中，顶夸克是最重的。它的质量是带不同电荷的底夸克的40倍，是带相同电荷的上夸克的30000倍^[58]还多。

顶夸克重到它们的衰变产物会留下不同的轨迹。当一种较轻的夸克衰变时，因为衰变产物与初始粒子相似，以接近光速运行，所以它们一起运动，看起来是一个单独的喷射流——哪怕喷射流来源于两个或更多不同的衰变产物。另一方面，除非顶夸克是极端高能的，它们会衰变成底夸克和W玻色子（带电的弱规范玻色子），而且可以通过发现这两者来确认顶夸克。因为顶夸克的重质量意味着，它与希格斯粒子和其他我们希望能尽快理解的弱尺度物理中的粒子相互作用的紧密联系。顶夸克和它们相互作用的性质可能为标准模型下的物理理论提供有用的信息。

发现弱相互作用力载体

在我们结束讨论如何确认标准模型粒子之前，最后要考虑的粒子是弱规范玻色子——两种W玻色子和Z玻色子，它们是传递弱相互作用力的媒介。弱规范玻色子性质特殊，与光子和胶子不同，它们质量非零。传递弱相互作用力的规范玻色子的质量，造成了标准模型中主要基本疑难问题的产生。W与Z玻色子的质量（以及本章讨论的其他基本粒子的质量）来源于希格斯机制，我们马上介绍。

W与Z玻色子很重，所以这些规范玻色子会衰变。这意味着，W与Z玻色子以及顶夸克与其他不稳定的重粒子，可以通过寻找它们衰变产生的粒子得以确认。由于重的新粒子很可能也是不稳定的，因此我们将使用弱规范玻色子来举例说明衰变粒子的另一种性质。

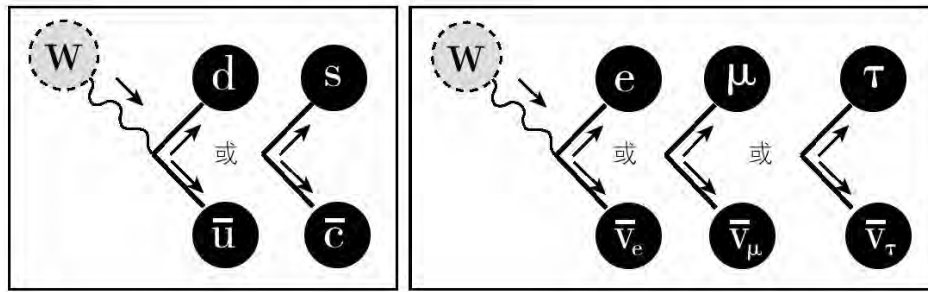


图14-5 W玻色子可以衰变成带电轻子以及它伴随的中微子，或者上夸克-下夸克，或者粲夸克-奇异夸克。实际上，物理粒子是不同类型夸克或者中微子的叠加态，这允许W玻色子有时可以同时衰变成不同代的粒子。

W玻色子会与所有感受到弱作用的粒子（也即我们已经讨论的所有粒子）相互作用。这让W玻色子有多种衰变的选择。它可以衰变成任何带电轻子（电子、 μ 子与 τ 子）以及它们伴随的中微子。它也可以衰变成上夸克-下夸克或者粲夸克-奇异夸克对，如图14-5所示。

粒子的质量在决定其衰变方式时也很关键。一个粒子只能衰变成几个质量之和小于初始粒子的其他粒子。虽然W玻色子也与顶夸克和底夸克相互作用，但是顶夸克比W玻色子重，所以衰变是禁止的^[59]。

让我们假设W玻色子衰变成两个夸克，因为在这种情况下实验物理学家会测量两个衰变产物（这不适用于衰变成轻子和中微子的情况，因为中微子是测不到的）。因为能量、动量守恒，所以测量夸克的末态（final state）总能量和动量告诉我们，衰变成它们的初始粒子，即W玻色子的能量和动量。

此时，爱因斯坦的狭义相对论与量子力学使得故事变得更有趣。爱因斯坦的狭义相对论告诉我们质量如何与能量和动量联系。绝大多数人知道 $E=mc^2$ 。如果将m换成 m_0 （即当粒子静止时它的内禀质量），该公式对于静止的粒子也适用。一旦粒子运动，完整的公式变为 $E^2=p^2$

$E^2 = m^2 c^4$ 。 [60] 在这个公式下的能量和动量使得实验物理学家可以推导出粒子质量，即便初始粒子早已通过衰变消失了。实验物理学家只要将所有动量、能量加和并应用该方程，初始的质量就可以确定。

量子力学发挥作用的原因更微妙。一个粒子表面看起来并不总是有真实的质量。因为粒子可以衰变，量子力学不确定关系说需要用无限长时间来精确测量能量，它也告诉我们，任何粒子如果不能永远存在，那么它的能量也不能被精确知道。与能量精确值的偏离越大，衰变越快、寿命越低。这意味着在任何既定测量中，质量可以逼近（但不能等于）真实的平均值。只有经过很多次测量实验物理学家才能导出粒子的质量（最可能值以及平均收敛值）与寿命，因为这是粒子在衰变前存在的时间，它决定了粒子的质量展宽（见图14-6）。对于W玻色子来说是这样，对于其他衰变粒子也是如此。

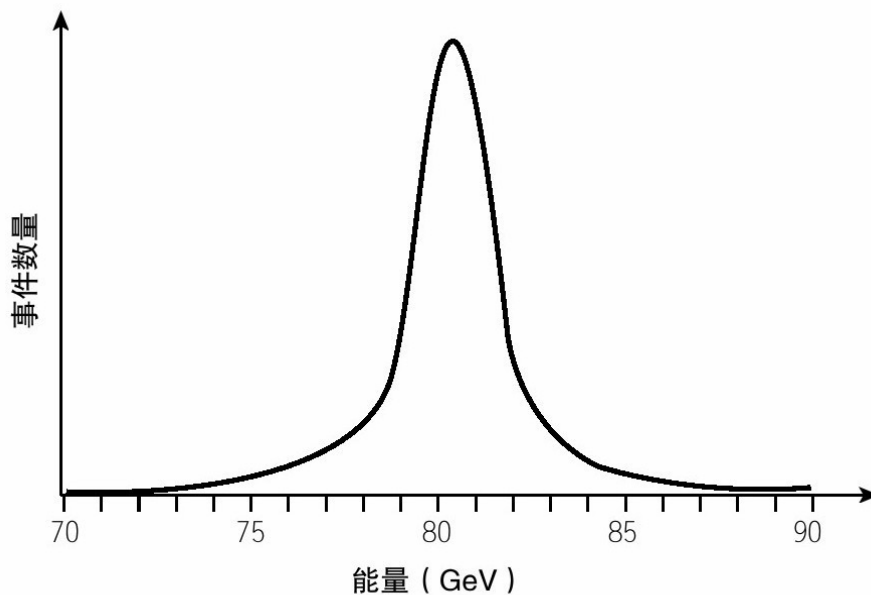


图14-6 衰变粒子的测量值以真实质量为中心，但是允许一个由其寿命决定的质量展宽值。该图刻画的是W规范玻色子。

当实验物理学家使用本章所描述的方法将他们所测量数据的整合起来时，他们可能发现标准模型粒子（见图14-7标准模型粒子及其性质总结）^[61]。但是他们也可能最终确认一些完全新的东西。大家的希望是，大型强子对撞机可以产生新的奇异粒子，并产生对物质本质、乃至时空本身的更深层次洞见。本书下一部分将探索一些更加有趣的可能性。

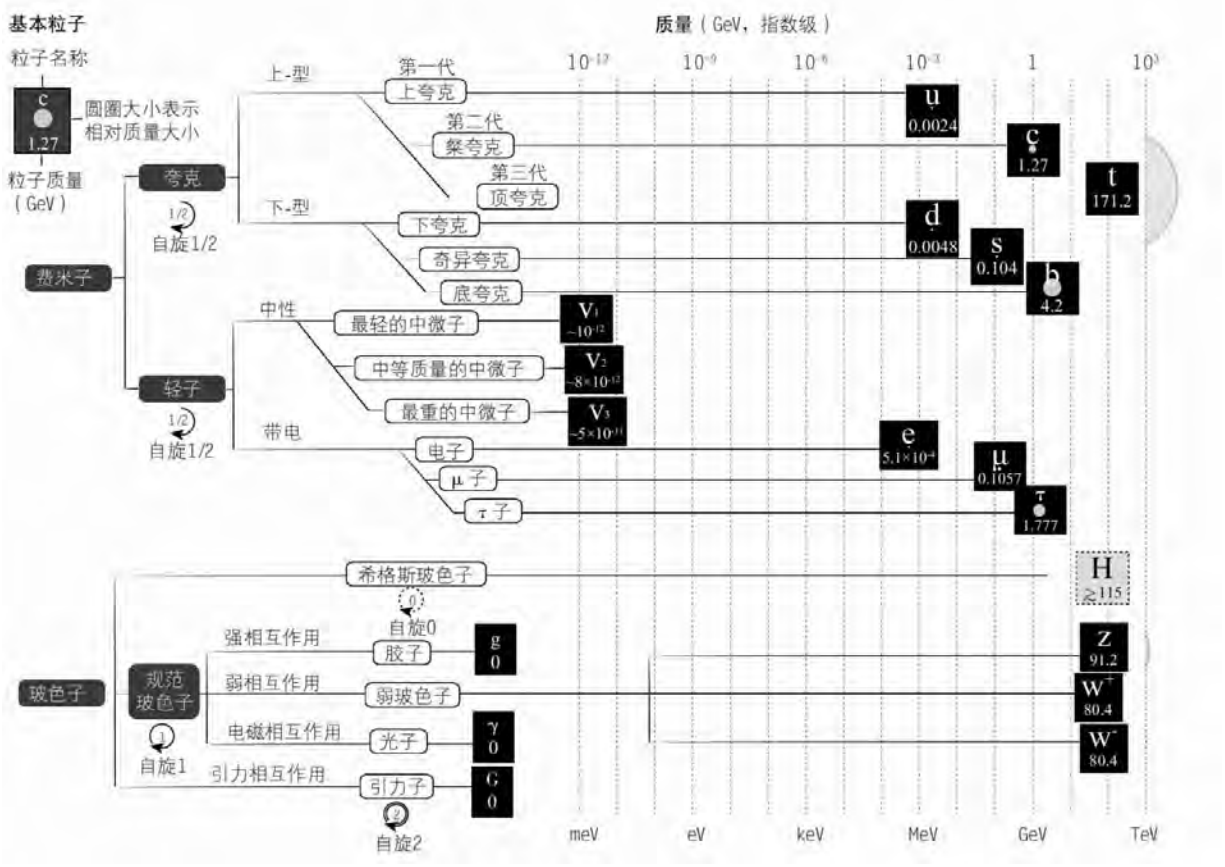


图14-7 根据类型和质量排列的标准模型粒子的总结表。灰色圆圈（有的在方框内）表示粒子质量。我们会从中发现标准模型元素的神奇变化。



第四部分

寻找“上帝粒子”

**KNOCKING ON
HEAVEN'S DOOR**



2007年2月，诺贝尔奖获得者、理论物理学家默里·盖尔曼在美国加州召开的TED大会上发表演讲。TED大会每年召开一次，与会者在各种主题上发表自己的新想法，是一次在科学、技术、文学、娱乐以及其他前沿领域中分享创新理念的盛宴。盖尔曼那万人空巷的演讲获得了全场观众的起立鼓掌，他演讲的主题是关于科学中的真与美。演讲的基本前提可以很好地用他的话总结——恰好呼应英国诗人约翰·济慈（John Keats）的诗：“真即是美，美即是真。”^[62]

盖尔曼有很好的理由相信这个伟大宣言。他的重要发现足以令他获得诺贝尔奖，其中有一些是关于夸克的，他掌握着可以把20世纪60年代实验中发现的看似随机的数据优美地组织起来的基本原理。在默里的亲身经历里，对美（或者至少是简洁性）的追寻也催生了真理。

听众中没有人质疑他的观点。毕竟绝大多数人喜欢美与真理相伴的想法，因此寻找一个往往会揭示另一个。但我总是觉得这个假设有一点不可靠。虽然每个人都相信伟大的科学理论的核心是美，并且真理永远在美学上是令人满意的，但是美从某种程度上来说，是一个主观标准而非真理的可靠裁决者。

在真理与美的评判中，基本的问题是：该判定不是一直成立的，只有时机合适时才成立。如果真理与美等价，那么“丑真理”就永远也不会进入我们的字典里。尽管这些话不是专门针对科学，我们对世界的观察也并不总是美的。达尔文的同事托马斯·赫胥黎（Thomas Huxley）很好地总结了这种观点。他说：“科学是整理好的常识，许多美的理论都被丑陋的事实给扼杀了。[分享微-信w s y y 5 4 3 7]” [1]

让事情变得更难的是，物理学家必须接受“宇宙和它的元素不全都是美的”这个令人不安的事实。我们观测到过很多我们想要理解的杂乱现象。在理想情况下，物理学家很想发现一个简单的理论，它仅仅需要很少的规则与最少可能的基本元素，就能揭示所有的结果。但是即使在寻找最简单、优美、统一的理论的过程中——该理论可以用来预测任何粒子物理学实验的结果，哪怕我们找到了这种理论，我们也需要众多步骤才能将它与现实世界相联系。

宇宙是如此复杂。在我们将一个简单、备用的构想联系到更复杂的周遭世界之前，我们还需要有新的元素和原理。这些附加的元素可能会破坏初始设定的构想中的美感，就像特殊条款经常会干扰国会法案理想的立法初衷。

由于存在失败的可能性，我们如何尝试去超越我们所知道的？如何诠释目前还没有解释的现象？本章将阐述关于美的概念以及审美标准在科学中担当的角色，以及将美作为指导原则的优缺点。本章同时会介绍建模知识，它采用一种自下而上的科学研究方法并且兼顾美学标准，来猜想接下来会被发现的理论。

美

我最近与一位艺术家聊天，他幽默地评论当代科学的一个巨大讽刺是，今天的研究者看起来比时下的艺术家更有可能将美作为他们的目标。当然艺术家们并没有遗弃美学标准，但是他们在讨论工作的同时还会讨论发现与发明。科学家们虽然也珍视其他属性，但他们同时致力于发现最引人注目的优美理论。

尽管许多科学家对于优雅的评价很高，他们关于简单与美却有不同定义。就像你与邻居就当今艺术家，例如达明安·赫斯特（Damien Hirst）的才华有着强烈的分歧一样，不同科学家对不同的科研方向也有着各自的喜好。

与志同道合者一样，我倾向于寻找能将观测到的表观不同的现象联系起来的基本原理。我的绝大多数弦理论同行研究具体的可解理论，他们会使用艰深的数学公式解决玩具问题（toy problem）——不必与真实的物理设置相关联的问题，那些只有以后才能被发现可以应用于观测到的物理现象的问题。其他物理学家则倾向于关注仅有简洁优雅公式的理论，这些理论可以产生许多他们可以进行系统计算的实验预测。

有趣的原理、高级数学与复杂的数值模拟都是物理学的一部分。绝大多数科学家很重视它们，但是我们根据自认为最满意或者最可能引领科学发展的事物选择最优的。事实上，我们常常按照哪一种方法与我们的独特偏好和天赋相契合，来选择它们。

不仅当前的审美观点会变。与艺术一样，随着时间流逝，连态度也会演变。默里·盖尔曼的专业——量子色动力学，就是一个极好的例子。

盖尔曼关于强相互作用力的猜想基于一个天才的洞察力——20世纪60年代不断发现的粒子可以被组织成一个合理模式，以此来解释它们的丰度和类型。他假设存在更多被称为夸克的基本粒子，认为它们带一种新型的电荷。这样，强相互作用力就会影响任何携带该种假想电荷的物质（就像电相互作用将电子和带电原子核结合起来，形成电中性原子）。如果这是真的话，那么所有被发现的粒子就可以被理解成这些夸克的束缚态（聚集在一起的不带净电荷的物质）。

盖尔曼意识到，如果有三种不同类型的夸克，其中每种夸克带一种不同的色荷（color charge），那么它们可以形成有许多种组合的中性束缚态。这些组合的确对应于实验发现的过多粒子。盖尔曼从而发现了看起来杂乱无章的粒子的一种优美解释方法。

然而，当盖尔曼以及另一位物理学家（后来的神经生物学家）乔治·茨威格（George Zweig）第一次提出这个理论时，人们甚至不相信这是一个正式的科学理论。其原因虽然是出于技术考虑，却也很有趣。粒子物理学的计算依赖于距离很远、没有相互作用的粒子，当粒子接近时，我们能计算其产生的相互作用的有限效应。在这种假设下，任何相互作用都可以完全被非常相近粒子的区域相互作用刻画。

另一方面，盖尔曼猜想的相互作用在粒子进一步分开时反而变得更强烈了。也就是说即使夸克相距较远时，它们也会相互作用。根据当时占统治地位的标准，盖尔曼的猜测甚至没有对应到一个可以用来做可靠计算的真正理论。因为夸克总是相互作用，即使所谓的渐近态（asymptotic state，夸克远离任何其他物质的状态）也非常复杂。在一

个明显的向丑让步的情况下，他们所提议的渐近态不是一个你希望看到的可计算理论中的简单粒子。

起初，没人知道如何在这些复杂的强束缚态中开展计算。然而，今天的物理学家关于强相互作用力的想法截然相反。我们现在的理解比该想法刚提出之时已经好得多了。戴维·格罗斯、戴维·普利泽（David Politzer）与弗兰克·韦尔切克（Frank Wilczek）因其名为“渐近自由”（asymptotic freedom）的研究获得诺贝尔奖。在高能时，强相互作用力并不比其他相互作用的耦合更强，前者的计算工作与后者类似。事实上，一些物理学家今天认为，在高能时耦合变弱的理论（例如强相互作用力）才是唯一合理定义的理论。因为相互作用的强度在高能时不会达到无穷大，而其他理论的强度则可能达到无穷大。

盖尔曼有关强相互作用力的理论是一个联系美学与科学准则的有趣例子。简洁性是他最初的指导原则。但在每个人都认同他的这个美的提议之前，还需要有艰难的科学计算和理论洞见。

当然，这不是唯一的例子。我们所相信的许多理论都有着表观丑陋和毫无说服力的方面，甚至受人尊敬、卓有成就的科学家最初也反对它们。量子场论（将量子力学和狭义相对论结合起来）是所有粒子物理学的基本理论。而诺贝尔奖获得者、意大利科学家恩里科·费米等一开始也反对它。对于他来说，尽管量子场论将所有的计算正规化与系统化并且作出了许多正确的预测，然而它涉及的许多计算技巧，哪怕在今天的一些物理学家看来也是非常怪异的。量子场论在不少方面都相当优美，且其洞见也引人注目。不过，我们必须忍受它的其他复杂、不甚让人满意的特点。

这种故事已经重复了很多次。美被认可经常是后验的。弱相互作用力违背了宇称守恒。这意味着，向左旋转的粒子的相互作用与向右

旋转的不同。这种左右等价的基本对称性的破缺令人烦恼，且毫无吸引力。然而正是这种不对称性催生了我们在世界上所看到的不同质量类别，它们对这个世界的构成、生命的形成也是必要的。起初它被认为是丑陋的，而现在我们知道它是必不可少的。虽然它自身可能很丑，但是宇称破缺让我们对所见之物有了更至关重要、更优美的解释。

美不是绝对的。一种想法对于创造它的人来说可能是有吸引力的，但从其他人的角度来看可能是麻烦的或是杂乱的。有时我会陶醉于自己想到的一个猜想，很大程度上是因为我知道以前人们所想到的其他想法都没有奏效。比以前出现的理论更好，却并不保证优美性。我让自己参与创建的模型满足了这个标准，但不甚熟悉该模型所表达主题的同事仍然对此表示怀疑和困惑。我想到了一个认定什么是好想法的更好标准：一个好的想法可以让甚至从来没有研究过相应问题的人都能欣赏到它的美。

事情的反面有时也会发生——好的想法被拒绝，是因为其发明者认为它们不够优美了。马克斯·普朗克不相信存在光子（他觉得这是一个让人讨厌的概念），即使是他引发的诸多逻辑导致了这种猜想。爱因斯坦认为根据其广义相对论得出的宇宙膨胀理论不可能正确，部分原因是它与他的美学和哲学倾向相违背。彼时，两者中没有哪一个想法看起来是最美妙的，然而物理定律与它们所应用的宇宙却并不会在意它们美不美。

看起来很好

鉴于美具有演变性和不确定性，值得考虑一些特征——它们可以使一个想法或者一个形象在某种形式上具有客观美丽性并具有普遍吸

引力。也许关于审美标准最基本的问题是：人是否拥有“何为美”的统一标准（在任何情景中），无论是艺术，还是科学。

没有人知道答案。毕竟，美与品位相关，而品位可以是很主观的判断。然而，我发现很难相信人们毫无审美标准的共识。例如，在特定的展览中哪些艺术展品是最好的，或者人们常选择去看哪些展览，人们的意见总是令人吃惊地一致。当然这并没有证明什么，因为我们处在相同的时间与空间中。很难将关于美的各种信仰从它们所产生的具体文化背景及时空背景中隔离开来，因此我们很难将浑然天成的属性从后天习得的价值与判断中分离出来。在一些极端情形中，人们可能会一致赞同某件事物看起来不错或者不好。而在极少情况下，每个人都同意一个想法的美妙之处。但是即使在这些鲜少的事例中，人们也不一定赞同所有的细节。

即使如此，一些审美标准看起来仍是统一的。任何艺术课刚开始的时候都要讲平衡。米开朗基罗的《大卫》雕塑就是这一原则的典范。大卫优雅地站着，他永远不会翻倒或者崩塌。人们可以从中寻找平衡与和谐。艺术、宗教与科学都给了人们接触这些品质的机会。当然平衡也可能只是一个规划的原则。当艺术打破了平衡感时，它同样可以让人陶醉，比如雕塑家理查德·塞拉（Richard Serra）的早期雕塑作品（见图15-1）。

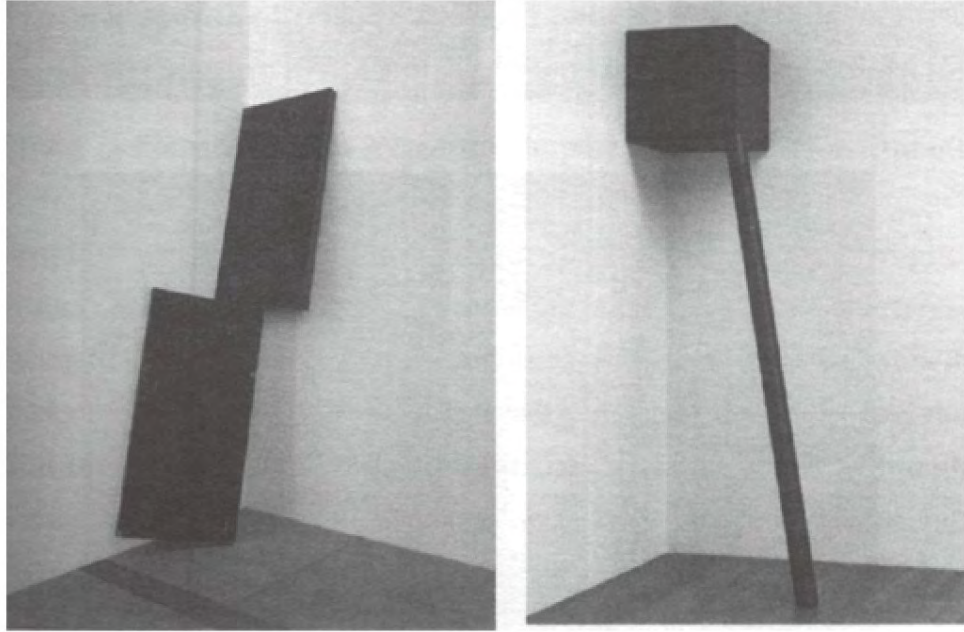


图15-1 理查德·塞拉的早期雕塑作品表明，有时一点失衡感反倒为艺术增添了几分趣味。（版权©2011，理查德·塞拉/艺术家权利协会，纽约）

对称性通常在美学的考虑中至关重要，艺术与建筑也经常展现出由对称产生的秩序。有些东西具有对称性是指当你改变它——例如当你转动它、用镜子反射它，或者交换它的部件时，经历这些改变后的体系仍与原来的体系没有什么差别。对称的和谐感可能是宗教符号常具有对称性的一个原因。基督教的十字架、犹太教的大卫之星、佛教的转经轮与伊斯兰教的新月等标志都是如此（见图15-2）。

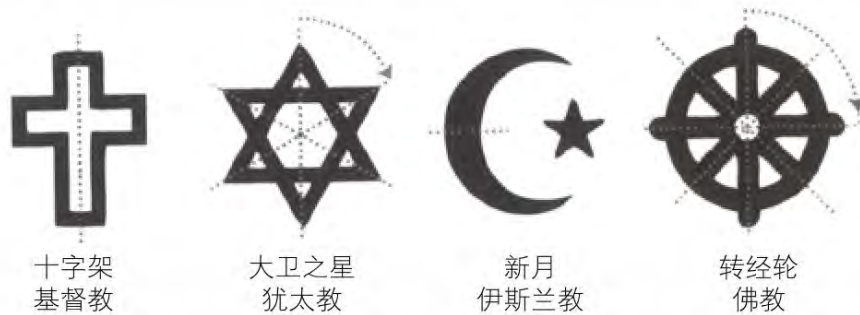


图15-2 宗教符号常常具有对称性。

更广泛地，伊斯兰艺术禁止象征，而依赖几何形式，尤以使用对称著称。例如，印度的泰姬陵是一个壮观的建筑。任何参观过泰姬陵的人，无一不被它高超的几何形状应用与对称性所吸引。位于西班牙南部的阿尔罕布拉宫（Alhambra）使用了摩尔人（Moorish）的艺术思想与有趣的对称图案，它可能是如今世界上目前保存完整的最美丽的建筑之一。

近代艺术，例如埃斯沃斯·凯利（Ellsworth Kelly）或布里奇特·莱利（Bridget Riley）的作品，展现了清晰的几何对称性。哥特式或文艺复兴时期的建筑，例如沙特尔大教堂与西斯廷教堂的穹顶，也精巧地使用了对称性（见图15-3）。

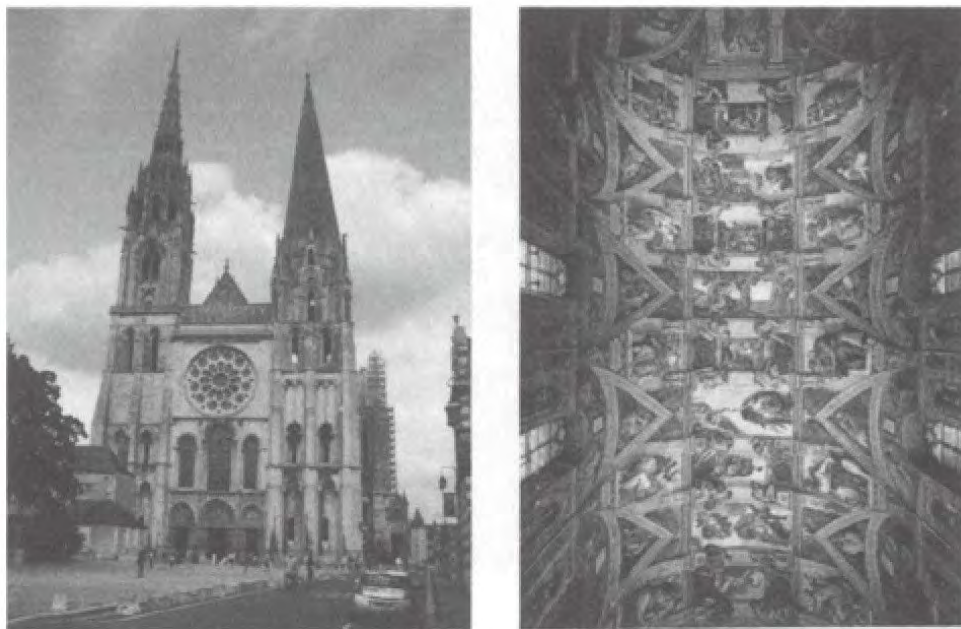


图15-3 沙特尔大教堂的建筑以及西斯廷教堂的穹顶都体现了对称性。

最美的艺术却常体现在其不对称性上。日本艺术以其优雅及其不对称性著称。日本绘画和丝绢印花有一个清晰的取向，可以将人的目光吸引到图画上（见图15-4）。

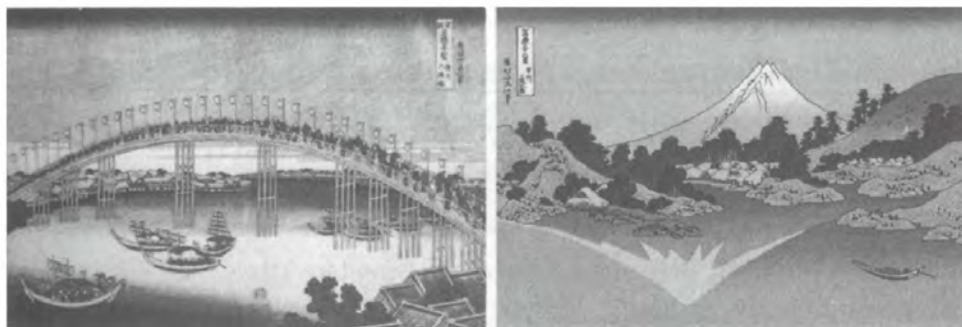


图15-4 日本艺术很有意思，某种程度上是因为它的不对称性。

简单性是另一个可以品评美感的标准。一些简单性出自对称性，但是背后的秩序可以被呈现出来，甚至在缺乏明显的对称时仍是如此。杰克逊·波洛克（Jackson Pollock）的画作在颜料的密度之中隐藏着简单性，尽管它给人的第一印象是杂乱无章。虽然每一次颜料的溅落看起来完全随机，但在他最著名、最成功的作品中，每种色彩都有比较均匀的分布。

艺术中的简单性往往具有欺骗性。我曾经试着临摹法国画家马蒂斯（Matisse）的几幅画，这几幅作品是他在年老体弱时所作。然而当我试图再现它们时，我意识到它们并非那么简单——至少对于不懂绘画技术的我来说如此。简单元素可以包含的结构，比我们表面上能观察到的要多。

无论如何，美不仅仅出现在简单基本的形式中。一些人崇尚的艺术作品，例如拉斐尔（Raphael）或者提香（Titian）的画包含许多内部元素的丰富复杂的画布。完全简单的东西会让人头脑麻木，因此当我们看待艺术时，我们倾向于一些有趣的、吸引眼球的东西。我们希望一些简单到可以理解而又不能太过简单以至无聊的东西。这似乎也反映了这个世界的建构规律。

科学中的美

审美标准很难确定。科学与艺术一样有统一的主题但没有绝对。即使科学的审美标准并没有被很好地定义，它们却大有用处，且无所不在。它们指导我们的研究，哪怕它们不能担保成功或者发现真理。

科学中的审美标准不像我们刚才罗列的艺术标准。它们帮助我们组织计算并会联系不同的现象。有趣的是，与艺术一样，对称性往往只是近似的。最好的科学描述常常在保留对称理论优美性的同时，结合必要的对称性破缺，来对世界作出预测。对称破缺丰富了对称性包含的想法，由此产生了更多元的解释。与艺术一样，具有对称性破缺的理论比那些完全对称的理论更美、更有趣。

希格斯机制——使基本粒子获得质量的机制，就是一个绝好的例子。希格斯机制巧妙地揭示了与弱相互作用力相联系的对称性，如何可以被轻微地破坏掉。我们目前还没有发现可以为该机制提供无可辩驳的证据的希格斯玻色子，希格斯机制美妙又独特地满足了实验和理论双方的要求，因而许多物理学家都相信希格斯粒子的存在。

简单性对理论物理学家来说是另一个重要的客观标准。我们有一种根深蒂固的观念：简单元素构成了我们所看到的复杂现象。寻找组成现实或者现实所代表的简单基本要素，从很久以前就开始了。在古希腊，柏拉图想象的完美形式——几何形态与理想存在，在地球上只有近似的存在。亚里士多德也相信理想形式，但他认为，由物理对象近似给出的理想将只能通过观察揭示。宗教也常常假设一个更完美、更统一的状态，远离但又以某种方式与现实相联系。甚至从伊甸园堕落的故事，它所预设的前提是一个理想化的先验世界。虽然现代物理学的问题和方法与我们祖先时代的方法已经截然不同，但是许多物理学家仍在寻找一个更简单的宇宙——不是在哲学意义和宗教意义上，而是在组成世界的基本要素上。

寻找隐含的科学真理往往需要寻找简单的要素，它建构出了我们所观察到的复杂而丰富的现象。这样的寻找通常包括试图找出有意义的模式或组织原则。只有简洁地实现一个简单和优美的想法，绝大多数科学家才能期望这个提案有正确的可能。出发点涉及最少的输入，它承诺最具预测力的效果越显著。当粒子物理学家考虑什么是标准模型的核心建议时，如果一个想法的实现变得非常烦琐，那么我们常会变得疑心重重起来。



与艺术一样，科学理论自身可以是简单的，或者它们可以由简单且可预料的元素构成的复杂组合。当然，即使初始组件甚至规则本身都是简单的，那结果也不一定是简单的。

这种追求的最极端版本是寻找一个仅仅由几个简单元素组成、遵从很少原则的统一理论。这一追求是一个颇有抱负（有人可能说太大胆）的任务。显然，阻止我们立即发现一个能完全解释所有观测的完美理论的一个明显障碍是：我们周遭的世界表明，一个理论只包含世界的一部分简单性。一个统一的理论，同时又是简单的、优美的，还必须在某种程度上容纳足够多的结构来与观测相符。我们愿意相信一个简单、优美、可预测的理论就是物理学的全部。然而宇宙不像理论那么纯粹、简单与有序。即使有一个潜在的统一表述，也仍需要大量研究，因为它们将理论与我们

在世界上所见的引人入胜的复杂现象联系了起来。

当然，我们可以在这些表征美与简单性上走得更远。教科学或数学的教授总会说已经研究透彻的现象“太容易”（trivial），不管它们有多复杂。教授已经知道答案并且对基本元素与逻辑很了解，但是对课堂上的学生来说情况并非如此。当学生们把问题约化成简单部分后，对他们来说也可以是“太容易”的了。但他们首先需要想出如何做到那一步。

美的标准并不唯一

像在生活中一样，美在科学中并非唯一标准。实验的约束以及直觉可以指导我们探索新的知识。艺术中与科学中的美可能有一些客观属性，但几乎所有的应用会与品味和主观性相关。

然而，对于科学家来说差别巨大。最终实验会决定我们的哪一种想法是正确的。科学的发展可能开发出更多审美标准，但是真正的科学进步还要求理解、预测与分析数据。不管一个理论看起来多么美丽，它也可能是错的——就此必须被丢弃。即使是最理智、最令人满意的理论，如果不能应用到真实世界中，同样也要被抛弃。

尽管如此，在我们达到更高能量，或者获得确定正确物理描述所需的足够参数之前，物理学家只能采用美学与理论思考来猜测超出标准模型之外的东西，除此之外别无选择。在此期间，只有有限的的数据，我们依赖现有的谜团与交织在一起的不同品位以及组织标准，来指导前行的道路。

理论上说，我们希望能够研究各种可能的结果。建模是我们做这种研究的通常方法。我和同事探讨各种粒子物理学模型，它们是构成

标准模型基础的物理理论的猜测。我们的目标是探索将复杂现象组织在一起的简单原理，那些复杂现象出现在更容易看到的尺度上，以至于我们可以通过我们的认识来解决当前的疑难。

物理模型的创建者采取有效的理论观点，理解越来越小尺度的渴望已经深深印在他们的心里。我们依据“自下而上”的方法，从已知的出发（我们可以解释的及我们感到迷惑的现象），并试图推导基本模型，从而可以解释观测到的基本粒子的性质以及它们相互作用之间的联系。

“模型”一词可能引发一个物理结构，如同房子的缩微版，用来展现和探索它的建筑结构。或者你可以想象在计算机上的数值模拟计算出已知物理原理的结果，这类似于气象建模或传染病传播的模型。

粒子物理学中的建模与上述两种定义都不同。粒子模型与杂志或者时装秀的模特的天赋有相似之处。模型（模特）在物理中（在T型台上）都展现了富有想象力的新想法。人们纷纷涌向那美丽的或者至少是更引人注目或更令人惊叹的模型（模特）。但最终，他们会被那些真正给人承诺的模型（美丽的模特）所吸引。

毋庸多言，相似性到此为止。

粒子物理学模型是关于什么可能隐藏在理论背后的猜测，这些理论预测已经被检验并且已经被我们理解。审美标准在决定哪一个想法值得探求时非常重要。想法的相容性与可检验性同样重要。模型所刻画的不同基本物理元素与原理适用在比现今实验尚未测量到的还要小的距离和尺度上。在模型的帮助下，我们可以确定不同理论假设的本质和结果。

模型是一种从已知事实出发，创造出说服力与综合性更强的理论的推断方法。它们是各种提案的组合，一旦实验允许我们深入到更小尺度或者更高能量上，我们就可以检验它们的基本假设和预测，它们则可能会、也可能不会被证明正确。

请记住“理论”与“模型”是不同的。“理论”这个词，我指的并不是粗略猜测——比如更常见的口头用语。已知粒子与它们遵从的已知物理定律是一个理论的组成部分，它是一个定义良好的元素与原理的集合，有法则和方程来预测元素之间如何相互作用。

即使我们完全理解了一种理论与它的启示，同一理论也还可以有不同的应用，而这些应用在真实世界中会有不同的物理后果。模型是一种抽样这些可能性的方法。我们将已知的物理原理和元素结合成为一个描述现实的备选者。

如果你将理论想象成幻灯片模版，那么模型就是你特定的演讲报告。理论允许连续的动画模拟，但模型只包含你需要用来概述你要点的部分。理论会给出标题与要点，而模型包含刚好你需要传达以及你期望能很好地应用到手边的任务。



物理中建模的本质已经随着物理学家试图解答的问题不同而发生变化。物理总是试图从最小数目的假设预测大量物理量，但那不意味着我们可以立即确定最基本的理论。物理学的进展常常发生在对所有事物最基本层面的理解获得之前。

19世纪，物理学家理解了温度与压力的定义，并将它们应用于热力学和引

擎设计中，远早于人们能够从更基本的大量原子分子随机运动的微观角度来诠释这些想法。20世纪初期，物理学家试图运用模型以电磁能量的方式来解释质量。虽然这些模型都来源于系统如何成功运行的共同信念，但是它们都被证明是错误的。后来，物理学家尼尔斯·玻尔构造了一个原子模型，来解释人们所观测到的发射光谱。不久，他的模型被更全面的量子力学理论取代，该理论不但吸收而且发展了玻尔的核心理念。

模型的创建者今天试图确立超越粒子物理学标准模型的理论。之所以现在被称为标准模型，是因为它已经得到了很好的验证和理解，但是它终究是一个猜测——如何将已知的观测与当时已经发展起来的理论配合起来。尽管如此，由于标准模型隐含着如何检测它的前提预言，实验最终可以证明它是正确的。

标准模型解释了迄今为止所有的观测，但物理学家却相当确信它是不完整的。特别是，它留下了悬而未决的问题：在希格斯区域中的元素，应该对基本粒子的质量负责的正确粒子与相互作用有哪些，以及为什么是这个区域中的有着特定质量的那些粒子。超越标准模型的理论应该阐述更深入的潜在连接与关系，从而解释这些问题。它们涉及基本假设与物理概念的特定选择，以及它们可以应用的尺度和能标。

我近来的许多工作都涉及新模型的思想，以及新的或者更细致的搜索策略——如若不然将会错过新现象。我思考自己提出的模型以及其他全方位的可能性。粒子物理学家知道元素的类型和可能涉及的法

则，比如粒子、力以及允许存在的相互作用。但是我们并不明确知道哪一种元素是现实成分的一部分。通过应用已知的理论成分，我们试图确认那些潜在的简单基本想法，它们可以被运用到最终的复杂理论中。

同样重要的是，模型为实验设定探索目标，在比目前物理学家实验所研究的更小的距离上探讨粒子的行为。测量为我们提供线索，帮助我们区分相互竞争的候选者。我们还不知道新的基本理论是什么，但是可以描述可能的偏离标准模型的性质。通过考虑反映基本现实与结果的候选模型，如果模型正确，那么我们就可以预测大型强子对撞机可能揭示的东西。使用模型可以确认我们想法的性质，认识到与现有数据相一致的过多的可能性，并解释至今仍令人费解的现象。只有一些模型会被证明是正确的，但创造与理解它们是确定选项的最佳方法，是建立一个令人信服的要素资源的手段。

探索模型与模型给出的详细结果，有助于我们确定可以令实验物理学家搜寻的可能存在的东西。模型为实验物理学家勾勒出新物理理论的有趣面貌，这使得他们可以检验模型的创建者是否已经正确地给出了物理元素或物理原理，指导体系之间的关系与相互作用。新物理定律的任何模型应用在可测的能量上，将预测新粒子与它们之间的联系。观测从对撞产生的粒子及它们的性质有助于确定存在的粒子类型、质量以及它们的相互作用。在发现新粒子或者测量不同相互作用的过程中，实验物理学家将确认或者排除已提出的模型，为更好的模型做好铺垫。

获得足够的的数据后，实验物理学家将决定哪一个基本模型是正确的，至少在我们可以研究的精度、距离与能量上做到这一点。我们希望在大型强子对撞机能标可以探测的最小尺度上，基本理论法则足够简单，这让我们可以推导和计算相关物理定律的效应。

物理学家常常热烈地讨论，哪些是最好的研究模型，以及从实验上寻找最好的获取它们的方法。我经常跟实验物理学同事们一起坐下来，讨论如何最好地使用模型来指导他们的研究。例如，具体参数在具体模型中的基准点是否太特殊？有没有更好的方法来覆盖所有的可能性？

大型强子对撞机实验太具挑战性了，以至于如果没有确定的搜索目标，实验结果就将被标准模型的背景所掩盖。实验不仅会根据已经存在的模型进行设计及优化，而且它们也会寻找更广泛的可能性。实验物理学家意识到，构造一个应用广泛的模型非常关键，因其遍布了所有可能出现的新迹象，这尤为重要，因为没人想要一个有太过偏见的具体模型。

理论物理学家与实验物理学家都在努力工作，以确定我们没有漏掉任何重要的东西。我们不知道不同提议中哪一个是正确的，一直到实验验证的那一刻。所提出的模型可能是现实的正确描述，但即使它们不是，它们也启发了我们的搜索策略，并告诉我们至今尚未发现的新物质的不同面貌。希望大型强子对撞机会给我们一个结果——不管结果如何，我们都要做好准备。



2010年3月30日，我在雪花般的电子邮件中醒来。邮件都是有关前一天晚上欧洲核子研究中心的7TeV对撞实验。它的成功，标志着大型强子对撞机实实在在的物理项目的开始。而2009年接近年底时的加速与对撞实验已经成为关键的技术里程碑。这些事件对大型强子对撞机的实验物理学家来说非常重要，因为他们终于可以通过来自大型强子对撞机真实对撞实验的数据，来校准并更好地理解探测器，而不是仅仅采用偶然穿过仪器的宇宙射线数据。在接下来的一年半时间里，欧洲核子研究中心的探测器会记录真实数据，使得物理学家可以证实或证伪他们的模型。最终，经过了起起伏伏的实验，大型强子对撞机的物理项目终于在排除万难后开始了征程。

7TeV的对撞能量只是大型强子对撞机设定能标的一半。真实的目标能标——14TeV在几年之内还无法达到。同时，在7TeV的一轮运行上目标锁定的光度（每秒相互碰撞的质子的数目），远低于实验设计者最初计划的数目。但是无论如何，有了这些对撞结果，大型强子对撞机的实验终于在经历重重困难之后步入了轨道。我们终于可以相信我们对物质内部本质的理解将很快得到提高。如果一切顺利的话，那么两年后仪器将经历完全关闭、整装待发、再重新以满负荷运转的阶段，来为我们提供我们所期盼的真实结果。

其中一个最重要的目标是，研究基本粒子如何获得质量的问题。为什么所有粒子不都像那些零质量粒子一样，以光速穿过？回答这个问题的关键是一系列被统称为希格斯区的粒子，包括希格斯玻色子。本章将介绍为什么成功的粒子搜寻过程可以告诉我们关于基本粒子质量的想法是否正确，以及一旦大型强子对撞机以更高强度和更高能量华丽地回归，重新启动搜寻工作，那么它终将告诉我们隐藏在这重要而又引人注目的现象之下的粒子和相互作用的本质。

希格斯机制，给出基本粒子质量的唯一方法

没有物理学家质疑在现今我们所能研究的能量上的标准模型。实验物理学家已经测试了很多它的预测，在1%的精度上，实验结果都与预测符合得很好。

然而标准模型建立在一个还没有观测到的因素上面。以英国物理学家彼得·希格斯命名的希格斯机制是我们所知的给出基本粒子质量的唯一方法。根据最初版本标准模型的基本前提，传递相互作用的规范玻色子没有质量，对标准模型来说非常重要的基本粒子，诸如夸克和轻子，也都没有质量。然而对物理现象的测量明显证明它们具有质量。基本粒子的质量对于理解原子与粒子的物理现象——例如原子中电子轨道半径或者弱相互作用的极微小范围，至关重要，更不用说它对宇宙结构的影响了。根据质能方程 $E=mc^2$ ，质量也决定了有多少能量是需要用来产生基本粒子的。然而对于一个没有希格斯机制的标准模型来说，基本粒子的质量将是一个谜题。这样的理论是不被允许的。

粒子对其质量丧失了不可剥夺的权利，这听起来似乎有点专制。你可能还期望粒子有权选择获得非零质量。然而标准模型与所有相互作用理论的精妙结构，恰恰表明它是非常独断专行的。该结构限制了

允许的质量类型。规范玻色子的质量解释与费米子的有些不同，但是两者的基本逻辑都与处于所有相互作用理论核心的对称性相关。

粒子物理学的标准模型涵盖了电磁、弱相互作用力以及强相互作用力，并且每一种相互作用都伴随着一种对称性。如果没有这些对称性，那么由量子力学与狭义相对论所预言的传递相互作用的规范玻色子其振动模式就会出现很多冗余。没有对称性的理论，其计算结果将产生荒谬的预测，例如赝振动（spurious oscillation）模式的高能相互作用的概率大于1。在任何对自然的准确描述中，这种非真实的物理粒子，也即在错误方向振动的非真实存在的粒子，显然需要被剪除。

从这个意义上说，对称性就像垃圾邮件过滤器或者质量监督管理局。例如，质量规定只有平衡性能^[63]良好的汽车才能被留下来，这使得所有从工厂出产的汽车能像预期的一样，良好地行车。对称性在任何相互作用的理论中都会排除不良因素。这是因为发生在不想保留的非物理粒子上的相互作用不遵从对称性，反之，那些遵守对称性的粒子则以它们被允许的方式振动。因此对称性确保了理论预测中只涉及物理粒子，这样理论才能合理并与实验相符。

对称性使得相互作用理论拥有了一种优美构想。与其通过计算将非物理模型逐一排除，不如利用对称性一举将所有非物理理论全部消灭。总之，任何具有对称性的相互作用理论只涉及物理振动模式，它们的行为是我们想描述的。

对任何理论来说，只要传递相互作用的粒子的质量为零（比如电磁或者强相互作用力），这种方法就很奏效。也就是说，对称理论对于它们的高能相互作用的预测都合情合理，而且的确只有自然中存在的物理模式才被涵盖。对于零质量的规范玻色子，高能相互作用问题

的解决相对来说很直接，这是因为相应的对称性限制和排除了理论中的非物理、表现不良的模式。

对称性解决了两类问题：第一，排除了非物理模式；第二，解决了伴随非物理模式而来的错误的高能预测。然而，非零质量的规范玻色子有一个额外的、物理的（即存在于自然中的）振动模式。参与弱相互作用力的规范玻色子就属于这一类别，对称性会过多地剪除它们的振动模式。因此，如果没有新要素被添加进来，那么弱规范玻色子的质量就将使它不能遵从标准模型的对称性。更进一步，对于非零质量的规范玻色子，我们别无选择，只能保留它们的一种不良模式，这也意味着解决高能的不良结果的问题将变得困难起来。无论如何，理论需要一些新要素来产生合理的高能相互作用。

此外，在没有希格斯粒子的情况下，没有一种标准模型的基本粒子可以既获得非零质量，又同时遵从最原始相互作用理论的对称性。由于相互作用的对称性，在没有希格斯粒子的标准模型中，夸克与轻子也不会具有非零质量。其原因看似与规范玻色子的逻辑无关，但是也可以追根溯源到对称性上面。

在第14章，我们展示了一张包括左手与右手费米子的图（即图14-1），图中非零质量的粒子成对出现。当夸克与轻子质量非零时，它们引入了将左手与右手费米子相交换的相互作用。但是为了左手与右手费米子可以互换，它们必须参与相同的相互作用。然而实验表明，弱相互作用力在左手夸克和轻子与右手夸克和轻子上的作用并不相同，但是后者可以由前者转换而来。这种对宇称守恒的破坏（宇称守恒的意思是左手与右手遵从相同的物理定律）令每个第一次听闻它的人瞠目结舌。毕竟其他已知的自然定律都不区分左与右。但是这个显著的特性，也即弱相互作用力区别对待左与右，已经被实验证明并且是标准模型的重要特征。

左手与右手的夸克和轻子有着不同相互作用告诉我们，如果没有一些新要素，非零质量的夸克和轻子将与已知的物理定律不相容。并且非零质量将把带弱荷的粒子与不带弱荷的联系起来。

换言之，因为只有左手粒子带弱荷，那么这种荷也可以失去。也即荷可以明显地消失于真空——宇宙中不含任何粒子的状态。通常这是绝不可能发生的，因为荷必须守恒。如果荷可以无中生有，又可以凭空消失，那么与之相伴的相互作用的对称性一定会破缺，并且那个荒谬的高能规范玻色子相互作用的概率超过1的预测又会重新出现。因此，假如真空确实空无一物（没有粒子和场），那么荷永远不会神奇地消失。

假如“真空不空”，而是包含了可以向真空提供弱荷的希格斯场，那么荷就可以产生或者消失了。一个希格斯场，尽管可以向真空提供荷，却不是由真实粒子组成的。它本质上是一个分布——只有当场自身取得非零值时，宇宙中弱荷才出现的一个分布。当希格斯场非零时，就好像宇宙有一个无穷供给的弱荷源。想象一下你有数不尽的金山银山，你可以任意借出和收回金钱，并且你将一直拥有无穷的钱财供你使用。与这个比喻类似，希格斯场将无穷的弱荷投放到真空中。这样，它就破坏了相互作用中的对称性，并且让荷流进与流出真空而使粒子出现了非零质量，却没有产生任何问题。

一种考虑希格斯机制以及质量产生的方法是，让真空表现得像带有弱荷的黏滞流体（希格斯场渗透布满真空）。带有这种荷的粒子，例如弱规范玻色子与标准模型中的夸克和轻子，可以与这种流体相互作用，从而使其运动速度减慢。这种减慢效应对应于粒子获得了质量，因为零质量的粒子将以光速在真空中运动。

这个巧妙的基本粒子获得质量的过程就是希格斯机制。它不仅告诉了我们基本粒子如何获得质量，还告诉了我们很多这些质量的性

质。例如，这个机制解释了为什么一些粒子重而另一些轻。简单地说，粒子与希格斯场的相互作用越多、质量越大，反之则质量越小。顶夸克是最重的夸克，因此它的相互作用最强。电子或者上夸克的质量相对较小，相互作用就比较弱。



希格斯机制也提供了一个关于电磁以及传递该种相互作用的光子的深刻洞见。它告诉我们：只有这些相互作用的媒介粒子与分布于真空中的弱荷相互耦合，这些粒子才获得质量。W规范玻色子与Z玻色子都与这些弱荷耦合，所以它们具有非零质量。然而布满了真空的希格斯场虽然携带弱荷，却是电中性的。光子与弱荷没有相互作用，所以它的质量保持为零。于是，光子被单独地拣选出来。如果没有希格斯机制，那么将有三种零质量的弱规范玻色子以及另一种零质量的相互作用媒介粒子——即被称作超荷（hypercharge）的规范玻色子。那样，根本没有人会提出光子的概念了。但是希格斯场的出现，使得只有这种超荷规范玻色子与三种弱规范玻色子之一的唯一组合方式，可以给出作为传递电磁相互作用的媒介光子。光子具有零质量，对电磁理论的现象来说至关重要。它解释了为什么电磁波可以远程传播，而相反，弱相互作用力只能在一个极小范围内传播。正是因为希格斯

场带有弱荷而没有电荷，所以光子可以以光速传播（它因此得名），而弱相互作用力的媒介则是重的粒子，所以不能以光速传播。

不要混淆^[64]，光子才是基本粒子。但从某种意义上说，最初理论中的规范玻色子被错认了，因为它们并不对应于具有正确质量（可能是零质量）的物理粒子，而且它们在真空中的传播毫不受阻。在从希格斯机制得知遍布真空的弱荷之前，我们并没有办法从中确定哪些粒子有非零质量、哪些有零质量。由于希格斯机制，真空被带上弱荷，因为超荷规范玻色子与弱规范玻色子可以于真空中传播时相互转换，所以我们不能赋予它们确定的质量。在真空具有弱荷的前提下，只有光子和Z玻色子于真空传播时保持不变，其中Z玻色子有非零质量，而光子没有。于是希格斯机制可以将特殊的光子挑选出来，而与它相应的荷是它所传递的电磁作用的电荷。

希格斯机制解释了为什么光子而非其他相互作用媒介粒子具有零质量。它也解释了质量的另一个性质。这个问题更巧妙一些，却给了我们为什么希格斯机制中质量与合理的高能预测相容的洞见。如果将希格斯场考虑成一种流体，那么我们可以想象它的密度将对粒子的质量产生不同的影响。进一步说，假如我们认为它的密度来源于相隔一定距离的弱荷，那么有的粒子在很小距离上的传播使得它们不会与任何一个弱荷相撞，它们的运动方式就好像其质量为零；然而有的粒子在长距离上传播，它们不可避免地会撞到弱荷上而反弹回来，速度就会降低^[65]。

这对应一个事实：希格斯机制伴随着自发破缺（spontaneous breaking）的弱相互作用力的对称性，而该对称破缺发生在一个确定的尺度上。当一种对称性在自然定律（比如作为相互作用的理论）

中出现，却被系统的真实状态破坏时，我们称这种对称性发生了自发破缺。如我们所讨论的，对称性存在的原因与理论中粒子的高能行为相关。唯一的解释是：对称性存在，但是它们自发破缺了，从而使得弱规范玻色子可以获得质量而又规避了不良的高能行为。

希格斯机制背后的想法是：对称性的确是理论的一部分。物理定律总有对称性，然而世界的真实状态不保持该对称性。考虑一支铅笔尖端着地倒立，然后在一个特定的方向上倒下来。当铅笔还是直立着的时候，所有环绕着它的方向都是等同的。因此，倒下的铅笔自发地破坏了当它还是直立状态时所具有的旋转对称性。

类似地，希格斯机制自发地破坏了弱相互作用力的对称性。这意味着物理定律保持对称性，而它被真空充满弱荷的状态所破坏。希格斯场在宇宙中渗透的方式是不对称的，才使得基本粒子获得了质量。因为它破坏了弱相互作用力的对称性，也就是说，如果没有希格斯场，该对称性仍然存在。相互作用理论保持了与该弱相互作用力相伴的对称性，但是该对称性被充满真空的希格斯场破坏了。

通过将弱荷放进真空，希格斯机制使得与弱相互作用力相伴的对称性被破坏，并且发生在一个特定的尺度上。该尺度由真空中荷的分布所决定。在高能或者量子力学意义上等价的小尺度上，粒子不经受任何弱荷，因此它们体现出零质量的性质。因此在这种情形下，对称性得到了保持。然而，在大尺度时，弱荷在某些方面表现得像摩擦力，减慢了粒子的运动速度。只有在低能量，或者说大尺度时，希格斯场才会给粒子质量。

这正是我们所需要的。那些有害的、对于非零质量粒子无意义的相互作用只适用于高能情况。在低能时，根据实验，粒子可以，而且必须带有一定的质量。希格斯机制是我们所知道的唯一可以通过自发破缺弱相互作用力对称性的方法。

虽然我们还没有观测到在希格斯机制中负责提供给基本粒子质量的粒子，但是我们的确有希格斯机制在自然中应用的实验证据，并且在完全不同的领域——即超导材料中也发现了它的应用。超导发生在当电子结成对并且电子对充满了整个材料时。所谓的超导体中的凝聚（condensate）由电子对组成，它与希格斯场的作用一样。

不过超导体中的凝聚携带的是电荷而非弱荷，因此凝聚为超导材料中传递电磁相互作用的光子提供质量。这质量将电荷“屏蔽”起来，意味着在超导体内部，电场、磁场不能达到很远的距离。相互作用在很短的距离上很快衰减。量子力学与狭义相对论告诉我们，超导体中的屏蔽距离（screening distance）是仅在超导基质中出现的光子质量的直接结果。在这些材料中，电场的穿透深度不能比屏蔽距离还要大，这是因为从遍布超导体的电子对碰撞反弹的光子获得了质量。

希格斯机制的运作方式与此相似。但我们预测希格斯场（带弱荷）布满真空而不是电子对（带电荷）布满基质。在此情形下，我们发现获得质量的弱规范玻色子屏蔽了弱荷，而不是获得质量的光子屏蔽了电荷。因为弱规范玻色子有非零质量，所以弱相互作用力只在亚原子尺度的短距离上有效。

这是赋予规范玻色子质量的唯一自洽方法，所以物理学家都相当确信希格斯机制在自然中存在。并且我们希望它不仅是规范玻色子获得质量的原因，也是所有基本粒子获得质量的原因。除此之外，我们不知道还有什么其他自洽理论可以让标准模型中带弱荷的粒子获得质量。

本章有一些抽象概念，因此可能非常难理解。希格斯机制和希格斯场的概念本质上与量子场论和粒子物理学相联系，与我们所能看见的现象相去甚远。所以让我来简要总结一些关键点。

- 首先，没有希格斯机制，我们不得不放弃易受影响的高能预测或者非零的粒子质量。然而这两者对于正确的理论来说都至关重要。
- 其次，该问题的解决方法是存在于自然定律中的对称性，然而它可以在非零希格斯场的出现下自发破缺。真空的对称性破缺允许标准模型粒子获得非零质量。
- 再次，因为对称性自发破缺与能量（或者说长度）标度相关，所以破缺效应只与低能——基本粒子质量所对应的能标，以及更低的能标相关（或者说只与弱相互作用力尺度以及更大的尺度相关）。
- 最后，在所有这些能量与质量的考虑中，引力的效应都可以忽略，标准模型（包括粒子的质量）正确地描述了粒子物理学的实验。然而，对称性仍在自然定律中呈现，它允许合理的高能预测。
- 另外，作为一个副产品，希格斯机制还解释了光子的零质量，其原因是光子与遍及宇宙的希格斯场没有相互作用。

虽然理论已经取得了极大成功，但我们还没有实验证据来证明我们的想法。甚至连彼得·希格斯本人也承认这些检测的重要性。他在2007年提到：“该理论的数学结构已经非常令人满意”，但是，“假如没有实验的证实，那么它也仅仅是一个游戏而已。”^[1]我们希望彼得·希格斯的理论是正确的，所以我们期待接下来几年能有激动人心的发现。大型强子对撞机将把单个粒子或者多个粒子的证据展现给我们，并且在该想法最简单的应用中，其证据毋庸置疑将是希格斯玻色子。

实验证据的搜寻

“希格斯”既指一个人、一个机制，又指一个公认的粒子。希格斯玻色子是标准模型缺少的关键环节。^[66]这是希格斯机制有望遗留的痕迹，我们希望通过大型强子对撞机实验可以发现它。它的发现将肯定其理论，并告诉我们希格斯场的确遍布真空。我们有很好的理由相信，希格斯机制在宇宙中是有效的，因为如果没有它，没有人知道如何可以构造一个可以给出基本粒子质量的合理理论。我们也相信它的一些证据将很快在大型强子对撞机将要探索的能标上出现，而这个证据很可能就是希格斯玻色子。

作为希格斯机制一部分的希格斯场，以及作为真实粒子的希格斯玻色子之间的关系非常微妙，这与电磁场和光子之间的关系相似。比如你可以从手中靠近冰箱的一块磁铁感觉到经典电磁场的效应，哪怕没有真实的物理光子被制造出来。经典希格斯场（甚至在量子效应时也存在）遍布真空并且取非零值，可以使粒子获得质量。但是即使在空间中没有任何真实粒子时，那个非零的场也存在。

然而如果有一些东西给场“挠痒痒”，也即增加一点点能量，那么该能量在场中所产生的振动会导致粒子的产生。在电磁场中，可以产生的粒子是光子。而在希格斯场中，相应的粒子就是希格斯玻色子。希格斯场充满了真空，是电弱对称性破缺的原因。另一方面，希格斯玻色子从带有能量的希格斯场，例如大型强子对撞机中的希格斯场中产生出来。希格斯场存在的原因仅仅是基本粒子带有非零质量。大型强子对撞机中（或者任何其他来源）的希格斯玻色子的发现，将证实我们关于希格斯机制是粒子质量之源的信念。

有时媒体称希格斯玻色子为“上帝粒子”，这个名字令许多人觉得很好奇。记者热衷于此词的原因恰好是它能吸引人眼球，这也是物理学家利昂

• 莱德曼 (Leon Lederman) 率先使用它的原因。但这也只是一个名称而已。希格斯玻色子将会是一个卓越的发现，而不会成为一个空头名号。

虽然也许下面的内容听起来太过理论化，但是，认为存在一种执行希格斯玻色子功能的新粒子的逻辑也非常合理。除了上面提到的理论缘由，含有非零质量粒子的标准模型理论的自洽性也要求这种新粒子存在。假设只有非零质量粒子作为基本理论的一部分，而没有希格斯机制来解释这些质量，那么本章前文提到的高能粒子的相互作用就是不合理的——甚至给出了概率超过1的荒谬结果。当然，我们不相信这样的预测。没有附加结构的标准模型是不完整的。因此引入附加的粒子和相互作用是唯一的解决方法。

有着希格斯玻色子的理论优雅地避开了高能情形的问题。有了它的相互作用不仅改变了高能相互作用的预测，而且完全消除了高能情形的不良结果。当然这不是一种偶然，而是由希格斯机制保证的。虽然我们还不确切地知道我们所正确预言的真实希格斯机制在自然中的应用，但是物理学家相当地确信一种或者多种新粒子将在弱尺度^[67]上显现。

基于这些考虑，我们知道无论由谁来挽救该理论，是新粒子也好，新的相互作用也好，它们都不能太重或者发生在太高的能标上。在缺失了附加粒子的情况下，荒谬的预测可能已经在大约1TeV的能标上产生了。因此不仅希格斯玻色子（或者其他起着相同作用的物质）必须存在，它还必须足够轻，让大型强子对撞机可以找到。^[68]更准确地说，除非希格斯玻色子的质量小于800GeV，标准模型将给出不可能正确的高能相互作用的预测。

事实上，我们预期希格斯玻色子的质量远比那个值来得小。当前的理论倾向于认为希格斯玻色子的质量很小，绝大多数理论所给出的线索都指向同一个数值，其略微超过当前的质量阈值114GeV（20世纪9

0年代大型正负电子对撞机实验所发现的质量)。它曾是大型正负电子对撞机所能产生并探测的希格斯玻色子的质量上限，过去许多人认为他们可能正处在发现它的边缘。当今的多数物理学家预期希格斯玻色子质量仍然很接近那个数值，并且很可能不会超过140GeV。

关于所预期的希格斯玻色子质量比较轻的最强论据来源于实验数据，不仅包括希格斯玻色子本身的搜寻，而且包括其他标准模型物理量的测量。标准模型的预测与测量结果惊人地相符，甚至小小的修正也会影响这种一致性。希格斯玻色子在标准模型中的贡献是通过量子效应体现的。如果它太重，这些效应就会太大，从而会破坏理论预测与实验结果之间的一致性。

量子力学告诉我们，虚粒子对任何相互作用都有贡献。虚粒子从任何初始的状态中产生或者湮灭，并且对总的相互作用有贡献。因此，即使许多标准模型的过程根本没有涉及希格斯玻色子，相互交换的希格斯玻色子也影响所有标准模型的预测，例如Z规范玻色子衰变成夸克和轻子的速率以及W与Z规范玻色子的质量比。希格斯玻色子对精密电弱（precision electroweak）测试的虚效应的大小依赖于它的质量。而结果是，只有在希格斯玻色子的质量不是那么大时这些预测才适用。

第二个（并且更具推测性的）倾向于轻希格斯玻色子的原因与超对称理论相关。许多物理学家相信超对称在自然中存在，并且根据该理论，希格斯玻色子的质量与所测到的Z规范玻色子的质量相近，因此算是轻的。

因此，假设希格斯玻色子质量不是很大，那么你可以合理地追问：为什么我们发现了标准模型中的几乎所有粒子，却还没有发现希格斯玻色子？该问题的答案在于希格斯玻色子的性质。即使一种粒子很轻，我们也可能看不见它，除非对撞机可以制造并探测它，能否

这样做依赖于其性质。毕竟，一种根本不参与相互作用的粒子不管多么轻，都是永远也无法看到的。



我们已经了解了很多希格斯玻色子相互作用的形式，因为尽管希格斯场和希格斯玻色子与其他基本粒子有着不同的实质，但其相互作用与它们的类似。因此，通过其他基本粒子的质量大小，我们可以知道希格斯场与它们的相互作用。又因为希格斯机制是基本粒子质量产生的原因，我们知道希格斯场和最重的粒子的相互作用最强。并且希格斯玻色子从该场中产生，我们也知道了粒子的相互作用。故此，希格斯玻色子（和场一样）与标准模型中最重粒子的相互作用最强。

越强的相互作用发生在越重的粒子与希格斯玻色子之间，这表明如果你可以从越重的粒子出发，使它们对撞，那么你就越可能制造出希格斯玻色子。然而不巧的是，在希格斯玻色子的制造过程中，我们并不是从最重的粒子对撞开始的。考虑一下大型强子对撞机如何制造希格斯玻色子——或者任何可以产生它的粒子。大型强子对撞机对撞涉及较轻的粒子。它们较小的质量告诉我们希格斯玻色子的相互作用是多么微小，以至于如果没有其他粒子参与制造希格斯玻色子，那么其产生的概率将非常低，以至于任何我们迄今为止所造的探测器什么都探测不到。

幸运的是，量子力学提出了另一种可能。粒子对撞机中希格斯玻色子的产生，有一个涉及重的虚粒子的微妙过程。当轻的夸克互相碰

撞时，它们可以制造重的粒子，紧接着发射出一个希格斯玻色子。例如，轻夸克可以对撞产生一个虚W玻色子，该虚粒子接着发射出一个希格斯玻色子（这个产生模式见图16-1第一张图）。因为W玻色子远比质子中的上夸克或下夸克重得多，W玻色子与希格斯玻色子的作用相当大。当碰撞的质子足够多时，这种方式可以产生希格斯玻色子。

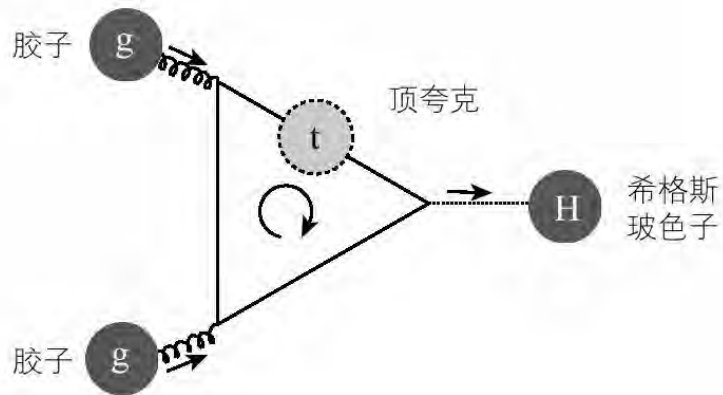
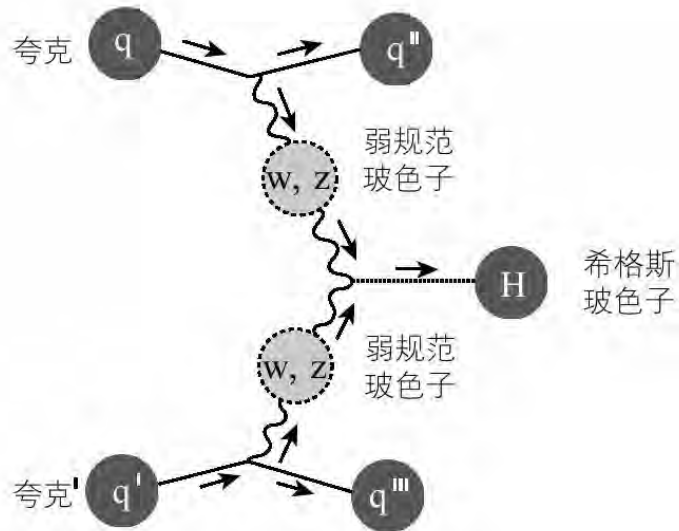
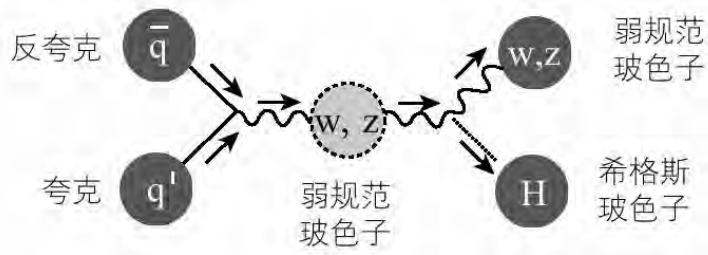


图16-1 三种产生希格斯玻色子的模式，从上到下依次是：希格斯辐射、W-Z融合以及胶子-胶子融合。

第二种希格斯玻色子产生的模式是：当夸克发射两个虚的弱规范玻色子时，它们接着碰撞出单一的希格斯玻色子（见图16-1第二张图）在这种情形下，当弱规范玻色子射出时，希格斯玻色子产生，同时两个与夸克相伴的喷射流散开射来。第二种与前一种模式的机制产生了希格斯玻色子以及其他粒子。在第一种情况下，希格斯玻色子产生于与规范玻色子相交的结点处。而后一种情况，这是大型强子对撞机更重要的一种模式，希格斯玻色子的产生伴随着喷射流。

希格斯玻色子也可以单独产生。这发生在胶子的碰撞中，其产生一个顶夸克和一个反顶夸克，它们湮灭产生一个希格斯玻色子（见图16-1第三张图）。事实上，顶夸克与其反夸克都是虚夸克，因此存在的时间不长，但是量子力学告诉我们这个过程相当常见，因为顶夸克与希格斯玻色子的相互作用很强。这种模式的产生机制与前两种不同，它没有留下任何除了希格斯玻色子以外别的迹象，而希格斯玻色子接下来也衰变了。

因此即使希格斯玻色子本身不必很重（它的质量可能与弱规范玻色子相当，并且比顶夸克的质量略小），重的粒子如规范玻色子或者顶夸克也可能参与到它的产生过程中。因此，高能碰撞（例如大型强子对撞机的那些）以及有着大量粒子碰撞比率的实验也一样，都可以帮助产生希格斯玻色子。

即使产生比率很高，另一个影响希格斯玻色子观测的挑战，是它的衰变方式。与其他重粒子一样，希格斯玻色子也是不稳定的。注意，发生衰变的是希格斯粒子而不是希格斯场。遍布真空的希格斯场给基本粒子提供质量，并且场不会消失。而希格斯玻色子是一种真实粒子，它是可以用来检验希格斯机制的一个实验结果。与其他粒子一

样，它可以在对撞机中产生。并且与其他不稳定粒子相似，它不能永远存在。因为本质上来说衰变来得太快，所以唯一发现它的方法是寻找它的衰变产物。希格斯玻色子可以衰变成与它有相互作用的那些粒子，也即，所有可以通过希格斯机制得到质量的粒子，而且是那些质量足够小到可以从它产生的粒子。当一个粒子和它的反粒子从希格斯玻色子的衰变中产生时，每个粒子的质量必须都小于希格斯玻色子质量的一半，以确保能量可以守恒。鉴于这个要求，希格斯粒子首先衰变成它能产生的最重的粒子。问题是，这意味着相当轻的希格斯玻色子极少衰变成容易鉴定和观测的粒子。

假如希格斯玻色子有违期待，质量不小，重于W玻色子的两倍（而是顶夸克质量的1/2）时，那么对它的搜寻将相当简单。实际上，希格斯玻色子就可以凭借其较大的质量总是衰变成W或者Z玻色子（衰变成一对W粒子，见图16-2）。实验物理学家知道如何检测剩下的W或者Z玻色子，因此发现希格斯玻色子并不难。

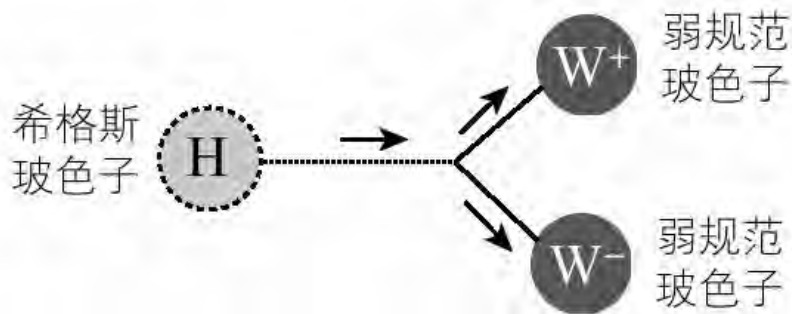


图16-2 较重的希格斯玻色子将衰变成W规范玻色子。

在这个相对较重的希格斯方案中，下一个最可能的衰变模式涉及底夸克与它的反粒子。然而衰变成底夸克及其反粒子的比率非常低，因为底夸克的质量比W规范玻色子小得多，所以它与希格斯玻色子的相互作用比W与希格斯玻色子的小得多。一个重得可以衰变成一对W玻色

子的希格斯玻色子衰变成底夸克与反粒子对的概率低于1%；衰变成更轻粒子的概率就更低。因此，如果希格斯玻色子相当重（比我们想象得重），它就会衰变成弱规范玻色子。而这些衰变将很容易被人们看到。

然而，正如前面暗示的，与实验数据相结合的标准模型的理论告诉我们：希格斯玻色子很可能轻到不能衰变成弱规范玻色子。在这种情况下，发生最多的衰变将会得到底夸克与它的反粒子（反底夸克）这一对粒子（见图16-3），而对这种衰变的观测是一个挑战。

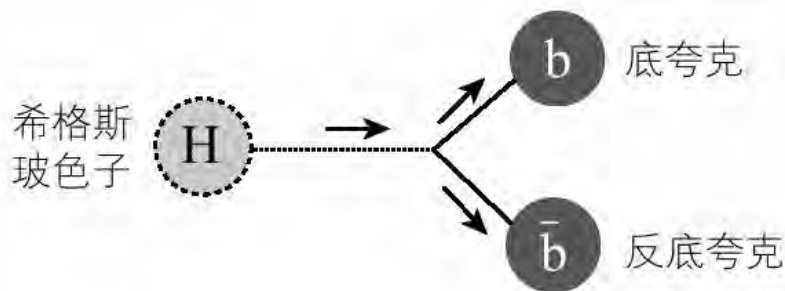


图16-3 较轻的希格斯玻色子将主要衰变成底夸克。

其中一个问题是，当质子对撞时，产生了大量参与强相互作用的夸克和胶子。这些粒子极易与假想的希格斯玻色子衰变出来的极少量底夸克混淆。除此之外，大型强子对撞机将产生如此之多的顶夸克，它们衰变成底夸克时也将隐藏希格斯的信号。理论物理学家和实验物理学家努力探索，寻找控制希格斯衰变的末态底夸克-反底夸克的方法。即使这样，尽管这种衰变的发生概率最大，尽管理论物理学家和实验物理学家可能寻找到有效利用它的方法，它也可能不是最有希望被用来发现大型强子对撞机中希格斯玻色子的模式。

因此，实验物理学家必须考察其他从希格斯玻色子衰变得来的末态，即使它们的发生概率较低。其中最有前途的候选者是 τ 子-反 τ 子

对，或者光子对。 τ 子是三种带电轻子中最重的一种，并且它也是希格斯衰变可以得到的粒子中，是除了底夸克之外最重的粒子。衰变成光子对的比率低得多（希格斯玻色子只能通过量子力学虚效应衰变成光子），但是光子的观测相当容易。虽然这种模式很难，但是实验物理学家可以如此之好地探测光子的性质，一旦有足够多的希格斯玻色子衰变，那么他们将确实能分辨出哪些是由希格斯玻色子衰变得来的。

事实上，因为希格斯粒子发现的重要性，紧凑 μ 子线圈实验和超环面仪器实验投入了复杂而细致的研究策略来发现光子与 τ 子，并且两个实验的探测器都以探测希格斯粒子为核心来构建。第13章介绍的电磁量能器被设计成可以精确地探测光子，而 μ 子探测器辅助记录更重的 τ 子的衰变。这两种模式的结合有望确立希格斯玻色子的存在性，而一旦有足够多的希格斯玻色子被检测到，我们就可以进而研究它的性质。

产生与衰变都给希格斯玻色子的发现设置了极大的挑战。但是理论物理学家与实验物理学家以及大型强子对撞机本身必须能够面对这些挑战。物理学家希望在接下来的几年，我们可以欢声庆祝希格斯玻色子的发现，并了解更多它的性质。

希格斯区，对撞出足够多的质子

我们期望能尽快找到希格斯玻色子。从原则上看，它能在只有大型强子对撞机最初所预期能标的一半的第一轮运转中产生，因为这一能标级别对于产生该粒子来说已经足够。然而，我们看到希格斯玻色子从质子对撞中产生的概率很小。这就是说，只有当足够多质子发生碰撞——光度很高时，希格斯玻色子才能产生。大型强子对撞机为了预备以目标能标运行将停机一年半，在此之前，计划发生的对撞数目

很可能太小以至于不能产生足够多可以被探测到的希格斯玻色子。但是大型强子对撞机当时计划于2012年全年运行，它可能可以探测到隐藏的希格斯玻色子。当然，当大型强子对撞机以满负荷运转时，光度将足够高，那时搜寻希格斯玻色子将会是它的一个主要目标。

假如我们已经很确信希格斯玻色子存在（并且如果探求太困难），这种搜寻看似有些多余。但它值得一试。首先，也是最重要的一个原因——理论预言只能带领我们到这里了。许多人只相信通过观测验证的科学成果，这是非常合情合理的。希格斯玻色子非常不同于其他任何已经被发现的粒子。它将是观测到的唯一的基本标量粒子（scalar）。与夸克和规范玻色子不同，标量粒子的自旋为零，也就是说，当你转动或者推动体系时，它保持不变。目前所观测到的自旋为零的粒子都是许多自旋非零粒子（诸如夸克）的束缚态。我们不确定希格斯标量粒子是否存在，除非它出现并在探测器中留下可见证据。

其次，即便我们发现了希格斯玻色子、确信了它的存在，我们也想知道它的性质。质量是所有未知量里面最重要的，但是了解它的衰变也同样重要。我们很清楚自己所期望的结果，然而我们需要测量这些数据是否与预测相符。所有这些将告诉我们现有的关于希格斯场的简单理论是否正确，或者它从属于一个更复杂的理论。因此通过测量希格斯玻色子的性质，我们可以获得超出标准模型的更基础理论的洞见。

举例来说，如果并非只有一种希格斯场，而有两种希格斯场承担了电弱对称性破缺的职责，那么我们所观测到的希格斯玻色子的相互作用将发生显著改变。在另一种模型中，希格斯玻色子的产生速率将不同于我们现在所期望的。并且如果存在其他载有标准模型相互作用的荷的粒子，它们将影响希格斯玻色子衰变成不同种可能的末态粒子的相对衰变率。

于是，我们有了第三种理由来研究希格斯玻色子——我们还不知道希格斯机制的真实蕴涵。事实上，最简单的模型（本章到现在一直所关注的）告诉我们实验将给出单一类型希格斯玻色子的信号。然而，哪怕我们相信希格斯机制是基本粒子质量的原因，我们也还不确定具体是哪些粒子集合参与了它的应用。目前大多数人仍认为我们可能发现一种较轻的希格斯玻色子，如果我们的确发现了，那么它将是一个重要想法的确证。

而其他涉及更复杂希格斯区的模型具有更丰富的预测结果。例如，超对称模型预言了更多的希格斯区的粒子。我们仍然期望发现希格斯玻色子，只不过它的相互作用不同于只含单一希格斯粒子的模型。最重要的是，希格斯区的其他粒子可以给出它们自己独有的有趣信号，假如它们轻到可以被制造出来的话。

一些模型甚至推测基本希格斯标量粒子并不存在，希格斯机制是通过一种更复杂的基本粒子的束缚态而非基本粒子得以实现的，这类似于超导材料中给予光子质量的电子对。如果真是这样，那么希格斯粒子束缚态就会出人意料地重，并且具有其他有别于基本粒子的相互作用。这些模型当前不被人看好，因为它们难以与所有实验观测相符。不管怎样，大型强子对撞机的实验物理学家都会努力寻找以便确信这件事情。

为什么引力如此微弱

希格斯玻色子只是大型强子对撞机所能发现物质的冰山一隅。希格斯玻色子与其他发现一样有趣，它并非大型强子对撞机实验的唯一搜寻目标。可能研究弱尺度最主要的原因是没有人认为希格斯玻色子是遗留的唯一问题。物理学家期望希格斯玻色子只是一个有着更多内

涵的模型中的一个元素，而该模型可能告诉我们关于物质乃至时空本身更多的性质。

等级问题

为什么决定基本粒子质量的弱能量标度比与引力相关的普朗克能标小16个数量级？这就是粒子物理学的等级问题。

因为只存在希格斯玻色子而没有其他新元素导致了另一个巨大的谜团，即所谓的等级问题。等级问题关心的是，为什么粒子的质量（特别是希格斯粒子的质量）是其实际所具有的那个数值。决定基本粒子质量的弱能量标度是另一个质量标度——决定引力强度的普朗克质量的一亿亿分之一（见图16-4）。

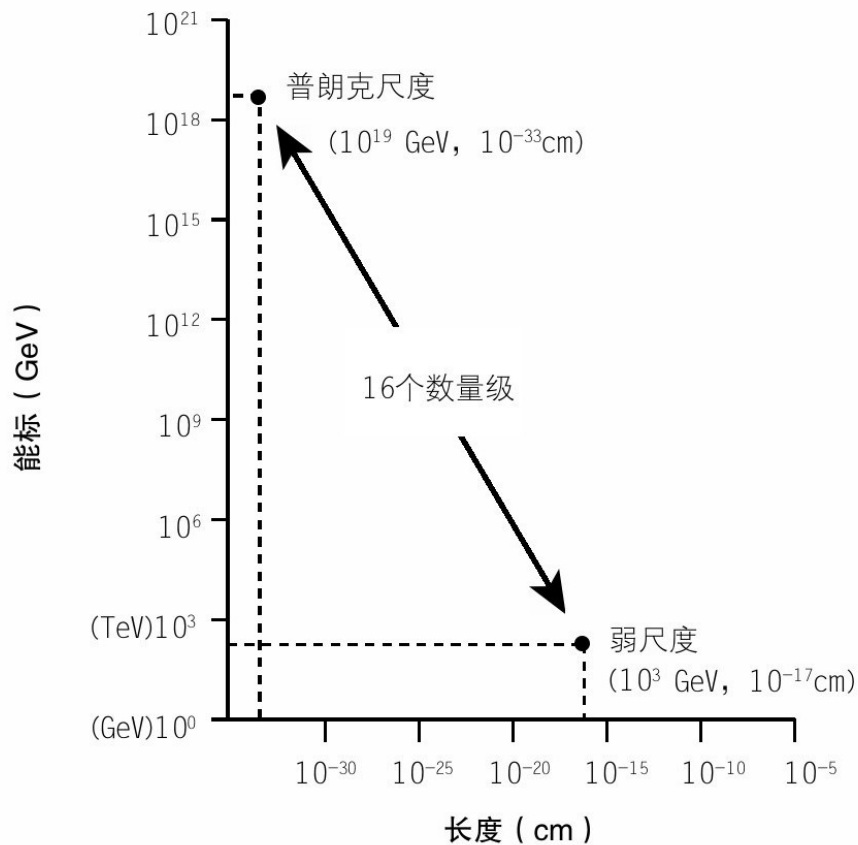


图16-4 粒子物理学等级问题：弱能量标度比与引力相关的普朗克能标小16个数量级。普朗克长度相应地比大型强子对撞机可以探测的尺度小得多。

普朗克质量与弱作用的质量标度相差如此巨大，这是引力如此微弱的原因。引力相互作用依赖于普朗克质量的倒数。如果普朗克质量如我们所知的那样巨大，引力必然极其微弱。

事实上，引力是迄今为止人们所知道的最弱的相互作用力。引力之所以看起来没有那么微弱，是因为整个地球的质量都在吸引着你。如果你换位思考，两个电子之间的引力，你会发现它们之间的电磁力比引力大了43个数量级。也就是说，电磁力的大小胜过引力一千亿亿亿倍。引力对于基本粒子的效应完全可以忽略不计。等级问题考虑的是：为什么引力比其他我们所知的基本相互作用力微弱了如此之多？

粒子物理学家不喜欢解释大数字，比如普朗克质量与弱作用质量之比。但是该问题比仅从审美的角度反对神秘的大数字来得严重。根据量子场论（将量子力学与狭义相对论结合起来的理论），这里不应该有任何差别。解决等级问题的紧迫性至少对所有理论物理学家来说都是心照不宣的。而量子场论却表明弱质量与普朗克质量应该相差无几。

在量子场论中，普朗克能标具有重要意义。不仅因为在该能标上引力变得极强，而且因为在此能标上引力与量子力学都是至关重要的，而我们已知的物理定律都失效了。然而，在低能标时，我们知道如何使用量子场论进行粒子物理学计算，而量子场论也做了很多成功的预测，让物理学家们相信它是正确的。实际上，所有科学领域中测量值与预言值符合得最好的结果，也是来自量子场论。这样的一致性绝非出于偶然。

然而，当我们将相同原理应用于结合量子力学的虚粒子贡献，给出希格斯质量时，结果却是出人意料地混乱。理论中任何粒子的虚效应看似可以给出希格斯粒子一个与普朗克质量相当的质量。中介粒子可以是质量很大的粒子，例如具有GUT能标质量的粒子（见图16-5左图）；或者是普通的标准模型粒子，例如顶夸克（见图16-5右图）。任何一种方式的虚效应修正都可以使希格斯玻色子质量非常巨大。这里的问题是：中间交换的虚粒子允许具有的能量可以和普朗克质量一样巨大。如果这个说法正确的话，希格斯质量也可以同样极其巨大。在这种情形下，弱相互作用所伴随的对称性自发破缺的能标也将是普朗克能标，也即比现在实验能达到的能标高16个数量级（一亿亿倍）。

等级问题对只含一个希格斯玻色子的标准模型来说是极其重要的问题。从技术上来讲，这里面的确存在一个漏洞。在没有虚效应时希格斯质量可以极其巨大，其数值刚好与由于虚效应带来的贡献相抵消，从而得到我们所需要的能量水平上的精确值。这里的问题在于：虽然原则上可能，这将意味着有16个数位的数字必须严格相消^[69]——那将是多么大的一种偶然啊。



图16-5 重粒子对希格斯玻色子质量的量子修正。例如大统一理论（GUT）尺度下的某个重粒子与它的反粒子（左图），以及一个虚的顶夸克与它的反粒子（右图）。

因此物理学家都不相信这种胡言乱语或者精细调节。我们认为不同质量标度之间差别的等级问题，暗示了一个更大更好的基本理论。看起来没有一个简单的模型可以完全给出该问题的答案。唯一有希望

的方案是将标准模型扩展，使其包含一些显著的性质。在落实希格斯机制的情况下，等级问题的解决是大型强子对撞机主要的搜索目标，也是接下来一章的主题。



2010年1月，各路英雄齐聚于南加州的一个会议，商讨大型强子对撞机纪元中暗物质的搜寻事宜与粒子物理学的新进展。会议的组织者、紧凑 μ 子线圈实验的成员之一、加州理工学院物理系教授——玛利亚·斯皮罗普鲁（Maria Spiropulu）邀请我做开场报告，并介绍大型强子对撞机的主要任务与近期将要研究的物理目标纲要。

玛利亚希望开一个互动性较强的会议，因此她建议我们三个开场报告的报告人来一次“决斗”（duel），假设“决斗”这个词可以用于三个人而不至于使人困惑的话。而受邀的听众则形成了更大的挑战，因为他们的人员组成从相同领域的专家到加州对技术领域感兴趣的观察家。玛利亚让我深刻挖掘、仔细寻找并且忽略当前理论与实验的一些特征。而其中一位与会者——丹尼·希利斯（Danny Hillis），他是Applied Minds公司一位优秀的非物理学人士，他建议我把一切讲得尽可能地基础，让非专业人士也可以理解。

在面对如此多的矛盾、分歧以及众口难调的意见时，我做了一件任何很多人都会做的事情——拖延。我在网上搜索的内容构成了我的第一张幻灯片（见图17-1），结果这个关于“小错误与一切”（typo and all）的主题成为丹尼斯·奥弗拜（Dennis Overbye）发表在《纽约时报》上文章的主题。

战场：决斗1

丽莎·兰道尔



图17-1 几种候选模型，这是我在会上展示的一张幻灯片。

这些主题指的是我和后续的演讲者计划要涉及的话题。但是伴随着决斗的两只猫的入场，我利用音效制造了一种幽默（这里我却不能重现该场面），其寓意乃是反映两个模型所具有的既令人热血沸腾又让人捉摸不定的性质。与会的每个人无论多么强烈地相信自己正在研究的内容，都清楚地知道实验数据马上就会浮出水面，而数据将是谁能笑到最后（或者得到诺贝尔奖）的最终裁判。

大型强子对撞机呈现给了我们唯一一个创造新理解与新知识的机会。粒子物理学家希望能尽快回答我们已经思考了很久的艰深问题：

- 为什么粒子具有实验观测到的如此质量？
- 什么是暗物质？
- 额外维度能解决等级问题吗？
- 时空有额外的对称性吗？
- 现今的工作有完全无法预见的东西吗？

被提议的解决方案包括一些模型，例如超对称、技术色、额外维度。当然最终答案也可能完全不同于任何期望的东西，但是模型可以为我们的搜寻提供具体目标。本章接下来展示一些候选模型，并讨论它们对于等级问题的回答，以及介绍它们为大型强子对撞机将进行的探测类型所做的添砖加瓦的工作。为这些或者其他模型所进行的搜寻工作即将展开，无论哪一种结果将构成自然的真实理论，这些工作都将为我们提供宝贵的见解。

模型一：超对称

我们将从一种离奇的对称性——超对称以及与它相结合模型出发。如果你在理论粒子物理学中做一个问卷调查，那么一大部分人可能会说：超对称可以解决等级问题。更进一步，如果你询问实验物理学家他们期望寻找什么新物质，那么也有一大部分人会提议超对称。

从20世纪70年代开始，许多物理学家已经认为具有超对称的理论是非常美妙、神奇，所以他们相信它也一定存在于自然当中。他们更进一步得出：超对称模型中的三种相互作用在高能标时具有相同的强

度。因为在标准模型中这些强度只是接近，所以超对称模型是标准模型的一个发展，使得统一理论成为可能。许多理论物理学家也发现超对称是等级问题最有说服力的解决方法，哪怕将所有细节与我们所知道的事情联系起来很困难。

超对称模型假定标准模型中的任何基本粒子，例如电子、夸克等等，都有自己的伙伴——有着相似的相互作用但不同的量子力学性质的粒子。如果世界果真是超对称的，那么它一定存在许多未知的粒子——每一种已知粒子的超对称伙伴都在等待我们去发现（见图17-2）。

超对称模型可以帮助解决等级问题，而且一旦它这么做了，它必定做得非常地超凡脱俗。在一个严格超对称的模型中，来自粒子及其超对称伙伴的虚效应严格相消。也就是说，如果你把来自标准模型的每种粒子的所有量子力学贡献全部加合起来，并统计它们对希格斯玻色子质量的效应，那么你将发现它们加起来全等于零。在超对称模型中，希格斯玻色子将是零质量或者质量很小，哪怕出现了量子力学虚效应的修正。因此在一个真实的超对称理论中，粒子与其超对称伙伴的贡献将会严格相消（见图17-3）。

可能这听起来很神奇，但由于超对称是一种非常特殊的对称性，所以这是有保证的。超对称是一种时空对称性，类似于你所熟悉的转动与平移对称性，但它将时空拓展到了量子领域。

标准模型粒子				超对称伙伴			
夸克		轻子		超夸克		超轻子	
u 上夸克	d 下夸克	ν_1 最轻的中微子	e 电子	\tilde{u} 超上夸克	\tilde{d} 超下夸克	$\tilde{\nu}_1$ 型超中微子	\tilde{e} 超电子
c 粲夸克	s 奇异夸克	ν_2 中间质量的中微子	μ 子	\tilde{c} 超粲夸克	\tilde{s} 超奇异夸克	$\tilde{\nu}_2$ 型超中微子	$\tilde{\mu}$ 超 μ 子
t 顶夸克	b 底夸克	ν_3 最重的中微子	τ 子	\tilde{t} 超顶夸克	\tilde{b} 超底夸克	$\tilde{\nu}_3$ 最重的中微子	$\tilde{\tau}$ 超 τ 子
电弱规范玻色子				W微子, 希格斯微子, 荷微子			
W^+ W^- Z γ				χ_1^+ χ_1^- χ_2^+ χ_2^-			
希格斯玻色子				希格斯微子, 电中性微子			
H	A	H^0	H^\pm	$\tilde{\chi}_1^0$	$\tilde{\chi}_2^0$	$\tilde{\chi}_3^0$	$\tilde{\chi}_4^0$
胶子				胶微子			
g				\tilde{g}			

图17-2 在超对称理论中，每一种标准模型的粒子都有一个超对称伙伴。该理论的希格斯区超出了标准模型。



图17-3 在超对称模型中，来自虚的超对称粒子的贡献与来自标准模型粒子的贡献，对于希格斯玻色子的计算严格相消。例如，图中粒子的贡献之和为零。

量子力学将物质分成两个完全不同的类别——玻色子与费米子。费米子是具有半整数自旋的粒子，而自旋是一个量子数，本质上，它告诉我们粒子在一种类似于旋转的操作中的行为如何。半整数意味着

1/2、3/2、5/2，等等。标准模型的夸克与轻子是费米子的一些例子，它们都有1/2自旋。玻色子的例子有：传递相互作用的规范玻色子或者正在找寻的希格斯玻色子，它们具有整数自旋，也就是整数0、1、2，等等。

费米子与玻色子不仅自旋有别，而且当两个或者更多个同类型的粒子在一起时，它们的行为也完全不同。例如，有着相同属性的全同费米子永远不能待在同一个位置上。这是泡利不相容原理（Pauli exclusion principle，以物理学家沃尔夫冈·泡利命名）告诉我们的。费米子是元素周期表形成的基本原因这一事实告诉我们，多个电子除非具有不同的量子数，必须占有围绕核子运转的不同轨道。这也是我的椅子没有落向地球中心的原因，因为椅子中的费米子不能同时与地心物质居于同一个位置上。

泡利不相容原理

微观粒子运动的基本规律之一。它认为，两个电子或两个任何其他类型的费米子，都不可能占据完全相同的量子态。

另一方面，玻色子的行为则完全相反。事实上，它们非常可能居于同一个位置上。玻色子可以堆积在彼此上面——有点像鳄鱼趴在彼此上面，这也是为什么诸如玻色凝聚（Bose condensate）一类现象存在的原因，它需要许多粒子堆积在量子力学的同一状态上。激光也依赖于作为玻色子的光子汇聚到一起；强光束是通过许多全同光子一起发射而产生的。

在超对称模型中，我们认为完全不同的玻色子和费米子可以相互转换，其最后的结果与转换之前的理论给出的结果相同。每一种粒子都有一种处于相反量子力学类型，但是带有完全相同质量与荷的伙伴粒子。新粒子的命名方法有点让人忍俊不禁，每当我在公众场合谈到这个话题时，总是会引来观众咯咯的笑声。例如，作为费米子的电

子与作为玻色子的超电子（selectron）互为超对称伙伴。一个玻色性的光子与一个费米性的光微子（photino）是一对，W规范玻色子与W微子（Wino）也是一对。^[70]新粒子与其在标准模型中的粒子伙伴有相关的相互作用，但是它们的量子力学性质（玻色性、费米性）却相反。

在超对称理论中，每种玻色子的性质都与它的超对称伙伴费米子的性质相联系，反之亦然。因为每个粒子都有伙伴并且相互作用是平行的，该理论允许这种奇异的交换费米子和玻色子的对称性。

虚效应对希格斯粒子的质量产生了神奇的相消，其原因是，超对称将任何玻色子关联到与之相伴的费米子上。特别是，超对称将希格斯玻色子与希格斯费米子——希格斯微子（Higgsino）结成伙伴。即使量子力学的贡献强烈地影响了玻色子的质量，费米子的质量却从来不比其经典质量大多少，也即哪怕添加了量子力学修正，它也跟你在考虑量子修正前所使用的质量一样。

这里的逻辑很微妙，但没有产生大的修正的原因是，费米子的质量涉及左手和右手粒子，质量项使得它们之间可以相互转换。如果没有经典质量项，并且在没有加入量子力学虚效应时它们不能彼此转换，那么即便考虑了量子力学效应它们也不会彼此转换。如果一个费米子没有初始质量（经典质量），那么它在加入量子力学贡献后仍然没有质量。

这个论断对玻色子并不适用。例如，希格斯玻色子的自旋为零，因此我们无法说希格斯玻色子的自旋是左手的还是右手的。但是超对称告诉我们：玻色子质量与其超对称伙伴（费米子）一样。因此，如果希格斯微子的质量为零（或者很小），那么其伙伴希格斯玻色子的质量也必须一样，哪怕量子力学的修正被考虑了进来。

我们不知道这对于等级的稳定性来说是不是一个相当优美的解释，以及是否给出了希格斯质量的巨大修正相抵消的解释。但是如果超对称确实回答了等级问题，那么我们就知道我们可以从大型强子对撞机中期盼些什么。我们就会知道存在什么样的新粒子，因为每一种已知粒子都应该有一个伙伴。最重要的是，我们可以估计新的超对称粒子的质量。

当然，前提是超对称在自然中是严格保持的，那么我们就知道所有超对称伙伴的质量，因为它们与其伙伴粒子有着相同的质量。然而，这些超对称伙伴还从未被观测到，这告诉我们即使超对称可以被应用在自然中，它也不是严格的。假设它是严格的，那么我们应该早就发现超电子、超夸克以及所有其他超对称理论预言的超对称粒子了。

因此，超对称一定是破缺的，意即超对称理论所预言的关系（虽然可能是近似的）是不严格的。在一个超对称破缺的理论中，每种粒子仍具有超对称伙伴，但是这些超对称伙伴不具有与标准模型中的粒子一样的质量。

然而，如果超对称破缺得太多，那么它对等级问题就毫无助益，因为世界就会看似完全没有超对称这种对称性。因此超对称必须破缺得刚刚好：我们还没能观测到超对称的迹象，但是希格斯质量又受其保护，而不会因为巨大的量子力学贡献而给出巨大的质量。

这告诉我们超对称粒子应该具有弱尺度的质量。比此更轻的话，它们就应该已经被观测到了；比此更重的话，我们就应该指望希格斯质量也更大。我们不知道准确的数值应该是多少，因为我们只知道希格斯质量的近似值。但是如果质量过重，那么等级问题仍将存在。

我们的结论是：如果超对称在自然中存在并且解释了等级问题，那么许多质量介于几百GeV到几个TeV之间的新粒子应该存在。这恰恰是大型强子对撞机锁定的质量区域的搜索目标。有着14TeV能量的大型强子对撞机应该能够制造这些粒子，哪怕只有一部分质子的能量可以转变为夸克和胶子进行对撞而产生新粒子。

大型强子对撞机所能制造的最简单的粒子应该是在强相互作用力中载荷的超对称粒子。当质子（或者更准确地说，它们中的夸克和胶子）对撞时，这些粒子将大量产生。当对撞发生时，参与强相互作用力的新的超对称粒子可以率先产生。如果是这种情况，那么它们将在探测器中留下明显的、标志性的证据。

这些信号（粒子留下的实验证据）依赖于当它们产生后所发生的事情。绝大多数超对称粒子都会衰变。这是因为一般来说，存在一些较轻的粒子（如标准模型粒子），其荷总和等于一个较重的超对称粒子的荷。那么，在荷守恒的前提下，较重的超对称粒子会衰变成这些较轻标准模型粒子的组合。实验物理学家进而可以探测这些标准模型粒子。

这可能仍不足以确定超对称。在几乎所有超对称模型中，超对称粒子不单单衰变成标准模型粒子。一个较轻的超对称粒子会存留到衰变结束后，因为超对称粒子总是成对出现或消失。因此，一个超对称粒子必须存留到某个超对称粒子衰变结束之后——一个超对称粒子不可能变成零个超对称粒子。因此，最轻的粒子必须是稳定的。这个最轻的粒子（已经不能再衰变成其他粒子）被物理学家称为最轻的超对称粒子（lightest supersymmetric particle, LSP）。

从实验的优越角度来看，超对称粒子的衰变是显著的，因为即便当衰变完成时最轻的中性超对称粒子也还存在。宇宙学约束告诉我们最轻的超对称粒子不带荷，因此它不会与任何探测器的元素相互作用。

用。这意味着一旦超对称粒子被制造出来并发生衰变，动量和能量将出现缺少的现象。最轻的超对称粒子将从探测器中消失并带走动量和能量，而不会被记录下来，因此留下能量缺失的信号。能量缺失并不是超对称独有的标志，但因为我们已经知道许多超对称质量谱，所以我们知道我们能或者不能观测到的东西。

例如，假定某个超夸克（夸克的超对称伙伴）产生了。它可以衰变成何种粒子依赖于何种粒子更轻。一种可能的衰变模型是：超夸克总是衰变成夸克和最轻的超对称粒子（见图17-4）。衰变总是很快发生，探测器记录的仅仅是衰变产物。如果这样一个超夸克发生了衰变，那么探测器将在追踪器中记录到夸克的轨迹，并且在强子量能器中测到由于强相互作用力粒子截停产生的能量沉积。但是实验物理学家还将测得缺少的能量和动量，他们应该能分辨出哪些动量缺失了，就像他们在测量实验产生的中微子一样。他们也会测量出垂直于粒子束方向上的动量，并发现其加合起来不等于零。实验物理学家所面临的一个最大的挑战将是：如何毫不含糊地确认这些缺失的能量。毕竟，任何没有探测到的东西看似丢失了。如果实验过程中出现了一些错误或者测量错误，即便是很小一部分能量没有被探测到，那么缺失的动量也可能加合起来伪装出一个逃离超对称粒子的信号，哪怕实际上实验并没有产生任何奇异的物质。

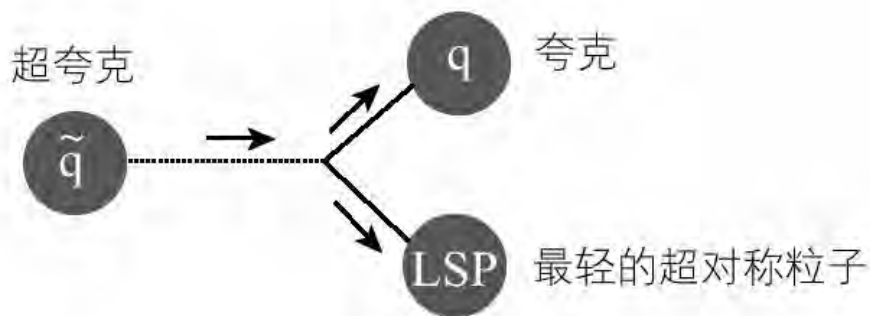


图17-4 超夸克可以衰变成一个夸克和一个最轻的超对称粒子。

事实上，超夸克不会单独产生，而是产生于与另一个强相互作用力物质（如另一个超夸克或者反超夸克）的相交处，实验物理学家将测量到至少两个喷射流（例如图17-5）。如果两个超夸克在质子对撞中产生，那么它们可以给出探测器能测量到的两个夸克。总的能量动量缺失将逃离而不被探测到，但这个缺失将被记录下来，并为新粒子提供证据。

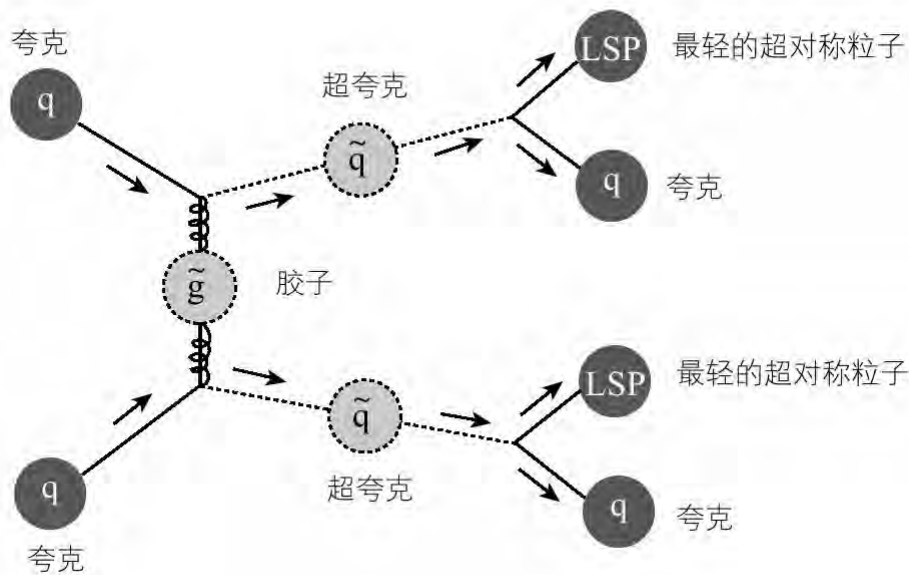


图17-5 大型强子对撞机可能产生两个超夸克，但是它们衰变成夸克和最轻的超对称粒子，留下能量缺失的信号。

大型强子对撞机日程安排上的各种延迟，却产生了一个主要优势：实验物理学家可以有足够的时间，来全面了解这些探测器。他们校准这些仪器，使得在仪器开始运行的那一天，测量可以足够精确，因此缺失的能量也可以准确测量。另一方面，理论物理学家也可以有时间思考更多关于超对称和其他模型的搜寻策略。例如，在与来自威廉姆斯学院（Williams College）的理论物理学家戴夫·塔克-史密斯（Dave Tucker-Smith）的合作中，我们发现一个不同于（但是相关的）此前所介绍的超夸克衰变的搜索方法。这个方法依赖于仅仅测量

从事件中产生的夸克的动量和能量，而不需要具体测量那些可能捉摸不定的缺失的动量。而近来大型强子对撞机令人欣喜的一个重要的发现是：许多紧凑 μ 子线圈的实验物理学家立即在此想法下操作实验，发现它不仅适用，而且他们还在几个月内推广和改进了该想法。现在它已成为标准的超对称搜索策略，而且紧凑 μ 子线圈首批超对称搜寻就是应用了我们所提议的技术。^[1]

沿着这条道路走下去，即使超对称被发现了，实验物理学家也不会停滞不前。他们将致力于发现全部的超对称粒子谱，而理论物理学家则将研究这些结果的意义。许多有趣的理论隐藏在超对称理论以及可以自发破坏该对称性的粒子当中。如果超对称与等级问题相联系，那么我们就知道应该存在哪些超对称粒子。但是我们还不知道它们的具体质量，以及这些质量产生的原因。

不同的质量谱会使大型强子对撞机可以看到的粒子产生巨大差异。粒子只能衰变成其他更轻的粒子。衰变的链条——超对称粒子可能发生的一系列衰变，依赖于质量，也即哪些粒子更重、哪些更轻。各个衰变过程的速率也依赖于粒子的质量。一般较重的粒子衰变得更快一些，并且它们通常也更难产生，因为只有能量很高的对撞才能创造它们。将各种结果综合起来，我们可以得到一些比标准模型更基本以及处于更高能标的理论的启示。任何关于新物理理论的分析都可能被我们发现，这是毋庸置疑的。

无论如何，我们需要牢记：尽管超对称对物理学家来说是一个热门方向，但这里仍存在着一些关于它是否真能应用于等级问题乃至真实世界的顾虑，有如下几个方面。

首先，这可能也是最令人担忧的，我们到现在还没有任何实验证据。如果超对称存在，那么唯一能解释我们至今还没有观测到任何迹象的原因是，超对称伙伴的质量太大。但是，等级问题的自然

解答需要超对称伙伴具有较轻的、合理的质量。超对称伙伴的质量越大，超对称作为等级问题的解就越不适宜。该解答的荒谬程度由希格斯玻色子的质量与超对称破缺的能标之比决定。该数值越大，这个理论就越需要精细调节。

目前还没有观测到希格斯玻色子的这个事实加深了问题的难度。在超对称模型中，为了使希格斯玻色子重得连实验也探测不到，唯一方法来自于重的超对称伙伴的巨大量子力学修正。但是同样地，这些极大的质量使得等级变得更不自然，哪怕是在超对称理论中也是如此。

超对称的另一大挑战是，能否发现一个完整、自洽而又包含超对称破缺的模型。超对称是一种非常具体的对称性，它将许多相互作用联系起来，并且排除了许多被量子力学认可的相互作用。一旦超对称破缺了，“无秩序原理”（anarchic principle）将占主导地位，即任何可能发生的都将发生。大多数模型会预测出在自然中从未见过的衰变或者极少见过的衰变来与预言结果符合。因此当超对称破缺时，由于量子力学的限制，很多非常丑陋的模型都将出现。

物理学家也很可能错过正确答案。我们当然不能肯定地说好的模型不存在或者一种小的精细调节不会发生。当然，如果超对称是等级问题的正解，那么我们将很快于大型强子对撞机中发现它的证据，因此这值得一试。超对称的发现意味着，这个奇异的新时空对称性不仅适用于演算纸上的理论推导，也适用于现实世界。然而，在缺乏证据的情况下，尝试一些其他理论也是值得的。第一个我们要考虑的是技术色理论。

模型二：技术色理论

回溯到20世纪70年代，物理学家先考虑了另一种关于等级问题的可能解——技术色理论。此理论构建的模型中粒子通过一种新型力强烈地相互作用，人们戏谑地称之为技术色相互作用力（technicolor force）。这项提案是，技术色的作用与强相互作用力之间相互作用（物理学家也称之为色相互作用）类似，但是它在弱能标而非质子质量标度上将粒子束缚起来。

如果技术色的确是等级问题的答案，那么大型强子对撞机不会只制造单一的基本希格斯粒子。相反，它会制造一个束缚态（类似于强子的物质），而该束缚态起到希格斯粒子的作用。支持技术色的实验证据将是大量束缚态粒子以及出现在弱尺度能标及以上的许多强烈的作用。它非常像我们所熟悉的质子，但是却出现在高得多的能标上。

目前还没有任何证据，这为技术色模型添加了很多限制。如果该理论真的是等级问题的解，那么我们预期的证据已经出现了，当然我们也可能错过了一些微妙的事物。

最重要的是，技术色模型的建立比超对称还难。找到一个与我们所有的观测相符的模型已经极其困难，而目前还没有发现一个完全合适的模型。

无论如何，实验物理学家一直保持着开放的头脑并尽力在寻找技术色和其他新型强作用的证据，但是它们存在的可能性不是很高。然而，假如技术色的确是世界的基本理论，或许以后当我再敲进“technicolor”这个词时，可能Word程序会停止自动纠错以及停止自动将首字母“t”改成大写字母“T”了。

模型三：额外维度

超对称或者技术色都不能为等级问题提供完美的解答。超对称理论还没有准备好容纳从实验中体现出来的超对称破缺，而由技术色理论的推导来预测正确的夸克和轻子质量则更困难。因此物理学家决定着眼于更远的地方并考虑一些表观上更具猜测性的其他想法。别忘了，即使一个想法起初看似丑陋或者不明显，在完全理解它的内涵以后，我们也可能觉得它是最优美的，而且更重要的是，它是正确的。

20世纪90年代，物理学家对于弦理论及其组成的理解越来越好，这导出了解决等级问题的新方案。这些想法是由弦理论的元素推动的（虽然不见得可以从它非常受限的结构直接推导出来），并且涉及空间的额外维度。如果额外维度存在（我们有理由相信这是可能的），那么它们可能是解决等级问题的关键。如果确实如此，那么它们将在大型强子对撞机中产生其存在的实验证据。

更多的空间维度是一种奇异的观点。如果宇宙有这些维度，那么空间将非常不同于我们每天生活中所观察到的。除了三个方向——左右、上下、前后，或者另一种描述方法——经度、纬度、高度，空间还可能在没有人可以看到的方向上延展。

显然，因为我们看不见它们，所以这些新的空间维度一定是隐藏起来的。就像物理学家奥斯卡·克莱因（Oskar Klein）于1926年所提议的，可能因为它们太小，以至于不能影响任何我们可以看到的東西。这种想法是说，我们受制于有限的分辨率，这些维度可能小得无法让人察觉。我们可能看不到某个卷曲的维度，我们不能在该维度上穿行——就像走钢丝的人会认为他的道路是一维的（见图17-6）。^[71]

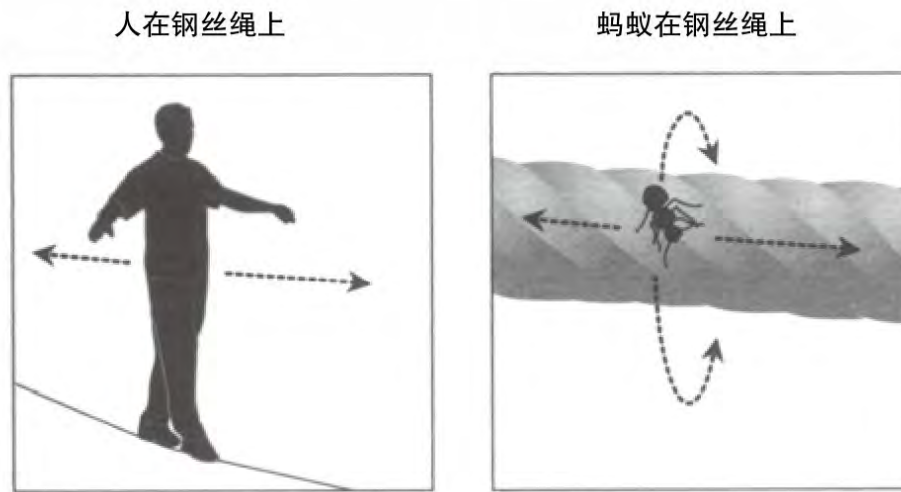


图17-6 一个人与一只小蚂蚁在一根钢丝绳上的体验是不同的。对人来说，绳子看起来是一维的，而对蚂蚁来说则是二维的。

另一种可能性使维度可以隐藏起来，因为时空是卷曲的或者弯曲的，就像爱因斯坦指出的在出现能量之后所发生的现象。如果弯曲得足够巨大，那么额外维度的效应是不明显的，正如拉曼·桑卓姆与我在1999年所提议的那样。^[2]我在《弯曲的旅行》中详细分析了这一点，这意味着弯曲几何可能为隐藏维度提供了一种方法。

假如我们永远无法看到额外维度，为什么我们会认为它们可以存在呢？物理学史上有过许多发现了我们看不见之物的例子——没有人可以“看见”原子，没有人可以“看见”夸克，而我们现在有它们存在的实验确证。

没有哪一个物理定律告诉我们空间只能有三维。爱因斯坦的广义相对论在任何维度都成立。事实上，在爱因斯坦完成他的引力理论之后不久，西奥多·卡鲁扎（Theodor Kaluza）推广了爱因斯坦的想法，提议存在第四个空间维度。5年以后，奥斯卡·克莱因提出了这个推广的维度应该如何卷曲起来，使之有别于另三个我们熟知的维度。

作为一个率先将量子力学与广义相对论结合起来的理论，弦理论是物理学家现在醉心于额外维度的一个原因。弦理论没有明显导出我们所熟悉的引力理论，该理论必须包含空间的额外维度。

时常有人问我存在于宇宙中的维度是多少维。我们不知道。弦理论建议有6个或7个额外维度，但是模型的创建者的眼界开阔。可以想象，不同的弦理论模型给出的可能性不同。无论如何，模型创建者关心的仅仅是那些足够弯曲或者足够影响物理预测的维度。对于那些与粒子物理学现象相关的维度，比它们更小的维度也可能存在，但是我们忽略这些如此细小的东西。我们再次采用有效理论的方法，略去对测量影响太小或者不可见的东西。

弦理论也引进了其他的元素——主要是膜，使得宇宙的几何更加丰富，假如宇宙果真包含额外维度的话。在20世纪90年代，弦理论学家约瑟夫·波尔钦斯基（Joseph Polchinski）所构造的弦理论不只是占据一个空间维度的被称为弦的物质的理论。与其他许多人一起，他们证明了：更高维度的、被称为膜的物质对该理论来说也是至关重要的。

“Brane”（膜）这个词来自于“membrane”（薄膜）。与存在于三维空间中的二维曲面的薄膜类似，膜是存在于更高维空间中的较低维度的曲面。这些膜可以将粒子或者力束缚在它们上面，使得粒子或者力不能在全部的高维度空间中传播。在高维空间中的膜有点像浴室中的浴帘，它是三维房间中的一个二维曲面（见图17-7）。水滴只能沿



着浴帘的二维表面运动，很像粒子和力被束缚在二维膜的“表面”一样。

膜

膜是存在于更高维空间中的较低维度的曲面。这些膜可以将粒子或者力束缚在它们上面，使得粒子或者力不能在全部的高维度空间中传播。

浴帘“膜”

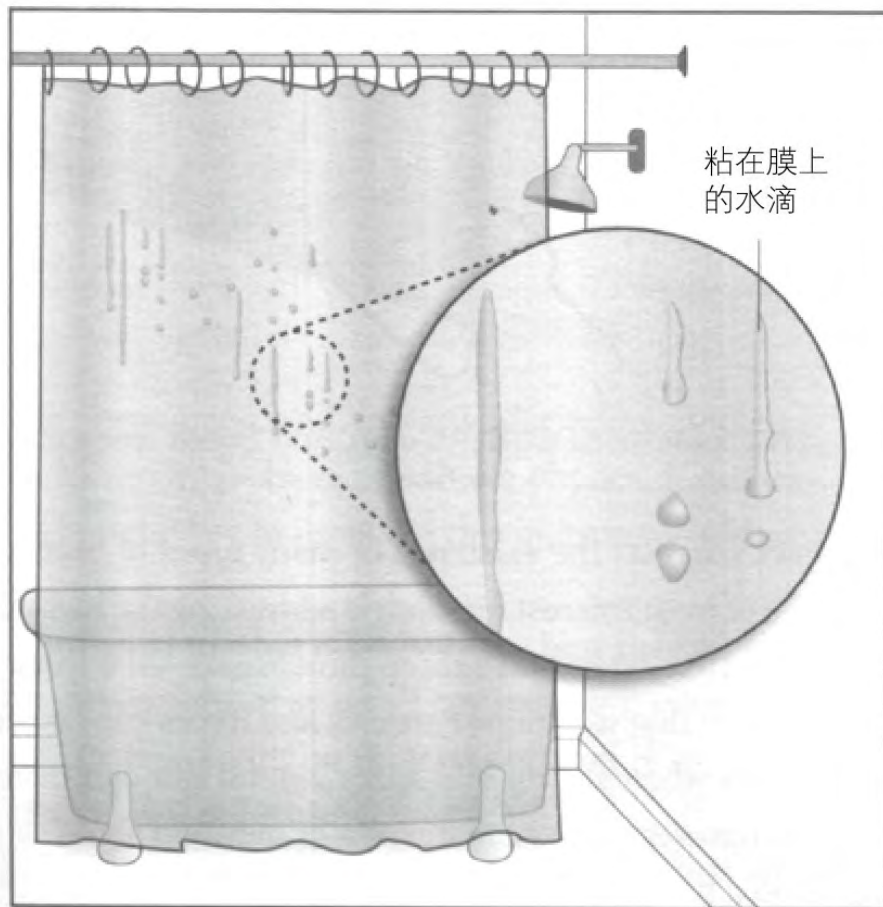


图17-7 膜将粒子和力束缚，它们只能停留在它的上面而不能离开，就像水滴只能沿着浴帘下滑而不能离开一样。

从广义上讲，存在两种弦：开弦（open string，有两个端点），与闭弦（closed string，形成一个闭合的圈），它们就像扎头发的橡

皮筋（见图17-8）。弦理论学家在20世纪90年代意识到：开弦的两个端点不能随心所欲地待在空间中的任意一个位置上，相反，它们必须待在膜上。当粒子从驻足于膜的开弦的振动产生出来时，它们也被限制在膜上。粒子，即弦的振动模式，也被束缚住了。就像浴帘上的水滴一样，粒子也只能沿着膜所伸展的维度方向运动，而不能脱离膜。

开弦/闭弦

弦理论认为，自然界的基本单元不是电子、光子、夸克等点状粒子，而是很小的线状“弦”，它包括有端点的开弦和圈状的闭弦。



图17-8 开弦有两个端点，闭弦没有端点。

弦理论提议存在许多种膜，而最有趣的一种解释等级问题模型中的膜有三个空间维度——刚好是我们所知空间的三个物理维度。粒子和力被束缚在这些膜上，哪怕引力和空间伸展到了更多的维度上面（图17-9系统表述了人和磁铁所处的一个膜世界，而引力则遍布该世界及其外面）。



图17-9 标准模型的粒子与相互作用可以被束缚在一个处于高维空间的膜世界中。在这种情况下，我们所知的物质和星辰、相互作用（如电磁作用），以及我们的星系和宇宙都生活在膜的三个空间维

度上。而引力则总是在所有空间维度中传播。（图片由马蒂·罗森伯格[Marty Rosenberg]友情提供）

弦理论的额外维度可能对可观测的世界，也就是三维膜有物理意义。也许研究额外维度最重要的原因是：它们可能影响那些我们可见的现象，特别是解释一些重大疑难问题，诸如粒子物理学的等级问题。额外维度与膜可能是解决该问题，也即解释引力因何如此微弱的关键。这是将我们引向考虑空间额外维度的最好原因。它们可能对我们正在试图理解的现象造成影响，如果的确如此，我们可能在不久的将来看到确实的证据。

别忘了，我们可以用两种方法表述等级问题。我们可以问希格斯质量：弱尺度为什么比普朗克质量小如此之多？这是我们考虑超对称和技术色的一个问题。我们也可以问一个等价的问题：为什么引力与其他相互作用相比如此微弱？引力强度依赖于普朗克能标，它是弱尺度的一亿亿倍。普朗克质量越大，引力越弱。只有当质量在达到或者接近普朗克能标时，引力才变得强大。只要粒子质量比普朗克能标低很多，引力就会极其微弱。

因此，引力如此微弱的疑难等价于等级问题，其中一者的解决意味着另外一者也能得到解决。但是即使问题是等价的，将等级问题用引力来表述有助于引导我们的思绪朝向额外维度。我们现在就来探究两种领导思潮。

模型四：大尺度的额外维度与等级

自从人们开始思考等级问题以来，物理学家就认为它的解决方法必定存在于弱能标尺度（约1TeV）上，修改粒子的相互作用。如果只有标准模型粒子，那么量子效应对希格斯粒子质量的修正就太大了。一些物质必须介入来调低量子力学对希格斯粒子质量的贡献。

超对称与技术色是两个例子——新的重粒子参加了高能相互作用，并且抵消或者从一开始就禁止了那些巨大的贡献。直到20世纪90年代，所有关于等级问题解的提案都被分类，新的粒子与相互作用以及新的对称性在弱能标上产生了。

1998年，尼玛·阿卡尼-哈米德（Nima Arkani-Hamed）、萨瓦斯·迪莫普洛斯（Savas Dimopoulos）与贾·德瓦利（Gia Dvali）^[3]提出了另一种解决这个问题的方法。他们指出：这个问题不仅涉及弱能标，而且涉及弱能标以及与引力相关的普朗克能标之间的比例，也许这个问题来源于一种关于引力本质的错误理解。

他们三人认为，实际上在质量上并没有等级——至少针对引力的基本尺度与弱尺度而言。也许反过来，是因为引力在额外维度宇宙中强得多，但只是在我们3+1维的世界测量中才变得如此微弱，因为它在我们看不见的那些维度中被弱化了。他们猜想，在额外维度宇宙中引力很强的那个能标，事实上就是弱能标。在这种情形下，我们测到的引力很小，不是因为它本质上很弱，而是因为它遍布在看不见的维度上面。

一种理解这一点的方法是想象一个类似于灌溉使用的旋转喷头。水从喷头里涌出。如果水只在我们的维度里铺开，那么它的效果依赖于水从龙头里面涌出的总量及其喷射距离。在离出水口固定的距离上，如果水还能喷洒到我们看不见的维度里的话，那我们得到的水量将比应该得到的少得多（见图17-10）。

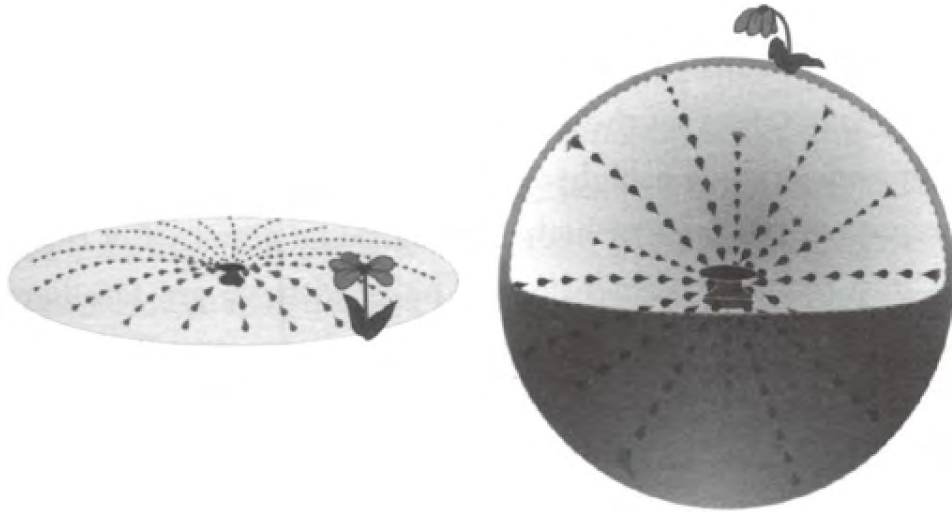


图17-10 相互作用的强度随着距离的增加而减弱，并且在高维空间比在低维空间减低得快。在高维旋转喷头的类比中，水在高维空间中随着距离的增加而稀疏得更快。水在三维空间比在二维空间的覆盖面广。在本图中，只有从低维空间的喷头获得水的花得到了充分浇灌。

如果额外维度的尺度有限，那么水会到达额外维度的边界而不再流到外面。但是在固定位置处的水量将比一开始没有这些维度、水也不会散布到这些维度里面的那种情况获得的水量少得多。

类似地，引力会传播到其他维度里。即使这些维度的尺度有限，引力不会无限地在其中伸展，大尺度的额外维度也弱化了我们在三维空间中感受到的引力。如果额外维度足够大，我们感受到的引力就非常弱，哪怕高维空间中引力的基本强度并不小。我们需要记住的是：为了使这种想法成立，额外维度必须比理论设想所预期的巨大得多，因为三维中的引力的确看起来太弱了。

无论如何，大型强子对撞机将这种想法带到了实验的台面上。尽管这个想法看似不太可能，但现实（而非我们对模型的感受）才是评判真理的唯一标准。如果该想法在世界上得到了实现，那么这些模型将导致截然不同的特征信号。因为更高维的引力在弱尺度能标（大型

强子对撞机可以创造的能标)上很强,粒子可以碰撞进而产生更高维的引力子(graviton)——传递高维引力的媒介粒子,而引力子可以在这些额外维度传播。我们所熟悉的引力非常微弱,以至于只有三个维度的空间无法产生引力子。但是在新情景中,更高维引力足够强,以至于其可以在大型强子对撞机所能达到的能标上产生引力子。

结果,卡鲁扎-克莱因粒子(即KK粒子)可以被制造出来,它们是更高维引力在三维空间的表示。KK粒子以西奥多·卡鲁扎和奥斯卡·克莱因两人的名字命名,因为他们最先研究额外维度。KK粒子有着与已知粒子相类似的相互作用,但是其质量更大。这些大质量是在额外维度中它们拥有的额外动量的结果。如果KK模式与额外维度的引力子(正如大尺度额外维度情景所预测的)相结合,那么它将在探测器中消失。它转瞬即逝的造访证据将从缺失的能量上体现出来(见图17-1, KK粒子产生并带走了无法直接看见的能量和动量)。

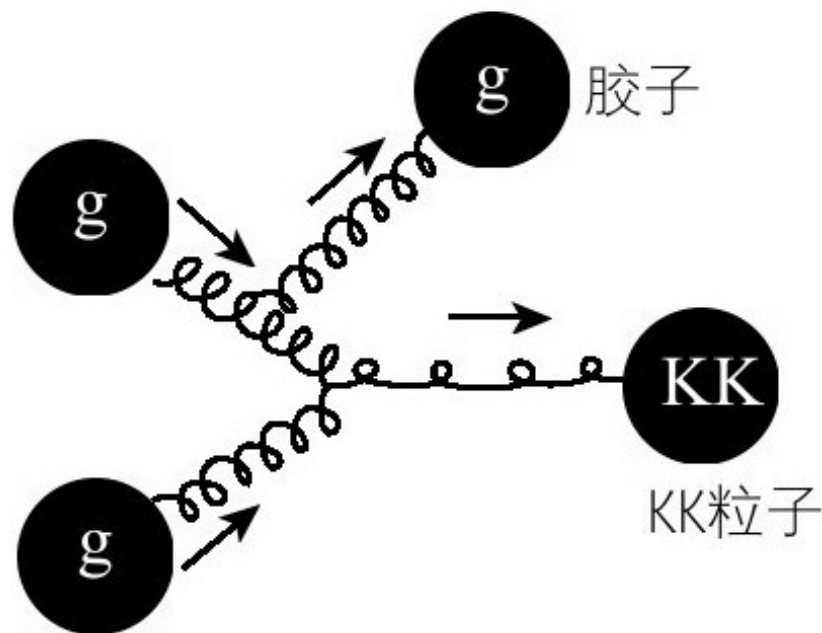


图17-11 在大尺度额外维度的情景中，引力子的一个具有动量的K伙伴粒子可以被制造出来。如果发生这种情况，那么它将从探测器中消失，只留下缺失的能量和动量的证据。

当然，缺失的能量也可以是超对称模型的一个特征。而这个信号看起来如此相似，以至于如果的确出现这样的发现，那么额外维度和超对称两大阵营的人们都可能用此数据来解释他们所支持的理论预期。但是伴随着对两种模型结果和预测细节的了解，如果两者之间必有一者是正确的话，那么我们就能够决定正确的那一方。我们建模的一个目标是，将实验信号与两种想法的真实内涵匹配起来。一旦我们能够将不同的可能性分门别类，也就能知道接踵而来的信号的速率和性质，那么我们就可以使用微妙的性质来区分它们。

不管怎样，目前我与大多数同事都怀疑大尺度额外维度的情景是否真的就是等级问题的解，就算我们将很快看到一个非常不一样的额外维度例子，而它看起来更有希望是正确的。一方面，我们不希望额外维度太大，因为额外维度若是比其他尺度大得太多也会造成新问题。那样，即使弱尺度与引力尺度的等级问题原则上消除了，一个涉及新维度尺度的新等级问题也会因此产生。

而更令人担忧的是，在这种情景中，宇宙的演化将非常有别于我们现今所观测到的。其问题在于，这些非常大尺度的维度可以随着宇宙的扩张而扩展，使得宇宙的温度降得很低。若是一个模型有可能成为现实的候选者，它所预言的宇宙演化就必须模拟出与我们所观测到的三维空间相一致的结果，这为大尺度额外维度的情景设置了一个艰难的挑战。

这些挑战还不足以明确地排除这些想法，更聪明的模型创建者可能会发现解决大多数问题的方法。但是为了与观测相符，模型倾向于变得过度复杂和纠结。绝大多数物理学家出于美学的考虑，对这种想

法持怀疑态度。许多人因此转向其他更有希望的额外维度想法（例如下文将介绍的）。即便如此，只有实验才能确切地告诉我们，有着额外维度的模型能否应用于真实的世界。

模型五：弯曲的额外维度

大尺度的额外维度并非唯一有潜质解答等级问题的方案，甚至在额外维度宇宙的框架下也是如此。自从通往额外维度想法的大门打开之后，我和拉曼·桑卓姆找到了一种看起来更好的解答方法^[4]——大多数物理学家都会同意其更有可能在自然中存在的方法。请注意，这并不意味着大多数物理学家都认为它是对的。许多人怀疑是否真的有人可以如此幸运，能正确预言大型强子对撞机将揭示的东西，或者能给出一个无须更多实验证据的、完全正确的模型。但是作为一个占尽先机的想法，它很可能是正确的，并且与许多好的模型一样呈现出清晰的策略，使得理论物理学家和实验物理学家可以更全面地开发大型强子对撞机的可能性，甚至可能发现该方法的证明。

我和桑卓姆所提的模型只涉及一个额外维度，它的尺度不必很大，因此没有引入新的涉及该维度尺度的等级。与大额外维度的情景相反，宇宙的演化自动与最近的宇宙观测相符。虽然我们的注意力只在这个单一的新维度上面，但是空间的其他附加维度还是可能存在的。但是在我们的情景中，在解释粒子性质时，它们没有发挥可以让人察觉的作用。因此当我们探究等级问题的答案时，为了配合有效理论的方法，可以合理地忽略这些维度，而只关注单一额外维度的效用。

如果我与桑卓姆的想法是正确的话，那么大型强子对撞机很快就将给出关于时空本质令人目眩的性质。原来，我们所假设的宇宙是高度弯曲的，这与爱因斯坦提出的出现了物质与能量的时空的情形相

符。其意思是，我们从爱因斯坦方程导出的几何是“弯曲”的（这确实是一个预先存在的术语）。这意味着空间与时间沿着我们感兴趣的那个单一额外维度的变化而发生改变。它改变的方式是：当你从额外维度的一个位置移动到另一个位置时，时空以及质量、能量都发生放缩（见图17-12）。

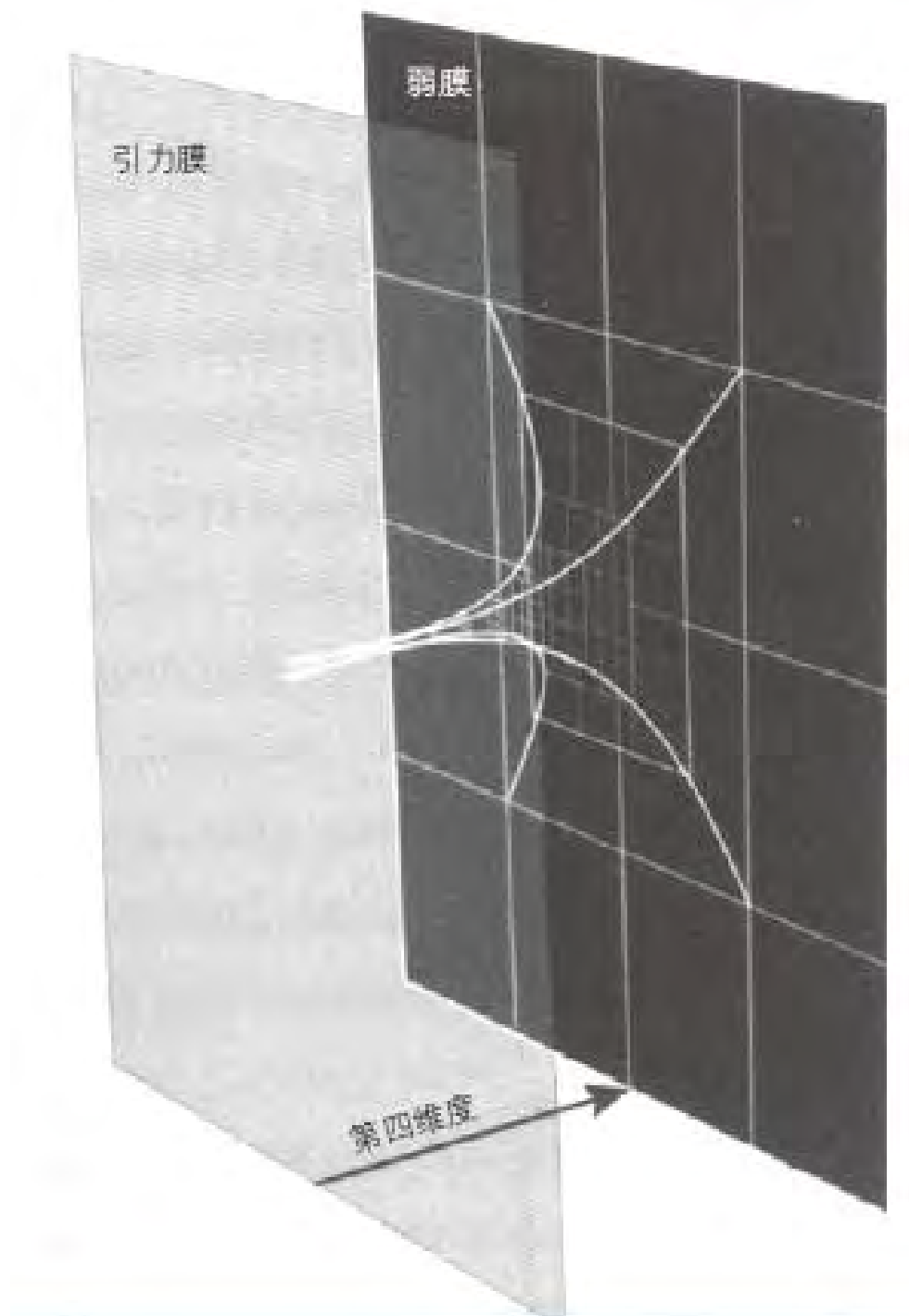


图17-12 另一种理解弯曲几何解决了等级问题的方法是通过几何本身。当你从一张膜移动到另一张膜上时，空间、时间、能量与质量都按照指数函数缩放。在这种情景中，希格斯质量是将普朗克质量按指数函数缩小，这种解释是非常自然的。

弯曲时空几何的一个重要结果是，虽然希格斯粒子质量在额外维度的其他位置上变得很重，但是它在我们所处的位置上将具有弱尺度质量（和应该表现出的质量完全一致）。可能这听起来具有任意性，其实不然。根据我们的情景，我们生活在一张膜，即弱膜上，而第二张膜（引力所聚集的膜）称为引力膜，或者有些物理学家也把它称为普朗克膜。这张膜包含另一个从额外维度来看与我们隔离开来的宇宙（见图17-13）。在这种情景中，第二张膜可能实际上就恰好跟我们相邻，只与我们相隔一个无穷小的距离——比如一亿亿亿亿分之一米。

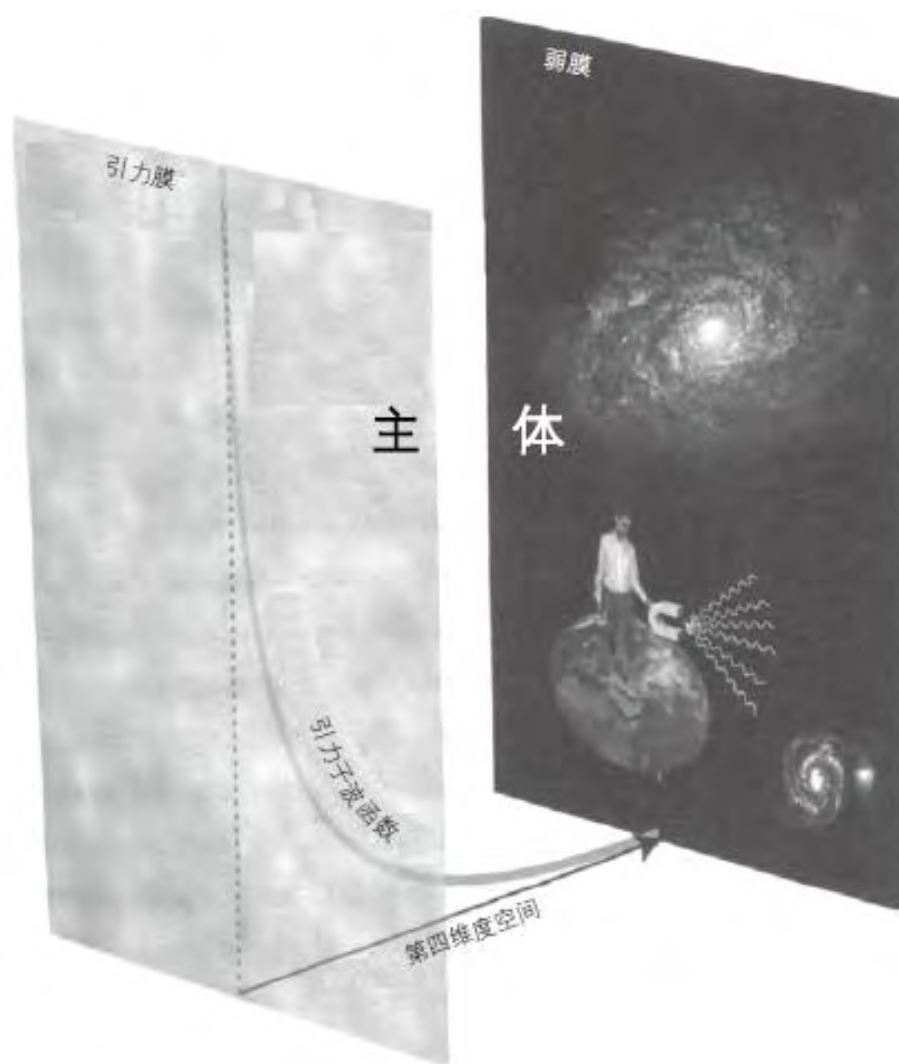


图17-13 兰道尔-桑卓姆模型包含两张膜，它们限制了第四维度空间（时空的第五维度）。在该维度空间中，引力子波函数（gravit

on wavefunction, 描述在任何空间点找到引力子概率的函数) 从引力膜到弱膜上呈指数级衰减。

图17-13中, 从弯曲几何而来的最显著的性质是引力子——传递引力的媒介粒子, 在另一张膜上的质量远比在我们这张上的大。这将使得引力在额外维度的其他位置上很强, 而在我们所处的这个位置上很弱。事实上, 我和桑卓姆发现引力的强度在距离我们很近的地方比在另一张膜上小, 并且呈指数级衰减, 因此给出了引力如此微弱的一个自然解释。

另一种解释这种结果的方法是通过时空几何, 系统地显示在图17-12中。时空尺度依赖于第四个空间维度的位置。质量也呈指数函数缩放——这样做使得希格斯玻色子质量呈现出它所需要的数值。虽然有人可能争辩说我们的模型依赖一个假设, 即两个巨大平坦的膜限定了额外维度宇宙, 但是一旦你给定由膜和称为“主体”(bulk)的额外维度空间所承载的能量, 那么该几何可以直接从爱因斯坦引力理论推导出来。当这样做时, 我们即可发现此前所提到的几何——即卷起来的弯曲空间, 其中的质量按照解决等级问题所需的方式进行缩放。

与大尺度额外维度模型不同, 基于弯曲几何的模型不会将老的等级问题谜题换成新的等级问题(即为什么额外维度如此巨大)。在弯曲几何中, 额外维度并不大。巨大的数值来自一个指数缩放的空间和时间。指数缩放使得尺度(以及质量)的比例是一个巨大的数字, 甚至当这些物体在额外维度中相隔的距离不大时, 也是如此。

指数函数不是编造出来的, 它来自我们所提议的情景中爱因斯坦方程的唯一解。我和桑卓姆计算出了在弯曲几何中, 引力与弱相互作用的比是两张膜距离的指数函数。如果两张膜的间隔是一个合理数值(大约几十倍于引力所设定的距离), 那么质量与相互作用强度的正确等级就会自然出现。

在弯曲几何中，我们所经受的引力如此微弱的原因，不是由于它在大尺度额外维度中被弱化了，而是因为它被聚集到其他地方——另一张膜上。我们的引力由位于额外维度另一个位置处，某个很强的相互作用的指数衰减的尾巴来决定。



我们之所以没有看到位于另一张膜上的另一个宇宙，是因为我们两个世界所共有的作用只有引力，而引力在我们附近已经太弱，以至于无法传递可以察觉的信号。事实上，这种情景可以看成多重宇宙的例子。在多重宇宙中，我们世界的物质和元素与另一个世界的物质之间的作用非常微弱，或者在某些情况下根本没有相互作用。绝大多数猜想都不能被检测而只能停留在想象的空间中。毕竟，如果物质如此遥远，连从那里来的光在宇宙有限的寿命中都不能到达地球，因此我们是不能探测到它的。然而，我和桑卓姆所提议的多重宇宙的情景不是一般的提议，因为共有的引力可以导致实验上可探测的结果。我们不是直接接触另一个世界，而是在更高维度内部空间中传播的粒子来造访我们。

额外维度世界最明显的效应（在缺乏诸如大型强子对撞机之类的仔细搜寻时）将是粒子物理学理论所需要的质量等级的解释，以便能成功解释观测到的现象。当然这对我们来说不足以有效解释这个世界，因为它与其他解释并没有区别。

大型强子对撞机即将展开的高能实验可以帮助我们确定，额外维度仅仅是一个天马行空的想法，还是一个关于宇宙的真实元素。如果我们的理论正确，那么我们将预期大型强子对撞机产生KK模式。因为与等级问题的联系，我们的模型寻找KK模式所需的能标，大型强子对撞机是可以达到的。它们应该大约在万亿电子伏的量级，即弱尺度能标上。一旦能量达到如此之高，这些重粒子可能产生出来。KK粒子的发现将为我们提供关键的确证，给我们提供扩张的世界的启示。

事实上，弯曲几何中的KK模式有一个重要且特别的性质。虽然引力子本身的强度极其微弱（毕竟它传播的是极其微弱的引力），但是引力子相互作用的KK模式比它强得多——几乎与弱作用的强度一样，它是引力强度的亿万倍。

KK引力子具有如此出人意料的相互作用强度的原因在于它们所处的弯曲几何。由于时空强烈地弯曲，引力子KK模式的相互作用比我们感受到的引力子传播的引力作用强得多。在弯曲几何中，不仅质量被缩放，引力强度也被缩放。计算表明在弯曲几何中，KK引力子的相互作用强度可以与弱尺度上的粒子的相互作用相当。

这意味着不同于超对称模型，也不同于大尺度额外维度模型，我们的模型的实验证据不是来自有趣的粒子逃脱而造成的能量缺失。相反，该证据将是更干净也更容易确认的信号——探测器中的粒子衰变成标准模型粒子而留下的可见轨迹（见图17-14，KK粒子产生并且衰变成电子-正电子对）。

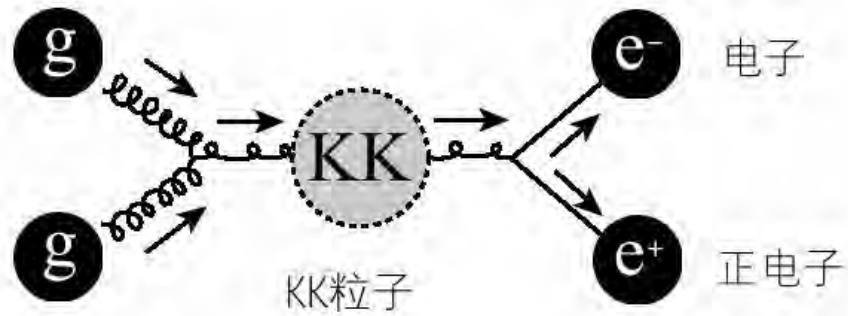


图17-14 在兰道尔-桑卓姆模型中，KK引力子可以在探测器内部产生以及衰变成可见粒子，如电子与正电子。

这实际上就是目前实验物理学家找寻所有新的重粒子的方法，他们并不能直接看到粒子，但却可以观测到那些粒子衰变之后的产物。从原则上说，这可以提供比缺失能量更多的信息。通过研究这些衰变产物的性质，实验物理学家可以得出最初出现的粒子的性质。

如果弯曲几何情景是正确的，那么我们将很快可以看到KK引力子模式衰变出来的粒子对。通过测量末态粒子的能量、荷以及其他性质，实验物理学家将可以推导出KK粒子的性质。这些鉴别特征以及粒子衰变成各种末态的比率，将有助于实验物理学家断定他们是否已经发现了KK引力子或者其他新的奇异元素。这些模型让我们知道需要找寻的粒子的本质，也让物理学家可以以此来分辨这些可能性并作出预测。

我的一个朋友（一个同时歌颂和嘲讽人之极端本性的剧作家）不理解，明知未来的发现可能具有无穷的可能性，我为什么不会坐立不安地等待结果。每当我见到他，他都一直追问我：“这会不会是一个改变命运的结果？他们能证实你的理论吗？”“你为什么不待在那儿（日内瓦）一直与别人讨论？”

当然，在某种意义上讲，我那位朋友的直觉是对的。但是实验物理学家已经知道应该找寻什么，因为理论物理学家已经做了这么多工

作。当我们对寻找的事物有了新想法时，总会进行交流。我们不必亲自待在欧洲核子研究中心或者甚至处于同一个房间才能做这件事情。为了交流，我们可以在全美境内甚至全球范围内找到实验物理学家。远程交流的效果也非常好——这也可以归功于许多年前欧洲核子研究中心的蒂姆·伯纳斯-李创建万维网的眼光。

我已经充分了解这些探索中的挑战在哪里，甚至当大型强子对撞机开始完全运行之后，也是如此。我知道我们可能需要等上一段时间。但我们很幸运，此前介绍的KK模式是实验物理学家所能找寻的一个最直接证据。KK引力子衰变成各种粒子——毕竟，每种粒子都受到引力的作用，因此实验物理学家可以关注那些他们容易鉴别的末态。

有两个注意事项：搜寻的难度可能比最初预想的更大的两个原因；哪怕最初的想法是正确的，但为了新发现，我们可能要等上相当长的一段时间。

搜寻难度大的一个原因是，其他关于弯曲几何的候选模型可能会导致更杂乱无章的实验信号，从而导致实验更难发现新粒子。模型表述了基本理论框架——涉及额外维度与膜。他们建议了此框架实体中的一般原理的特定应用。我们最初的情景提议只有引力能穿越更高维的“主体”的空间。但是我们的一些同事后来研究了另外的应用。在另一种情景中，所有粒子不是都处于膜上。这意味着将会有更多的KK粒子，因为每个主体中的粒子都有其相应的KK模式。但是这些KK粒子将很难被发现，它带来的挑战已经促使许多研究考虑如何探索这些更多的难以捉摸的情景。因此，在KK粒子与新模型可能出现的高能粒子的搜寻中，接下来的调查都是非常重要的。

搜寻可能十分困难的另一个原因是，KK粒子可能比我们预期的要重。我们知道KK粒子预期质量的范围，但是我们还不知道它的具体数值。如果KK粒子足够轻的话，那么大型强子对撞机将能够大量制造

它们，而发现它们就比较容易。但是如果这些粒子很重，那么大型强子对撞机可能只能制造少量粒子。更有甚者，要是它们实际上非常重，那么大型强子对撞机根本不会产生任何这种粒子。换言之，新粒子与新的作用很可能只在大型强子对撞机不能企及的能量以上产生和出现。这是大型强子对撞机的顾虑，因为它的隧道尺度是固定的，所能达到的最高能量已经被限定了。

作为理论物理学家，我所能做的仅此而已，大型强子对撞机的能量也仅此而已。但是即使KK模式很重，我们也可以努力找到关于额外维度的精妙线索。当我和帕特里克·米德计算可能的高维黑洞的产生速率时，我们不仅关注负面结果——黑洞产生的速率远低于最初人们所宣称的，而且我们也会思考，如果更高维引力很强，那么会发生什么（即便没有产生黑洞）。我们探求大型强子对撞机是否可以产生任何关于高维引力的有趣信号，并发现了即使没有发现新的粒子或者奇异的类似于黑洞的物质，实验物理学家也能观测到与标准模型所预想的有一些偏离的东西。新的发现并没有得到保证，但是实验物理学家将利用已有的仪器和技术尽其所能地去探求。在其他更高深的研究中，我的同事们还考虑探索KK模式的改良方法，即使那些标准模型粒子还存在于主体中。

还有一种可能是：我们可能很幸运，新粒子与相互作用出现的能标也许比我们期望的低一些。如果的确如此，我们不仅很快会看到KK模式，还将看到新的现象。如果弦理论是大自然的基本理论，而新物理的能标又足够低，那么大型强子对撞机甚至可以产生除了KK粒子与新相互作用之外的、由基础弦振动产生的附加粒子。这些粒子在传统假设之下的质量太大，而不可能被制造出来。但是由于弯曲几何，希望是：一些弦的模式可以比预期的低很多，从而可以在较低的能标出现。

这里显然有几种有趣的弯曲几何的可能性，所以我们迫切地期待实验结果的出台。如果几何的结果被发现，那么它们可能改变我们对宇宙本质的看法。但是我们只能在大型强子对撞机完成它的探索之后，才能知道哪些可能性（如果的确存在的话）在自然中是现实的。

回归，静候一个不同凡响的宇宙

大型强子对撞机的实验物理学家目前正在检验本章所提的各种想法。我们希望如果任何一种模型是正确的，那么提示将很快出现。这些可能会是坚实的证据，如KK模式，或者可能是对标准模型的微妙修正。不管怎样，理论物理学家与实验物理学家都将保持心明眼亮，并翘首以待。每一次大型强子对撞机看见或者没有看见，都会进而限制一些可能性。如果我们幸运，那么前面讲述的想法之一将被证明是正确的。当我们从大型强子对撞机了解更多探测器工作的原理以及产生机制，我们将了解如何将它的研究扩展来检验更大范围的可能性。当数据出来之后，理论物理学家将把数据与他们的理论结合起来。

我们不知道还需要多久才可以开始获取答案，因为我们不知道新的东西及其质量与相互作用可能是什么。一些发现可能在一两年内就可以获得，其他的则可能需要花费10年以上，还有一些甚至需要比大型强子对撞机所能提供的更高的能量。这种等待虽然会令人不安，但其结果可能颠覆我们以往的想法。因此，我们所迈出的每一个艰难步伐都是值得的。它们可能会改变我们对大自然及宇宙本质的看法，或者至少会改变我们对物质组成的看法。当结果出来时，新的世界也将出现。在有生之年，我们有可能看到一个不同凡响的宇宙。



实验结果的确定性无可替代。而物理学家在过去25年里，并非无所事事地坐等大型强子对撞机的出现，进而产生有意义的数据，而是努力思考实验应该寻找什么，其数据会具有什么意义。我们也会研究从目前的实验中得来的结果，它们会为我们提供关于已知粒子与相互作用的细节信息，有助于我们整理思绪。

这段期间也给了我们很多机会，来深入思考一些想法——至少是短期内更容易用数据排除的想法。一些更有趣、更具想象力的模型与理论启示，来源于过去25年的数学探索。如果数据更丰富，我怀疑我更可能会从数学角度思考额外维度或者超对称。即使所做的测量最终支持这些想法，如若没有之前充足的数学探求，其应用的获得也将耗费不少时间。

实验与数学都能引领科学的进步，但是通往发展的道路是不清晰的，而物理学家已经根据最好的发展策略做了划分。模型创建者使用“自下而上”的方法，从已知的实验事实出发，解释剩余的未知谜题——通常采用更多理论的以及数学的方法。第17章介绍了一些具体模型的例子，以及它们如何影响大型强子对撞机的实验物理学家来开展进一步的搜索。

其他人（主要是弦理论学家）则采取“自上而下”的思考方法，从他们认为正确的理论（即弦理论）出发，试图使用它的基本观念来构造一种自洽的量子引力理论。自上而下的理论是定义在高能标和小尺度上的。它指的是一种理论观念：任何东西都可以从定义在高能标上的基本假设推导出来。虽然该名称可能很具有迷惑性，但是因为高能标对应于小尺度，别忘了在小尺度上的元素都是物质的基本组成要素。在这样的思考模式中，每一样东西都可以从基本原理和基本元素（这些都定义在小尺度和高能标上）推导出来，因此称为“自上而下”。

本章将要介绍“自上而下”与“自下而上”两种方法，并将它们进行对比。我们将探索它们的差别，也将思考它们如何可以结合在一起，以产生惊人的见解。

优美的弦理论

与模型创建者不同，更倾向于使用数学的物理学家试图从纯理论出发构造理论。他们的期望是，从一个优美的理论出发，推导出可能的结果，并将该想法用于解释实验数据。绝大多数建立统一理论的尝试都采用“自上而下”的研究方法，弦理论也许是这些例子中最有希望的一个。它是一个猜想，是其他所有已知物理现象所追随的最终的基本框架。

弦理论学家在物理学尺度上迈出了一大步，他们试图征服从弱能标到普朗克能标（在该能标上引力变得很强）的跨越。实验物理学家可能无法在可见的未来直接检验这些想法（额外维度模型也许是一个例外）。即使弦理论本身太难检验，弦理论的元素却为一些与之相结合且有机会观测到的模型提供了想法。

物理学家所问的，是有关模型建立与弦理论之间的取舍问题——是需要按照柏拉图式的方法，试图从一些更基本的理论事实中获得预见；还是按照亚里士多德的方法，扎根于经验的观测。你愿意采取“自上而下”还是“自下而上”？这个选择也可以重新表述成“老爱因斯坦与小爱因斯坦”的选择。爱因斯坦最初确实认为实验是物理研究的基础，尽管他同时也给予“美”与“雅”高度的评价。甚至当一个实验结果与他的狭义相对论相矛盾时，他也确信（最终被证明是对的）是实验发生了错误，因为否则相对论的应用就太丑陋，以至于不能被信赖了。

当数学最终帮助爱因斯坦完成了广义相对论的研究时，他变得更具有数学倾向。数学进展对于完成他的理论至关重要，所以在职业生涯的后期，爱因斯坦更加相信理论方法。然而，爱因斯坦没有解决这个问题。尽管他将数学成功应用到广义相对论中，然而他后期对于统一理论的数学研究也从来没有取得成果。

哈沃德·乔吉与谢尔登·格拉肖提出了大统一理论，它也是一种“自上而下”的想法。如我们所知，大统一理论基于数据，他们猜想的灵感来源于标准模型中特殊集合的粒子与相互作用，以及它们相互作用的强度，但是该理论却从我们所知道的东西中，选择出可能结果，它发生在距离现在遥不可及的高能标上。

有趣的是，即使统一理论可能出现在远超今天的粒子加速器所能达到的能标上，大统一理论初始模型作出的预测却是有可能可以观测到的。大统一理论模型预测，质子可以衰变。该衰变会花费很长时间，但是实验物理学家建造了巨大的池子，其中装满材料，希望其中至少有一个质子可以衰变而留下可见信号。然而当它并没有发生时，初始的大统一理论模型被否决了。

自那以后，乔吉与格拉肖都没有再研究任何有如此大能量跨度的“自上而下”的理论，即从我们目前在加速器中使用的能量，到那些高到没有它们，对目前的实验也只会造成微妙影响（或者一点影响都没有）的能量。他们认为将目前我们所能理解的能量与尺度改变这么多数量级，从而猜想出一个正确的理论，那简直太不可思议。

尽管乔吉与格拉肖有所保留，其他许多物理学家却认定，“自上而下”才是唯一能够攻破难题的方法。弦理论学家选择了一条不归路——不是传统的科学，却带来了丰富（虽然存在许多争议）的想法。他们理解理论的一些方面，却仍在努力把它们拼接起来，在形成与发展他们激进的想法时寻找底层的关键原理。

弦理论的目标是建立一个不是来自于实验数据，而是来自于理论疑难的引力理论。它提供了一种引力子的自然解释，量子力学告诉我们，传递引力的粒子必须存在。当前量子引力首当其冲的、自洽的候选理论是弦理论，而量子引力是将量子力学与爱因斯坦的广义相对论相结合的理论，因此它的适用范围是在相当高的能标上。

物理学家可以使用已知理论合理地小尺度（例如在原子内部）上作出预测，量子力学起重要作用，引力可以被忽略。引力对相当于原子质量大小的粒子的影响极其微弱，所以我们可以使用量子力学而放心地忽略引力。物理学家也可以在大尺度（例如在星系内部）上作出预言，引力占据主导地位而量子力学可以被忽略。

然而，我们没有一种可以将量子力学与引力结合起来的理论，即在任何可能的能标和尺度上都适用的理论。我们不知道在极高能标、极小尺度的地方——与普朗克能标或尺度相当的地方，如何开展计算。因为引力的影响对超重或者超高能粒子更大，所以引力对于具有普朗克质量的粒子起着非常重要的作用。在极小的普朗克长度上，量子力学的效应也很显著。

虽然这个问题没有损害任何可观测到的现象（至少在大型强子对撞机上没有），但是它意味着理论物理是不完备的。物理学家还不知道如何自洽地将量子力学与引力结合在极高能标与极小尺度的地方——两者对预测而言都相当重要且不能被忽略。我们理解中的一个重要缺口可能反而会为我们指出一条道路。许多人认为弦理论是它的解决方法。

“弦理论”名称的由来是振动的基础弦，它是最初理论构想的核心。粒子存在于弦理论中，但是它们从弦的振动而来。不同的粒子对应不同的振动，非常像是从振动的小提琴琴弦产生的不同音符。原则上看，弦理论的实验证据应该包含那些相应于弦振动产生的模式的新粒子。

然而，绝大多数这种粒子很可能因为太重而不可能被观测到，因此这是实验很难检验弦理论在自然中是否存在的原因。弦理论方程描述的物质如此微小且具有如此之高的能标，以至于任何我们可以想象出的探测器都不可能探测到它们。并且该理论定义在一亿亿倍于人类现今的仪器所能达到的能标之上。目前，我们甚至还不知道当碰撞粒子的能标提高10倍时所能发生的事情。

弦理论学家不能对实验可及的能量上的现象给出唯一预言，因为粒子组成与其他性质依赖于该理论中还没有确定的基本元素。弦理论在自然中的结果依赖于这些元素的安排。在当前的构造中，弦理论包含了太多的粒子、太多的相互作用、太多的维度，远远超出真实世界所能观测到的。是什么原因导致这些粒子、作用与维度有别于我们能看见的那些？

例如，弦理论的空间不一定是我们周遭有着三个空间维度的空间。相反，弦理论的引力描述了具有6个或者7个额外维度的空间。弦理论的一种有效版本解释了这些看不见的额外维度如何有别于我们能

看到的那三维。弦理论是如此引人注目，以至于类似额外维度等迷惑人的性质模糊了它与可见宇宙之间的联系。

为了从高能弦理论推断目前可以测量的能量上的结果，我们需要推导，将原始理论中的极重粒子去除后所得到的结果。然而，弦理论在低能量时，有许多种可能的实现方法，我们也不知道如何区分这些极大范围内的可能性，或者如何找出与现实世界相似的那一个。问题是，我们还没有充分地理解弦理论，以推导出它在所见能标上能给出的结果。理论的预测被它的复杂性所削弱。不仅它的数学性是一个艰难的挑战，而且我们都不清楚如何组织弦理论的元素，以及确定哪些数学问题需要被解决。

最重要的是，我们现在知道，弦理论比物理学家最初认为的复杂得多，并且涉及了许多其他有着不同维度的要素——主要是膜。弦理论这个名称还是普遍地适用，但是物理学家也可能指的是M-理论，虽然没有人真正知道“M”代表什么意思。

弦理论是一种宏伟的理论，它已经导致了意义深远的物理与数学的洞见，并且很可能包含正确的描述自然的要素。然而，一个巨大的理论沟壑将目前所理解的理论 with 描述现实世界的预测远远地隔离开来。

最终，如果弦理论正确，所有描述真实世界现象的模型应该可以从它的基本前提出发推导出来。但是它的初始构造非常抽象，而且它与可观测现象相去甚远。如果能找到所有正确的物理原理，使弦理论的预测可以与自然相符，那我们将是多么幸运啊！这是弦理论的终极目标，也是一个令人望而却步的任务。

虽然优雅与简单是一个正确理论的标志，但是我们只能在我们已经完全理解一种理论如何运行之后，才能真正判断它的优美。关于自

然如何、因何将弦理论的额外维度隐藏起来的发现将是一次令人惊叹的成功，物理学家想要找出让它得以实现的方法。

宇宙景观

正如我在《弯曲的旅行》一书中所开的玩笑，绝大多数企图将弦理论现实化的想法都与整形手术有点相同。为了让弦理论与我们的世界一致，理论物理学家不得不找到某种把不应该存在的部分隐藏起来的方法，将粒子移除并将维度消融。尽管粒子集合的结果与现实惊人地接近，你也能分辨出它们并非完全正确。

大多数将弦理论真实化的企图都有点选角面试的味道。虽然大部分天真无邪的少年演得不是很好，面部表情有些僵硬而无法表露丰富的感情，但是经过足够多次的试镜后，一个美貌、有天赋的演员就有可能出现。

类似地，弦理论的一些想法也依赖于我们的宇宙：它是弦理论成分的理想实现。即使弦理论最终会统一所有已知的相互作用与粒子，它也可能包含代表了特殊集合的粒子、力与相互作用的单一稳定的盆地；或者更可能的是，包含一个复杂得多的有着山川、峡谷以及一系列可能出现的景观。

根据近来的研究，弦理论可以在许多可能的宇宙中实现，该情景对应于一个“多重宇宙”。不同的宇宙可以离得非常远，以至于在它们的一生中，彼此也不能相互影响（甚至通过引力也不行）。在那种情形下，完全不同的演化可以在各自的宇宙中发生，而我们只能处于其中之一。

假如这些宇宙存在而没有办法把人送到它们里面去，那么我们就可以合理地忽略它们，只留下我们自己的宇宙。但是宇宙演化提供了

产生所有宇宙的方法，并且不同宇宙可以拥有不同的性质，不同的物质、相互作用以及不同的能量。

一些物理学家采用景观的方法结合“人择原理”（anthropic principle）来解释弦理论于粒子物理学中的棘手问题。人择原理告诉我们：因为人类存在于一个有着星系与生命的宇宙中，某些参数必须取特定的数值或与之相接近的数值，否则人类就不可能存在，并在这里讨论什么问题了。例如，宇宙不能有太多的能量，否则它就会膨胀过快，物质就不会坍缩形成宇宙的结构。

人择原理

正是因为人类的存在，才能解释宇宙的种种特性。宇宙若不是这个样子，就不会有人类这样的智慧生命来谈论它了。

如果真是这样的话，那么我们需要决定什么样的物理性质（假如存在的话）适合粒子、相互作用与能量的构型。我们还不知道哪些性质可以预测；哪些对于我们今天能够在这里讨论科学来说，是必须的；哪些性质有基本解释，而哪些只是一个由于所处位置造成的意外。

我认为，由许多可能的构型形成的景观，合理而有前途，因为不同组合的引力方程有许多可能的解集合，我不认为我们观测到的解就一定是所有问题的解。我发现人择原理作为一种对观测到的现象的解释，也是一种非常令人不满的解释。问题是，我们永远不知道是否只要有人择原理就够了。我们能唯一预测哪些现象，哪些现象只能决定到差不多的程度？最重要的是，人择的解释不能检验，它可能是正确的。但是如果有更基本的从第一性原理而来的解释出现，那么人择原理必定会被抛弃。

重归实地

弦理论很可能包含一些深刻且有希望成功的想法，它已经给了我们一些量子引力与数学的洞见，并为模型创建者提供了有趣的素材。但是在通过回答我们最希望解决的问题，从而创立理论之前，很可能仍需要经历很长的一段时间。从演草纸中直接推导出弦理论关于真实世界的结果太困难了。即使成功的模型最终来自于弦理论，这些过多的元素也使得其结果很难显现出来。

物理中模型创建的方法来自一种直觉，因为弦理论可以作出预言的能标远远高出我们可以观测的能标。由于在不同能标上的许多现象有不同的表述，而在粒子物理学中，在相应的能量上用来解释问题的机制是这里面研究得最好的。

物理学家拥有共同的目标，但他们对于如何最好地达到目标，却有着不同的期盼。我倾向于建模的方法，因为它在不久的将来更能获得实验的指导。我和同事可能使用弦理论的想法，我们的一些研究可能有弦理论的理念，但是应用弦理论不是我的首要目标，理解可以检测的现象才是。模型可以描述和经受实验的检验，甚至在与任何基本理论发生关系之前也可以如此。

模型的创建者非常务实地承认，我们不能立即导出所有东西。一个模型的假设可能是终极基本理论的一部分，或者它们仅仅阐述了一些新的关系，有待更深层次的理论来解释。模型都是有效理论。一旦某个模型被证实了，它可以为弦理论学家，或者任何想要采用“自上而下”方法的人，提供研究方向。事实上，模型已经从弦理论的丰富观点中获益良多，但是其最主要的关注是在低能标与低能标的实验上。

超出标准模型之外的模型理论结合了标准模型的要素，以及已经可以探索的能量上的结果，但是它们也包含新力、新粒子以及在更小尺度上的新相互作用。即便如此，为了将我们所知道的东西全部整合

起来，我和其他人研究出的具体模型通常都会失去其原先所具有的优秀性质。由于这个原因，模型的创建者往往需要具备开阔的视野。

人们常常不理解我因何需要研究不同的模型，因为我明白这些模型不可能都是正确的，而大型强子对撞机能给我们提供更多线索。人们会感到更加吃惊，因为我解释说，我不必对任何我所思考的特殊模型抱持更多期望。无论如何，我选择的是阐述真正新的原理或者新型实验研究的项目。我所考虑的模型一般具有一些有趣的性质，或者对于一些奇异现象具有一定解释力的机制。鉴于这许多未知因素（对于研究进程有无帮助的不确定因素），预测和诠释现实世界的设想中有许多难以想象的挑战。如果从一开始就能取得成功，那简直就可以堪称奇迹了。



额外维度理论最优美的一面是：从“自上而下”与“自下而上”两种研究策略中而来的想法走到了一起，得出了该理论。弦理论学家明白其理论构造中膜起的关键作用；模型创建者意识到通过将等级问题解释成引力问题，就能找到其他解决方法^[72]。

大型强子对撞机现在正在检验这些想法。不管大型强子对撞机探索到什么，它都将指导以及限制将来要构造的模型。当其高能实验的结果出来以后，我们可以把所有观测结果综合起来，进而决定哪一个理论是正确的。即使观测不能与任何一个具体的建议理论相符，我们也会从构造模型中得到启示，它可

以帮助我们最终可能正确的理论的范围缩小。

建模帮助我们意识到各种可能性，指导实验的探索以及数据出来以后解释这些数据。我们可能足够幸运，恰好给出了正确的模型。建模还将启发我们应该寻找什么证据。即便没有哪一个特定模型的解释会完全正确，它们也能帮助我们导出新实验结果的含义。结果会将众多的想法区分开，断定哪一种（如果存在的话）特定的实现是真实世界的表述。也可能当前所有提议都无效，那么数据则责无旁贷地成为我们建立可能正确的新模型的依据。

高能实验不仅仅搜寻新粒子，还寻找有更多解释力的基本物理定律的结构。在实验协助给出答案之前，我们只能提供各种猜想。目前我们应用美学标准来有意偏向某些模型。但是当实验达到一定的能标（及尺度）并且给出可以区分不同模型的必要统计之后，我们就可以了解得更多。实验结果将决定我们的猜想中哪些是正确的，从而帮助我们建立真实世界的基本理论。



第五部分

缩放宇宙，让我们极目远眺

**KNOCKING ON
HEAVEN'S DOOR**



读小学时，有一天早上醒来后，我看到一则令人困惑的新闻，说我们的宇宙（至少以我当时的理解）突然变老了一倍。我被这个改变吓坏了。怎么可能有像宇宙年龄这么重大的事情会随便发生如此巨大的改变，却没有将我们所知道的一切都毁灭呢？

现在，让我惊讶的事情发生了改变。我对人类能够精确测量宇宙及其历史的纵深程度感到惊讶。现在我们不仅比以往都更准确地知道了宇宙的年龄，还知道了宇宙如何随时间演化、核子如何产生、星系与星系团如何演化等。以前，我们只有关于这些演化的定性图像，但现在我们已经有了准确的科学图像。

宇宙学已经进入了一个卓越时代，实验与理论的革命性进展已经取得了更广泛和更详细的解释，这对生活在哪怕只有20年前的任何人来说也是难以置信的。通过将革新的实验方法与扎根于广义相对论和粒子物理学的计算结合起来，物理学家已经建立了一个关于宇宙早期面貌以及它如何演变成今日之貌的具体图像。

迄今为止本书主要关注于小尺度，在此之上我们研究物质的本质。到达了向物质内部探索的极限之后，现在我们继续第5章的内容，完成此次尺度上的旅行。我们将把注意力转向外部，考虑外部空间中物质的尺度。

我们需要警惕本次旅行中的宇宙尺度与前文的巨大差别，因为我们不再仅仅根据尺度来划分宇宙的各种事物。观测不只记录今天的宇宙。因为光的传播速度有限，所以观测也回顾了过去的宇宙。我们今天所观测到的结构可能是早期宇宙所具有的结构，它所发出的光线通过几十亿年才传播到我们的望远镜中。当前急剧膨胀的宇宙围绕在早期宇宙的外面，是其规模的很多倍。

然而，尺度在对当前宇宙及其历史观测的分类中起着至关重要的作用，本章将在这两个方面展开讨论。后文，我们将考虑整体的宇宙演化，从初始的小尺度到我们现在观测到的巨大尺度。但是首先我们将着眼于宇宙现在的样子，以此来熟悉一些能将周遭事物进行分类的尺度。然后我们将走向大尺度和远距离的物体——地球上的以及宇宙中的，来对外部空间中我们要探索的巨大类型的结构形成一个概念。这种大尺度旅行比我们前文在物质内部的旅行要简短。尽管宇宙结构非常丰富，绝大多数我们所看见的都可以被物理定律解释，而不涉及更基本、更新颖的定律。恒星与星系的形成依赖于已知的化学和电磁学定律，它们是我们已经讨论过的基于小尺度上的科学。然而，引力现在也起着至关重要的作用，最好的叙述依赖于它所作用的物体的速度和密度，因此也导致了理论表述的多样性。

宇宙的冒险之旅

《十的次方》（*Power of Ten*）这本书与同名电影^[1]，是一次关于距离尺度的标志性旅行，它从有一对情侣坐着的芝加哥格兰特公园（Grant Park）出发，最后又回到该地，对于我们来说这也是一个很好的出发点。我们先将镜头在地面上停留片刻，看一看我们周边熟悉的长度和尺度。在短暂地回顾了人的尺度——大约两米左右之后，让我们离开这个舒适的位置，向更大的高度与尺度迈进（关于本章的尺度示例见图19-1）。

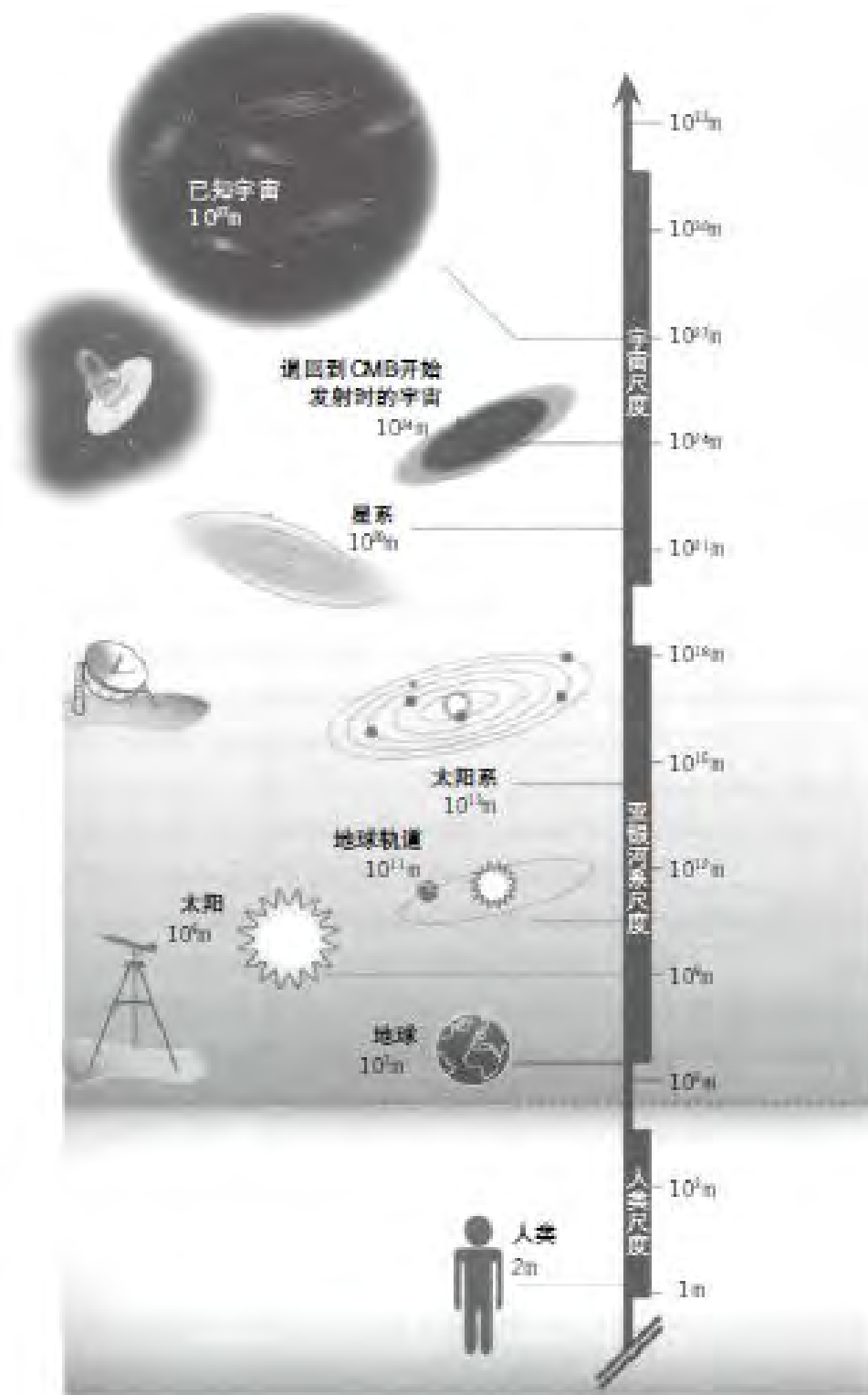


图19-1 一个关于各种尺度的旅行，以及用来描述它们的长度单位。

我曾见过的最吸引眼球的一次演出，是伊丽莎白·斯特布的舞蹈公司关于人体对高度的反应的表现。她的舞蹈演员（“行为工程师”）从一个横梁上跳下，腹部落地，而横梁上升得越来越高，直到最后一名舞蹈演员从9米高的地方跳下。那绝对超出了我们感受的舒适区，现场许多的观众都发出重重的惊讶声。因为人们往往不会从那么高的地方跳下，更不会面部朝下地跳下。

虽然绝大多数高耸入云的建筑给人的感觉不比上述表演更加摄人心魄，但它们也给人带来了从敬畏到恐惧的强烈震撼感。建筑师所面对的一个挑战是，赋予比我们大得多的物体一种人性化的结构。建筑与结构在尺度和形状上的差异很大，但是我们对于它们的回应反映了我们对尺度的生理和心理态度。

世界上最大的人造结构是位于迪拜的哈利法塔（Burj Khalifa Tower），它高828米。虽然它高得离奇，但是它的大部分都是空的。电影《碟中谍4》所赋予它的文化感可能并不如《金刚》赋予帝国大厦的文化感。虽然纽约这座高381米的标志性建筑比哈利法塔矮一半还要多，但是它的知名度却高很多。

我们生活在一个周围有着巨大自然实体的世界当中，它们中许多还激起了我们的敬畏之心。珠穆朗玛峰海拔8800多米，是世界第一高峰。欧洲的最高峰勃朗峰只有它的一半高，但是我却很开心我能登顶此峰，尽管我和朋友在登顶后拍摄的照片中都看上去很惨。马里亚纳海沟深11000米，它是海洋中最深的地方，也是地壳表面最低处。而詹姆斯·卡梅隆执导的3D电影《阿凡达》中则有一个异世界的海沟。

大自然所创造出来的雄伟之作在地球表面伸展到很广的区域。太平洋宽20000公里，而俄罗斯领土宽大约为8000公里，几乎是太平洋宽度的一半。地球的形状接近球形，直径约12000公里。美国领土宽4200公里，占地球周长的1/10，但是比月球的直径（3600公里）还长。

外部空间的物体尺度也很大。例如小行星之间的大小差异很大，有些小到只有鹅卵石那么大，而有些大到比任何地表所呈现的面貌还要大。太阳直径约为10亿米，该直径大约是地球的100倍。我所采取的太阳系尺度的计算方法是按照从太阳到冥王星（不管它是否还属于行星）的距离，这个距离是太阳半径的7000倍。

日地距离相对小很多（约1000亿米），是一光年的十万分之一。一光年指的是光线一年所走过的路程——3亿米每秒（光速）乘以3000万秒（一年中的秒数）。因为光速有限，我们看到的太阳亮光事实上是8分钟以前从太阳发出的光。

许多可见结构有着不同的形状和尺度，在我们的广博空间中存在着。天文学家根据绝大多数天体的类型对其归类。给定一些标准尺度，星系的大小通常是直径3万光年或者 3×10^{20} 米。包括我们的星系——银河系，它的尺度大约是这个大小的3倍。星系团一般含有从几十到几千个星系，尺度大约是 10^{23} 米，或者1000万光年。也就是说，光线从一个星系团的一端传播到另一端需要1000万年。

尽管这些尺度巨大，绝大多数天体也还遵循牛顿定律。月球轨道、冥王星轨道或者地球轨道都可以通过牛顿引力定律解释。基于行星离太阳的远近，它的轨道可以由牛顿引力定律预测。这个定律同样是使牛顿的苹果落向地面的原因。

尽管如此，行星轨道更精确的测量显示，牛顿定律不是终极理论。我们需要用广义相对论来解释水星的近日点进动，它的绕日运动轨道随时间的推移可以被观测到。广义相对论是更复杂的理论，当能量低、速度慢时它也会包含牛顿定律，但是在此范围之外广义相对论也仍然有效。

然而，不必使用广义相对论来描述绝大多数物体的运动。但它的效应可以累积起来，并且当物体足够致密（如黑洞）时，它的效应就会非常显著。我们的星系中心是一个半径约为 10^{13} 米的黑洞，它所包含的质量非常巨大，约等于太阳质量的400万倍，与其他黑洞一样需要用广义相对论来解释它的引力性质。

目前整个宇宙的直径是1000亿光年，即 10^{27} 米，是我们星系的100万倍。这个距离太大，令人震惊，因为它甚至比从大爆炸开始我们实际能观测到的距离——137.5亿光年更大。没有什么可以比光传播的速度更快，因此就宇宙年龄是137.5亿年来看，宇宙目前的尺度看似不可能。

这并不矛盾。宇宙作为一个整体，比在其有限生命中的一个信号传播的距离尺度大，其原因在于宇宙曾经经历膨胀的阶段。广义相对论在理解该现象时起到了重要作用。其方程告诉了我们膨胀的时空结构。我们可以观测宇宙中相隔很远的位置，即使它们并不能看见彼此。

鉴于光速有限，宇宙的年龄有限，这一节将我们带到有限的观测区域的边界。可见宇宙是指我们的望远镜所能达到的地方。然而宇宙的尺度却不受制于我们的观测极限。与小尺度情形相似，在超过我们实验水平之上，我们只能猜想可能存在的东西。对我们能思考的最大尺度事物的唯一限制，是我们的想象力，以及猜想那些我们无望取得任何观测结果的结构时的耐心。

我们真的不知道视界（horizon，视界是我们所能观测到的宇宙边界）之后有什么东西。我们观测的极限使得在其之外的新奇与特异的现象成为可能。不同的结构、不同的维度，甚至不同的物理定律原则上可能存在，只要它们没有与我们所观测到的东西发生矛盾。但这并不意味着每一种可能性都存在于自然之中，就像我的天体物理学同事

麦克斯·泰格马克（Max Tegmark）的观点那般。然而，这却意味着可能性的多样化。



我们还不知道额外维度或者另外的宇宙是否存在。我们甚至不能肯定地说宇宙作为一个整体是有限的还是无限的，尽管很多人倾向于后者。测量没有显示出任何它的边界，但是测量只能达到那么远。从原则上说，宇宙可能有限或者甚至会具有类似球体或者气球的形状，但目前理论与实验都没有给我们提供任何该方向的线索。

绝大多数物理学家宁愿不去想那么多关于可见宇宙区域之外的东西，因为我们不太可能可以想象出来那里所存在的东西。然而，引力或者量子引力的理论为我们提供了数学工具，我们至少可以思考那里可能具备的几何。基于空间额外维度的理论方法与想法，物理学家有时会思考其他奇异的宇宙，在我们宇宙的一生之中它们或者与我们没有交流，或者仅仅通过引力来相互作用。如第18章介绍的，弦理论学家和其他人猜想存在多重宇宙，它包含许多不连通的独立宇宙，它们与弦理论的方程相容。有时将这些想法与人择原理结合起来，可以得出可能存在的多个宇宙的丰富结构。有人甚至试图寻找将来这种多重宇宙的可观测信号。如第17章介绍的，在一种不同的情景中，两个膜的“多重宇宙”也许可以帮助我们理解粒子物理学的方程，并给出可以观测的结果。绝大多数的额外宇宙，尽管可以想象它们的存在，但在我们可以预见的未来，也只能停留在实验可以检验的范围之外。它们将继续作为一些理论抽象出来的可能性而存在。

大爆炸

现在我们的冒险已经到达最大尺度了——我们在可观测宇宙的意义上进行观测或者讨论的范围，并一直达到我们所能看到的、所能想象的极限位置。我们探索了宇宙如何随时间演化，并最终呈现今日之貌。大爆炸理论告诉我们，宇宙在它137.5亿年的生命中是如何从一个很小的规模增长到现在的疆域——1000亿光年的宽度。著名天文学家弗雷德·霍伊尔（Fred Hoyle）开玩笑并且怀疑地以“最初的爆炸”为该理论命名，它从一团炽热密集的气体开始，扩展成为今天所见的大规模的恒星和结构——宇宙不停地增长，物质变得稀薄，然后逐渐冷凝。

我们显然不知道的一件事情是：最初是什么导致了宇宙的爆炸以及它是如何发生的，我们甚至不知道宇宙最初所拥有的严格尺度是多少。尽管我们对宇宙后来的演化有所了解，可是它的起点仍然隐藏在迷雾中。无论如何，虽然大爆炸理论没有告诉我们宇宙初始时刻的所有事情，但它仍然是一个成功的理论，它告诉了我们宇宙接下来的历史。目前的观测与大爆炸理论为我们提供了很多关于宇宙如何演化的信息。

20世纪之前，人们并不知道宇宙正在扩张。当爱德温·哈勃率先凝视天空时，人们还一无所知。哈罗·沙普利（Harlow Shapley）首先测得银河系的尺度为30万光年，但他很确信银河系就是宇宙的唯一。到了20世纪20年代，哈勃发现了一些星云，沙普利认为它们是由尘埃组成的云团（因此有了“星云”这个平凡之名），事实上它们却是星系，不过是在几百万光年之外。

在确认了这些星系之后，哈勃作出了他的第二个惊人发现——宇宙正在膨胀。1929年，他观测到星系的红移，也就是说这里存在一个多普勒效应，距离越远的物体，其光波会向波长越长的方向移动。红移现象证明了星系在远离我们而去，这与救护车远离我们时，它尖锐

的警报声频率降低的效果类似（见图19-2）。哈勃所确认的星系相对我们的位置不是静止的，而是都朝着远离我们的方向运动。这是一个证明我们的宇宙在膨胀的证据，在此情况下，星系之间的距离越来越大。

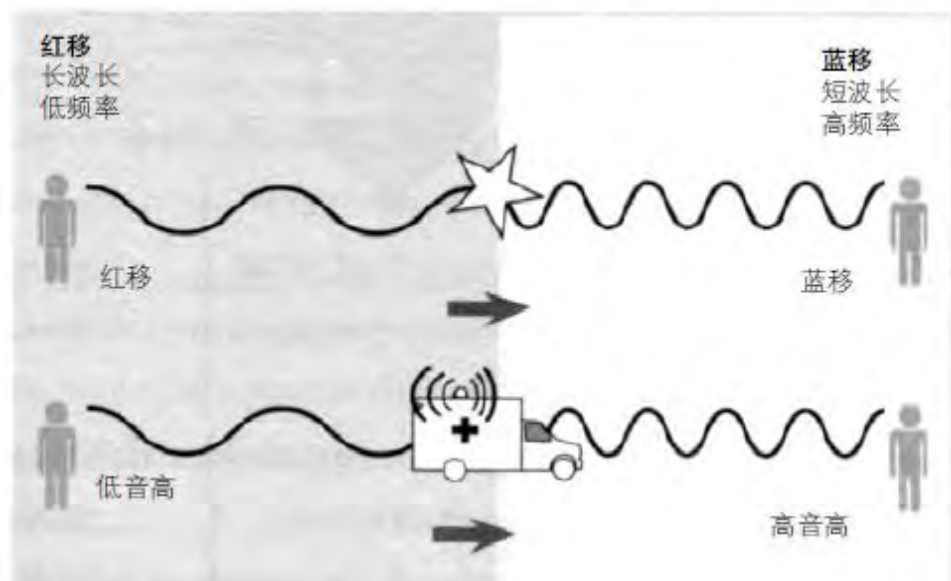


图19-2 从远离我们的物体传来的光线，光的频率变低或者说光谱向红端移动；而从靠近我们的物体传来的光线^[73]，光的频率变高或者说蓝移（blue shift）。这与从一个警报器传来的响声相似——当救护车离开时音高变低，而当它靠近时音高变高。

宇宙的膨胀与我们可能首先想到的图像是不同的，因为宇宙并不是在原本就存在的空间中扩展。宇宙就是所有的存在。没有任何东西可以让宇宙在其里面膨胀，宇宙与空间本身一起膨胀。随着时间流逝，它里面的任意两点分离得越来越远。其他星系远离我们而去，但是我们的位置并不特殊，它也会移动，与其他星系彼此远离。

想象宇宙是一个气球的表面。假设你在气球表面标记两点。随着气球胀大，表面拉伸开来，这两点也随之远离（见图19-3）。事实上，当宇宙膨胀时，其中任意两点的情况与此相似。两点或者说两个星系之间的距离增加。

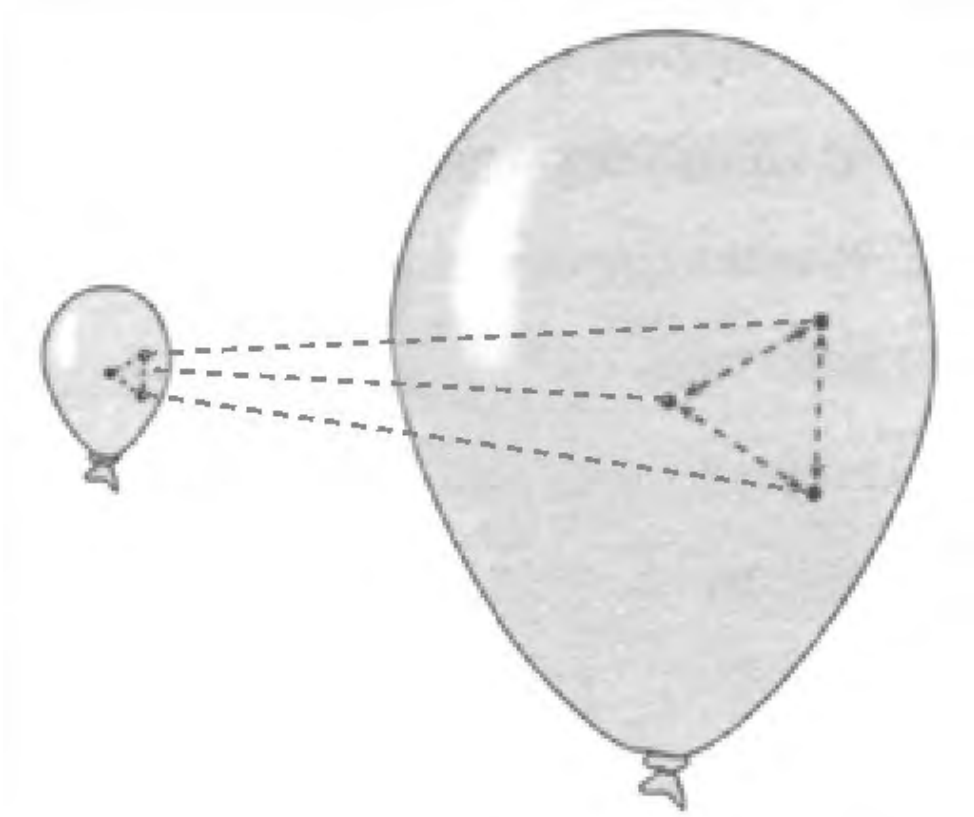


图19-3 “气球宇宙”展现了当气球膨胀时，气球上的标记点如何彼此远离。

注意，在我们的类比中，气球上的点并不会扩大，只是它们之间的距离会增大。这也是膨胀宇宙中的情景。例如，原子是由电磁力紧紧约束而成的，它们不会变得更大。相对密集的紧紧束缚在一起的结构（如星系）也是一样。驱使宇宙膨胀的作用也施加在它们身上，但是因为其他作用的效力还在，星系本身没有随着宇宙的膨胀而增长。它们经受很强的吸引力，因此其尺度会保持不变，而相互之间的距离越来越大。

当然气球的类比不够完美，因为宇宙有三个空间维度，而不是两个。并且，宇宙很大而且可能无穷大，而不像气球很小而且表面是弯曲的。最重要的是，气球处于宇宙中并且气球向外扩展的空间已经存在。而宇宙不同，它充满了空间，并不在已经存在的东西内部扩展。

尽管有此区别，气球表面是空间扩展的一个很好的示意，特别是每一个点同时彼此远离。

气球的类比（现在指的是其内部）也对理解宇宙如何从初始的一团炽热火球冷凝下来很有帮助。想象一下一个极端炽热的气球，假如你允许它膨胀到一个很大的尺度。虽然开始它可能太烫很难握在手中，但是一旦膨胀，它里面的气体就会变凉，不久之后再触碰它就不难了。大爆炸理论预言：最初炽热、致密的宇宙不断膨胀，因而同时它也不断地冷却了下来。

事实上，爱因斯坦曾经从广义相对论方程推导出宇宙正在膨胀这个结论。然而，那时没有人能测量到宇宙的膨胀，因此他本人也不相信这个预测。爱因斯坦引入一种新的能量来源，试图修改他的理论以给出静态宇宙。在哈勃的测量之后，爱因斯坦放弃了他开始所做的这种修正，并称之为“他最大的错误”。然而这种修正并非完全错误。我们将很快看到最近有测量显示，爱因斯坦所增加的宇宙学常数项事实上对最近的观测很有必要，尽管所测量到的对近来确认的加速膨胀负有责任的数值，比爱因斯坦所提的数字大了一个数量级。

宇宙膨胀是一个“自上而下”与“自下而上”两种物理方法汇合的很好例子。爱因斯坦的引力理论预示了宇宙的膨胀，然而只有当膨胀被发现了，物理学家才确信他们的研究步入正轨。

今天我们将决定当前宇宙膨胀速度的一个数值定义为哈勃常数（Hubble constant）。从“空间任何一个部分的膨胀都相同”这个意义上讲，它确实是一个常数。然而，哈勃这个参数却不是时间的常数。在更早时期，当宇宙还是炽热、致密的状态时，它膨胀的速度快得多。

精确测量哈勃常数是很困难的，因为我们面临的是前面提到的将过去和现在的纠缠解开的问题。我们需要知道红移的星系有多远，因为红移量依赖于哈勃参数与距离。测量的不精确是本章开头所提的宇宙年龄相差了两倍的原因。如果哈勃参数测量相差两倍，那么宇宙年龄也一样会差两倍。

这个矛盾现在基本已经解决了。哈勃参数已经由史密松森天体物理台（SAO）的温迪·弗里德曼（Wendy Freedman）与她的同事以及其他测量，膨胀速度对100万光年以外的星系来说大约是每秒22公里。基于这个数值，我们现在知道宇宙的年龄大约是137.5亿年，上下可能有2亿年的误差，但绝对不会相差两倍。虽然这听起来可能还是有许多不确定性，但这个范围对于我们目前的理解来说已经小到不会造成什么太大的差别了。

另外两个关键的观测与预测相符得很好，进而确定了大爆炸理论。一种测量依赖于粒子物理学和广义相对论，进而成为这两种理论的确据，即宇宙中各种元素的密度，比如氦和锂。大爆炸预言的这些元素的数量与测量相一致。在某种意义上来说，这是间接证据，需要用基于核物理和宇宙学的具体计算来给出这些数值。即使如此，除非物理学家和天文学家已经走在正确的道路上，各种不同元素的丰度在理论和测量上的一致性并不太可能只是一种巧合。

1964年，美国的罗伯特·威尔逊（Robert Wilson）与出生于德国的阿诺·彭齐亚斯（Arno Penzias）偶然发现了2.7K微波背景辐射，后来证实了大爆炸理论。从温度的角度来看，没有什么东西比绝对零度还冷，即开尔文单位制的0度。宇宙辐射温度比绝对零度高了不到3K。

威尔逊与彭齐亚斯的合作与冒险（因此他们在1978年获得了诺贝尔奖），是科学与技术琴瑟和鸣取得举世瞩目成就的一个实例。回到AT&T还是电信寡头的年代，AT&T做了一件了不起的事情——创建贝尔实验室，它为纯理论与应用的并肩发展提供了一个卓越的研究环境。

作为电子技术方面的极客，颇具远见的威尔逊与彭齐亚斯，都在贝尔实验室工作过，他们一起使用和完善了射电望远镜。威尔逊与彭齐亚斯都对科学与技术感兴趣，而AT&T也对通信业有着相当的兴趣，天空中的射电波对每一个相关的人来说，都很重要。

在追寻射电天文学的一个具体目标时，威尔逊与彭齐亚斯发现了他们起初无法解释，并认为神秘的讨厌东西。它看似一种均匀的背景噪声——本质上是静态的。它不是从太阳来的，也与之前的核实验无关。他们尝试了每一种他们能想到的解释方法，包括著名的鸽子粪，耗时9个月，努力找寻答案。在考虑了所有可以想象的可能性之后，清理了鸽子粪（或者彭齐亚斯所称的“白色介电材料”），甚至射杀了鸽子之后，噪声依然存在。

威尔逊告诉我他们有多幸运，巧就巧在他们发现的时间点。他们并不知道大爆炸理论，但是普林斯顿大学的罗伯特·迪克（Robert Dicke）与吉姆·皮泊斯（Jim Peebles）却知道。那里的物理学家刚刚意识到该理论的一个弦外之音就是，它将产生一个微波波段的辐射的残余。这些物理学家正试图设计一个实验来测量这个辐射，而他们却发现该辐射已经被发布了——贝尔实验室的科学家已经找到了它，只是还不知道发现的是什么。威尔逊与彭齐亚斯很走运，威尔逊对我说，麻省理工学院的天文学家伯尼·伯克（Bernie Burke）就像那个时代的网络，既知道普林斯顿的研究，又知道威尔逊与彭齐亚斯的发现。伯克把两者结合起来，使得相关的科研人员之间取得了联系并带来了丰硕的成果。

这也是一个科学采取行动的成功案例。研究本身是出于具体的科学目的，而该目的也带来了额外的技术和科学成就。天文学家并没有在寻找他们所发现的东西，但是其技术和科学技能娴熟。当发现新东西时，他们知道不能错过它。虽然开始寻求的是小现象的研究，但是结果却导致了有着深刻内涵的巨大发现，他们能有所发现是因为他们与其他人都同时在思考一幅大图景。虽然贝尔实验室科学家的发现是偶然的，但它却永远改变了宇宙学的进程。

宇宙辐射被证明是一个非常强大的工具，不仅证实了大爆炸理论，同时也将宇宙学转变成了一门具体的科学。宇宙微波背景（cosmic microwave background, CMB）辐射为我们提供了一种观测过去的不同方法，它不同于传统天文学测量的方法。

以前，天文学家观测太空中的物体时，试图确定它们的年龄，尝试推断它们的演化史。在宇宙微波背景中，科学家可以直接看向过去，甚至达到诸如恒星与星系等结构形成之前。他们所观测到的光线来自很早以前——宇宙演化的早期。当微波背景被发射出来之时，宇宙仅有现在规模的千分之一。

虽然宇宙最初充满了各种粒子（带电的与不带电的），一旦它冷却下来，在其演化的40万年之后，带电粒子组合形成中性原子。一旦发生这个结合，光线就不再被散射。地面上或者卫星上的望远镜所观测到的宇宙微波背景辐射（既没有被隐藏也没有被阻隔），直接来自宇宙演化之后的大约40万年的那一刻。威尔逊与彭齐亚斯发现的背景辐射与宇宙历史早期的辐射相同，但是它已经在宇宙膨胀的过程中被稀释和冷却了。望远镜所搜集到的辐射在它们传播的路途中，并没有被任何中介的带电粒子散射和阻隔。这些光线为我们提供了观测过去的直接和精确的视窗。

宇宙微波背景探测器（Cosmic Microwave Background Explorer, CMBE）是一个1989年发射的、为期4年的卫星任务。它在测量背景辐射时非常准确，该任务的科学家发现他们的测量与理论预测的相符程度高达千分之一。但是宇宙微波背景探测器也测量其他新东西。到目前为止，宇宙微波背景探测器测量到的最有意思的事情，是天空温度微小的非均匀性。虽然宇宙极其均匀光滑，但是微小的非均匀性在早期宇宙小于万分之一的级别上增长起来，变得越来越大，并对结构的形成起着至关重要的作用。这种非均匀性来源于微小的距离尺度，但是被拉伸到可以与天文测量和结构相比拟的尺度。引力造成了致密的区域——那些微扰特别大的地方，变成高度集中以至于形成我们现在观测到的致密物体。恒星、星系与星系团都是这些初始的微小量子力学涨落与引力演化的结果。

微波背景测量对于我们理解宇宙的演化不可或缺，它的作用就好像一扇直接朝着早期宇宙打开的窗户。近来，与一些传统方法相结合，宇宙微波背景测量为科学实验提供了另外几个神秘现象——宇宙暴胀、暗物质与暗能量的洞见，我们将在后文介绍。



当我还在麻省理工学院当教授时，物理系三楼已经没有多余办公室给粒子物理学家了。因此我搬到二楼阿兰·古斯（Alan Guth）办公室旁边的一间专门接待理论天文学家和宇宙学家的开放办公室中。虽然阿兰最早从事的是粒子物理学研究，但是他现在却是最优秀的宇宙学家之一。在我换办公室时，我已经对粒子物理学与宇宙学之间的联系展开了一些探索。当你的邻居和你有着相同的兴趣时，继续开展这样的研究会容易得多。而这些相同的兴趣还包括他与你一样，办公室也是乱糟糟的，你会觉得到他的办公室就好像到家里一样舒服。

许多粒子物理学家则迈得比一层楼远得多了，他们已经穿越到了各种不同的科研领域中。全球生物技术产生巨擘生物基因爱迪克公司（Biogen Idec）的创始人之一沃利·吉尔伯特（Wally Gilbert），也是粒子物理学家出身，但他后来转到生物学领域，而他获得诺贝尔奖的原因却是化学研究。许多粒子物理学出身的人也追随他的脚步转行到生物学领域。另一方面，我的许多研究生离开粒子物理学领域后去了华尔街做宽客（quant）^[74]，他们能为将来的市场变化做分析投注。他们只是选择了正确的时间采取这样的行动，因为用新的金融工具来对冲这种投注刚好适时地发展了起来。在转向生物学领域的过程中，一些思考和组织问题的方法可以平移过去；而在转向金融的过程中，一些方法和方程也是一样的。

粒子物理学与宇宙学的交叉比上面所说的两种情况更深刻、更丰富。一个对于宇宙在不同尺度上的更细致观察，为小尺度上的基本粒子与大尺度上的宇宙建立了许多联系。毕竟宇宙是独一无二的，它包含了一切。粒子物理学家看向物质的内部时，他们询问物质核心的基本物质组成是什么类型；而宇宙学家看向物质的外部时，他们思考世界的外部是什么、它们是怎样演化出来的。宇宙的奥秘（主要是它由什么组成）对于宇宙学家和粒子物理学家都一样重要。

两种研究人员都探索基本结构、应用基本的物理定律。每一方都要考虑另一方的研究成果。粒子物理学家所研究的宇宙成分也是宇宙学家的一个重要研究方向。更进一步地，将广义相对论与粒子物理学结合起来的自然定律描述了宇宙的演化，因为两种理论都正确而且都适用于宇宙。同时，弄清宇宙的演化可以限制存在于宇宙中的物质的性质，避免那些与观测到的历史相违背的假设。宇宙从某种意义上说，是第一个也是最强有力的粒子加速器。在其演化的早期，宇宙的能量与温度都非常高，今天加速器可以达到的高能量就是为了能在地球上重现宇宙当时的一些条件。

近几年各种研究兴趣都不约而同地汇聚到一起，产生了许多研究硕果并引发了对将来工作的启示，而这种势头还将持续下去。本章思考的是一些宇宙学中的重大问题，它们令宇宙学家和粒子物理学家一起来探索。这些交叉领域包括：宇宙暴胀、暗物质以及暗能量。我们将考虑这些现象中我们所知道的方面（更重要的是抛砖引玉），以及那些我们未知的方面。

宇宙暴胀理论

即使还不能知道宇宙最初发生了什么，由于需要一个包容性很强的理论来涵盖量子力学与引力，我们也可以相当确定地断言在非常早

的时期（也许在宇宙开始演化的前10-39秒），一个被称为“宇宙暴胀”的现象开始了。

1980年，阿兰·古斯率先提议这种情景，他认为非常早期的宇宙必不可少地经历了向外爆炸式的扩张。有趣的是，他最初试图解决的是一个粒子物理学问题，其涉及大统一理论的宇宙学效应。由于具有粒子物理学背景，阿兰所采用的方法根源于场论——将狭义相对论与量子力学结合起来的理论，粒子物理学家都采用该理论来做计算。他却推导出了一个全新的理论，革新了我们关于宇宙学的思考方法。暴胀是如何以及何时产生的仍处于猜想阶段，但是在这种爆炸式扩张驱动之下的宇宙应该留下清晰的证据，而许多证据现在已经被找到。



在标准的大爆炸情景中，早期宇宙的增长是温和而平稳的，例如，随着它的年龄增长4倍，它的尺度会翻倍。但是在一个暴胀时期，空间经受了令人难以置信的迅速膨胀期——随着时间增长，空间尺度以指数增长。宇宙在给定的一段时间内尺度翻倍，然后在同样的时间段尺度再翻倍，并且在后续的至少90个这样的时间区域上持续翻倍90次^[75]，一直到暴胀期结束，而当时的宇宙已经如我们今天所见到的一样均匀了。这种指数增长的膨胀意味着如果宇宙的年龄增长60倍，宇宙的尺度就会增长一万亿亿亿亿倍。而没有暴胀，尺度只会是原来的8倍。在某种意义上讲，暴胀是这个从小到大演化故事的起源，它至少是我们可以有机会通过观测来理解的部分。初始巨大暴胀式的扩张会将组成宇宙的物质和辐射稀释到几乎为零。因此我们现今所观测到的宇宙的一切必定来自于暴胀之后，驱动暴胀的能量重新转化成物质和

辐射。在该时间点上，传统的大爆炸演化才开始掌权，而宇宙开始了它的进一步膨胀，直到今天给出我们所见的结构。

宇宙暴胀

该理论认为，宇宙初期经历了一个暴胀时期，其速度非常快，使得宇宙的尺度在极短的时间内增大了几十个数量级。

我们可以将暴胀想成是“爆炸”，根据传统大爆炸理论，它是宇宙演化的前锋，它不是真正的开端。我们其实不知道量子引力起作用时发生了什么，但它标志着大爆炸的演化阶段开始——物质开始冷却以及最终聚合起来。

暴胀也部分地回答了为何“存在一些事物”而非“任何事物都不存在”。巨大能量密度中的一部分储存在暴胀中，其按照 $E=mc^2$ 公式转化成物质，这些物质后来演化成我们今天看到的物质。在本章结尾时我会提到，物理学家仍希望知道为什么宇宙中的物质比反物质多得多。但是不管对此问题的解答如何，根据大爆炸理论的预测，我们知道的物质在宇宙暴胀结束之后就立即开始演化了。

暴胀是一个“自下而上”的理论。它解决了传统大爆炸理论中的重要问题，但只有很少人相信有关它是如何出现的任何模型。目前，还没有令人信服的高能理论能明显蕴含暴胀理论。因为建立一个可信的模型实在是太具有挑战意味，所以许多物理学家（包括那些当我还是研究生时就已经在哈佛大学做研究的人）怀疑这种想法。另一方面，在斯坦福大学任教的俄裔美籍物理学家安德烈·林德（Andrei Linde），是最早从事暴胀理论研究的人员之一，他认为暴胀理论必然是正确的，仅仅因为迄今为止还没有人能发现其他任何可以解决宇宙尺度、形状与均匀性的理论，而暴胀理论都能给出解释。

暴胀是一个关于真与美连接或其缺乏连接的有趣实例。宇宙的指数级增长优美而简洁地解释了许多有关宇宙演化的现象，而在理论上研究那些可以自然导出指数级增长的模型又不怎么优美。然而最近，绝大多数物理学家，即使还不满足于大多数模型，也开始相信暴胀理论或者某些类似于暴胀的理论。近几年的观测已经明确了大爆炸宇宙学图像的早期呈现暴胀。许多物理学家现在相信大爆炸演化和暴胀都发生了，因为基于这些理论的预测已经得到了可观的精确鉴定。暴胀理论背后的真实模型还没有得到解决，但是目前指数级增长已经得到了很多证据支持。

宇宙暴胀的一种证据与偏离完全均匀性有关，它由第19章介绍的宇宙微波背景辐射给出。背景辐射可以告诉我们的不仅仅是大爆炸。它的美在于，它本质上是宇宙非常早期（比恒星的形成更早）的一张快照，让我们可以直接回顾到宇宙仍然光滑的初始结构。宇宙微波背景测量也揭示了一些与完全均匀性的偏离。暴胀预言了该偏离，由于量子涨落使得暴胀在宇宙不同区域结束的时间有些许差别，因此造就了对于绝对均匀的微小偏差。设置在卫星上的威尔金森微波各向异性探测器（Wilkinson Microwave Anisotropy Probe, WMAP），它以开创了此项目的普林斯顿物理学家大卫·威尔金森（David Wilkinson）命名。它作出了细致的测量，使得暴胀的预测可以与其他可能性区分开来。尽管暴胀发生在很久以前，那时宇宙的温度极高，基于暴胀宇宙学的理论却预言了温度变化规律的精确统计性质，它在今天的天空上印刻下了辐射的图像。威尔金森微波各向异性探测器测量在温度与能量密度上微小的非均匀性，它比以往的探测的精度更高，角度范围更小。测量得到的图案与暴胀理论的预期一致。

威尔金森微波各向异性探测器提供给的关于暴胀的主要确证是关于宇宙极端平坦性的测量结果。爱因斯坦告诉我们，空间可以是弯曲的（见图20-1中的弯曲二维表面）。曲率由宇宙的能量密度决定。当

暴胀最早被提出时，宇宙被认为膨胀得太快以至于任何曲率都会被消耗殆尽，然而那时的测量精度远不能检验暴胀的预言。微波背景测量现在证明，宇宙在1%的精度上是平坦的，如果没有一个基本的物理解释，那这一点令人难以置信。

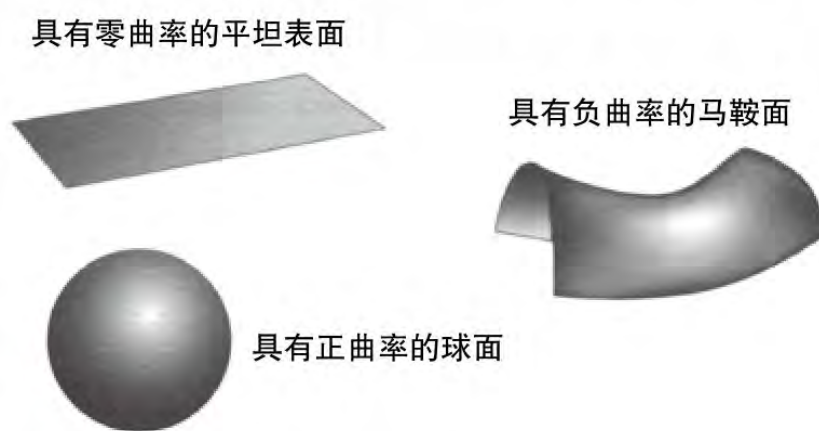


图20-1 具有零、正、负曲率二维表面。宇宙也可以是弯曲的，但它具有四维时空，很难画在图上。

宇宙的平坦性是暴胀宇宙学取得的一个巨大胜利。如果宇宙不是那么平坦的，那么暴胀理论就已经被否决了。威尔金森微波各向异性探测器测量也是科学史上的一次胜利。当理论物理学家首先提出具体测量微波背景的方案——即最终可以告诉我们宇宙几何的方法，每一个人都认为它的趣味性足够引起科学界的注意，然而在短期内要获得成果从技术上讲太困难了。10年间，伴随着对所有期望的困惑，观测宇宙学家终于作出了必要的测量，并且提供给了我们宇宙如何演化的惊人启示。威尔金森微波各向异性探测器仍将提供新的结果，对全天温度变化展开细致的测量。而现在运行的普朗克卫星正在以更高精度测量这些量子涨落。宇宙微波背景的测量已经被证明是我们洞察早期宇宙的主要资源，而且它还将继续测量下去。

近来宇宙学家细致地研究了全天遗留的宇宙辐射，使得我们对宇宙及其演化的量化知识发生了又一次巨大的跳跃。辐射的细节已经提供了关于我们周遭的物质与能量的丰富信息。宇宙微波背景除了告诉我们光线开始向我们发射的条件，还告诉我们光线所穿越的宇宙信息。假如宇宙在过去137.5亿年有所改变，或者它的能量与我们所期望的不同，那么相对论告诉我们，它会影响光线所走的路线，并且影响所测到的辐射的结果。因为它是一个十分敏感的宇宙能量探针，微波背景向我们提供了宇宙组成的信息。这些组成还包括我们接下来要介绍的暗物质与暗能量。

暗黑之心


宇宙微波背景除了成功地确证了暴胀理论，它的测量还呈现给宇宙学家、天文学家和粒子物理学家几个主要的谜题。暴胀告诉我们，宇宙应该是平坦的，但它没有告诉我们使宇宙平坦的能量在哪里。不管怎样，基于广义相对论的爱因斯坦方程，我们可以计算出使宇宙平坦的能量，然而已知的可见物质只提供了所需能量的4%。

另一个谜题是，能够解释对宇宙微波背景探测器测量的温度与密度的微小涨落的新事物。只有可见物质以及如此微小的扰动，宇宙不可能得以坚持如此长时间，并使这些扰动可以增长得如此巨大以至于足够产生现在的结构。星系、星系团与测量到的微小涨落联合起来共同指向一种人类所未见的物质存在方式。

事实上，远在宇宙微波背景探测器微波背景结果出来之前，科学家已经知道新的被称为暗物质的物质类型应该存在。我们即将介绍的其他观测已经暗示了另一种不可见物质必然存在。这种神秘物质，现在被称作暗物质，通过引力相互作用，但是与光没有相互作用。因为它既不发射光也不吸收光，它是不可见的而并非是暗的。暗物质（我

们还将使用这个名称) 目前除了引力效应(并且还很弱), 没有其他可以触及它本质的性质特征。

更有甚者, 引力效应和测量显示, 存在一种比暗物质还要更神秘的物质, 我们称其为暗能量。这种能量弥漫在整个宇宙中, 但是又不像一般的物质可以塌缩成团或者当扩张时变得稀薄。它很像使暴胀停止的能量, 但是它今天的密度也比那时低了许多。



虽然我们正活在一个宇宙学复兴时期, 理论与观测已经发展到可以精确检验一个想法的程度, 但是我们也同样生活在一个暗黑年代。大约23%的宇宙能量是暗物质, 另外73%是神秘的暗能量(见图20-2)。

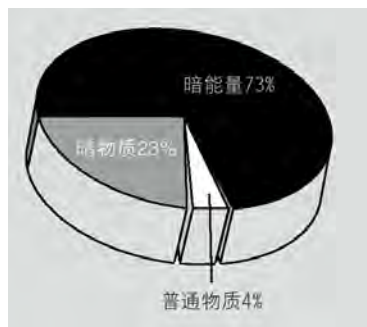


图20-2 宇宙组成中普通物质、暗物质与暗能量的比例。

上一次物理中出现一个被冠以“暗”字的东西还是在19世纪中期。当时, 法国的勒维耶(Urbain Jean Joseph Le Verrier)提出, 有一个不可见的暗星球, 并将其称为瓦肯星(Vulcan)。勒维耶的目标是解释水星的奇特轨道。勒维耶与英国的约翰·亚当斯(John C. Adams)此前曾经用海王星对天王星的影响推导出海王星的存在。然而对

于水星的考虑，勒维耶完全错了。水星那奇异轨道的原因比存在另一个行星更令人激动。该解释只有在爱因斯坦建立相对论理论之后才被找到。事实上，广义相对论的第一个证据就是它可以正确预测水星的轨道。

暗物质与暗能量可能是一些已知理论的结果。但是也可能这些宇宙失落的元素预示了重大的图景转变。只有时间能告诉我们什么样的选择将解决暗物质与暗能量的问题。

即使如此，我还是认为暗物质很可能有更多的传统解释，与我们现在所知的物理定律相容。即使新物质按照我们所熟知的力的定律来相互作用，为什么所有物质都会呈现出与我们熟悉的普通物质一样的性质呢？简而言之，为什么所有物质都会与光有相互作用呢？如果说科学史曾经给过我们何种启示，那么相信“所见即所得”这种想法就是极其短视的。

许多人却不这样认为。他们发现暗物质的存在非常神秘，并疑问：怎么可能绝大多数物质（是我们所能见到的物质的6倍）是我们无法用常规的望远镜观测到的？某些人甚至怀疑暗物质只是某种错误观念。我个人认为恰恰相反（虽然我承认不是全部的物理学家都能这样看待）。如果我们能看到的竟然是所有存在的物质，那么这可能更加令人吃惊。为什么我们就偏偏有最完美的感官来直接观测到一切的事物呢？并且物理学几个世纪的教训告诉我们：有多少东西曾被隐藏而未被人类所看见。从这个角度看，为什么我们所知道的物质竟然占据了所有物质能量的1/6之多？这是一个巧合吗？目前我和同事正努力想要理解这一点。

我们知道，有些具有暗物质特性的东西必须存在。虽然我们并没有严格“看见”它，却探测到了暗物质的引力效应。我们知道暗物质存在，因为不断拓展的观测证据表明它的引力效应存在于宇宙中。其存

在性的第一个线索来自星系团中恒星的转动速度。1933年，弗里茨·兹威基（Fritz Zwicky）观测到星系团中的星系转动的速度比星系中只有可见质量造成的转动要快。很快，奥尔特（Jan Oort）观测到银河系也有类似的现象。兹威基很相信他的工作，他猜想宇宙中存在人们无法直接看到的暗物质。但是这两个观测结果并不是确凿无疑的。一个错误的测量或者一些其他星系的动力学看似更可能对此作出合理的解释，这比一些创造出来仅仅为了提供更多引力效应的不可见物质更合理。

在兹威基做此测量时，他还没有足够的分辨率来看到单个的星体。更多有关暗物质的坚实证据来自观测天文学家薇拉·鲁宾（Vera Rubin），20世纪60年代末70年代初，鲁宾对星系中的恒星转动做了非常细致的定量测量。起初看似无聊的研究结果，第一次作为暗物质存在的坚实证据出现。而鲁宾的研究在当时与其他天文学活动相比，属于鲜少有人问津的领域。鲁宾与肯特·福特（Kent Ford）的观测为兹威基早年的结论提供了确凿的证据。

你也许会好奇，人们如何通过望远镜来看到暗的东西。答案是，它可以看到这些物质的引力效应。星系的性质，例如它里面的恒星的转动速率，是由星系所含的物质决定的。如果只有可见物质，那么人们将认为，位于星系外围的恒星对星系引力的敏感程度会降低。然而，与位于中心附近的亮物质相比，远上10倍的恒星，仍然以与靠近中心的恒星一样的速度转动。这意味着质量密度并没有随着距离的增加而降低，至少在从星系中心到发光物质10倍远的距离上并没有降低。天文学家因此认为，星系主要由不可见的暗物质组成。我们可以看见的明亮物质固然是一个相当可观的部分，但星系中绝大多数物质都是不可见的——至少从字面意义上说是这样。

我们现在有许多其他补充证据来证明暗物质的存在。一些最直接的证据来自透镜效应（见图20-3）。透镜效应是光线掠过一个质量巨大的物体时发生的现象。即使该物体本身不发光，它也对光线施加引力。引力会造成物体后面（从我们的角度来看）一个发光物体所发出的光线的弯曲。光线在不同方向的弯曲取决于它掠过暗物质的路径，透镜会在天空呈现原始物体的多个镜像。通过推导使这些可见光弯曲所需的引力，这些镜像让我们可以“看见”暗黑物体或者至少暗示了它的存在和性质。

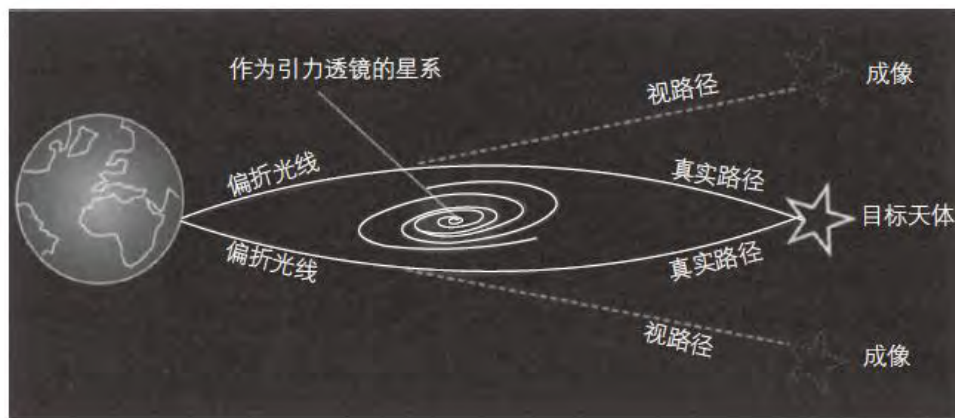


图20-3 光线掠过一个重的物质时，会发生弯曲，从一个观测者的角度，出现了多个原始天体的像。

也许目前关于暗物质最强有力的证据（而非修改引力理论）解释了来自子弹星系团（Bullet cluster）的现象，它涉及两个星系团的碰撞（见图20-4）。它们的碰撞证明，明星系团包含恒星、气体以及暗物质。星系团中的炽热气体强烈地相互作用，强到以至于气体始终

聚集在碰撞的中心。而暗物质则不发生相互作用，至少作用不强。因此暗物质仅仅会彼此穿过。透镜测量显示，暗物质与炽热气体是分开的，模型给出的暗物质的相互作用非常弱，而普通物质的相互作用很强。



图20-4 子弹星系团暗示，星系团中包含暗物质，它们的动力学不太可能被修正的引力定律解释。我们可以看到，当两个星系团碰撞时，两团强烈地相互作用的普通物质被困在中心，而相互作用弱得多的暗物质（通过引力透镜探测到的）则会明显地互相穿过，它们两者之间有一个分离。

我们还有暗物质存在进一步的证据，它来自前面提到的宇宙微波背景辐射。与透镜效应不同，辐射的测量没有告诉我们任何暗物质的分布。相反，它们告诉了我们暗物质的总能量组成（见图20-2）。

宇宙微波背景测量告诉了我们关于早期宇宙的大量信息，并给我们提供了有关它性质的很多细节。这些测量不是单为了暗物质而设的，它们也为暗能量的存在提供了证据。根据爱因斯坦的广义相对论

方程，宇宙只能在拥有了正确的数额的能量时，才能是平坦的。物质（哪怕将暗物质也计算在内）仍然不够给出威尔金森微波各向异性探测器和气球探测器所观测到的平坦度。因此其他能量必须存在。暗能量是既解决了宇宙平坦性——三维空间没有可以观测到的曲率，又与目前所有其他观测相符的唯一方法。

暗能量占据了宇宙主体能量的大约70%，比暗物质更令人费解。让物理学界相信暗能量存在的证据是宇宙目前正加速膨胀的发现，类似于宇宙早期的暴胀但速度慢很多。在20世纪90年代，两个独立的研究组超新星宇宙学计划（Supernova Cosmology Project）和高红移超新星搜索队（High-z Supernova Team）发现，宇宙膨胀速度不是降低而实际上是在增加，这个发现令整个物理学界震惊。

在超新星测量之前，也有几个指向存在能量缺失的暗示，但这些证据都太弱了。20世纪90年代，精确测量显示远距离的超新星比预期的暗淡很多。因为特殊类型的超新星有相当均匀的和可预期的发光度，上述观测必须通过一种新的机制来解释。这种新机制就是宇宙的加速膨胀——即膨胀的速度越来越快。

加速度不可能来自普通物质，因其引力吸引将减缓宇宙的膨胀速度。而唯一的解释只能是宇宙的行为像暴胀时期的行为，只是能量比其在暴胀时期所具有的能量低了很多。这种加速度应该是源于某种类似于爱因斯坦曾经提出的宇宙学常数或者说暗能量。

与暗物质不同，暗能量对于其所处的环境施加负压。普通的正压将导致向内塌缩，而负压则导致向外加速膨胀。^[1]负压最明显的候选者（与迄今为止的测量相符）是爱因斯坦的宇宙学常数，代表了充满宇宙而物质中没有的能量和压强。暗能量是我们所使用的更一般的术语，它允许一种可能性——宇宙学常数所假设的能量与压强之间的关系不是完全正确而只是近似正确。

暗能量在宇宙能量中占主导地位。这更为显著，因为暗能量密度极小。暗能量在过去的几十亿年都起着主导作用。早期宇宙演化中，辐射和物质分别占主导地位，但两者在宇宙不断增大的体积中不断被稀释。然而即使宇宙在不断增大，暗能量却保持不变。当宇宙的寿命达到一定长度时，辐射与物质的能量密度已经降低得很厉害，暗能量由于没有被冲淡，终于占据了主导地位。尽管暗能量不可思议的小，最终它也变成了主导因素。在100亿年缓慢增长的膨胀中，暗能量的影响也最终积累起来而使得宇宙开始它的加速膨胀。最终宇宙将终止于空无一物的状态而只剩下真空能，并且它的膨胀还会相应地加速下去（见图20-5）。这个温和的能量可能不会接管地球，但它却在逐渐地接管整个宇宙。

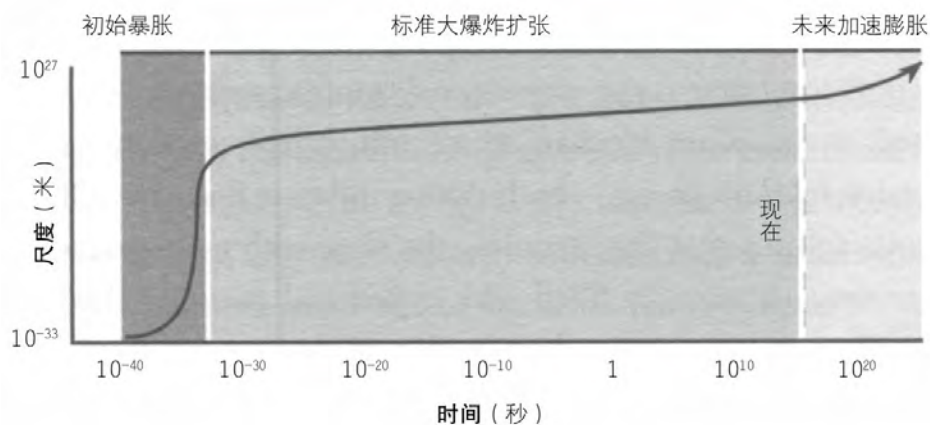


图20-5 宇宙在不同时期的膨胀不同。在暴胀期，它按指数快速增长；暴胀结束后，传统的大爆炸式膨胀开始；暗能量现在又重新使膨胀速度加速增长。

能量、宇宙物质与暗物质谜题

研究暗能量与暗物质的必要性告诉我们，我们不能因为宇宙学理论与宇宙学数据的暗示惊人地相符，就沾沾自喜于我们对宇宙演化的

理解。宇宙中绝大多数物质的本质还都是一个迷。20年后的人们可能会嘲笑我们今天的无知。

这些还不是由宇宙的能量引出的唯一谜题。特别是，暗能量的数值实际上是一个更大谜题的细枝末节：为什么充满了宇宙的能量如此小？暗能量的数量如果大一些的话，它就会在宇宙演化的更早时期取代物质与辐射的主导地位，而结构（以及生命）都不会有足够的时间来形成。最重要的是，没有人知道早期引发和提供了暴胀的巨大能量密度是从哪里来的。但是宇宙能量最大的问题是宇宙学常数问题。

基于量子力学，我们期望的是一个数值上大得多的暗能量，无论在暴胀期间还是现在。量子力学告诉我们真空（没有永恒的粒子的状态）实际上充满了转瞬即逝的粒子，它们随时可以产生和湮灭。这些寿命极短的粒子可以具有任何能量。它们有时能量大到其引力效应不能再被忽视。这些高能粒子向真空贡献了极高的能量——比经历了漫长演化的宇宙所允许的大得多。为了让宇宙看起来与我们所看到的相同，真空能的大小需要比量子力学所允许的数值惊人地小上120个数量级。

关于这个问题还有进一步的挑战。为什么我们恰好生活在物质、暗物质与暗能量可以相提并论的年代？暗能量比物质^[76]的比例高，但是相差少于三倍。鉴于原则上能量有完全不同的起源，并且任何一方都曾经占据过主导地位，它们的密度现在却显得如此接近，这个事实非常不可思议。这个巧合的奇异性非常显著，因为它只有（粗略地说）在我们的时代，才发生了这样的巧合。在宇宙更早时期，暗能量所占的比例非常小；而以后它所占的比例又将变得很大。只有现在这三个部分：普通物质、暗物质与暗能量的比例是相当的。

不过，为什么能量密度极小、为什么不同的能量来源贡献了比例相当的部分，这些问题目前完全没有解释。事实上，一些物理学家相

信不存在真正的解释。他们认为我们生活的宇宙有如此令人难以置信的真空能，稍大一点就会阻止宇宙中星系与结构（包括人类在内）的形成。那么我们今天也不会存在于这里，在这里询问具有稍微大一点宇宙学常数的任何宇宙的能量值问题。那些物理学家相信有许多宇宙，而且每一个宇宙都具有不同的暗能量。在许多可能的宇宙中，只有能给出现有结构的这一个才可以创造出人类。该宇宙中的能量值超乎寻常得小，但是我们只能存在于恰好具有这么小能量值的宇宙中。它的原因就是我们在第18章考虑的人择原理。就像在第18章说过的，我并没有被说服。无论如何，我或者其他人都没有更好的解答。暗能量的值可能是粒子物理学家和宇宙学家目前所面临的最主要的疑难。



除了能量之谜，我们还有关于物质的一个更进一步的宇宙学疑难：为什么宇宙中会有物质？我们的问题源于物质与反物质有着相同的基础。物质与反物质在碰到对方时同时湮灭，两者都消失。当宇宙降温时，物质与反物质都不应该保留下来。

虽然暗物质的相互作用很微弱，因此可以在宇宙中飘荡，但是普通物质通过强相互作用力的相互作用却很强。如果没有对于标准模型的额外添加，几乎所有的普通物质都会在宇宙冷却到现在的温度时消失殆尽。物质可以存留下来的唯一原因是物质比反物质多了很多。我们的理论中最早没有这样的设定。我们需要找到质子存在而反质子不存在的

原因。因此必须引入一种物质-反物质的不对称性。

剩下来的物质总量比暗物质的少，但它仍然是宇宙中可观的一部分，更不用说它还是我们所知和所爱的一切东西的来源。何时、如何产生了物质-反物质的不对称性？这是粒子物理学家与宇宙学家非常想要解决的另一个大问题。

当然，是什么组成了暗物质也是至关重要的一个问题。也许我们最终能发现将暗物质与物质密度相关联的基本模型，如同近来的一些研究所提议的。无论如何，我们希望可以尽快从实验了解更多有关暗物质的问题，这些实验是我们现在探索的一个样本。



大型强子对撞机的首席工程师林恩·埃文斯在2010年1月加州大型强子对撞机/暗物质会议中所做报告的结束语是，在过去的20年，“你们理论物理学家在暗黑（区）中瞎白忙活了一场”，以此来逗乐听众。他还顺便附加了一个说明，“现在我理解了为什么我过去花了15年来建造大型强子对撞机”。林恩的评语暗指在过去的年头里一直缺乏高能物理实验数据。但他也暗示了大型强子对撞机的发现有可能将暗物质显现出来。

粒子物理学与宇宙学存在许多联系，其中一个最令人好奇的方面是，暗物质也许可以在大型强子对撞机能探索到的能标上被制造出来。如果存在一种具有弱能标质量的稳定粒子，这种粒子从早期宇宙存留到今天，那么这种粒子刚好带有可以作为暗物质的等值能量。从最初炽热而渐渐冷却下来的宇宙中残留下来的暗物质，对它的计算有可能证明这个想法是正确的。这不仅意味着暗物质真实存在，其身份也可以得到证明。如果暗物质确实由这种弱能标质量的粒子组成，那大型强子对撞机也许不仅能提供有关粒子物理学问题的启示，也能解答一些与宇宙学有关的问题，如宇宙中存在的物质以及它们是如何形成的。

大型强子对撞机实验不是唯一研究暗物质的方法。物理学现在进入了一个有潜质的数据时代，不仅针对粒子物理学，也针对天文学与宇宙学。本章将介绍实验如何采取三管齐下的方式寻找暗物质（见图2 1-1）。首先，探索为什么具有弱能标质量的暗物质粒子最受青睐。其次，探索如果这个假设正确，那么为什么大型强子对撞机可以制造和确认暗物质粒子。我们接着会考虑，为寻找暗物质粒子而量身定制的实验如何在它们到达地球时发现它们，并记录下它们微弱而又可以探测到的相互作用。最后，我们将考虑在地面上与在太空中的望远镜以及探测器的运作方式，它们如何寻找在空中湮灭的暗物质粒子的产物。

透明物质

我们知道暗物质的密度，它的温度低（也即它运动的速度相比光速而言很低），它的相互作用极其微弱，与光也没有明显的相互作用。暗物质可以说是透明的。暗物质的质量，它是否具有除引力之外的其他相互作用，它如何在早期宇宙中产生，这些我们都不知道。我们只知道暗物质的平均密度，在我们的星系中可能每立方厘米有一个质子质量的暗物质，或者在一个致密物体中该数值也可能为1000万亿。两种情况都可以给出暗物质相同的平均密度，但两种情况都不能给出关于结构形成的任何信息。

因此，即使知道暗物质存在，我们也不知道暗物质的本质。它可以是小型黑洞或者来自额外维度。最可能的情况是，它仅仅是一种新型基本粒子，不具有标准模型的相互作用；它或许是一种稳定的电中性残留物质，会在一个即将被发现的弱能标物理理论中出现。即使的确如此，我们也想知道暗物质粒子的物理性质：质量、相互作用、它是不是一种更广泛的粒子族群的一员。

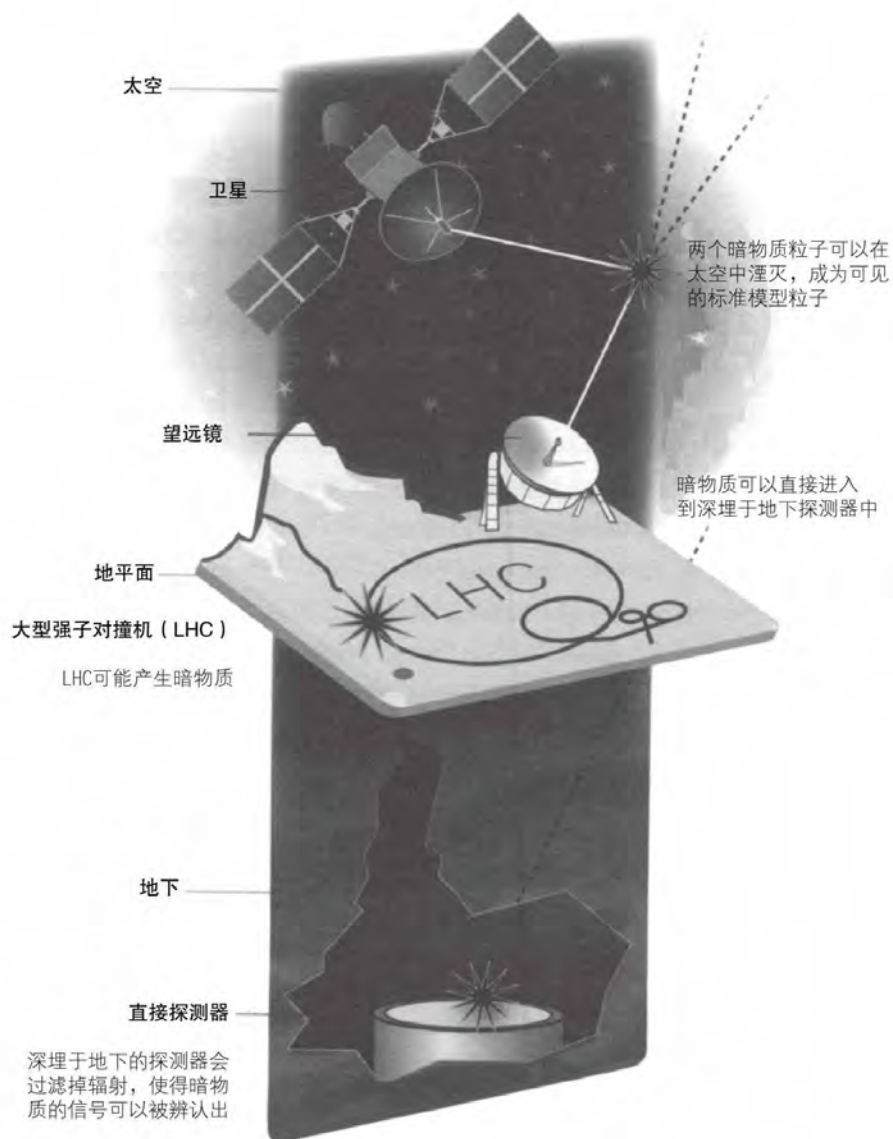


图21-1 三管齐下搜寻暗物质。地下探测器主要寻找直接撞击靶核的暗物质；大型强子对撞机也可能产生暗物质，因此可能在实验装置中留下证据；卫星和望远镜则可能发现由于太空中的暗物质湮灭而产生可见物质的证据。

基本粒子这种解释受到人们青睐的一个原因是：上文所暗示的暗物质的丰度，它所携带的能量份额对于这种假设是一个支持。令人惊诧的一个事实是：一种稳定粒子的质量（根据 $E=mc^2$ ）如果处在大型强

子对撞机将开始探索的弱能标的量级，那么它的残留密度也即宇宙中该粒子所携带的能量，恰好与暗物质处于同一区域。

这里的逻辑如下。随着宇宙的演化，温度将会下降。当宇宙处于炽热状态时，重的粒子曾一度丰富。随着宇宙冷却下来，这些粒子逐渐被消耗，因为低温状态的能量不足以产生它们。一旦温度下降得足够低，这些重粒子与其反粒子湮灭使得两者均消失，但是反过来的过程——即它们成对产生，却不再以一个可观的速度出现。因此，由于湮灭，随着宇宙冷却，重粒子的粒子数密度会急剧下降。

当然，为了湮灭，粒子与反粒子首先必须“兵戎相见”。^[77]但是当数目降低时，它们变得非常稀缺，湮灭也变得更困难。结果粒子湮灭也随着宇宙的演化变得更不可能，因为这要求至少两个粒子必须处在相同的位置。

结果是，弱能标粒子基本更稳定，可以存留到今天，而不像纯粹使用热力学所揭示的那样，原因就是从小时刻开始粒子与反粒子已经非常稀薄，它们不能相遇并湮灭。到如今还有多少粒子遗留下来取决于假定的暗物质候选者的质量与相互作用。如果物理学家能知道这些性质，那么我们就可以知道如何计算残留丰度。而令人困惑而又显著的事实是，剩余下来的稳定的弱能标粒子刚好给出了与暗物质一样的丰度。

我们既不知道粒子精确的质量，也不知道它们的相互作用（更不用说包含稳定粒子的模型），因而我们还不知道前面说到的数目是否有用。然而在浮出水面的两种截然不同现象的两个数据之间，这种偶然的一致性（虽然很粗糙）的确令人很困惑，但它又可能是弱尺度物理可以被用来解释宇宙中暗物质的一个信号。

广为人知的暗物质的候选者是一种被称为WIMP的粒子，其全称是大质量弱相互作用粒子（Weakly Interacting Massive Particle）。这里“弱”只起到描述的作用而不特指弱相互作用，大质量弱相互作用粒子可能具有的相互作用比标准模型中参与弱相互作用的中微子的作用还要更弱。如果没有关于暗物质及其性质更直接的证据（比如大型强子对撞机可以揭示的），那么我们就不能知道是否暗物质的确由大质量弱相互作用粒子组成。因此实验搜寻是我们接下来所关注的话题。

大型强子对撞机中的暗物质

宇宙学家很好奇弱能标上的物理以及大型强子对撞机可能发现的东西，产生暗物质是一种有趣的可能。大型强子对撞机的能量水平刚好能寻找大质量弱相互作用粒子。如果暗物质的确如计算所提议的由弱能标上的粒子组成，那么它也许能在大型强子对撞机中产生。

即便如此，暗物质粒子也不一定被发现。毕竟，暗物质不怎么与其他物质发生相互作用。由于它们与标准模型物质的相互作用有限，暗物质粒子不能直接产生或者直接被探测到。即使产生了，它们也仅会穿过探测器。然而，不是所有的粒子都会逃掉（哪怕暗物质粒子将会逃掉）。关于等级问题的任何解释都包含除了标准模型的其他粒子——它们绝大多数具有的相互作用都很强。一些粒子可能产生得很多，接着它们可以衰变成暗物质粒子，进而带走无法测量到的动量和能量。

超对称模型是这一类型中研究得最透彻的弱能标模型，它自然包含一种切实可行的暗物质候选者。如果超对称在自然中存在，那么最轻的超对称粒子可能可以组成暗物质。这种不带电荷的最轻粒子，相互作用太弱以至于靠自己无法产生足够多的粒子来被发现。然而，胶

微子（传递强相互作用力的胶子的超对称伙伴）和超夸克（夸克的超对称伙伴）假如存在，那么可以在正确的能标上产生。如第17章所介绍的，这些超对称粒子最终都会衰变成最轻的超对称粒子。因此，即使暗物质粒子不能直接被制造出来，其他激增的粒子也能够以可以观测到的速度产生出最轻的超对称粒子。

其他弱能标暗物质的情景如果有可以探测的效应的，基本上也需要以相同的方法产生以及被“测量”。暗物质粒子的质量应该大约在大型强子对撞机可以研究的弱能标。因为粒子微弱的相互作用强度使得它们不会直接被产生出来，但是许多模型都包含其他可以衰变成它们的新型粒子。由此，我们也许可以得知暗物质粒子的存在，并且通过它们带走的能量，还可能知道它们的质量。

若大型强子对撞机发现暗物质，这当然会是一个重大的成功。如果真的发现了，那么实验物理学家甚至可以研究它的具体性质。然而，要确信大型强子对撞机发现的粒子的确组成暗物质还需要附加的证据。它们也许可以由地面上的以及太空中探测器提供。

直接探测暗物质的实验

大型强子对撞机有制造暗物质的潜力这件事的确很吸引人，但是许多宇宙学实验不是在加速器中做的。在地面与太空中的天文实验与暗物质搜寻，才是解释和推进我们对宇宙学问题理解的最主要手段。

当然，暗物质与普通物质的相互作用非常微弱，因此当前的寻找基于一个信仰——暗物质虽然几乎不可见，与已知物质（包括探测器也是由它们所造的）的作用微弱，但却不是没有作用。这并非仅仅是一个美好的愿望。它其实基于我们前面提到的关于残留密度的计算，该计算显示，如果暗物质与解释等级问题的模型相连，那么遗留下来

的粒子密度恰好可以给出暗物质观测的解释。这种计算提议的许多大质量弱相互作用粒子暗物质候选者与标准模型粒子发生相互作用的速率，也许用目前的暗物质探测仪可以探测到。

即使如此，因为暗物质的微弱相互作用，这种搜寻需要使用地面上的巨型探测器或者非常灵敏的仪器，以在地面上或者太空中寻找暗物质作用、湮灭以及产生新粒子与反粒子的产物。如果你只买一张彩票的话，那么你很可能中不了奖，但是如果你可以买到超过一半的彩票，那么你的胜算将很高。类似地，许多大型探测器增加了寻找暗物质的概率，哪怕暗物质与探测器中的任何单一核子的相互作用极其微小。

对于暗物质探测器来说，一个巨大的挑战是检测中性（不带电）的暗物质粒子，而且还要将它们与宇宙射线或者其他背景辐射区分开来。不带电的粒子与探测器没有传统意义上的相互作用。暗物质穿过探测器的唯一踪迹将是当它撞击探测器中的核子时，使核子的能量发生一点点微小的改变。这是唯一可以观测到的效应，所以探测器除了寻找暗物质粒子通过时产生的微小热量或者反冲能量的证据之外，别无他法。因此探测器设计成要么温度极低要么非常灵敏，就是为了记录暗物质粒子反弹时的一点点热量或者能量沉积。

这种极冷装置被称为低温探测器（cryogenic detector），当暗物质粒子进入仪器时，它们可以探测其发出的很少量的热量。一个小额的热量进入已经很热的探测器中将很难被发现，但是如果进入的是经过特殊设计的很冷的探测器的话，情况则截然不同，极其微小的热量仍然可以被吸收和记录。低温探测器是采用诸如锗的晶体吸收器所制成的。这类型的实验包括：低温暗物质搜寻计划（Cryogenic Dark Matter Search, CDMS）、CRESST与EDELWEISS^[78]。

其他类型的直接探测实验涉及惰性液体（noble liquid）探测器。即使暗物质不直接与光发生作用，由于碰撞加入到氙（xenon）或者氙原子中的能量也会引起一种特征发光。有关氙的实验包括XENON100和LUX^[79]，另一个惰性液体实验包括ZEPLIN计划^[80]与氩暗物质实验（ArDM）。

理论与实验物理学界的每一个人都盼望着知道这些实验的新结果将是怎样。2009年，我有幸参加了在加州大学圣塔芭芭拉分校卡弗里理论物理研究所（KITP）召开的一次暗物质会议。两位领头的暗物质专家——道格拉斯·芬克拜纳（Douglas Finkbeiner）与尼尔·韦纳（Neal Weiner）也参与了会议。那时低温暗物质搜寻计划是最灵敏的暗物质探测实验之一，正要公布最新结果。除了同时作为个头很高的年轻一代且一起博士毕业于伯克利大学之外，他们还都对暗物质实验及其可能的含义有着深刻的理解。尼尔有更深的粒子物理学背景，而道格拉斯则做了更多天体物理的研究，但当暗物质的研究表明它涉猎这两个领域，最终两人的研究走到了一起。在会议上，他们搜集了这个领域的理论与实验上最前沿的专业知识。

我到达的那天上午有一个最精彩的报告。加州大学圣塔芭芭拉分校的哈利·尼尔森（Harry Nelson）教授介绍了低温暗物质搜寻计划的结果。你可能不理解为什么一个讲述老结果的报告会引人注目。原因是会议的每个人都知道三天以后该实验将公布新的数据。一些传言说，低温暗物质搜寻计划的科学家实际上看到了对于某个探索的令人信服的证据，因此每个人都想更好地理解该实验。许多年来理论物理学家一直在听说关于暗物质的探测，但主要是听说它们的结果，并且只对细节有肤浅的关注。但是在可以想象的即将出现的暗物质探测结果出来之前，理论物理学家很渴望了解更多。之前一个星期，结果公布了，却令对此抱有极大希望的听众非常失望。但是在哈利的报告期间，每个人都聚精会神。尽管有许多关于即将发表的结果的尖锐问题，他仍可以坚定地讲述他的报告。

因为那是一个时长两小时的非正式演讲，与会者可以根据理解的需要随时提问题。该报告很好地总结了听众——主要是粒子物理学家，可能会

感到困惑的问题。哈利是粒子物理学家而不是天文学家出身，他讲的方式我们很受用。

在这些极其艰难的暗物质实验中，最折磨人的是这些细节。哈利将这一点表述得非常清楚。低温暗物质搜寻计划基于一种高端的低能标物理技术——往往更多地被凝聚态物理或者固态物理学家使用的那种技术。哈利告诉我们在参加到此项合作之前，他根本无法相信这种精巧的探测竟然可能奏效，并开玩笑地说，他的实验学同事应该庆幸他不是这项原始议案的评审。

低温暗物质搜寻计划的运行与闪烁氙（scintillating xenon）和碘化钠（sodium iodide）探测实验不同。它有一个材料为锗或者硅的曲棍球大小的部件，其顶端是一个精妙的记录装置，这是一个声子传感器。探测器在很低的温度下运作，温度介于超导态与非超导态之间。如果哪怕只有一点点声子（phonon）的能量，那么声音的最小单位所携带的能量在通过锗或者硅时（就像光的最小单位光子撞击探测器时），可以使该装置离开超导态，从而通过一个超导量子干涉仪（Superconducting Quantum Interference Device, SQUID）记录到一个可能的暗物质事件。这些装置极其灵敏，它们可以非常精确地测量能量沉积。

然而，记录一个事件并不代表故事的终结。实验物理学家需要确认记录到的是暗物质而不是背景辐射。问题是任何东西都发出辐射。我们发出辐射，我正在打字的计算机也发出辐射，你正读的书（或者电子设备）也发出辐射。从一个实验物理学家的手指流出的汗也足以将暗物质的信号淹没。而这些还没有包括所有原始的和人造的发射源物质。环境和空气以及探测器自身都携带一定的辐射。宇宙射线也可以击中探测器。岩石中的低能中子也可以模仿暗物质。宇宙射线 μ 子可以击中岩石从而飞溅出一堆的物质，其中包含的中子也可以模仿暗物质。与信号事件相似的起码有超过一千种背景电磁事件，甚至在关

于暗物质质量和相互作用强度方面有着合理与乐观的假设的情况下，也是如此。

为暗物质实验所设计的游戏名称是“屏蔽和区分”（shielding and discrimination，这是天文学家的术语。粒子物理学家使用更个性化的术语“粒子身份” [particle ID]，虽然现在我也不知道哪一种方法更好）。实验物理学家需要将探测器尽可能地屏蔽起来，将辐射隔离在外面，并且将潜在的暗物质事件从我们不感兴趣的辐射散射中分隔开来。通过将实验安置于地下的深井中，屏蔽的目的完成了一部分。该想法就是让宇宙射线击中探测器周边的岩石而不击中探测器。暗物质因为有着更低的相互作用，可以不受岩石阻碍到达探测器。

对于暗物质的探测来说幸运的是，世界上恰好有许多矿井和隧道。DAMA实验与XENON10和它的加强版XENON100以及CRESST都是用钨做成的探测器，同属大萨索山（Gran Sasso）实验室，坐落于意大利一条3000米深的隧道中。位于美国南达科他州霍姆斯特克（Homestake）矿井中一个1500米深的洞穴，最初是掘金时代修建的，将会成为另一个基于氙的实验——LUX实验的所在地。在此洞穴中，雷·戴维斯（Ray Davis）曾经发现了来自太阳的核反应的中微子。低温暗物质搜寻计划则在苏丹的一个矿井中，大约在750米的地下。

所有在矿井与隧道之上的岩石不足以确保探测器完全不受辐射，实验还需要进一步采取各种不同的屏蔽方法。低温暗物质搜寻计划有一层聚乙烯环绕，如果有什么强烈相互作用的东西表明是暗物质从外面进来，那么它就会发亮。更加让人印象深刻的是，周围环绕的铅则来自18世纪法国的沉船。这些年代久远的铅材料已经在水下躺了几个世纪，因此它们的放射性都已消失殆尽。铅是一种致密的吸收材料，它可以完美地将探测器与外面来的放射性隔离开来。

即使有了这些保护措施，许多电磁辐射仍然存在。辐射与潜在的暗物质候选者的区分还需要进一步甄别。暗物质的相互作用与中子击中靶物质时的核反应相似。因此，与声子读数系统相对的是一个更为传统的粒子探测系统，它可以测量当推断的暗物质粒子穿过锗或者硅时产生的电离度。这两种测量——电离度与声子能量，可以区分那些可能来自暗物质过程的核反应事件，以及仅仅来自放射性的电子事件。

低温暗物质搜寻计划其他优异的性质还包括，绝佳的位置和其绝妙的时间测量。因为虽然只能在两个方向上直接测量，但是声子的定时给出了位置的第三个坐标。因此实验物理学家可以严格定位事件发生的地点并将背景事件排除。另一个良好的性质是，实验被划分成一个个曲棍球大小的探测单元。一个真实的事件只会在其中一个探测单元上发生。局域性可以减少辐射，而另一方面又不会将探测限制在单一的探测单元上面。有了这些性质，并且今后还会有更好的设计出来，低温暗物质搜寻计划有很大的机会发现暗物质。

尽管低温暗物质搜寻计划非常震撼，但它却不是唯一的暗物质探测实验，并且低温仪器也不是仅此一种。后来，氙实验创始人之一的艾琳娜·阿普里勒（Elena Aprile）详细地介绍了她的实验（XENON10，XENON100），以及其他使用惰性液体的实验。因为这些实验很快也将成为最灵敏的暗物质探测器，所以听众也都全神贯注地聆听了她的报告。

氙实验通过闪光记录暗物质事件。液态氙很稠密以及均匀，每个原子的质量较大（增加了暗物质的反应速度），闪烁也很强，当遇到能量沉积时电离很迅速，因此上面描述的两种信号可以有效地与电磁信号区分开来，而且它还是一种相对其他有潜质材料而言更为廉价的材料，尽管在10年中它的价格上下变动有6倍之多。这一类型的惰性气

体实验随着它们使用得越来越多而效果越来越好，并且这样的趋势还将持续下去。有了更多材料，不仅探测变得更有希望，而且探测器外层也将更为有效的屏蔽内层，为实验结果的准确性提供保障。

通过测量离子化和最初的闪烁，实验物理学家可以将它们的信号从背景辐射中区分开来。XENON100实验使用非常特殊的光电管（photo tube），它被设计成在低温高压的环境下测量闪烁。将来氩探测器通过以闪烁脉冲的具体形状作为时间的函数来使用，也许可以提供更好的闪烁信息，因此也将把有用的信号从杂乱无章的信号里面分离出来。

现今的奇异状态来自一个闪烁实验（虽然可能马上会改变）——位于意大利大萨索山实验室的DAMA实验，它曾真的发现过一个信号。DAMA与我前面提到的那些实验不同，在信号与背景之间它不存在内在的区分。相反，它鉴别暗物质信号的方式仅仅来自它们的时间依赖性，一种使用地球绕日轨道独特的速度依赖。

入射的暗物质粒子的速度是其中一个相关因素，原因在于：它决定了有多少能量可以沉积在探测器中。如果能量太低，那么实验不可能灵敏到可以知道在那里的事物。而越多的能量表明实验也越可能记录事件。由于地球的轨道速度，暗物质相对我们的速度（也即能量沉积）依赖于一年中的不同时间，这使得一年中的某些时间（夏季）比另一些时间（冬季）更容易看见信号。DAMA实验寻找年度事件发生速率的变动，以确定它是否与预测符合。数据显示，它们已经发现了这样的信号（DAMA振荡数据见图21-2）。

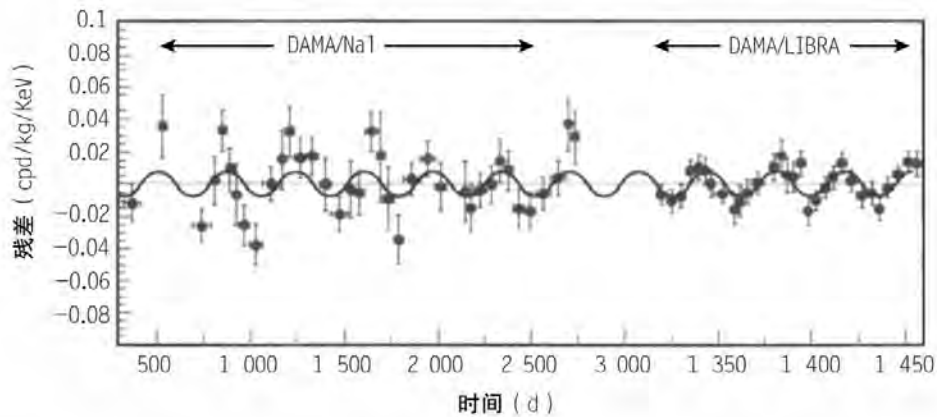


图21-2 DAMA实验数据所显示的随时间调制的信号。

还没有人可以确切地知道是否DAMA信号代表了暗物质，抑或其他可能来自探测器或环境的误解。人们对此表示怀疑，因为其他实验都还没有检测到任何东西。其他信号的缺失，与绝大多数暗物质模型的预测相违背。

虽然现在很令人困惑，但这也是一种令科学变得有趣的因素。实验结果推动我们思考可能存在的不同种类的暗物质，以及是否暗物质可能具有使DAMA比其他暗物质探测实验更容易作出探测的性质。这些结果也迫使我们更好地了解探测器，使我们可以确认多余的信号，理解实验数据到底意味着什么。

全球很多实验都在尽力获得更好的灵敏度。它们或者排除或者肯定DAMA暗物质的发现，或者可能独立地发现不同种类的暗物质。如果有至少有一个实验肯定了DAMA的发现，那么人们才会赞同暗物质被发现了这个结论。但目前它还没有发生。不管怎样，答案快要揭晓了。即使实验结果在你看这本书的时候已经过时了，实验的本质也极有可能不会改变。

间接探测暗物质的实验

大型强子对撞机实验与地下的低温或者惰性液体探测器是两种了解暗物质本质的途径。第三种并且也是最后一种方法是，在太空中或者在地面上间接测量暗物质。

暗物质很稀薄，然而它偶尔会与同类型的粒子或者它的反粒子发生湮灭。虽然这件事的发生次数没有多到可以显著影响整体的密度，但是它也许可以多到足够产生一个可测量的信号。这是因为当暗物质粒子湮灭时，会产生新的粒子，并带走它们的能量。根据性质，暗物质湮灭时，有时会产生标准模型粒子与反粒子，如电子与正电子，或一对光子，并且可以被检测到。天体物理探测器测量粒子或光子也许可以看到这些湮灭的迹象。

搜寻暗物质的湮灭而给出标准模型产物的装置，最初却不是为此目标设计的。它们是在外太空中或者在地面上的望远镜或者探测器，用来探测光或者粒子，使得我们可以更好地理解太空中有什么。通过观测从恒星与星系以及处在它们之间的奇异物体发射出的物质类型，天文学家可以了解这些天体的化学组成，并推断恒星的本质。

哲学家奥古斯特·孔德（Auguste Comte）于1835年错误地评论，“我们永无可能用任何方法查考恒星的化学组成”，他认为恒星已经超出了人类可以获得认知的范围。然而在他说了这些话的不久以后，对太阳光谱的发现和诠释（光线的发射和吸收）告诉了我们太阳的组成，证明了他的话完全错误。

当今天的实验致力于分析其他天体的组成时，它们继续着这样的任务。现今的望远镜灵敏度很高，每过几个月，我们对宇宙中存在的东西就了解得更多一些。

在对暗物质的搜寻中幸运的是，这些已经展开的实验中光和粒子的观测也许已经昭示了暗物质的本质。因为宇宙中的反粒子相对很稀少，光子能量的分布可以展示不同的以及专门的性质，例如探测也许

最终可以与暗物质相关联。粒子的空间分布也可能帮助我们从天文物理的背景中区分出湮灭的产物。

在纳米比亚的高能立体望远镜系统（High Energy Stereoscopic System, HESS），与在美国亚利桑那州的超高能辐射成像望远镜阵列系统（Very Energetic Radiation Imaging Telescopic Array System, VERITAS）都是地面上的大型望远镜阵列系统，它们的目标都是寻找来自星系中心的高能光子。下一代超高能 γ 射线天文台——切伦科夫望远镜阵列（Cherenkov Telescope Array, CTA）将有更高的灵敏度。另一方面，费米 γ 射线空间望远镜（Fermi Gamma-ray Space Telescope）安装在距离地面550公里的卫星上，该卫星于2008年底发射，每95分钟绕地球运行一周。地面上的光子探测器的优点是，它们具有巨大的采集信号区域；而费米卫星上的精密仪器则有更好的能量分辨率和更直接的信息，它们对低能光子很敏感，并且它们有200倍的视野。

这其中任何一种实验都可以看到从湮灭的暗物质中发出的光子，或者从暗物质湮灭产生的电子与正电子所产生的辐射。如果我们看到了两者之一，那么我们已然对暗物质的身份和性质有了更多了解。

其他探测器主要找寻电子的反粒子正电子。物理学家在意大利主导建设的卫星实验PAMELA中已经报道了他们的发现，而他们没有找到预测中的任何东西（PAMELA的结果见图21-3）。这个实验的全称是绕口的“负载为反物质探测和轻核天体物理载荷”（Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics, PAMELA）。我们还不知道PAMELA超载的事件是出于暗物质，还是出于诸如脉冲星（pulsar）等天文物体的错误估计。但无论如何，此结果引起了天文物理学家和粒子物理学家的关注。

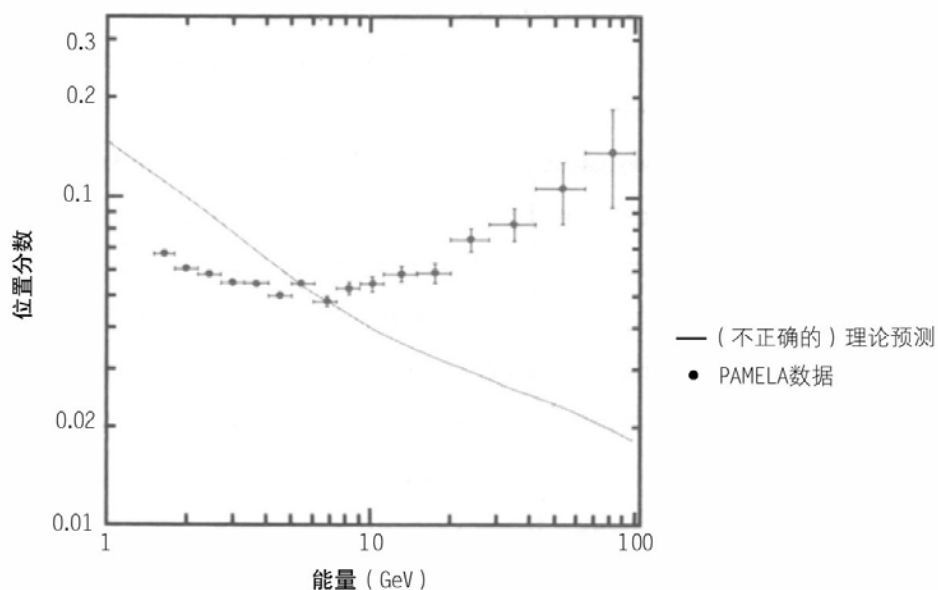


图21-3 PAMELA实验得出的数据显示，实验数据（十字形线）与理论预测（点线）非常不一致。

暗物质也能湮灭成质子与反质子。事实上，许多模型预言，如果暗物质粒子的确相互碰面以及发生湮灭，那么这也是最常发生的情形。然而，大量源于其他已知天文过程的反质子在星系里面飘荡，可能会掩盖暗物质的信号。我们也许仍然有机会通过反氦核（antideutrons）——反质子与反中子的弱束缚态，来发现暗物质，它们可能在暗物质湮灭时形成。位于国际空间站上面的 α 磁谱仪（Alpha Magnetic Spectrometer, AMS-02），以及专用卫星实验，例如通用型反粒子谱仪（General Antiparticle Spectrometer, GAPS），也许最终可以发现这些反氦核，进而发现暗物质。

最终，不带电的中微子只能通过弱相互作用来与其他粒子作用，它也可能是暗物质间接测量的关键。暗物质可能被困在太阳或者地球的中心。在这种情况下可以从中逃逸的唯一信号会是中微子，因为与

其他粒子不同，中微子在逃逸时不会由于相互作用而被截停。名为AMANDA、IceCube、ANTARES的探测器都在找寻这些高能中微子。

如果上述任何一种信号被观测到（或者即使它们都没有观测到），我们都将对暗物质的相互作用及其质量增加一些了解。同时，物理学家已经根据从各种可能的暗物质模型的预测，哪些信号值得期待。当然，我们会问，任何现有的测量意味着什么。暗物质的测量非常具有挑战性，因为它的相互作用如此微弱。我们希望，由于目前正在运行的许多不同类型的暗物质实验，暗物质可能在不远的将来被探测到，并且大型强子对撞机与其他实验的结果，将对宇宙中有什么以及它们如何配合在一起，提供一个更好的理解。



第六部分

宇宙的探索仍在继续

**KNOCKING ON
HEAVEN'S DOOR**



本书对人类智力如何探索大至宇宙极致、小至物质结构做了一些展示。在这两个方向的追求中，已故的哈佛大学教授西德尼·科尔曼（Sidney Coleman）是公认的最聪明的物理学家之一。学生们听到的故事是：当西德尼博士毕业后申请博士后资助时，除了理查德·费曼的推荐信，几乎所有推荐信都将他描述成“他们所认识的最聪明的物理学家”。费曼将西德尼评价为“最好的物理学家”（显然费曼没有将自己考虑进去）。

在庆祝西德尼60岁生日的聚会上，许多同时代的著名物理学家都作了演讲。哈沃德·乔吉是西德尼在哈佛大学的老同事，也是同时代著名的粒子物理学家，他很惊喜地观察到，参会的优秀理论物理学家的报告如此出色，而且思考方法千差万别。

乔吉的判断很对。每一个演讲者都有一种特殊的探索科学的方法，而且都已经通过他自己（的确他们都是男性）独特的技能取得了显著的成果。一些是可视化的，一些是有数学天赋的，还有一些是具有兼收并蓄的惊人能力。“自上而下”与“自下而上”的品味都于其中有所展示，从对物质内部强相互作用力的理解到由弦理论推导出来的数学，各行各业都取得了成功。

伟大诗人普希金说得对，他曾经写道：“几何中所需要的灵感，不亚于诗歌中所需要的灵感。”创造力对于粒子物理学、宇宙学、数学以及任何其他科学领域来说，都至关重要，以至于对更受广泛认同的获益者——艺术和人文学科也是一样。科学体现了额外的丰富内涵，足以激发那些发生在特定条件中的创造性努力。涉及的灵感和想象隐没在逻辑规条之中，很容易被忽视。然而，数学与技术本身就是由如下两种人发现和创造的：一类人是，在归纳理念方面具有创造性思维的人；另一类人是，偶然发现了某个有趣的结果，且能以开创性的敏锐意识到其价值的人。

在过去几年中，我很幸运地获得了各种各样的机会，得以在不同的生活曲线上与有创意的人们相遇并合作。回想他们的分享是一件很有趣的事情。科学家、作家、艺术家与音乐家可能外表看起来非常不同，但是他们的技能本质、天赋与气质并不像你所想象的那样天差地别。我将以具备某些显著品质的科学与科学思维的故事作为本书的总结。

另类才能

科学家与艺术家在做一件很重要的事情时，都不大可能思考创造力本身。没有几个（如果有的话）成功人士会坐在桌子旁边决定：“我今天将很有创造力。”相反，他们的注意力是在一个问题之上。并且当我说“关注”时，我指的是聚精会神，除此之外别无他念，有意专注于他们的工作的那种专注。

我们通常只看见创造性努力的结果，而看不见为此付出的巨大奉献，以及暗含于其中的技术专长。

我曾观看2008年的电影《走钢丝的人》（*Man on Wire*）。该电影是为了纪念菲利普·珀蒂（Philippe Petit）于1974年在世贸中心的双子塔之

间400米长的距离间高空行走。这在当时是一个壮举，吸引了多数如我一样的纽约客，并且这在全球也是轰动一时。我很欣赏菲利普的冒险精神、他的表演与技巧。但是菲利普并不是仅仅用螺钉在双塔的墙壁上将钢丝拴一拴，然后就在上面摇摇晃晃地走起来。编舞者伊丽莎白·斯特布向我展示了一本约3厘米厚的书，里面记录了菲利普在安装钢丝之前在她的工作室所进行的绘图和计算。这样我才理解了确保菲利普行走计划稳定性中的准备和重点。菲利普是一个“自学成才的工程师”，他这样开玩笑地称呼自己。只有在细致地研究和应用已知的物理定律，来理解他所采用的材料的性质之后，他才准备好走钢丝。当然在成功之前，菲利普不能确定他已经把所有因素都考虑周全，但是毫无疑问他所考虑的所有东西是足够的。

如果你发现这种程度的专注很难让人相信，那么你可以环顾四周。人们往往被那些不论是意义大还是意义小的活动所吸引。你的邻居喜欢做填字游戏，你的朋友为运动节目所着迷，地铁上的人因为看书太专心而坐过站，更不用说你也可能会花上无数小时的时间来打游戏了。

那些心系研究的人却幸运得多，因为他们所从事的赖以谋生的事业刚好与他们所心仪的兴趣（或者说可以说至少他们不愿意错过的东西）相重合。在这个领域中的专业人员一般都有舒心的想法（尽管可能是虚幻的），那就是他们所研究的东西可能具有永恒的意义。科学家倾向于认为我们肩负着揭示世界真理的使命。我们也许没有时间玩填字游戏，但是我们非常希望能花更多的时间在一个研究课题上面，特别是一个与更大的图像和更大的目标相关的一个课题。真实的工作可能包含与玩游戏或者观看体育比赛相同的关注力。^[1]但是科学家很可能在开车时或者夜晚躺上床以后继续思考研究的问题。这种花在项目上的经年累月的付出当然与其认为“研究很重要”的信仰无法分开——哪怕很少有人能够理解（至少在研究初始之时）他们，哪怕他们的研究最终可能会被证明是错误的。

最近出现了一个热门的话题：质疑与生俱来的创造力和天赋，并将成功归于早期的经历和训练。在《纽约时报》的一个专栏中，大卫·布鲁克斯（David Brooks）总结了近来关于此话题的几本书：“我们相信，莫扎特拥有的是与老虎伍兹一样的能力——长时间的专注力与进一步提升技能的意愿。”^[2]毕加索是他引用的另一个例子。毕加索是一个古典艺术家的儿子，在得天独厚的环境中，他从小就在绘画方面崭露头角。比尔·盖茨也拥有卓尔不凡的机会。马尔科姆·格拉德威尔（Malcolm Gladwell）在他的书《异类》（*Outliers*）^[3]中提到，比尔·盖茨就读的西雅图高中是极少的一所具有计算机俱乐部的中学，以及盖茨后来在华盛顿大学如何有机会长时间地使用计算机终端。格拉德威尔暗示，盖茨的机会比他的驱动力和天赋更重要，并促成了他的成功。

的确，早期的专注力和练习使得方法和技巧成为许多创造性的基础。如果你有疑难需要解决，你希望花尽可能少的时间在那些基础上面。而一旦技能（数学或者知识）成为第二本质，你可以在需要它们时马上轻易地召唤出它们。这种深埋的技能常常在底层层面上不断运作，甚至在它们向你的头脑推出什么好想法之前也是如此。不单单只有一个人曾经在睡梦中解决了问题。拉里·佩奇（Larry Page）告诉我，建立谷歌的想法在脑海中形成的那一刻就发生在某次睡梦中，不过那已经是他在考虑该问题的数月之后的事情了。人们总是将灵感归功于“直觉”，而没有意识到在那些欢呼雀跃时刻的背后，隐藏了多少准备的时间以及多少细致的研究。

因此，布鲁克斯与格拉德威尔毫无疑问在某些方面是正确的。如果没有付出与练习带来的对技能与强度的磨炼，即使有技能与天赋，它们也不会让你走得更远。但是在年轻时的机遇与系统训练并不是全部，这种观点忽略了强有力的专注与练习的能力本身就是一种技能。能从他们之前所做的事情中学习，并能将积累的经验置于头脑之中的

特殊人群，更可能从学习与重复中学到更多的东西。不管是在科学研究中还是在其他领域的创造性追求中，这种倾向使得专心与关注最终会获得成效。

时装品牌CK (Calvin Klein) 最早的香水名称是“Obsession”，这绝非偶然。正因其创造人卡尔文·克莱恩的痴迷，才获得了成功。即使高尔夫职业选手在经过无数次重复练习之后能够完美地挥动球杆，我也不相信每个人都可以做到一千次击球而不会变得厌烦与沮丧。我一个爱登山的朋友凯·津恩 (Kai Zinn) 会走不同的艰难路线（用行话来说，是硬5.13s级^[81]），他记忆细节与行动的速度都比我快得多。当他走过同一条路线10次以后，他为此获益良多。这让他的忍耐力更强。我在重复中会变得非常厌烦，因此停留在中级登山水平上；而凯知道应该从重复中学习什么，因此他的登山水平不断提高。布封（原名乔治-勒克莱尔 [Georges-Louis Leclerc]）是18世纪的博物学家、数学家与作家，他简洁地总结了这种能力：“天才只不过是更大的耐心。”不过我还要补充一条：天才同时也根源于对缺乏进展的不耐。

一沙一世界

练习、专业训练与动力对于科学研究来说都是至关重要的，但是它们不构成所有必需的要素。自闭症患者（更不用说一些学者以及更多的官僚主义者）常常展现出极高水平的技术技能，然而却缺乏创造力与想象力。要想见证没有其他品质支撑的驱动与技术成果的局限性，只需去看场电影就知道了。动画片中的人物（或物体）与人物（或物体）的搏击场景处理得非常连贯，给人的印象深刻，但是它们鲜少有创造力来吸引人，因此即使伴随着灯光与响声，我也常常中途睡着。

于我来说，最能吸引我的是那些提出了重大问题与真实想法，却又将它们应用在我们可以欣赏与领会的小例子中的电影。电影《卡萨布兰卡》是关于爱国主义与爱情、战争与忠诚的影片，即便主人公里克（Rick）提醒伊尔莎（Ilsa）“不难看出问题所在。在这个疯狂的世界中，三个小人物撑不起一座山”，这三个小人物也是我被该片深深吸引的原因。

在科学中同样如此，正确的问题常常总是来自大景象与小图像两个方面。我们想回答的是大问题，而我们可以处理的是小问题。确定出大问题往往还不够，因为通常解决了较小的问题才能更进一步。正如盐湖城尺度会议的会名（见第3章）对我们的提醒，伽利略很早就领会到了的“一沙一世界”（出自威廉·布莱克的诗《天真的预言》）。

对于每一个有创造力的人来说，不可或缺的一种本领是：能够提出正确问题的能力。他们可以分辨出对于取得进步来说有潜力的、刺激的、最重要的以及可行的方法，最终正确地将问题建构出来。最好的科学往往综合了覆盖面宽广以及显著的问题，同时又集中在一些人们非常想解决的明显的细节或者具体问题上面。有时这些小问题或者小矛盾恰好是取得关键性进展的线索。

达尔文进化论的想法源于鸟类与植物学的一些细小观测。水星的近日点进动也不是一个实验错误，相反它暗示了牛顿物理定律是有局限性的。该测量最后成为爱因斯坦引力理论的一个确证。这些断层与矛盾也许对某些人来说看似太小或者太模糊，但是它们对于那些找对了问题的人来说，却成了新观念与新想法的入口。

爱因斯坦甚至没有一开始就着手理解引力。他曾试图理解那时正处于发展阶段的电磁理论的含义。他关注那些与所有人认为的时空对称性特异的或者甚至矛盾的方面，结果颠覆了我们的思考方法。爱因

斯坦相信该结果是合理的，他凭着远见与毅力从中探索可能是正确的东西。

近来的研究也显示了这样的关系。理解为什么某些相互作用不会在超对称理论中产生，对某些人来说可能看似非常技术化。我的同事大卫·卡普兰（David B. Kaplan）于20世纪80年代在欧洲讨论这种问题时，就常被人取笑。但是该问题结果成为超对称与超对称破缺新灵感的丰富来源，它甚至引领了新的想法，而大型强子对撞机的实验物理学家现在正准备对其进行检验。

我很坚信宇宙是自洽的，任何偏差意味着新东西将被发现。在华盛顿特区举行的创意基金（Creativity Foundation）的会议中，我表达了这种观点，之后一个博客友好地将此解释成我标准很高。但实际上，对宇宙自洽性的信仰，可能是许多科学家在决定研究主题时的动力源泉。

我认识的许多有创造力的人都有能力同时提出几个问题与想法。任何人都可以用谷歌查找资料，但是除非你能将事实与想法巧妙地结合起来，否则你可能不会发现新东西。来自不同方向稍微有所争执的想法，通常会催生新的联系或者灵感或者诗歌（这恐怕是“创造力”一词最早被应用的领域）。

很多人倾向于一条道路走到底。但是这也意味着一旦他们被困住或者发现道路不明确，他们的追寻也就终止了。与许多作家以及艺术家一样，科学家也时常在迂回曲折的道路上取得进展，他们通常并不能一条道路走到底。我们也许理解一个难题的某些部分，暂时将我们不理解的搁置一旁，希望以后再来把漏洞填补。只有很少人能通过持续的阅读来理解一个理论的全部。我们必须相信自己最终可以将理解的内容拼接起来，因此我们可以忍受一开始跳过一些内容，以后当我们掌握了更多知识或者有更广博的眼界时，再回过头来考虑它们。文

献或者结果可能最初看似不可理喻，但是无论如何我们会坚持读完。当发现自己无法理解的一些东西时，我们先忽略它们而一直读到末尾，把疑问先拣选出来，最后再回过头来看那些我们不理解的地方。为此我们必须有足够的耐性来坚持研究我们所知道的与不知道的领域。

托马斯·爱迪生的一句名言是：“天才是百分之一的灵感，加百分之九十九的汗水。”近代微生物学奠基人路易·巴斯德（Louis Pasteur）也说过：“在观察事物之际，机遇偏爱有准备的头脑。”献身科学事业的科学家有时也因此而发现他们正在找寻的答案。但是他们也可能发现问题的答案与最初的目标相去甚远。亚历山大·弗莱明（Alexander Fleming）一开始并没有寻找治疗炎症的药物。他发现，一种霉菌可以把他正在研究的葡萄球菌的菌群杀死，从而意识到它潜在的医学效用，尽管他和很多人花了10年才将盘尼西林（即青霉素）研制成一种强有力的药物，而这种药物改变了世界。

附加效益常常来自一个范围宽广的基础问题的储备中。当我和拉曼·桑卓姆一起研究超对称时，我们结果发现了弯曲的额外维度，它可以解决等级问题。之后在刻苦钻研此问题，并将之放置于一个更广阔的情景中时，我们又发现了可以存在一维无穷大的额外维度弯曲空间，而且与任何已知的物理定律或者观测都不相矛盾。我们此前已经研究了粒子物理学——这可是完全不同的题目，但是我们把大景象和小图像都放在脑海中。甚至在专注于更抽象的问题，例如理解标准模型弱能标的等级问题时，我们也保持着对空间本质这个大问题的警觉。

这种特殊研究的另一个重要特征是，桑卓姆和我都不是相对论专家，因此我们在展开研究时，思维很开阔。我们都没有（其他人也没有）猜想过爱因斯坦的引力理论允许一个不可见的无穷大维度的存

在，一直到方程让我们看到了其中的可能。我们千辛万苦地寻求方程的结果，而没有意识到以前无穷大的额外维度曾被认为是不可可能的。

即使如此，我们也没有马上相信我们的工作是正确的。桑卓姆和我没有盲目投身到额外维度这个疯狂的想法中去。只有当我们与其他许多科学家已经尝试了许多传统的方法之后，跳脱出经典时空观的藩篱才合情合理。虽然额外维度是一个奇异的、全新的提议，但爱因斯坦的相对论理论仍然适用。因此我们凭着方程与数学方法来理解我们所假想的宇宙的可能情况。

人们接着从该研究中假设的额外维度的结论出发，发现了一些新的物理想法，甚至可能在没有额外维度的宇宙中也适用。通过与此正交的（orthogonal）思考方式，物理学家意识到他们以前没有完全想到的可能性。这帮助他们跳出了三维空间的思维定势。

任何人在面对新基础时，在完全处理一个问题之前不得不面对一些不确定性。即便从一个已经存在的坚实的知识平台出发，人们在考察新现象的过程中会在所难免地遭遇未知和不确定因素——虽然不及高空走钢丝那样危险。然而这些空中冒险家与艺术家和科学家一样，致力于“勇踏前人未至之境”^[82]。但是这种勇气并非意气用事或者匹夫之勇，它没有忽视以前的成功法门，尤其是当新的领域涉及起初显得不可能的新想法或者看似疯狂的实验时。研究人员竭尽所能地做好充分的准备。规则、方程、直觉对于理论的自洽性很有帮助。这些线索有助于我们跨越到新的领域之中。

我的一个同事马克·卡米科维斯基（Marc Kamionkowski）说过：“有雄心和有远见是好的。”但是在这之前的诀窍是确定可行的目标。一个获奖的商科专业学生在创意基金活动期间报告说，最近经济增长升级成经济泡沫的部分原因正是出于一个创意。但是他也注意到，缺乏相应的约束也造成了泡沫的崩溃。

过去一些最具开创性的研究充分体现了自信和审慎的矛盾冲突。科学作家加里·陶布斯（Gary Taubes）曾经对我说过，在他所认识的人中，学者是最自信而又最缺乏安全感的人。这种矛盾推动他们——他们不断前行的信念与他们用来确保正确的严格标准交织在一起。有创造力的人们必须相信他们是独一无二的，然而又要随时谨记，存在许多因素使得其他人可能已经想到或者反驳了类似的想法。

科学家在思考问题时非常具有冒险精神，但在展示想法时却又非常小心谨慎。两个最有影响力的代表——牛顿和达尔文，在将他们的伟大想法公之于众之前，考虑了很长时间。达尔文的研究跨越很多年，直到他完成了广泛的观测研究之后才发表《物种起源》。牛顿的《自然哲学的数学原理》则展示了他为此发展了十年的引力理论。牛顿迟迟未发表，一直到他取得了完全满意的证明——任意具有空间形状的物体（不仅仅是点状物体）遵从平方反比定律。该定律，即引力随着到物体中心的距离的平方而减弱的证明，促使牛顿发展了数学中的微积分理论。

有时，需要为一个问题创建一个新体系，才能正确地看待它、重新界定它的边界，从而找到问题的解答，而如果仅仅停留在表面，那么一切都是不可能的。毅力和信念往往会对结果产生意想不到的效果，这种信念不是宗教信仰，而是对于答案必然存在所抱持的信念。成功的科学家（以及其他各行各业有创造力的人们）拒绝钻牛角尖。如果我们不能以一种方法解决问题，那么就要另辟蹊径。如果遇到绕不开的路障，那么我们就挖隧道过去，或者再找别的方向，或者飞越过去。这就是想象和疯狂的想法一展身手的时间了。我们必须相信现实中存在答案，而且相信世界有一个可以被我们最终发现的内在的自洽逻辑。如果从正确的方面思考问题，那么我们时常会发现内在联系，以及那些我们可能会错过的联系。

“跳出思维框”这个说法不是从你的工作间之外来的（我曾以为是这种情况），而是从一个“9点问题”来的，它让你如何用4条直线连接9个点，而笔尖不能离开纸面（见图22-1）。如果你把笔限制在方

块区域之内，那么你是找不到答案的。事实上没有人要求你那样做。走到“盒子之外”带来了解决方法（见图22-2）。你也可以设计许多其他方法改换该问题。如果你使用大黑点，那么你可以只用三条线就可以了；如果你把纸片叠起来（或者使用一根极粗的线条——一个小女孩对该问题的设计者如此建议），那么你还可以只用一条线就完成题目。

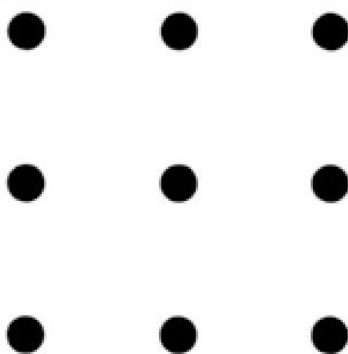


图22-1 “9点问题”：
如何用一笔将图中所有
点用4条直线连接起来。

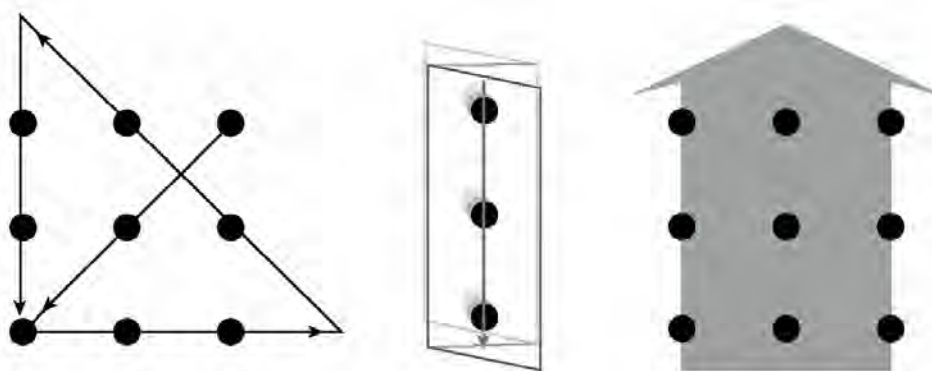


图22-2 “9点问题”的一些创意解释，“跳出思维框”，或者将纸折叠以使点重合，或者使用粗笔尖的笔。

这些解释并非作弊。除非是在你有附加的限制条件下，它们才是作弊。然而，教育有时鼓励学生不仅要学习如何解决问题，还要研习

老师的意图——将正确的范围缩小甚至可能把学生的想法也缩小了。

在《夸克与美洲豹》（*The Quark and the Jaguar*）^[4]一书中，默里·盖尔曼引述了华盛顿大学物理学教授亚历山大·卡兰德拉（Alexander Calandra）的“气压计故事”^[5]。一个老师说他知道是否应该给一个学生分数。该老师问他的学生们如何使用气压计来推断一栋楼的高度。这个学生回答说，可以用一根绳子系在气压计的末端，将绳子从楼顶垂到地面，看绳子共有多长。当被告知必须使用物理时，这个学生建议可以测量气压计从楼顶落到地面的时间，或者在一天中的一个特定时间测量气压计在太阳下影子的长度。这个学生甚至还主动提出了非物理的方法：跟管理员做交易，把气压计给他，让他告诉楼的高度。这些答案可能都不是这位老师所期望的答案。但是这个学生机灵且幽默地意识到老师没有对问题做任何限制。

当我和其他物理学家在20世纪90年代开始思考额外维度空间时，我们不仅跳出了思想框，甚至跳出了三维空间本身。我们想到了一个世界，在解决问题的每一阶段，它都会变得比我们原先设想的大得多。渐渐地，我们发现该问题潜在的解决方案实际上已经困扰了粒子物理学家许多年。

即使如此，研究也不是无中生有。它被前人的许多想法与观点所丰富。优秀的科学家互相倾听。有时我们可以通过仔细听、观察或者阅读其他人的工作成果来发现正确的问题或者答案。我们常常会与有着不同才能的人合作，并同时保持真实的自己。

即使每个人都希望成为第一个解决重要问题的人，科学家却仍然互相学习、分享经验，并且钻研同样的课题。有时，其他科学家甚至在不知不觉中说出了某些有趣的问题和答案的线索。科学家也许有自己的灵感，但他们也相互交换想法以研究出结果，并随时调整想法或

者当原来的想法不奏效时再重新开始。想出新方法，并从中有所摒弃，这就是我们周而复始的生活，这也是我们如何取得进步的途径。

作为研究生导师，我所能扮演的一个最重要的角色就是对学生的好想法保持警觉，甚至当他们还没有学会如何表达自己观点时，我给他们的建议中如果有存在漏洞，如果学生指出来，我仍然会虚心听取。这种往复式的互动也许是最好的教学方法，至少它培养了创造力。

在科学领域以及其他各种有关创造力的活动中，竞争也起了重要作用。在讨论创造力时，艺术家杰夫·昆斯（Jeff Koons）告诉我们，当他还小时，他姐姐从事艺术工作，那时他就意识到他可以做得更好。一个年轻的电影制片人解释，竞争如何鼓励他与他的同事吸收彼此的技术和想法，以此来完善和提高自己。名厨张大卫（David Chang）直言不讳地表达了相似的想法。在进到一家新餐厅之时，他的反应是：“菜真是太好吃了。我为什么没有想到呢？”

牛顿一直等到他的结果全部完成才发表他的著作。但是他可能也曾经警惕他的对手罗伯特·胡克——胡克也知道平方反比定律，但是胡克却缺乏足够的计算来证明该定律。无论如何，牛顿著作的出版非常及时，部分原因也是由于他知道胡克的研究与他的有所重叠。达尔文也显然出于认识到阿尔弗雷德·华莱士（Alfred Russel Wallace）正在研究类似的进化理论，才发表他的研究成果——如果他继续保持沉默，他的惊世之见就有可能被华莱士抢先发表。达尔文和牛顿都想在发表他们革命性成果之前拥有整个理论，并不断完善它们直到他们都极其确信自己的理论是正确的，或者至少直到他们认为它们应该是正确的。

宇宙一再向我们展示它是如何比人类聪明。方程和观测打开了人类的思维空间，并且只有那些拥有创造力的开阔头脑的质疑，才能揭开未来深藏不露的现象之谜。没有确凿的证据，科学家不能建立量子力学。我怀疑，那些认定精确的DNA结构和无数现象可以构成生命的期望，它们几乎是不可能的，除非我们可以弄清那些现象或者解开那些方程，它们告诉我们什么东西会存在。希格斯机制太巧妙了，原子的

内部运行方式与粒子的行为也同样巧妙，它们隐藏在我们可以看见的现象背后。

科研是一个系统的过程。我们不必总是知道我们走向何方，但是实验和理论是我们宝贵的向导。准备与技能、专心与耐心、问对问题，以及审慎地信赖我们的想象，都将帮助我们寻求理解、寻找答案。因此我们要开放头脑，保持与其他人的交流，致力于比我们的前辈和先贤做得更好，并且相信答案存在。不管出于何种目的，哪些独立的特别技能可以投入进来，科学家将继续向内、向外搜寻，期望了解宇宙所蕴藏的更多奥妙机制。

结语

寻路宇宙边缘

当我第一次看到德国媒体对我的物理研究或者说我的书《弯曲的旅行》^[83]的报道时，我惊讶于短语“宇宙边缘”重复出现的次数。对于看似合理而又似乎在文章中随机出现的这个短语，最初我找不到清晰的解释。后来我才发现，这个短语是电脑将我的姓“Randall”从德语直接翻译出来的结果。^[84]

我们的确正处于宇宙边缘，在小尺度与大尺度上都是如此。科学家实验上能探索的范围从弱尺度 10^{-17} 厘米到宇宙尺度 10^{30} 厘米。我们不能确定什么尺度限定了将来研究范式的转变，但是许多科学的考虑聚焦于弱尺度——即实验正在探索的大型强子对撞机与暗物质的尺度。同时，在理论方面，物理学家还在继续从弱能标到普朗克能标的探索，我们都在试图填补我们理解中的空白。认为人类已经看透了自然中所有事物，是一种傲慢的想法。新的发现无疑还在等着人类去探索。

现代科学的纪元仅仅代表了历史时间线上的一个闪光点。从17世纪科学的诞生开始，从技术与数学的进步中得来的卓越见解，一直带领我们走在一条通向理解世界的漫长道路上。

本书探讨了一些问题：当今的高能物理学家和宇宙学家如何决定他们的研究道路，以及如何将理论与实验结合起来，以使它们可以揭示更深刻、更基本的问题。大爆炸理论为宇宙当前的膨胀作出了解

释，但是它没有回答宇宙更早时期的问题——暗能量与暗物质的本质是什么。标准模型预言了基本粒子的相互作用，但是它也没有解决为什么它的性质就是我们所见到的那样。暗物质和希格斯玻色子的答案可能近在眼前——就如同新的时空对称性和新的空间维度一样。我们也可能足够幸运可以尽快回答这些问题。或者如果相关的质量太重或者粒子的相互作用太弱，我们还需要等上一段时间。无论如何，只有先提出问题并努力寻找，我们才可能知道答案。

我也呈现了一些更难于证实的猜想。虽然它们拓展了我们的想象而且可能最终与现实相连，但它们也可能仅仅停留在哲学与宗教的范围之内。科学不能证伪多重宇宙的景观（就这一点而言，这想法就像上帝一样），但是要想证实它们也是不可能的。即使如此，某些多重宇宙的性质，例如可以解释等级问题等，确实具有可以检验的效应。科学家的任务是把这些效应挖掘出来。

《叩响天堂之门》的另一个主要内容是探索科学思维的一些概念，例如尺度、不确定性、创造力、理性和批判性推理等。我们相信，科学将朝着问题答案的方向迈进，而复杂性也将与时俱进，甚至在我们获得有血有肉的解释之前它仍会如此。答案可能是错综复杂的，但这不能成为我们放弃信仰的理由。

理解自然、生命以及宇宙是极其困难的问题。我们都愿意更好地了解我们是谁，我们从哪里来，我们要往哪里去，进而将精力集中在比我们自身更大、比最新的小发明和风尚更持久的事情上面。我们很容易看出为什么有些人转而向宗教寻求解释。如若没有事实也没有启发人的解释来展现那些令人吃惊的关联，科学家目前所给出的答案在以前看来就是极难猜测到的。用科学方法思考的人们拓展了我们对世界的认知。我们面临的挑战是如何尽可能多地理解新事物，而好奇心（不受教条的束缚）正是我们所需要的。

合理的质疑与傲慢之间的分界线也许对某些人来说是一个问题，但在回答关于宇宙组成的问题上，最终，批判性的科学思维才是唯一可信赖的方法。当前一些宗教运动中的极端反智主义组织与传统的基督教传承水火不容，更不用说与进步和科学相容了，但是所幸它们并不代表所有宗教或者智力的思潮。许多思考方法（甚至宗教的）将已经存在的范式与挑战结合起来，并且允许想法发生演变。于我们而言，进步牵涉到将错误的想法摒弃、将新想法建立在正确的思想之上。

我很欣赏美国国家科学院前院长、现任《科学》杂志主编布鲁斯·阿尔贝茨（Bruce Alberts）最近一场演讲中的观点，他强调，对创造力、理性、开放性以及容忍性的需求，是科学与生俱来的本性。这种众多素养的强力组合，被印度第一任总理贾瓦哈拉尔·尼赫鲁（Jawaharlal Nehru）称为“科学的本性”。^[1]在现今世界中，科学的思考方法是很重要的，它为处理许多社会、实践与政治中的棘手问题提供了必要的工具。我愿意用与科学和科学思维方式相关的更进一步的思考来作为本书的结束。

今天，某些错综复杂的难题可以解释成为技术、大量信息与原始计算能力的组合。但是许多主要进步（科学的或者其他的）仅仅需要来自于大量个体或者小规模群体的思考，他们是受启发在相应难题上研究了很长时间的个人。虽然本书集中介绍了自然和基础科学的价值，但是纯粹的、求知欲所驱使的研究（与科学本身的进步一起），已经带来了技术的突破，并且完全改变了我们的生活方式。基础科学除了为我们提供解决疑难的重要方法，还促进了技术的不断进步；技术则结合了更多的科学思考从而吸纳其创造性和原理性，而技术的进步又反过来帮助我们发展明日的科学。

现在的问题是，如何在现有框架下提出更大的问题。我们如何为技术设定一个长期发展的深远目标？甚至在一个技术的世界里面，我们需要想法与激励齐头并进。公司在制造一些必需的小电子产品时可以做得非常成功，也很容易陷入对新产品的追逐中。但是仅仅保持这种做法会掩盖我们对技术的真实需求。虽然 iPod 能给我们的生活带来乐趣，但是有 iPod 的生活方式并不能帮助我们解决今天世界上的重大问题。

在一个关于技术与发展的会议的小组讨论中，《连线》杂志的创始人之一凯文·凯利（Kevin Kelly）说过：

技术是宇宙最强劲的推动力。

如果果真如此，那么科学就该对这个最强劲的推动力负责，因为基础科学是技术日新的根源。电子的发现并不存在隐秘的动机，而电子决定了我们世界的面貌。电力也只是一个纯粹的智力发现，而我们的星球现在却随着电线、光缆而脉动。甚至量子力学——原子中的深奥理论，也成为贝尔实验室的科学家们开发晶体管（技术革新的基本硬件）的关键。然而，早期原子的研究者却并没有人相信他们正在做的研究将来会有任何的应用价值，更不用说制造出了像计算机这么伟大的机器以及它所带来的信息革命了。基础科学知识和科学思维方法对于深入研究现实的本质都是必要的，最终将引导我们在认识上的突破。

没有任何计算功能或者社交网络可以帮助爱因斯坦，使他能够更快一点发展出相对论，同样我们也不能让科学家对量子力学的理解更快一些。但我们并不是不承认，一旦一种好的想法或者关于一个现象的新理解出现，它马上就会加速技术的进步，而且某些问题的确也需要通过大量数据来推动。但通常，一个核心想法非常关键，对于现实本质的洞察乃是科学的用武之地，它最终会导致革命性的突破乃至以

无法预期的方式来影响我们。可见“吾将上下而求索”是何等重要了。

不可否认，“技术就是核心”。我们的确可以说绝大多数新发展都强烈地依靠技术，但是我还想增加一点，它是中心并不是指它由始至终都在中心，而是指在把事情做成和将发展关联起来的方式上是关键。我们想怎样使用技术，是我们的选择。而解决问题或者为新的发展提供启示，完全也可以来自许多种有创造性的想法。

技术的确使我们每个人成了个人世界（比如MapQuest和各种社交网络）的中心。但是真实世界中的问题则广阔得多、普遍得多。技术可以帮助解答问题，但是它们只有被清晰、有创造性的思想激励之后才更加可能成功。

过去，美国对科学与技术的关注，对它们需要作出长期承诺与坚持的认识，已经被证明是一个使美国保持在发展与思想前沿的战略方针。我们现在正面临着失去这些让我们曾经做得非常好的价值观。我们需要重新承认这些我们所追寻的原则——不仅为了短期利益，而且也要从长远的角度来理解这些付出与成效。

理性地质疑世界的做法应该得到更多肯定，因为我们可以使用它来解决一些重要的难题。布鲁斯·阿尔贝茨在他的演讲中也建议，科学思维可以作为帮助人们反对胡言乱语、断章取义的新闻以及过度主观导向的广播的一种方法。我们不希望人们被科学的方法吓跑，因为该方法在获取有意义的结论方面非常重要，特别是对今天社会中必须处理的许多复杂系统——如金融系统、环境、风险评估与卫生保健等。

无论在科学中还是其他事情中，取得进展与解决问题的一个关键要素是对度的认识。根据度来分类已经观察到的以及已经理解的事

物，它带领我们在理解物理学与理解世界中走了如此之远——无论这里的“度”指的是什么，不管是物理尺度、人口群体，还是时间区域。不仅是科学家，还包括政治家、经济学家以及政策制定者，他们都需要具有这样的观念。

最高法院大法官安东尼·肯尼迪（Anthony Kennedy）在对第九巡回法院（Ninth Judicial Circuit）的一次演讲中指出，不仅是科学思维具有重要性，“微观”与“宏观”思考的重要对比也很重要。“微观”与“宏观”的字眼既可应用在描述宇宙中小尺度与大尺度的元素上，也可应用在我们对世界思考的细致与笼统的方式上。如同我们在本书中所看到的，解决科学问题、实际问题与政治问题中的一个因素是思想在微观与宏观两个方面的相互结合。意识到“微观”与“宏观”这两个方面，是获得创造性思维的一个重要方面。

大法官肯尼迪也指出他所喜欢的科学的一个特质是，“最不可思议的答案结果往往就是正确的那一个。”的确有时情况就是这样。尽管如此，好的科学甚至当它导致了表观牵强或者有悖常理的结论时，也扎根于那些能够证实它们的测量中，或者那些我们猜想是真实的然而却有着疯狂结论的问题中。

许多因素结合起来构成了一种良好科学思维的基础。在《叩响天堂之门》一书中，我尝试传达理性的科学思维与其唯物主义前提的重要性，以及科学思维通过实验检测想法，并将与测量不符的想法抛弃的理念。科学思维认为，不确定性并不是一种失败，它合理地评估了风险并考虑了短期与长期的影响。它允许在寻找答案过程中的创造性思维方式。这些是思维的所有模式，可以确保在实验室和办公室内外都能取得进展。科学方法有助于我们理解宇宙的边际，也引导我们对自己所生存的世界作出重要的决定。我们的社会需要吸收这些原则，并将它们传承下去。

我们不应该害怕问大问题或者思考宏伟的观念。我的一个物理学同事马修·约翰逊（Matthew Johnson）说的很对：“历史上还从没有哪一个时期出现了与现在一样多的想法。”但是我们还不知道答案，并在静待实验结果。有时答案出现得比预想的快，例如宇宙微波背景告诉了我们早期呈指数级增长的宇宙；有时答案出来得慢，例如我们仍在等待的大型强子对撞机的结果。

我们将很快知道宇宙的组成和相互作用，以及为什么物质具有我们所观测到的特性。我们希望了解更多被称为“暗”的缺失物质。因此作为“前传”的结束，让我们回到甲壳虫乐队的一句歌词，它也是《弯曲的旅行》中的一句话：“因他难得一见，你得将自己精心装扮。”（Got to be good-looking' cause he's so hard to see.）探索新现象与发现新见解也许会是一个巨大的挑战，但为此付出的等待与努力都是值得的。

致谢

本书涵盖了大量基础知识，我有幸得到了许多慷慨的、头脑丰富的人士的指导。在本书尚未羽翼丰满之时，我已经知道我有如此之多的才思敏捷的头脑可以依靠，这成了我努力前进的强劲动力。我特别要感谢Andreas Machl、Luboš Motl以及Cormac McCarthy。他们不止一遍地阅读本书初稿，在不同的阶段提出了各种宝贵的反馈意见。Cormac McCarthy的高标准、耐心以及对这本书的信心，Luboš Motl作为物理学家的细致、与我进行科学交流的热情，以及Andreas Machl的智慧、热忱与始终如一的支持，都极其珍贵。

其他人的帮助与热情支持也对本书的完成起到了关键作用。Anna Christina Büchmann很有见地，且聪明、善良，她向我提出了不少建议；Jen Sacks以她的智慧与谨慎，帮助我度过了犹豫不决的时刻；Polly Shulman为我指明了重要的方向，并且从很早就开始鼓励我；Brad Farkas的兴趣以及犀利观点，帮助我完成了这项工作；我的编辑Will Sulkin则用他敏锐的眼光与过人的编辑技能，改进了书中的一些关键内容。感谢你们。当然，还要感谢Bob Cahn、Kevin Herwig、Dilani Kahawala、David Krohn与Jim Stone对本书的校正，以及他们在阅读本书最终草稿后所提的宝贵建议。

我也非常感激物理学家Fabiola Gianotti与Tiziano Camporesi。他们为我提供了很多有关大型强子对撞机以及超环面仪器、紧凑 μ 子线圈实验的细节，为我做了很多事情。另外，在通读我为大型强子对撞机及其历史所写的章节中，没有谁能比Lyn Evans做得更好。还要感

谢Doug Finkbeiner、Howie Haber、John Huth、Tom Imbo、Ami Katz、Matthew Kleban、Albion Lawrence、Joe Lykken、John Mason、Rene Ong、Brain Shuve、Robert Wilson、Fabio Zwirner，他们也都针对书中的物理学知识作出了中肯的评论。同时，也要感谢我2010级与2011级的哈佛新生研讨班，他们对大型强子对撞机的理解也融入本书中。

宗教与科学的关系对我来说是比较新的领域，在综合了Owen Gingerich、Linda Gregerson、Sam Haselby与Dave Thom的建议与智慧之后，我更加自信，也更游刃有余。感谢所有帮助我整理科学史的人——Ann Blair、Sofia Talas与Tom Levenson，他们让书中的故事更加准确。

诸如风险性与不确定性的主题可能本身就是有风险的（也是不确定的）。感谢Noah Feldman、Joe Fragola、Victoria Gray、Joe Kroll、Curt McMullen、Jamie Robins与Jeannie Suk，他们是哈佛法学院座谈会的与会者，特别是Jonathan Wiener，与我分享他们的专业知识。更早时期与Cass Sunstein的交流，也让我受益颇丰。创造性可能是另一个富有挑战性的话题，我很感谢Karen Barbarossa、Paul Graham、Lia Halloran、Gary Lauder、Liz Lerman、Peter Mays与Elizabeth Streb，与我分享他们的见解。也特别感谢Scott Derrickson，与他的交流内容是第1章的关键，他纠正了我的一些错误，因为他的记忆力比我好。感谢2010 Techonomy会议主办方邀请我加入演讲小组，为此付出的准备也构成了本书的结语部分。还要感谢本书所提到的其他与我交流的人。感谢Alfred Assin、Rodney Brooks、David Fenton、Kevin McGarvey、Sesha Pratap、Dana Randall、Andy Singleton、Kevin Slavin等人慷慨的反馈与意见，感谢A. M. Homes、Rick Kot的建议与鼓励。

同时感谢其他在我开始这一具有挑战性的工作之初时，对我大加鼓励的人。感谢John Brockman与Ecco的Dan Halpern的付出，他们使本书的出版成为可能；感谢Matt Weiland与他的助理Shanna Milkey帮助把本书的不同部分串联起来。同时还要感谢Ecco其他帮助使本书出版的每一个人，感谢Andrew Wylie为本书领航到最后阶段。我很高兴曾与优秀团队Tommy McCall、Ana Becker以及Richert Schnorr一起工作，他们将我书中复杂的思想用清晰与精确的照片传达了出来。

最后，感谢我的研究合作者与所有同事对我的悉心指导；感谢家人鼓励我对理性的爱；感谢朋友们的耐心与支持；感谢那些曾经帮助过我的人（无论我是否提到了他们的名字）；感谢所有一路走来帮助我不断完善想法的人。

注释

01 神奇的科学尺度

1. Fielding, Henry. *Tom Jones*. (Oxford: Oxford World Classics, 1986) .

02 伽利略的科学求索

1. Levenson, Tom. *Measure for Measure: A Musical History of Science* (Simon&Schuster, 1994) .
2. Hooke, Robert. *An Attempt to Prove the Motion of the Earth from Observations* (1674) , quoted in Owen Gingrich, *Truth in Science: Proof, Persuasion, and the Galileo Affair, Perspectives on Science and Christian Faith*, vol. 55.

03 生于物质世界

1. Doyle, Arthur Conan. *The Sign of the Four* (originally published in 1890 in Lippincott's Monthly Magazine, chapter 1) , in which Sherlock Holmes comments on Watson's pamphlet, "A Study in Scarlet."
2. Browne, Sir Thomas. *Religio Medici* (1643, pt.1, section 9) .
3. Augustine. *The Literal Meaning of Genesis*, vol.1, books 1-6, trans. and ed. by John Hammond Taylor, S.J. (New York: Newman Press, 1982) . Book 1, chapter 19, 38, pp.42-43.
4. Augustine. *On Christian Doctrine*, trans. by D.W. Robertson (Basingstoke: Macmillan, 1958) .
5. Augustine. *Confessions*, trans. by R.S. Pine-Coffin (Harmondsworth: Penguin, 1961) .
6. Stillman, Drake. *Discoveries and Opinions of Galileo* (Doubleday Anchor Books, 1957) p.181.
7. *Ibid.* , pp.179-180.
8. *Ibid.* , p.186.
9. Galileo, 1632. *Science&Religion: Opposing Viewpoints*, ed. Janelle Rohr (Greenhaven Press, 1988) , p.21.
10. 例如, 参见Gopnik, Alison. *The Philosophical Baby* (Picador, 2010) .

04 物中之妙

1. Blackwell, Richard J. *Galileo, Bellarmine, and the Bible* (University of Notre Dame Press, 1991) .
2. Gerald Holton, "Johannes Kepler's Universe: Its Physics and Metaphysics," *American Journal of Physics* 24 (May 1956) : 340-351.
3. Calvin, John. *Institutes of Christian Religion*, trans. by F.L. Battles in *A Reformation Reader*, Denis R. Janz, ed. (Minneapolis: Fortress Press, 1999) .

05 谜般的梦幻之旅

1. Gamow, George. *One, Two, Three.....Infinity: Facts and Speculations of Science* (Viking Adult, September 1947).

06 “眼见”为实

1. Feynman, Richard. *The QED Lecture at University of Auckland* (New Zealand, 1979). See also: *Richard Feynman Lectures, Proving the Obviously Untrue*.
2. 例子可见: Richard Rhodes, *The Making of the Atomic Bomb* (Simon&Schuster, 1986)。

07 寻找宇宙的答案

1. Overbye, Dennis. “Collider Sets Record and Europe Takes U.S. Lead.” *New York Times*, December 9, 2009.
2. 1997年, 欧洲物理学会认可了罗伯特·布绕特 (Robert Brout)、弗朗索瓦·恩格勒与彼得·希格斯三人的成就, 3人再次于2004年获得沃尔夫物理学奖 (Wolf Prize in Physics)。恩格勒、布绕特、希格斯、杰拉尔德·古拉尔尼克 (Gerald Guralnik)、哈根 (C.R.Hagen)、汤姆·基博尔 (Tom Kibble) 6人同时于2010年被美国物理学会授予了J.J. 樱井理论粒子物理学奖 (J.J. Sakurai Prize for Theoretical Particle Physics)。不过在本书全书中, 我会只提到希格斯玻色子与彼得·希格斯, 因为我叙述的重点是物理机制而不是名人。当然, 如果希格斯玻色子被发现了, 最多只能有三人获得诺贝尔奖, 所以排名问题很重要。对这种情况的概述请参考: Luis Álvarez-Gaumé and John Ellis, “Eyes on a Prize Particle,” *Nature Physics* 7 (January 2011)。

10 有黑洞, 还是没黑洞

1. *Physical Review D*, 035009 (2008)。
2. <http://lsag.web.cern.ch/lsag/LSAG-Report.pdf>.

11 这是一项风险差事

1. 例子可参见: Taibbi, Matt. “The Big Takeover: How Wall Street Insiders are Using the Bailout to Stage a Revolution,” *Rolling Stone*, March 2009.
2. 这一点在一些著作中也有提及: J.D.Graham and J.B.Wiener, *Risk vs. Risk: Tradeoffs in Protecting Health and Environment* (Harvard University Press, 1995), especially Chapter 11.
3. 例子可参见: Slovic, Paul. “Perception of Risk,” *Science* 236, 280-285, no. 4799 (1987). Tversky, Amos, and Daniel Kahneman, “Availability: A heuristic for judging frequency and probability,” *Cognitive Psychology* 5 (1973): 207-232. Sunstein, Cass R., and Timur Kuran. “Availability Cascades and Risk Regulation,” *Stanford Law Review* 51 (1999): 683-768. Slovic, Paul “If I Look at the Mass I Will Never Act: Psychic Numbing and Genocide,” *Judgment and Decision Making* 2, no. 2 (2007): 79-95.
4. 例子可参见: Kousky, Carolyn, and Roger Cooke. *The Unholy Trinity: Fat Tails, Tail Dependence, and Micro-Correlations*, RFF Discussion Paper 09-36-REV (November 2009). Kunreuther, Howard, and M. Useem. *Learning from Catastrophes: Strategies for Reaction and Response* (Upper Saddle River, NJ: Wharton School Publishing). Kunreuther, Howard. *Reflections and Guiding Principles for Dealing with Societal Risks*, in *The Irrational Economist: Overcoming Irrational Decisions in a Dangerous World*, E. Michel-Kerjan and P. Slovic, eds., New York Public Affairs Books 2010. Weitzman, Martin L., *On Modeling and Interpreting the Economics of Catastrophic Climate Change*, *Review of Economics and Statistics*, 2009.
5. 例子可参见: Joe Nocera’s cover story on “Risk Mismanagement” in the *New York Times Sunday Magazine*, January 4, 2009.
6. 很多经济学家都讨论过这种不可替代性的问题, 例如: Arrow, Kenneth J., and Anthony C. Fisher. “Environmental Preservation, Uncertainty, and Irreversibility,” *Quarterly Journal of Economics*, 88 (1974): 312-319. Gollier, Christian, and Nicolas Treich, “Decision Making under Uncertainty: The Economics of the Prec

autionary Principle, ” Journal of Risk and Uncertainty 27, no.7 (2003) .Wiener, Jonathan B. “Global Environmental Regulation, ” Yale Law Journal 108 (1999) : 677-800.

7.E.g., Richard Posner, Catastrophe: Risk and Response (Oxford University Press, 2004) .

8.Leonhardt, David. “The Fed Missed This Bubble: Will It See a New One?” New York Times, January 5, 2010.

12 一切都是概率

1.Kristof Nicholas. “New Alarm Bells About Chemicals and Cancer” New York Times, May 6, 2010.

15 真、美与其他科学错觉

1.Ian Stewart, Why Beauty Is Truth, Basic Books, 2007.

16 希格斯玻色子，宇宙万物为何产生

1.On WNYC’ s The Takeaway, March 31, 2007.

17 谁才是世界的下一个顶级模型

1.http://xxx.lanl.gov/PS_cache/arxiv/pdf/1101/1101.1628v1.pdf.

2.Lisa Randall and Raman Sundrum, Physical Review Letters 83 (1999) : 4690-4693.

3.Arkani-Hamed, Nima, Savas Dimopoulos, Gia Dvali, Physics Letters B429 (1998) : 263-272; Arkani-Hamed, Nima Savas Dimopoulos, Gia Dvali, Physical Review D59: 086004, 1999.

4.Randall, Lisa, and Raman Sundrum, Physical Review Letters 83 (1999) : 3370-3373.

19 内外翻转

1.Original short film Powers of Ten by Ray Eames and Charles Eames, 1968; Powers of Ten: A Flip Book by Charles and Ray Eames (W.H.Freeman Publishers, 1998) ; also Philip Morrison and Phylis Morrison and the office of Charles and Ray Eames, Powers of Ten: About the Relative Sizes of Things in the Universe (W.H.Freeman Publishers, 1982) .

20 君之须弥，我之芥子

1.更详细的介绍可以参考：Alan Guth, The Inflationary Universe.Perseus Books, 1997.

22 海阔天空与脚踏实地

1.米哈里·希斯赞特米哈伊 (Mihaly Csikszentmihalyi) 在其著作Flow: The Psychology of Optimal Experience中首先提出了“心流”的概念来描述这种现象。

2.Brooks, David. “Genius: The Modern View, ” New York Times, April 30, 2009.

3.Gladwell, Malcolm.Outliers: The Story of Success (Little Brown&Co., 2008.)

4.Gell-Mann, Murray.The Quark and the Jaguar: Adventures in the Simple and the Complex (W.H.Freeman&Company, 1994) .

5.Teacher’ s Edition of Current Science 49, no.14 (January 6-10, 1964) .

结语 寻路宇宙边缘

1. 其他例子见: Susan Jacoby, *The Age of American Unreason* (Pantheon, 2008) .

译者后记

众所周知，2012年，“科学家发现‘上帝粒子’”的新闻席卷了全球，为了验证希格斯玻色子存在性的大型强子对撞机也在一夜之间家喻户晓。公众对科学的热情也在与日俱增，一些“民间人士”也发表了对该科学发现的一些观点。现代意义下的科学应该从伽利略说起，因为公理化的理论体系以及实验验证的方法才是现代科学的特征。所以，在“什么是科学”这个问题上，为何不来听一听一线科学家的观点呢？

本书作者兰道尔教授绝非公众所认为的“传统科学家要么是nerd要么是geek”——这里我们不单指她那曾经在《时尚》杂志中占据一版的女性魅力，更指她的个人魅力：她对哲学、历史、宗教、文学、艺术和电影等方面，都有着深入的涉猎与思考，并且善于从科学的角度提出自己的见解。作为一位弦理论学家，她的代表性成果是兰道尔-桑卓姆模型。这项研究工作于21世纪初在弦理论和宇宙学界引起了一番轰动。在本书中，兰道尔教授也用一定的篇幅介绍了这项工作，并采用比喻、图示等方法以让读者更好地理解它。

近年来，兰道尔教授的主攻方向是对额外维度的研究。在本书中，她也将花大量篇幅介绍额外维度如何帮助我们解决宇宙学常数、希格斯等级问题、黑洞、暗物质等问题。同时，她也是直接参与大型强子对撞机工作的一线物理学家。她的叙述让我们了解了大型强子对撞机在建造、运行过程中的许多轶事，让我们不禁赞叹大型强子对撞机是多么浩大的一项工程。书中同时向我们指出该实验如何为物理理

论服务，譬如，可以用来检验什么理论模型。其中，她向我们展示的科学研究的苦痛与欢乐，只有身在其中方能体会。可以说，《叩响天堂之门》是一本内容丰富又有趣的科普作品。它将一位理论物理学家对生命、宇宙以及物理规律的理解娓娓道来。

记得当时湛庐文化的编辑找到我，问我认不认识具有物理学背景的专家学者，或者我自己是否感兴趣翻译兰道尔教授的几本著作时，我在几本相关资料中，对这本《叩响天堂之门》产生了极大的兴趣，也立即想到了杨洁老师。杨洁老师在加州理工学院获得博士学位，科研方向是拓扑弦理论，理论物理背景深厚，应是翻译此书的上佳人选。彼时，杨洁老师在美国做访问学者，学术工作繁忙，无法在编辑要求的时间内完成任务。于是，我请缨与她合作翻译本书，并且商定：本书前10章由我翻译，后面的章节由她翻译，并且交换审阅译稿。

在此我感谢湛庐文化的信任与支持；感谢兰道尔教授的平易近人与大家风范。当我们把翻译过程中遇到的30余个问题以电子邮件的形式咨询她时，她耐心地一一给出了解答与回应，并且修正了一些错误。

正如著名物理学家费曼在自己的《费曼物理学讲义》一书中所说：

一个诗人曾经说过，“这个宇宙就在一杯葡萄酒中”，我们大概永远也不会知道他是在什么背景下说这句话的。但是如果我们有足够细致地观看一杯葡萄酒，我们的确看到了这个宇宙。这里有物理学的东西：涡动的液体，其蒸发依赖于风和天气，酒杯玻璃上的反射，我们想象的原子。玻璃是地上岩石的提纯物，在其成分中我们看到了宇宙年龄和恒星演化的奥秘。葡萄酒中的种种化合物又有着怎样奇特的排列？这些化合物是怎样产生的？这里有种种酵素、酶、基质和它们的生成物。在葡萄酒中发现了伟大的结论：整个生命过程就是发酵。所有研究葡萄酒化学成分的人都会像

巴斯德那样，发现许多疾病的原因。红葡萄酒多么鲜艳！把它深深地铭刻在你脑海中吧！如果我们微弱的心智为了某种便利，把这杯葡萄酒——这个宇宙分为几部分：物理学、生物学、地质学、天文学、心理学等，那么让我们记住，大自然并不知道这种分法。所以让我们把所有这些都放到一起，别忘记这杯酒最终的用途。让它最后再给我们一次欢乐：干杯，忘了它！

让我们跟着兰道尔教授一起，饮下《叩响天堂之门》这杯“佳酿”吧！

如何阅读商业图书

商业图书与其他类型的图书，由于阅读目的和方式的不同，因此有其特定的阅读原则和阅读方法，先从一本书开始尝试，再熟练应用。

阅读原则1 二八原则

对商业图书来说，80%的精华价值可能仅占20%的页码。要根据自己的阅读能力，进行阅读时间的分配。

阅读原则2 集中优势精力原则

在一个特定的时间段内，集中突破20%的精华内容。也可以在一个时间段内，集中攻克一个主题的阅读。

阅读原则3 递进原则

高效率的阅读并不一定要按照页码顺序展开，可以挑选自己感兴趣的部分阅读，再从兴趣点扩展到其他部分。阅读商业图书切忌贪多，从一个小主题开始，先培养自己的阅读能力，了解文字风格、观点阐述以及案例描述的方法，目的在于对方法的掌握，这才是最重要的。

阅读原则4 好为人师原则

在朋友圈中主导、控制话题，引导话题向自己设计的方向去发展，可以让读书收获更加扎实、实用、有效。

阅读方法与阅读习惯的养成

- (1) 回想。阅读商业图书常常不会一口气读完，第二次拿起书时，至少用15分钟回想上次阅读的内容，不要翻看，实在想不起来再翻看。严格训练自己，一定要回想，坚持50次，会逐渐养成习惯。
- (2) 做笔记。不要试图让笔记具有很强的逻辑性和系统性，不需要有深刻的见解和思想，只要是文字，就是对大脑的锻炼。在空白处多写多画，随笔、符号、涂色、书签、便签、折页，甚至拆书都可以。
- (3) 读后感和PPT。坚持写读后感可以大幅度提高阅读能力，做PPT可以提高逻辑分析能力。从写读后感开始，写上5篇以后，再尝试做PPT。连续做上5个PPT，再重复写三次读后感。如此坚持，阅读能力将会大幅度提高。
- (4) 思想的超越。要养成上述阅读习惯，通常需要6个月的严格训练，至少完成4本书的阅读。你会慢慢发现，自己的思想开始跳脱出来，开始有了超越作者的感觉。比拟作者、超越作者、试图凌驾于作者之上思考问题，是阅读能力提高的必然结果。



扫码关注湛庐文化，
回复“阅读”
这5种方法，让读过的书变成你的影子

[特别感谢：营销及销售行为专家 孙路弘 智慧支持！]

老 我们出版的所有图书，封底和前勒口都有“湛庐文化”的标志



并归于两个品牌



老 找“小红帽”

为了便于读者在浩如烟海的书架陈列中清楚地找到湛庐，我们在每本图书的封面左上角，以及书脊上部 47mm 处，以红色作为标记——称之为“小红帽”。同时，封面左上角标记“湛庐文化 Slogan”，书脊上标记“湛庐文化 Logo”，且下方标注图书所属品牌。

湛庐文化主力打造两个品牌：**财富汇**，致力于为商界人士提供国内外优秀的经济管理类图书；**心视界**，旨在通过心理学大师、心灵导师的专业指导为读者提供改善生活和心境的通路。



老 阅读的最大成本

读者在选购图书的时候，往往把成本支出的焦点放在书价上，其实不然。

时间才是读者付出的最大阅读成本。

阅读的时间成本=选择花费的时间+阅读花费的时间+误读浪费的时间

湛庐希望成为一个“与思想有关”的组织，成为中国与世界思想交汇的聚集地。通过我们的工作和努力，潜移默化地改变中国人、商业组织的思维方式，与世界先进的理念接轨，帮助国内的企业和经理人，融入世界，这是我们的使命和价值。

我们知道，这项工作就像跑马拉松，是极其漫长和艰苦的。但是我们有决心和毅力去不断推动，在朝着我们目标前进的道路上，所有人都是同行者和推动者。希望更多的专家、学者、读者一起来加入我们的队伍，在当下改变未来。

湛庐文化获奖书目

《大数据时代》

国家图书馆“第九届文津奖”十本获奖图书之一
CCTV“2013中国好书”25本获奖图书之一
《光明日报》2013年度《光明书榜》入选图书
《第一财经日报》2013年第一财经金融价值榜“推荐财经图书奖”
2013年度和讯华文财经图书大奖
2013亚马逊年度图书排行榜经济管理类图书榜首
《中国企业家》年度好书经管类TOP10
《创业家》“5年来最值得创业者读的10本书”
《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
《中国新闻出版报》2013年度好书20本之一
2013百度网·中国好书榜·财经类TOP100榜首
2013蓝狮子·腾讯文学十大最佳商业图书和最受欢迎的数字阅读出版物
2013京东经管图书年度畅销榜上榜图书，综合排名第一，经济类榜榜首

《牛奶可乐经济学》

国家图书馆“第四届文津奖”十本获奖图书之一
搜狐、《第一财经日报》2008年十本最佳商业图书

《影响力》(经典版)

《商学院》“2013经理人阅读趣味年报·心理学和行为科学类最受关注图书”
2013亚马逊年度图书分类榜心理励志图书第八名
《财富》鼎力推荐的75本商业必读书之一

《人人时代》(原名《未来是湿的》)

CCTV《子午书简》·《中国图书商报》2009年度最值得一读的30本好书之“年度最佳财经图书”
《第一财经周刊》·蓝狮子读书会·新浪网2009年度十佳商业图书TOP5

《认知盈余》

《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
2011年度和讯华文财经图书大奖

《大而不倒》

《金融时报》·高盛2010年度最佳商业图书入选作品
美国《外交政策》杂志评选的全球思想家正在阅读的20本书之一
蓝狮子·新浪2010年度十大最佳商业图书，《智囊悦读》2010年度十大最具价值经管图书

《第一大亨》

普利策传记奖，美国国家图书奖
2013中国好书榜·财经类TOP100

《真实的幸福》

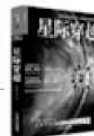
《第一财经周刊》2014年度商业图书TOP10
《职场》2010年度最具阅读价值的10本职场书籍

《星际穿越》

国家图书馆“第十一届文津奖”十本获奖图书之一
2015年全国优秀科普作品三等奖
《环球科学》2015最美科学阅读TOP10

《翻转课堂的可汗学院》

《中国教师报》2014年度“影响教师的100本书”TOP10
《第一财经周刊》2014年度商业图书TOP10



湛庐文化获奖书目

《爱哭鬼小牟》

国家图书馆“第九届文津奖”十本获奖图书之一
《新京报》2013年度童书
《中国教育报》2013年度教师推荐的10大童书
新阅读研究所“2013年度最佳童书”

《群体性孤独》

国家图书馆“第十届文津奖”十本获奖图书之一
2014“腾讯网·啖书局”TMT十大最佳图书

《用心教养》

国家新闻出版广电总局2014年度“大众喜爱的50种图书”生活与科普类TOP6

《正能量》

《新智囊》2012年经管类十大图书，京东2012好书榜年度新书

《正义之心》

《第一财经周刊》2014年度商业图书TOP10

《神话的力量》

《心理月刊》2011年度最佳图书奖

《当音乐停止之后》

《中欧商业评论》2014年度经管好书榜·经济金融类

《富足》

《哈佛商业评论》2015年最值得读的八本好书
2014“腾讯网·啖书局”TMT十大最佳图书

《稀缺》

《第一财经周刊》2014年度商业图书TOP10
《中欧商业评论》2014年度经管好书榜·企业管理类

《大爆炸式创新》

《中欧商业评论》2014年度经管好书榜·企业管理类

《技术的本质》

2014“腾讯网·啖书局”TMT十大最佳图书

《社交网络改变世界》

新华网、中国出版传媒2013年度中国影响力图书

《孵化Twitter》

2013年11月亚马逊(美国)月度最佳图书
《第一财经周刊》2014年度商业图书TOP10

《谁是谷歌想要的人才?》

《出版商务周报》2013年度风云图书·励志类上榜书籍

《卡普新生儿安抚法》(最快乐的宝宝1·0~1岁)

2013新浪“养育有道”年度论坛养育类图书推荐奖



延伸阅读

《叩响天堂之门》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂宇宙求索的漫漫历程。
- ◎ 宇宙如何起源？为什么我们要耗资巨额，建造史上最大型的科学仪器——大型强子对撞机？宇宙万物的真相又如何向我们徐徐展开？
- ◎ 科学小白与科学大V都不可错过的年度最佳科普巨作，韩涛、张双楠、陈学雷、朱进、苟利军、吴岩、万维钢、郝景芳等众多顶尖科学家与科学达人挚爱推荐。



扫码直达本书购买链接



《弯曲的旅行》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂神秘的额外维度。
- ◎ 我们了解宇宙吗？宇宙有哪些奥秘？宇宙隐藏着与我们想象中完全不同的维度吗？我们将怎样证实这些维度的存在？

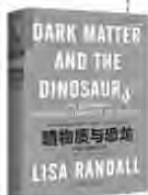


扫码直达本书购买链接



《暗物质与恐龙》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂暗物质以及恐龙灭绝背后的秘密。
- ◎ 暗物质是什么？它是如何让昔日的地球霸主毁灭的？宇宙万物又是如何在看似无关的情况下联系在一起，从而改变了世界的发展的？

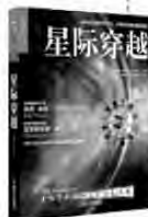


扫码直达本书购买链接



《星际穿越》

- ◎ 天体物理学巨擘，引力波领域大师，同名电影科学顾问基普·索恩巨著，媲美霍金《时间简史》。
- ◎ 国家天文台8位天体物理学科学家权威翻译。
- ◎ 国家图书馆“第十一届文津奖”科普奖获奖图书。



扫码直达本书购买链接



[0]在《弯曲的旅行》这本书中，兰道尔以轻松活泼、浅显易懂的文风，从小尺度出发，深入浅出地向读者介绍了宇宙的奥秘。在另一本著作《暗物质与恐龙》（*Dark Matter and the Dinosaurs*）中，兰道尔从大尺度出发，讨论了宇宙如何进化发展成今日之貌，并探索了暗物质与恐龙灭绝之间错综复杂的联系。《叩响天堂之门》《弯曲的旅行》《暗物质与恐龙》堪称兰道尔“宇宙三部曲”。该系列图书中文简体字版即将由湛庐文化策划出版，敬请期待。——编者注

[1]大型强子对撞机本身非常大，却被用来研究无限小尺度上的事物。在后续章节中，当我们讨论大型强子对撞机的细节时，会提到它体积如此之大的原因。

[2]1942年美国影片《卡萨布兰卡》的插曲《随时光流逝》（又译《年复一年》），由赫尔曼·赫普费尔德（*Herman Hupfield*）作词与谱曲。原书中英文歌词的大意为：时代让我们不安/日月如梭、知识飞速发展/无论是这些，还是四维空间/爱因斯坦先生的观念/让我们有些许疲倦。——译者注

[3]与影片中的版本不同，赫尔曼·赫普费尔德写于1931年的著名歌曲《随时光流逝》清晰地与人们当时所熟悉的物理学新进展相关内容开头：This day and age we're living in/Gives cause for apprehension/With speed and new invention/A nd things like fourth dimention/Yet we get a little weary from Mr.Einstein' s theory.

[4]这里是指量子力学的哥本哈根解释。——译者注

[5]这里指的是今日的科学明日可能被证伪，科学始终被怀疑。——译者注

[6]卢德派是19世纪初英国手工业者组成的团体，他们反对以机器为基础的工业化，在诺丁汉等地从事破坏机器的活动。——译者注

[7]量子力学在周密设计的系统、测量应用于高度统计学的情形或者装置足够精确以使微小的效应浮现出来时，可以产生宏观效果。然而，这并不意味着我们在处理一些平常现象时，经典理论作为一种良好的近似就不适用了，它决定于第12章提到的“精确度”。有效理论允许近似的存在，并且在近似不适用时提供正确的结论。

[8]有时，我会引入指数计数法以解释“我们处于10的幂次的中间位置”是指什么。宇宙的尺度是 10^{27} 米，这个数字是在1后面接27个0，或者说是一千兆兆。人类可想象的最小尺度是10-35米。这个数字是小数点后34个0接一个1，或者说是一百兆兆分之一（从这个描述你能发现指数计数法的方便性了吧）。我们日常的尺度是 10^1 。这里的指数是1，它处于27与-35之间的合理位置。

[9]稠密集是一个数学概念，说一个集合在某个拓扑空间内稠密，是指它“几乎占满了整个空间”。作者在这里使用这一数学概念来表达“原本的信息浩如烟海、阅之不尽，我们只能取其中的一些来处理”的意思。——译者注

[10]马丁·诺瓦克，哈佛大学数学与生物学教授、进化动力学中心（PED）主任，是与著名生物学家理查德·道金斯和爱德华·威尔逊齐名的科学巨星。继达尔文之后为进化论作出突破性贡献的第一人，他的研究成果告诉我们：合作是继突变和自然选择之后的第三个进化原则。推荐阅读其经典著作《超级合作者》，这是一部洞悉人类社会与行为的里程碑式科普著作。诺瓦克在书中从生物学、数学、社会学、计算机科学等多学科角度出发，深入剖析并阐述了生物之间“合作”得以达成的五种机制。该书中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[11]在宗教审判中，虽然第谷是路德教会的信仰者，然而天主教徒并没有针对第谷的著作提出批判，因为他们需要第谷与伽利略观测结果一致、能支持地球静止不动假说的理论框架。

[12]这里的原文是“A Universe in a Grain of Sand”。由于英国诗人威廉·布莱克（William Blake）的诗篇《天真的预言》（*Auguries of Innocence*）的首句为“To see a world in a grain of sand, and a heaven in a wild flower”，其通行译法为“一沙一世界，一花一天国”，出于对会议组织者诗人背景的考虑，译者推想他们在为会议命名时应该是借用了“a grain of sand”的典故，所以这里就也译为“一沙一世界”。——译者注

[13]这里作者用的原词是“spectrum”，即数理概念中的“谱”，大意指把一些连续量按照某种特征分类之后排列在一起，对这个类别的总称呼。这里作者所指的应该是一个“把不同思想以与科学、宗教的契合程度来分类而产生的排列方式”，所以这里把这一词汇译为“思想谱”。——译者注

[14]Rilke, Rainer Maria. Duino Elegies (1922)；这里使用的是醉心研究里克尔的林克教授的译法。——编者注

[15]即平行公设，指如果一条直线与其他两条直线相交，若在直线同侧的两个内角之和小于 180° ，则这两条直线经无限延长后，在这一侧一定会相交。——编者注

[16]凯撒大帝制定的日历，在西方国家一直使用至以阳历取代为止。凯撒的名字叫儒略·凯撒（Julia Ceasar），所以该历法名为“儒略历”。——编者注

[17]理查德·道金斯，英国著名演化生物学家、无神论者。他改变了无数普通人看待生命的方式，也受到了无数赞誉和攻击。想了解他更多传奇故事，推荐阅读其唯一中文版自传《道金斯传》，感受这位无神论斗士的传奇一生。该书中文简体字版已由湛庐文化策划、北京联合出版社出版。——编者注

[18]例如，在古希腊，长度单位斯塔蒂亚（stadia）的长度各地不一、四时不一。

[19]确切地说，应该是重力加速度。——译者注

[20]在低速时，动量近似于物体质量与速度的乘积。不过对接近相对论速度下运动的事物而言，动量等于能量除以光速。

[21]这里的说法不是很确切，能标可以存在简并的情形，同一个能量本征值也可以对应多个不同的波函数，即不同的电子轨道。——译者注

[22]两人共同提出了第一个大统一理论，其中谢尔登·格拉肖是热播美剧《生活大爆炸》（*The Big Bang Theory*）的主角之一，呆萌天才理论物理学家谢尔登·库珀（Sheldon Cooper）的人物原型。——译者注

[23]尽管非常相近，我们现在知道了，大统一并不会出现在标准模型中。然而，大统一却可以出现在标准模型的改进模型中，例如第17章将提到的超对称模型。

[24]波速=波长×频率，而根据波动力学创始人、物质波理论创立者路易·德布罗意（Louis V. de Broglie）的理论， $E=h\nu$ ， ν 为频率。——译者注

[25]即氦原子核。由于那时原子核尚未被发现，故称 α 粒子。——译者注

[26]根据质能方程 $E=mc^2$ ，在高能粒子中往往不区分质量与能量。而且实际研究中一般应用令 $c=\hbar=1$ 的自然单位制，这样 $E=m$ ， $c^2=1$ 。因此，这里说质量等于能量是没有问题的。——译者注

[27]粒子物理学家以电子伏为单位衡量能量，在本书中我也将通篇使用这一单位。1电子伏（eV）是一个自由电子被1伏特的电势差加速时，所能获得的能量。一般情况下，我会使用GeV与TeV这两个单位。1GeV即10亿电子伏，1TeV即1万亿电子伏。

[28]具有讽刺意味的是，在丹·布朗（Dan Brown，《达·芬奇密码》的作者）的著作《天使与魔鬼》中表达了对反物质的关注，而大型强子对撞机是欧洲核子研究中心使用纯粹的正物质进行对撞的第一个对撞机。

[29]本书写作于2011年，这个说法是作者在当时对客观情况的真实判断。2012年7月4日，欧洲核子研究中心宣布大型强子对撞机的紧凑 μ 子线圈探测到质量为 $125.3 \pm 0.6 \text{ GeV}$ 的新粒子（超过背景期望值4.9个标准差），超环面仪器测量到质量为 126.5 GeV 的新粒子。而在2013年3月14日，欧洲核子研究中心发布的新闻稿表示，先前探测到的新粒子正是希格斯玻色子。2013年，因为“次原子粒子质量的生成机制理论，促进了人类对这方面的理解，近来经欧洲核子研究组织属下大型强子对撞机的超环面仪器及紧凑 μ 子线圈探测器所发现基本粒子而获得证实”，彼得·希格斯（Peter Higgs）与弗朗索瓦·恩格勒（François Englert）共同获得诺贝尔物理学奖。彼时，希格斯已是84岁高龄，距他提出希格斯机制的1964年时隔近50年，闻讯后的希格斯泪流满面。他终于在有生之年赢得了其生命与实验结果的赛跑，见证了自己的理论最终得以验证。——译者注

[30]即静质量。——译者注

[31]标准模型到底是否应该包含可能存在的、在中微子质量中扮演重要角色的、非常重的右手征中微子，还是一件不明确的事情。

[32]绝对零度就是0K，等于 -273.15°C 。——译者注

[33]想了解更多有关引力波及宇宙现象的知识，推荐阅读天体物理学巨擘、引力波领域大师基普·索恩（Kip Thorne）的著作《星际穿越》。该书被誉为“媲美《时间简史》的巨著”，并荣获国家最高图书奖第十一届文津图书奖。该书中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[34]原文是intensity，译者认为这里应该是指粒子束的光度。——译者注

[35]SpbarpS最初的目的是加速质子与反质子，不过现在仅用来加速质子，就像大型强子对撞机中SPS（超级质子同步加速器）的功能。

[36]这里是指让束缚态电子的能级变高，变成自由电子的过程。——译者注

[37]这里是指无线电波段的电磁波。——译者注

[38]这里是指电场的方向发生变化，加速器使用的是交流电，当电流反向时，相应的电场不仅不会加速电子，反而会减速电子。——译者注

[39]大型强子对撞机命运多舛，自2007年起发生过多次停机事故，这里作者应该是指距离写作时间最近的一次。——译者注

[40]原文为“the best laid plans o' mice an' men gang aft agley”，出自苏格兰诗人罗伯特·彭斯（Robert Burns）的《致老鼠》（*To a Mouse*），原文应是苏格兰语。——译者注

[41]2009年11月，因为一只飞鸟把面包屑掉入了机器，造成了大型强子对撞机局部过热而再次导致了停机事故。作者在这里应该是指这次事件。——译者注

[42]401(K)计划的名称取自美国1978年《国内收入法》中的401(K)条款。它是美国的一种特殊退休储蓄计划，其深受欢迎的原因是可以享受税收优惠。——译者注

[43]这里的英文原文是“resistance is futile”，这是《星际迷航》中博格人（Borgs）的象征。——译者注

[44]“不知之不知”是美国前国防部部长拉姆斯菲尔德（Donald Rumsfeld）在2002年2月回应记者提问时的名句。——译者注

[45]零和博弈，又称零和游戏或零和赛局，与非零和博弈相对，是博弈论的一个概念，属非合作博弈。零和博弈表示，所有博弈方的利益之和为零或一个常数，即一方有所得，其他方必有所失。——译者注

[46]黑洞的中心处应该是引力奇点（gravitational singularity），黑洞的视界处才是表观奇点。——译者注

[47]金色降落伞（Golden Parachute，又称黄金降落伞），是一种补偿协议，它规定在目标公司被收购的情况下，公司高层管理人员无论是主动还是被迫离开公司，都可以得到一笔巨额安置补偿费用，金额可能高达数百万甚至数亿美元，这会使收购方的收购成本增加，成为抵御恶意收购的一种防御措施。但其弊端是，巨额补偿有可能诱导管理层低价出售企业。——译者注

[48]本书中我使用了“系统的不确定性”而非更常用的“系统误差”。误差常常伴随着错误，而不确定性指的是在实验装置给定的情况下，不可避免的不精确度量级。

[49]同理，人们通常使用“统计误差”来指由于有限统计而造成的测量不确定性。

[50]BPA（bisphenol A），双酚A，一种化工原料，是已知的内分泌干扰素。——译者注

[51]根据最新数据，不确定性的精确度已达到3000万年中差一秒，即千万亿分之一。——译者注

[52]“particle shower”也可翻译成粒子簇射，但此处及后文都将其译为“粒子雨”，这样更生动形象。——译者注

[53]这里指的是两个端帽部分。——译者注

[54]怀疑这里是作者笔误，译者认为应为2007年。——译者注

[55]图14-1中每一行代表一代，每一列的三个粒子表示相似粒子的不同版本。——译者注

[56]三种类型的中微子通过弱相互作用与三种带电轻子配对。然而一旦它们产生了，中微子还可以通过振荡变成其他类型的中微子，不再由此前与其配对的带电轻子唯一确定。中微子有时用相对质量的类别简单标记，有时用文中所述的带电轻子类别所标记。

[57]电影中的一个帮派的名称是喷射机（jet）。——译者注

[58]译者认为此处应该为70000倍。——译者注

[59]然而，在W玻色子、顶夸克与底夸克之间的相互作用，正是顶夸克可以衰变成底夸克和W玻色子的原因。

[60]我们也可以利用能量和动量推断出相对论性质量，其他条件不变。

[61]注意，图14-7区分了玻色子与费米子，即由量子力学分类的粒子。相互作用的媒介粒子和理论假设的希格斯粒子是玻色子。所有其他标准模型的粒子都是费米子。

[62]出自约翰·济慈所作诗歌《希腊古瓮颂》（*Ode on a Grecian Urn*）。——译者注

[63]平衡性本身就是一个对称性的体现形式，参见作者在第15章关于艺术中平衡性的描述。——译者注

[64]指上文所说的超荷规范玻色子与光子。——译者注

[65]也就不会以光速传播，换言之，质量非零。——译者注

[66]有时，人们也争论标准模型中是否包含右手中微子。哪怕右手中微子确实包含于其中，它们也可能极重，以至于在低能过程中并不是那么重要。

[67]弱尺度，全称为电弱尺度（electroweak scale），指电弱理论的典型能量标度，约为246GeV。——译者注

[68]大型强子对撞机运行在1TeV以上，而高能实验可以向下找寻低能粒子，反之则不行。——译者注

[69]例如13051-13050=1，前4个数位都抵消了。——译者注

[70]这里给出了两种命名超对称伙伴的方法。一种是费米子的超对称伙伴，在费米子名字前加s-，如selectron和squark。翻译成“超-”，如超电子、超夸克。另一种是玻色子的超对称伙伴，有的需要将该名词末尾的音节去掉，或者以元音字母结尾的，将末尾的元音字母去掉，再添加-ino，如photino、gravitino、gaugino、Wino、Higgsino。翻译成“-微子”，如光微子、引力微子、规范微子、W微子、希格斯微子。——译者注

[71]这些内容在《弯曲的旅行》中有更详细的介绍。

[72]这里作者指的是第17章所介绍的额外维度理论。——译者注

[73]原文为“moving away”，有误，译者认为应该是靠近（moving towards）。——译者注

[74]Quant是“量化（分析师）”（Quantitative[analyst]）的简称，网络上有时会用一个音近的词“矿工”来代替。——译者注

[75]也就是说尺度增大了 $2^{90} \sim 10^{27}$ 倍（比印度舍罕王棋盘上的麦粒 $2^{64} \sim 10^{19}$ 还要多）。——译者注

[76]包括暗物质。——译者注

[77]一些暗物质粒子是它们自己的反粒子，在此情形下，它们需要找到另一个相似的粒子。

[78]CRESST指超导温度计探测稀有事件（Cryogenic Rare Event Search with Superconducting Thermometers）；EDELWEISS指地下WIMP探测器（Expérience pour DEtecter Les Wimps En Site Souterrain）。——译者注

[79]XENON100是XENON暗物质计划三期实验；LUX指大型地下氙实验（Large Underground Xenon）。——译者注

[80]最新的一期是ZEPLIN-III探测器。——译者注

[81]关于登山难度级别的相关知识，读者可参考：[https://en.wikipedia.org/wiki/Grade_\(climbing\)](https://en.wikipedia.org/wiki/Grade_(climbing))。——译者注

[82]原文为“boldly go where no one has gone before”，是科幻电影剧集《星际旅行：原初系列》开场白中的一句话。——译者注

[83]这本书的书名德语写法是Verborgene Universen。

[84]德语中，“rand”的意思是“边缘”，而“all”的意思是“宇宙”。

all science
business 科学 与 思想 有关

理论物理学大师
丽莎·兰道尔宇宙三部曲

WARPED PASSAGES

UNRAVELING THE MYSTERIES
OF THE UNIVERSE'S
HIDDEN DIMENSIONS

向爱因斯坦宣战的理论物理学大师 丽莎·兰道尔 风靡世界之作

弯曲的旅行

隐秘的宇宙之维

[美] 丽莎·兰道尔 著 窦旭霞 译 李泳 审校

LISA
RANDALL

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社 于2016年10月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：弯曲的旅行

著者：[美]丽莎·兰道尔（Lisa Randall）

译者：窦旭霞

字数：420000

电子书定价：53.99美元

WARPED PASSAGES

Copyright © 2005, Lisa Randall

All rights reserved

目 录

版权信息

湛庐文化“科学素养”专家委员会 寄语

科学伴光与电前行，引领你我展翅翱翔

湛庐文化“科学素养”书系 专家委员会

重磅赞誉

推荐序1

看不见的第五维

推荐序2

发现的激情

中文版序

宇宙的故事

引言

宇宙中的隐秘之维

第一部分 三维之外

01 维度之谜

02 额外维度究竟有多大

03 我们生活在一个膜宇宙上吗

04 理论物理世界的奇幻旅行

第二部分 20世纪两大革命性宇宙观

05 相对论：弯曲的时空

06 量子力学：上帝不掷骰子

第三部分 探秘基本粒子物理

07 物质的已知最基本结构

08 高能粒子加速器的惊人新发现

09 对称：粒子获得质量的基本组织原则

10 基本粒子的质量来源：有趣的自发对称破缺和希格斯机制

11 大统一理论：不同尺度和能量的相互融合

12 等级问题：唯一有效的涓滴理论

13 超对称：超越标准模型的大飞跃

第四部分 弦理论，神秘膜宇宙的创世者

14 弦的律动：轰动世人的超弦革命

15 辅助通道：膜的狂飙突进

16 熙熙攘攘的通道：膜宇宙

第五部分 额外维度宇宙假说

17 人迹罕至的通道：多重宇宙与隔离

18 泄露秘密的通道：四维世界的高维来客

19 宽阔的通道：庞大的额外维度

20 弯曲的通道：一个五维世界

21 漫游弯曲空间王国的“爱丽丝”

22 深远的通道：不为人知的无穷大额外维度

23 曲折延伸的通道：我们生活在一个四维引力的孤岛上

第六部分 弯曲的旅行

24 额外维度：你是在里面，还是在外面

结语

即将被证实的宇宙真相

译者后记



你不是一个人在读书！
扫码进入湛庐“趋势与
科技”读者群，与小伙
伴“同读共进”！

湛庐文化“科学素养”专家委员会

寄语

科学伴光与电前行，引领你我展翅翱翔

欧阳自远

天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

当雷电第一次掠过富兰克林的风筝到达他的指尖；
当电流第一次流入爱迪生的钨丝电灯照亮整个房间；
当我们第一次从显微镜下观察到美丽的生命；
当我们第一次将望远镜指向苍茫闪耀的星空；
当我们第一次登上月球回望自己的蓝色星球；
当我们第一次用史上最大型的实验装置LHC对撞出“上帝粒子”；
……

回溯科学的整个历程，今时今日的我们，仍旧激情澎湃。

对科学家来说，几个世纪的求索，注定是一条充斥着寂寥、抗争、坚持与荣耀的道路：

我们走过迷茫与谬误，才踟蹰地进入欢呼雀跃的人群；

我们历经挑战与质疑，才渐渐寻获万物的部分答案；
我们失败过、落魄过，才在偶然的一瞬体会到峰回路转的惊喜。
在这泰山般的宇宙中，我们注定如愚公般地“挖山不止”。所以，

不是每一刻，我们都在获得新发现。但是，我们继续。
不是每一秒，我们都能洞悉万物的本质。
但是，我们继续。

我们日日夜夜地战斗在科学的第一线，在你们日常所不熟悉的粒子世界与茫茫大宇宙中上下求索。但是我们越来越发现，虽这一切与你们相距甚远，但却息息相关。所以，今时今日，我们愿把自己的所知、所感、所想、所为，传递给你们。

我们必须这样做。

所以，我们成立了这个“科学素养”专家委员会。我们有的来自中国科学院国家天文台，有的来自中国科学院高能物理研究所，有的来自国内物理学界知名学府清华大学、北京师范大学与中山大学，有的来自大洋彼岸的顶尖名校加州理工学院。我们汇集到一起，只愿把最前沿的科学成果传递给你们，将科学家真实的科研世界展现在你们面前。

不是每个人都能成为大人物，但是每个人都可以因为科学而成为圈子中最有趣的人。

不是每个人都能够成就恢弘伟业，但是每个人都可以成为孩子眼中最博学的父亲、母亲。

不是每个人都能身兼历史的重任，但是每个人都可以去了解自身被赋予的最伟大的天赋与奇迹。

科学是我们探求真理的向导，也是你们与下一代进步的天梯。

科学，将给予你们无限的未来。这是科学沉淀几个世纪以来，对人类最伟大的回馈。也是我们，这些科学共同体里的成员，今时今日想要告诉你们的故事。

我们期待，

每一个人都因这套书系，成为有趣而博学的人，成为明灯般指引着孩子前行的父母，成为了解自己、了解物质、生命和宇宙的智者。

同时，我们也期待，

更多的科学家加入我们的队伍，为中国的科普事业共同贡献力量。

同时，我们真诚地祝愿，

科技创新与科学普及双翼齐飞！中华必将腾飞！

刘阳自述

湛庐文化“科学素养”书系 专家委员会

主席

欧阳自远 天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

委员（按拼音排序）

陈学雷 国家杰出青年科学基金获得者，国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

陈雁北 加州理工学院物理学教授

苟利军 中国科学院国家天文台研究员，中国科学院大学教授

李淼 著名理论物理学家，中山大学教授，中山大学天文与空间科学研究所所长、物理与天文学院行政负责人

王青 清华大学物理系高能物理核物理研究所所长，中国物理学会高能物理分会常务理事

张双南 中国科学院高能物理研究所研究员和粒子天体物理中心主任，中国科学院粒子天体物理重点实验室主任，中国科学院国家天文台兼职研究员和空间科学研究部首席科学家

朱进 北京天文馆馆长，《天文爱好者》杂志主编

朱宗宏

北京师范大学天文系教授、博士生导师，教育部“长江学者”特聘教授，北京天文学会理事长

重磅赞誉

韩 涛 著名理论物理学家，美国匹兹堡大学物理天文系杰出教授
匹兹堡大学粒子物理、天体物理及宇宙学中心主任

人类真的生活在一个具有多维空间的膜宇宙之上吗？暗物质真的是毁灭“地球霸主”恐龙的“幕后黑手”？发现了“上帝粒子”希格斯玻色子的大型强子对撞机，以及未来的超级对撞机，会为这些玄妙的问题提供深刻的答案吗？听天才理论物理学家丽莎·兰道尔教授用妙趣横生的案例、通俗易懂的语言，对科学求索的真相与未来娓娓道来，让人欲罢不能。这是时下科学研究前沿最振聋发聩的声音！振奋人心，启迪心智！

张双南 中国科学院高能物理研究所和国家天文台双聘研究员
中国科学院粒子天体物理重点实验室主任

我们还没有探测到暗物质，但恐龙的灭绝竟然是暗物质造成的？兰道尔“宇宙三部曲”将告诉读者，想理解地球和人类的现在、历史与未来，我们必须搞清楚物质最深层次的结构和宇宙最大尺度的规律！唉，我真为其他想写类似主题的作家们担心，再写出这么出色的书恐怕很难了。

陈学雷 国家杰出青年科学基金奖获得者
国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直开展着最前沿的科

学研究。在这套科普书中，兰道尔教授介绍了物理学家们是如何研究、探索宇宙之谜的。

她并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家们现在正在进行的猜想和探索，使读者真切地欣赏到科学研究的丰富多彩和趣味，体验科学家们在构造假说、探索未知、获得新发现时所体验到的激情。我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。

朱进 北京天文馆馆长

在兰道尔教授的笔下，额外维度、暗物质、暗能量、对撞机，这些科学家的“烧脑伙伴”也变得平易近人起来。这套科普书系通俗易懂，与晦涩无缘，揭示了即使是门外汉都读得懂的宇宙真相。

苟利军 中国科学院国家天文台研究员，中国科学院大学教授“第十一届文津奖”获奖图书《星际穿越》译者

几千年来，人类一直在试图回答“宇宙是什么”这一古老问题。现代天文观测和研究揭示，宇宙包含了时空和普通物质以及很多神秘“角色”。作为世界知名的粒子物理学家，哈佛大学物理系教授丽莎·兰道尔在她的这套系列丛书中，以其渊博的知识、广阔的视野、通俗的语言，以及丰富有趣的事例，给我们讲述了宇宙的基本组成和包含万物的时空，非常值得一读。作者大胆推断，地球上恐龙的灭绝与银河系中的某种暗物质有关。如果这能够被证实，将颠覆我们对宇宙神秘物质的现有认识。

吴岩 科幻作家，北京师范大学教授

简明扼要、通俗易懂、内容独创。地球人非读不可！

万维钢（同人于野）

科学作家，畅销书《万万没想到》作者“得到”App《万维钢·精英日课》专栏作家

过去几十年来，理论物理学中最酷的话题已经从量子力学、相对论和黑洞变成了超弦、希格斯粒子和暗物质。如果说，黑洞让人着迷、量子力学让人困惑、相对论让人脑洞大开，那这些新概念则更难让人理解！不过一旦你理解了，就会获得更大的智力愉悦感。物理学家一直致力于在不用公式的情况下让公众理解物理学，丽莎·兰道尔正是这项事业的新晋翘楚。她用一贯的机智语言告诉我们，这一代的物理学正在发生什么。

郝景芳 2016年雨果奖获得者，《北京折叠》作者

在这个信息爆炸的时代，我们收到的碎片化信息太多，反而难以获得真知。碎片化文章看得再多，也不如读一本真正的好书，尤其是深入浅出、结构恢弘的好书。兰道尔“宇宙三部曲”就是难得一见的、视野辽阔的好书，每一本都选择了令人好奇的话题：宇宙结构、宇宙历史、宇宙物质，并且还与恐龙灭绝这样有趣的话题相结合，更加吸引人，让人读起来手不释卷。而最为难得的是，兰道尔的文笔简洁、优美，你在书中找不到像一般物理学科普图书那种艰深晦涩的语句。她用小说一样的文笔娓娓道来，让你理解人类对宇宙最全面的认知。

史蒂芬·平克 著名认知心理学家，科普作家
畅销书《心智探奇》《思想本质》《语言本能》作者

兰道尔说：“在物理界，有几个人从一开始就理解我们，而且相信我们是正确的。令我们感到无比幸运的是，史蒂芬·霍金就是其中一位。”兰道尔在《弯曲的旅行》一书中重在唤起读者的好奇心，而不是讲述专业知识。她耐心地引领读者从简单的直觉经验出发，穿越当代物理学的主要概念，直达令人兴奋的科学最前沿。

马丁·里斯爵士 英国皇家学会会长，剑桥大学天文学家，宇宙学大师

《弯曲的旅行》作者兰道尔可以说是最为杰出的宇宙学家之一。她告诉我们，在咫尺之外可能还有另外一个宇宙，可惜我们看不到它，实际上我们被束缚在了三维空间——这实在是太吸引人了！

戴维·格罗斯 诺贝尔物理学奖得主

作为一位杰出的探索额外维度的先驱，兰道尔不仅讲述了自己的探索历程，而且驾轻就熟地描述了粒子物理学早期发展的广阔领地。正是它们，引出了更多维度的假说。

推荐序1

看不见的第五维

韩 涛

美国匹兹堡大学物理天文系杰出教授

美国匹兹堡大学物理天文系杰出教授

匹兹堡粒子物理、天体物理与宇宙学中心主任

人类对自然现象的好奇心驱使着我们探索时间和空间（时空）的奥秘。时空是物质存在的形式，可是时空本身到底以什么形式展现、如何确切地描述，却是古今中外物理学家、数学家以及哲学家们煞费苦心研究的问题。自从科学诞生以来，对于时空性质的研究就几乎成了永久的课题。

本书作者丽莎·兰道尔现任教于哈佛大学。她曾是普林斯顿大学物理系的第一位终身女性教授，也是哈佛大学、麻省理工学院的第一位理论物理系的终身女性教授。从事高能理论物理研究工作近30年来，兰道尔获得了多项开创性的重要成果，是当前最有影响力的理论物理学家之一。她和拉曼·桑卓姆（Raman Sundrum）十几年前提出的额外维度模型，与另一个大额外维度理论一起，产生了划时代的意义。现在，兰道尔仍然活跃在科学前沿，致力于基本粒子理论和宇宙学的研究。

作为对额外维度理论的最新进展有开创性贡献的物理学家之一，兰道尔相信额外维度的存在。可不管这些卷曲的额外维度有多大，都

是我们肉眼所看不见的！那么，我们如何从实验中确认它们的存在呢？幸运的是，这种理论很可能在我们目前及未来的实验中留下了蛛丝马迹。现在，设在日内瓦欧洲核子研究中心（CERN）的大型强子对撞机（LHC）已经投入运行，而它也是迄今国际上最大的科学实验装置。

LHC可以使高能质子对撞，产生高能反应。然后，能量转变成质量，产生新的粒子。一种被称作卡鲁扎-克莱因的粒子就是额外维度理论预言的确证，也是LHC实验中希望发现的。如果LHC中的高能反应有幸达到强引力的范畴，那么还可能产生高维黑洞！

与许多高能物理学家一起，兰道尔也积极投身于LHC实验的研究，她近来关于卡鲁扎-克莱因粒子及高维黑洞现象的理论预言，都对LHC实验具有指导意义。高能物理领域热切期待着革命性的新发现！这种额外维度理论甚至对早期宇宙暴胀理论，以及暗物质的来源都会提供一种新视角。

从目前的观测看来，我们生活在一个均匀的、各向同性的三维空间加一维时间，即四维时空中，还没有实验证据揭示三维空间外的“额外维度”的存在。然而，自20世纪初以来就有理论物理学家提出额外维度存在的可能性。这些额外维度不像我们周围的三维空间那样延伸，而是卷曲到小得看不见。有希望统一引力与量子力学的弦理论，恰恰要求额外维度的存在。这个观点引起了物理学家及数学家的极大兴趣，甚至被认为是物理学领域的一次革命！传统的弦理论认为，我们生活的三维空间只不过是一个更高维空间中的一部分，但现代弦理论指出，额外维度可以微小到永远无法被观测的程度，即接近所谓的“普朗克长度”，约 10^{-33} 厘米。如果真是这样，额外维度的存在与否，似乎对我们的可观测世界也没有什么作用，因此其存在很难由实验来证实或证伪。

但是，理论物理学研究的最新结果改变了物理学家对额外维度的认识。其实，额外维度可以很大，有的甚至可以延伸到无穷大！它们可以不是平坦的，而是变形的、卷曲的。这种观点来源于爱因斯坦广义相对论中的一个概念。物质能量对时间和空间的制约导致了这些异常性质的出现。更引起物理学家重视的是，这些新理论还有其他方面的意义和应用。比如，与我们周围的核力和电磁力相比，引力如此之微弱。这种卷曲空间的新理论可以解释这个长期以来的困惑：在严重卷曲的空间里，引力可以在某些区域里表现得很强，而在其他区域里表现得很弱。进而，这种卷曲的额外维度理论还可以帮助理解弱核力得以稳定存在的事实。

在本书的最后一部分，兰道尔提出了一个让读者感到惊讶的问题：究竟什么是维度？“确定一个点所需要的独立坐标数就是维度”，不是吗？其实答案要比想象的复杂得多！物理现象都是由物质间的相互作用所决定的。如果存在很强的相互作用，那么维度的概念就会变得更加复杂。十几年来，理论物理学的最新进展表明，一个四维的强相互作用理论可以等效于一个高维的引力理论，即所谓的对偶性。在这种情况下，维度的概念不再明确。其实，弦理论学家早就发现了另外一个对偶性，称作T对偶。在弦理论里，卷曲空间的极小体积和极大体积产生的物理结果是一样的，这使得维度的含义再次变得模糊起来。这些属性实在太玄妙了！兰道尔以及许多顶尖的物理学家都认为，时间和空间一定还有更为基本的描述。

本书以艾克和阿西娜兄妹俩的对话开始引入维度，每一章都会以他们的梦境、经历或故事来代入主题，别有一番风味。本书语言轻松易懂，且图文并茂。书中多处以有趣的例子、图画和故事来描述与比喻深刻的理论：从在婴儿床里爬行时感受到的维度，到不粘锅中准晶体的高维结构，再到以浇花时水的流量比喻力的强度，又到用漏斗的形状来比喻卷曲的额外维度，等等，既生动又贴切，极具可读性。而

且该书的翻译严谨，语言通俗，适合于对现代前沿科学有兴趣的广大国内读者。

时空的性质的高能物理和宇宙学中最重要研究主题，那么额外维度真的存在吗？这样引人入胜的想法到底会不会得到实验观测的肯定呢？读者们还是自己从这本妙趣横生的书中寻找答案吧！

发现的激情

陈学雷

国家杰出青年科学基金获得者

国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

拿到这套书的样章，让我想起20多年前（1993年），我作为一名物理学研究生，参加了由李政道先生创办的中国高等科学技术中心组织的一个国际物理学会议。会议日程上列出的报告中有几位大名鼎鼎的学者，他们的名字，我们在粒子物理学教科书中早已熟悉。但当时还有一个我不很熟悉的名字“Lisa Randall”，而且在日程中排在十分显著的位置。会议开始后，我见到了她：一位面容美丽、身材苗条的女子。她看上去似乎比我大不了几岁，却十分高冷。而且，我听说她酷爱攀岩。然而在会议中，无论是演讲、问答还是讨论，她都显得学识渊博、机敏睿智、充满自信，与那些年龄、资历都老得多的学者辩论时，完全不落下风，成为会议的中心人物之一。这完全打破了我那时对女性物理学家的错误刻板印象。诚然，我从小遇到过很多成绩比我更优秀的女同学，但也许是因为女孩子们的谦让、文静和不好争辩，总让我怀疑她们不过是比我更用功、更擅长作业和考试而已。对于她们是否能深刻地思索或者作出创造性的发现，我内心总有一点儿怀疑。在物理学发展史上，女性物理学家，特别是理论物理学家，也确实屈指可数。然而，站在我面前的就是一位活生生的杰出的女性物理学家，这证明之前我错了。当然，自那之后，我有幸遇到过很多优秀的女性物理学家，其中也包括丽莎·兰道尔教授的一位中国女弟子

苏淑芳博士。她们都向我证明了，女性在物理学或者其他科学研究中，完全可以取得毫不逊色于男性的成就。

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直在进行着最前沿的科学研究。她有许多卓越的成就，其中最著名的是她与桑卓姆合作提出的“额外维度”模型。在这个模型中，我们所熟知的三维空间只是高维空间中的“膜”（参见《弯曲的旅行》一书）。兰道尔教授的这套科普书系，介绍了物理学家是如何研究、探索宇宙之谜的。《叩响天堂之门》一书不仅介绍了大型强子对撞机所进行的研究的意义，也用科学的道理和事实，澄清了人们对科学的各种误解；《弯曲的旅行》一书，重点是对高维空间的探索；《暗物质与恐龙》一书则介绍了作者提出的一种特别的暗物质模型，并就此提出了一个关于恐龙灭绝的有趣假说，借此又阐述了从宇宙起源到暗物质、从太阳系演化到恐龙等多方面的知识。这三部著作的共同特点是，作者并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家现在正在进行的猜想和探索，当然也清楚地说明了哪些仍仅仅是猜想和假说。这些猜想也许未必都正确，其中许多可能也会被未来的实验和观测所否定。但是，对这些内容的介绍更可以使读者真切地欣赏到科学研究的丰富多彩和趣味，体验到科学家们在构造假说、探索未知、获得新发现时所体验的激情。

我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。我也特别希望，这些书能鼓励那些喜爱科学、希望未来从事科学研究的女孩子们。

中文版序

宇宙的故事

得知我的三本书将在中国出版，我感到十分兴奋。不论在理论物理学还是在实验物理学的舞台上，中国都在扮演着日益重要的角色。

我有一些优秀的中国学生以及博士后，而且我也发现，近年来在中国这片土地上，人们对我研究方向的兴趣正在不断增长。不仅如此，中国的实验物理学也在近期取得了一些重要成果。例如，在大亚湾中微子实验室中对最轻、最重中微子混合的振荡测量，其结果震惊了世人，而且它比人们的预期早了至少一年。现有的暗物质探测器，包括PandaX与CDEX，标定了一些重要的能量范围，并且仍在不断探索，以揭示神秘的暗物质粒子的本质。展望未来，计划在中国建造的最大型的对撞机至关重要，它将成为国际主要的粒子物理学实验装置，并能够胜任探索超越已知领域的重任。

作为一位理论物理学家，我的研究领域涉猎甚广，小到物质的内部结构，大到宇宙、空间的本质。这些研究令人兴奋，然而又很难向他人解释清楚——在没有对应语境的情况下更难说清。这三本书给了我一次机会，不仅可以向世人解释我的研究，还可以同时解释作为我研究基础的量子力学、相对论、粒子物理学与天体物理学等物理学知识。我将乐于讲述一些展现这些领域中研究前沿的宏大故事。

我在《叩响天堂之门》一书中解释了科学的本质。它强调了尺度的重要性，也即如何在基本粒子、原子、普通物质或是宇宙的尺度上

思考科学问题。《叩响天堂之门》一书也探索了科学的发展历程、什么是“对”与“错”，以及创造力在科学发展中的意义。在这本书中，我还预测了大型强子对撞机（LHC）上的物理结果。大型强子对撞机是建造在日内瓦附近的大型加速器，能让高能质子对撞，以产生新的粒子与新的相互作用，它可以用来研究人类之前所不能及的更小尺度。这本书解释了大型强子对撞机如何运作，以及在这一实验中，科学家正在研究什么以及他们未来将要研究什么。

《弯曲的旅行》一书讲述了我对空间中可能存在的卷曲的额外维度的研究——额外维度是在我们容易观察到的三个维度（左-右、前-后、上-下）之外的某个维度。额外维度可能具有重要意义，它将解释基本粒子的质量，并为它们之间的相互作用引入新的理论可能性。这些卷曲也将容许空间具有一个无穷大的额外维度，它将与我们观测到的一切事物相容。为了讲述这个故事，我回顾了前沿研究中的量子力学、相对论、粒子物理学的基础（既有理论，也有实验），还回顾了弦理论。在这本书中，我会讲述我们是如何把所有研究领域联系在一起，我们是如何得到了这一切，以及我们已经走到了哪一步的大故事。

《暗物质与恐龙》一书，既向外审视宇宙的宏大图景，又向内一窥物质的内部结构。它解释了暗物质的本质及其在宇宙演化中扮演的角色——暗物质是宇宙中捉摸不定的物质，只与引力而不与光相互作用。《暗物质与恐龙》一书也强调了物质的基本性质与我们今日所见的地球、宇宙之间的联系。这本书的内容不仅涵盖了宇宙学，还涉及星系、太阳系和地球之间的相互作用，及其与周边环境的联系。在这一旅程中，我还将解释我对暗物质的新理念：暗物质可能包含了某个小组分，这个组分通过自身的媒介物质——光进行相互作用，而普通物质不与之相互作用。这可能会产生激动人心的结果，包括在银河系

平面上，暗物质盘将形成，其引力效应可能导致巨大的流星体撞击地球，从而最终导致恐龙的灭绝。

《叩响天堂之门》、《弯曲的旅行》与《暗物质与恐龙》三本书包含了粒子物理学的广阔思想领域，是对我的研究以及更宽泛的粒子物理学和宇宙学的一个简述。这三本书把许多新颖且多元的理念与科学领域结合在一起，给出了对今日科学家工作状态的一个感性认知。

在完成《暗物质与恐龙》这本书之后，我已继续投身于对暗物质的研究中。在许多已有研究的基础上，暗物质已经是一个比较成熟的研究主题了。我们在实验上有了许多直接的结果，也开始着手于更好地理解用来探索宇宙的天文望远镜、人造卫星是如何阐明“暗物质是什么”这一问题的。同时，理论也在不断发展，人们已经超越了对暗物质粒子非常狭隘的假设，并对“暗物质如何相互作用”有了更深刻的想法。我正在思考关于暗物质粒子全部带电（而不只是一小部分带电）的可能性，并自问这一假设可能导出什么结果。现有的研究忽视了某些使这个假设可行的重要结果，或许这可能只是因为：暗物质着实不像粒子物理学家之前所假设的那么无趣。

理解暗物质这种神秘物质的本质，是一个非常令人兴奋的研究主题，毕竟它占据了全部物质能量的85%。我希望中国的读者能享受这一旅程：跟随我一起探索我们是由什么组成的，宇宙中的相互作用是如何发生的，以及我们这些科学家是如何研究宇宙问题的。我确信，你们将会从中学到许多新知识，同时又能提出属于自己的问题和观点。



额外维度是什么？我们的三维世界之外真的存在一个膜宇宙吗？扫码关注“庐客汇”，回复“弯曲的旅行”听兰道尔教授用3分钟解答额外维度。

引言

宇宙中的隐秘之维

你必须要打扮得漂亮，以免被他忽视。

甲壳虫乐队 (The Beatles)

宇宙有着说不尽的秘密，空间的额外维度可能就是其中之一。如果真是这样，那一定是宇宙一直在隐藏着这些维度，小心地包裹着它们，以免被人发现。如果不仔细看，你还真想不到会有什么玄妙。

关于维度认知的错误信息源于婴儿时代，婴儿床首先引领我们进入了一个三维空间（见图0-1）。当我们在婴儿床里爬行时，是在一个二维的平面上；等我们能站起来往外爬时，又多了垂直的第三维。自那时起，物理定律（更不用说人的常识）就加强了三维理论的信念，排斥任何可能存在更多维度的设想。

然而，时空很可能与我们的想象大相径庭。我们所熟知的物理理论，没有任何一个明确指出时空就只能是三维的。不假思索地否认其他维度存在的可能性，未免显得草率。正如“上下”与“前后”、“左右”有着不同的方向，是一个不同的维度一样，在浩淼的宇宙中也可能存在其他全然不同的新维度，尽管看不见、摸不着，却合情合理。

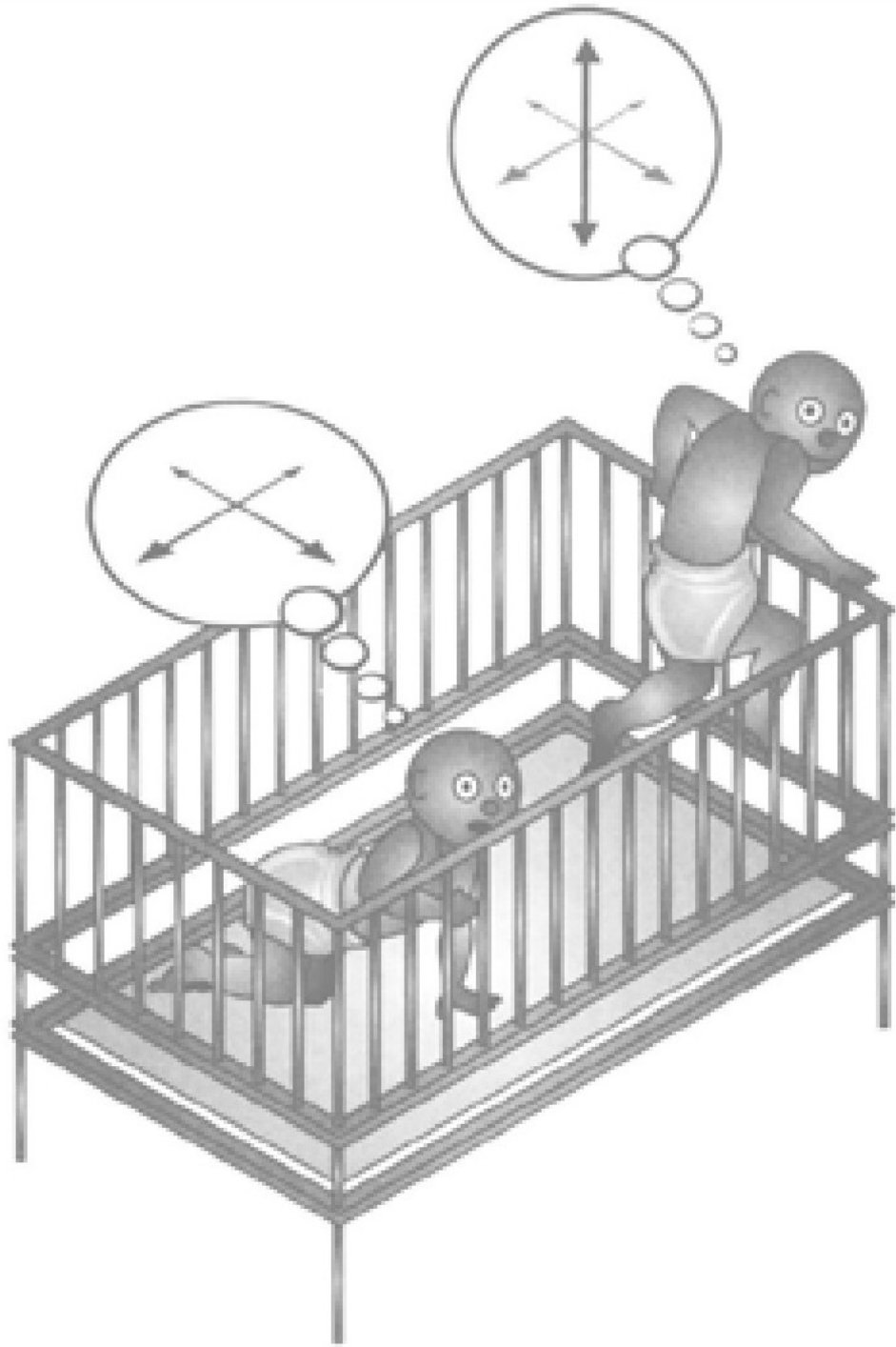


图0-1 一个婴儿的三维世界。

这种未知的不可见的维度，尚没有一个确切的名称。如果它们真的存在，事物就有了一个新的行动方向，因此，当需要一个名称来提

及某个额外维度时，我会称之为“通道”（阐释额外维度时，我用“通道”这个词）。这些通道可能很平坦，就像我们熟知的那些维度一样；也可能是弯曲的，就像哈哈镜里的影像一样。它们可能极其微小——比原子还要小得多：迄今为止，凡是相信额外维度存在的人都这么认为。但是，最新成果显示，这些额外维度也可能很大，甚至是无穷大，却仍然很难被看见。我们的感官只认知三个大的维度，突然多出一个无穷大的额外维度，听起来还真是令人难以置信，而这个不可见的无限维度正是宇宙中众多离奇的可能性之一。在本书中，我们就来一探究竟。

有关额外维度的研究还引出了其他一些引人注目的概念，它们足以满足一个科幻迷的科学幻想，诸如平行宇宙、弯曲几何、三维溶洞等，这些话题听起来更像是小说家的杜撰，甚或是痴人说梦，而不像真正的科学求索的主题。尽管如此，它们却是可能出现在额外维度世界中真正的科学景象。（现在不熟悉这些名词和概念没关系，后面我们会进行介绍和探究。）

不粘锅里的另一个世界

即便包含额外维度的物理学允许出现这些引人入胜的情景，或许你仍会纳闷，一向只专注于对可见现象作出预言的物理学家，为什么会对这些如此看重呢？答案一如额外维度本身一样耐人寻味。新的研究成果表明，尽管人们还没有切身感受到额外维度的存在，也依然对额外维度甚是迷惑，但它们却能够揭示宇宙中某些最本质的秘密。对于我们看到的世界，额外维度可能蕴含着某些深意，而有关它们的思想，或许最终会揭示我们在三维空间中错失的某些关联。

如果我们不能从时间上认识到因纽特人和中国人系属同一祖先，那么就不会明白为什么他们有着相似的体貌特征。同样地，与额外维

度相伴的联系有可能解释粒子物理学中令人费解的方方面面，帮助我们揭开几十年来的难题。因为当空间锁定在三维时，粒子属性和力的关系似乎难以解释；而在高维空间里，那些关系就自然逻辑自恰了。

我相信额外维度吗？我承认，我信。过去，对于可测量范围之外的物理学的猜想——包括我自己的观点，我虽然很着迷，却也存有一定程度的怀疑。我想正是这样，才使我对这些未知既没有失去兴趣，又保持了忠实的态度。不过有时候，某个想法似乎必定蕴含着真理的萌芽。

5年前的一天，正当我坐船穿过查尔斯河进入剑桥镇时，我突然意识到：我的确相信某种形式的额外维度一定存在。环顾四周，我思忖着那些看不见的维度。我对自己世界观的改变感到非常惊讶，这种感觉我之前也曾经历过：作为一个地道的纽约人，在棒球加时赛中，我突然发现，我竟然在为纽约洋基队的对手波士顿红袜队呐喊助威——那同样是一件我从未料到的事。

对于额外维度的深入了解更增强了我对其存在的信心，而同时那些反对的论点则漏洞百出，不能令人信服；离开额外维度，物理学理论就有很多问题无法解答；随着对额外维度的探索，我们发现，可能还有更多类似于我们这个宇宙的多重宇宙，这表明我们目前只看到了冰山一角！即便额外维度与我将做的描述并不十分一致，但我依然认为，它们极有可能以这样或那样的形态存在，其意义也必定是出人意料、发人深省的。

如果得知在你家厨房里就有可能隐藏着额外维度，那么你的好奇心或许立马就会被激发出来。这个痕迹就在不粘锅里：不粘锅上有一层准晶体的镀层。准晶体的结构十分奇妙，其基本秩序只有在额外维度中才能显现。晶体是一种原子和分子排列高度对称的网格结构，其中一个基本元素会不断重复出现。在三维空间中，我们知道晶

体会形成怎样的结构，会以哪种模式排列，可是，准晶体的原子和分子排列却不符合晶体的任何一种模式！

图0-2所示的是准晶体的一种排列模式。真正的晶体，结构精确、规则，就像是坐标纸上的网格；而在准晶体的图例中，我们却看不到这种精确的规则排列。对这些奇特材料的分子排列模式，最贴切的解释办法是借助高维晶体模式的投影——有点儿像三维的影子，以此来揭示它在高维空间里的对称性。三维空间中全然无从解释的模式，却反映了它在高维世界的有序结构。带有准晶体镀层的锅之所以不粘，正是利用了准晶体与常见食物的结构差异，锅里镀层的高维晶体投影与常见食物的一般三维结构是有差异的，原子排列模式的不同使得它们不会粘连在一起。原子的这种不同排列模式，既使我们依稀看到了额外维度的存在，也使我们对一些可见的物理现象作出了解释。

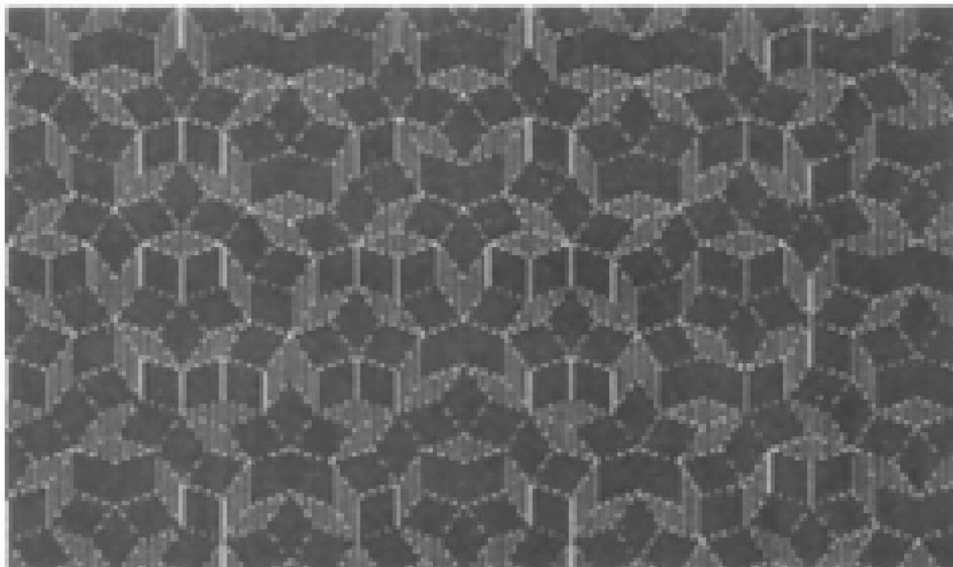


图0-2 彭罗斯拼图。该图是五维晶体结构在二维平面的投影。

相比其他作用力，引力为何不堪一击

额外维度帮助我们理解了一直令人费解的准晶体分子的排列模式，正因如此，现在物理学家们推想，额外维度理论也可以解释粒子物理学和宇宙学中的某些联系——那些只有在三维空间里令人费解的联系。

近40年来，物理学家们一直依赖于粒子物理学中的标准模型理论。标准模型理论（在第7章中我们会进一步讨论）讲述了物质的根本性质以及物质基本成分之间的相互作用力。物理学家创造了一些自宇宙最初几秒以来从未在我们的世界出现过的粒子，以此来检验标准模型。他们发现，标准模型对这些粒子的许多属性都作出了成功的解释。不过标准模型却无法解答一些根本性的问题，而这些问题又是如此直抵了事物的本质，它们的解答有望使我们对这个世界的建构基础及其相互作用产生新的认识。

本书讲述了我和其他科学家是如何探求标准模型难题的答案，以及怎样进入额外维度领域的。额外维度研究的新进展最终必然会成为万众瞩目的焦点，但在此之前，我要先介绍几个强有力的理论——它们是20世纪物理学的革命性进步，我随后讲述的新的理论观点正是建立在这些重大突破之上的。

我们要回顾的话题大致分为三类：20世纪早期物理学、粒子物理学以及弦理论。我们会研究相对论和量子力学的主要观点、粒子物理学的现状以及额外维度可能遇到的难题。我们还将思考弦理论的基本概念。许多物理学家认为，弦理论是现在最有希望将量子力学和万有引力统一起来的理论。弦理论指出，自然界的最基本单位不是粒子，而是基本的振动弦。因为这一理论需要一个超出三维的空间，所以就为额外维度的研究提供了强大动力。我还会讲到膜的作用，膜是薄膜状的物质，在弦理论中如同弦本身一样，是最基本的。我们还将探讨

这些理论的成功之处以及它们未能解答的问题，正是这些未解之谜推动了我们当今的研究。

其中，一个主要的未解之谜是：相比其他作用力，引力为何如此微弱？爬山时，你可能觉得地球引力并不微弱，但想想，那可是整个地球都在作用于你。但反过来看，一块小小的磁铁却能将一枚曲别针吸起，任地球庞大的质量正在将其拉向相反的方向。面对一块小小磁铁的小小拉力，地球引力何以如此不堪一击？在标准三维粒子物理中，微弱的地球引力实在是一个巨大的谜团，而额外维度就有可能成为这个关键的解。

1998年，我和同事拉曼·桑卓姆提出了一个可能的解释。我们的观点基于弯曲几何，这是源于爱因斯坦广义相对论的一个概念。据此理论，时间和空间被物质与能量融合在一起，形成了一个卷曲的或弯曲的时空结构。我和拉曼将此理论应用到了一种新的额外维度的模型中，然后发现了一种时空严重弯曲的结构，即便引力在一个空间区域内很强大，而在其他区域内也会变得微弱。

我们还有更加不同寻常的发现：尽管近90年来，物理学家一直假定额外维度一定是极其微小的，以此来解释我们为什么看不见它们。可是在1999年，我和拉曼不仅发现弯曲时空能够解释地球引力的微弱，还发现看不见的额外维度也可能延伸至无穷大——只要它在弯曲时空里发生适当的卷曲。额外维度可以是无穷大的，却依然隐藏着——这一观点并非所有物理学家都能立即接受，但我物理学领域之外的朋友很快就以为我发现了什么，这并非因为他们完全精通了物理学，而是因为在一次会议上我表述完自己的研究成果之后，在宴会上史蒂芬·霍金让我坐在了他旁边！

在书中，我将解释这些以及其他一些理论成果的基本物理学原理，还有能借以构造它们的空间新概念。然后，我们还会遇到更加离

奇的事：一年后，我和物理学家安德烈亚斯·卡奇（Andreas Karch）发现：尽管宇宙的其他部分都是多维的，但我们却可能生活在空间的一个三维口袋里。这个结果引出了大量可能的时空新结构，它可能由不同的区域组成，每个区域包含不同数量的维度。500多年前，哥白尼提出地球不是世界中心的观点震撼了整个世界，而今，我们不仅不是世界的中心，还有可能只不过生存在一个孤立的三维空间里，而它却也只是高维宇宙的一部分！

新发现的被称作膜的薄膜状物质是丰富的高维景象的重要组成部分。如果说额外维度是物理学家的运动场，那么膜宇宙——在这个虚拟宇宙中，我们就生存在一个膜上，就是一个可望而不可即、多层、多面的丛林体育馆^[1]。本书将带你进入弯曲、庞大、有着无限维度的膜宇宙中，其中一些宇宙只有一个膜，而另外一些有着多层膜，笼罩着看不见的世界。

这一切皆有可能。

未知中的兴奋

假想的膜宇宙是信念的理论性飞跃，它们包含的观点是猜测性的。就如股市一样，冒险可能失败，但也可能带来巨大的回报。

想象暴风雪过后的第一个晴日，你坐在上山的缆车上放眼望去，一片白茫茫的雪野，一个脚印儿都没有，雪地就在你的脚下，闪着银光诱惑着你。你知道，无论怎样，只要踏上这片雪地，这必然会是美妙的一天。有些滑道是陡峭的，崎岖不平；而有些则畅通无阻，如闲逛的巡游；还有些要穿越丛林，像走入迷宫一般。但是，即使偶尔会有错误的转向，偏离滑道，这一天仍将是非常美妙且值得留恋的。

对我来说，模型构建就如这雪地一般，有着同样难以抗拒的诱惑力。对于当前现象的基础理论探索，物理学家称为模型构建，它是在

概念和观点之间穿梭的“冒险旅行”：有时，新观点显而易见；而有时，却如捉迷藏般不可捉摸。但是，即使在我们并不知道这些有趣的新模型最终会把我们引向何方时，它们也常常开拓出不为人知且令人兴奋的新天地。

关于我们在宇宙中的位置，究竟哪个理论会作出正确的描述，我们还尚不知情，其中有一些，我们可能永远都不会了解。但令人难以置信的是，并非所有的额外维度理论都是如此。任何一个解释了引力微弱的额外维度理论，最令人兴奋的一个特征就是，如果它正确，我们很快就能发现。而LHC研究高能粒子的实验很可能会发现支持这些假设以及它们所包含的额外维度的证据。

LHC会让一些高能粒子发生碰撞，生成一些我们从未见过的新物质。如果有任何一个额外维度理论正确，它都可能会在LHC里留下可见的痕迹，这些证据将包括叫作卡鲁扎-克莱因模式的粒子的痕迹。卡鲁扎-克莱因模式是额外维度留在我们三维世界的脚印。如果足够幸运，实验还会记录其他线索，甚至可能会发现高维黑洞。

记录这些对象的探测仪是一个惊人的庞然大物——甚至需要穿戴好肩背带、头盔等登山装备才能爬上去工作。事实上，在瑞士靠近欧洲核子研究中心（CERN）附近，我就曾用这些装备进行了一次冰川旅行，CERN就是LHC安家的地方。这些庞大的探测仪会记录下粒子的性质，然后物理学家利用这些性质重构通过它们的粒子。

诚然，额外维度的证据可能并不直接，必须由我们将各种各样的线索拼接起来，不过几乎所有新近的物理学发现都是如此。20世纪，随着物理学的发展，我们的研究对象已不单纯是能由肉眼直接观察的事物，而是转向了只有通过测量和逻辑推理才能“看见”的事物。例如，我们中学时代就已熟悉的质子和中子的组成成分夸克，它就从未独立地出现过，在影响其他粒子时，它们会留下痕迹，正是跟

踪这些痕迹，我们才发现了它们。暗物质和暗能量的发现也是如此：我们不知道宇宙中大部分的能量来自哪里，也不知道宇宙所包含的大多数物质的本质是什么，但我们知道宇宙中存在着暗物质和暗能量，这并非因为我们直接探测到了它们，而是因为它们对周围的物质有着明显的影响。我们确定夸克或暗物质、暗能量的存在，只是通过间接的方式；同样地，额外维度也不会直接出现在我们面前。但是，即便额外维度存在的迹象并不直接，我们也能知道它们确实是存在的。

首先我得声明：并非所有的观点最终都能被证明是正确的，许多科学家对所有新理论都持怀疑态度，我这里讲的理论也不例外，但猜想是助推我们理解的唯一办法。最后，即使有些细节并不能完全与现实相符，但一个新的理论观点仍能阐明一些物理学原理，启发我们找到真正的宇宙理论。我确信，本书中我们将探讨的额外维度理论绝不仅仅是真理的一丝萌芽而已。

在从事未知事物的研究和推想新观点时，我们不禁会想起，基本结构的发现总是惊人的，而且总会遭遇怀疑和抵触。令人感到奇怪的是，不仅仅是普通大众，有时，就连提出那些基本结构的人，起初也是犹豫不决的。

例如，詹姆斯·克拉克·麦克斯韦创立了经典电磁理论，却并不相信基本电荷单位（如电子）的存在；乔治·斯托尼（George Stoney）于19世纪末提出电子是基本的电荷单位，却并不相信科学家可能会将电子从原子中分离出来，它们只是原子的组成部分（事实上，电子只携带热量或电场）；元素周期表的发现者德米特里·门捷列夫拒绝接受周期表所体现的化合价的概念；马克斯·普朗克提出由光携带的能量是不连贯的，却并不相信他自己观点里暗含的光量子的现实存在；阿尔伯特·爱因斯坦提出了光量子的概念，却不知道它们的力学属性可能使它们等同于粒子——即我们现在所知的光子。但是，并非每个有着正确新思想的人都会否认它们与现实的联系。许多观点，无论是否被承认，最终还是会被证明是确凿无疑的。

还有什么东西等待我们去发现吗？这个问题的答案，我们可以从卓越的核物理学家、科普作家乔治·伽莫夫（George Gamow）的一段“短命”名言中寻得一丝端倪。他在1945年写道：“现在，我们只剩下三种本质不同的实体：原子核、电子、中微子，而不是经典物理学中大量的‘不可见的原子’……如此一来，在物质基本构成的探索中，我们似乎已经找到了根源。”伽莫夫写这些的时候，根本不会想到原子核是由夸克构成的，而且不到30年，人们就发现了它！

我们在继续探求着更基本的结构，如果到我们这里便不再有新成果、新发现了，难道你不觉得很奇怪吗？你不觉得那简直令人难以置信吗？现有理论之间的矛盾透露给我们的是，它们一定不是最终结果。前辈物理学家们既没有当今物理学家的工具，也没有动机来探索本书中将描述的额外维度领域。额外维度，或是支持粒子物理学中标准模型的任何其他东西，都将是一项重大发现。

面对周围的世界，我们有什么理由不去探索呢？

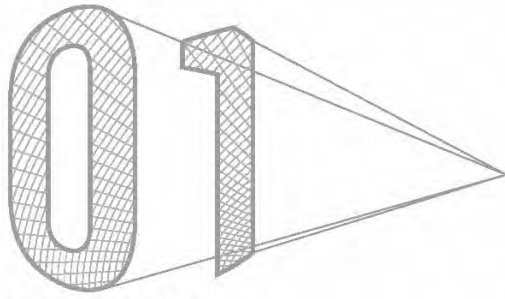


WARPED

第一部分
三维之外

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



维度之谜

WARPED PASSAGES

你可以走自己的路，走自己的路。

佛利伍麦克合唱团 (Fleetwood Mac)

不同维度，而不是异度空间

“艾克，我对自己写的这个故事不太有把握，我想加进更多层次，让它更加多维，你觉得这个主意怎么样？”

“阿西娜，你大哥我对于编故事可不太在行！可是，你要让它更加多维，这不会有什么害处。那么，你是要增加新的人设呢，还是要让现有人设的个性更丰满？”

“都不是，我不是这个意思。我要引进新的维度，类似于新的空间维度。”

“你不是开玩笑吧，你要写异度空间？就是人们会有异度灵魂体验，抑或人们死后或临死时会去的那些地方？^[2]我没想到你还会对那种东西感兴趣。”

“别呀，艾克，你当然知道我不感兴趣。我说的是不同的空间维度，不是异度灵魂世界。”

“可是不同的空间维度又能有什么差别？难道换一张不同边长的纸，比如说，把11厘米×8厘米换成12厘米×9厘米，会带来什么变化？”

“别闹了，你还没弄明白我说的意思。我是真的打算加上一些新的空间维度，就如我们看得见的维度一样，只是沿着完全不同的方向。”

“还有我们看不见的维度？我还以为三维就是世界的全部了呢。”

“耐住性子，艾克。我们很快就会弄明白的。”

“维度”这个词，就像许多描绘空间或是在空间中运动的词汇一样，可以有多种释义——我想目前为止我已经听到了全部。因为我们所看到的事物都有它在空间里的图像，因此在描述许多概念时，包括时间和思维，我们也常常使用一些空间术语。这就是说，应用于空间的许多词汇都有多重含义。当我们在专业领域使用这些词汇时，那些不同的含义会使它们的定义变得非常混乱。

额外维度

当我们将其应用于空间之中时，那个空间本身就超出了我们的感官体验。看不见的东西往往难以描述，而我们的生理机能天生就看不见超出三维的空间。我们无法直接观察额外维度，即使它们真的存在。

“额外维度”这个术语尤其令人迷惑，当我们将其应用于空间之中时，那个空间本身就超出了我们的感官体验。看不见的东西往往难以描述，而我们的生理机能天生就看不见三维之外的空间。光、引力以及所有用于观察的工具，给我们展现的世界都像是一个只含有三维的空间。

我们无法直接观察额外维度，即使它们真的存在。所以有些人会担心，妄图去领悟额外维度的尝试着实让人头疼，BBC的一位新闻主持人在一次采访中就是这么对我说的。但是，令人不安的不是想象额外维度，而是描绘它们。描绘高维世界的图像必然会引出复杂的问题。

而想象额外维度，完全是另外一回事，我们完全可以想象它们的存在。当我和同事们用到“维度”和“额外维度”这些词时，我们的大脑就会浮现出它们清晰的概念。因此，请不要急于弄清楚这些新概念在宇宙中会以什么样的图像呈现出来，在此之前，我首先要解释一

下“维度”和“额外维度”这两个词，以及在后面用到它们时，我所表达的意思。

我们很快就会发现，当不止有三个维度时，“一言胜千图”。当然，一个方程也如此。

从一维到多维

其实我们每天都在与高维空间打交道，只不过大多数人不会以此来考虑问题。想想看，当你做一个重要决定（如购置房产）时，会考虑多少因素？你可能要考虑房子的大小，附近有没有学校，离你想去的地方有多远，建筑风格是什么样的，周围是否有噪声，等等。你得将所有愿望和要求都罗列出来，然后在多维里作出最佳决策。

维数

维数就是完全确定空间一点所需知道的量的数目。高维空间可以是抽象的，也可以是具体的。

维数就是完全确定空间一点所需知道的量的数目。高维空间可以是抽象的，比如刚才用来描述你想买的房子特征的那个空间；也可以是具体的，比如我们即将探究的物理空间。不过，买房子时，你也可以把维数看作你在数据库里记录的条目数量，也就是值得你考虑的备选项的数目。



一个更为浅显易懂的例子是将维度应用于人。

如果你认定某人是一维的，那么在你脑海里就会有一个非常具体的理由：这个人只有一种兴趣。比如，山姆成天

宅在家里，除了看体育节目外什么都不干，那么描述他就只需要上述这一条信息。你甚至可以将山姆爱看体育节目的嗜好用一维图上的一点画出来（见图1-1）。在画图时，要明确单位，这样别人就可以明白沿着这条唯一轴线的距离代表什么。在图1-1中，山姆被表示为横轴上的一点，这张图表现的是山姆每周看体育节目的小时数。



图1-1 一维山姆定位图。

我们再深入地探讨这个概念：艾克，波士顿居民，是一个更为复杂的人。事实上，他是三维的：21岁，爱开快车，每月都会在旺德兰（波士顿附近一个有赛犬场的小镇）输钱。图1-2画出了艾克，尽管图是画在一张二维图纸上的，但三条轴却说明艾克绝对是三维的。^[3]

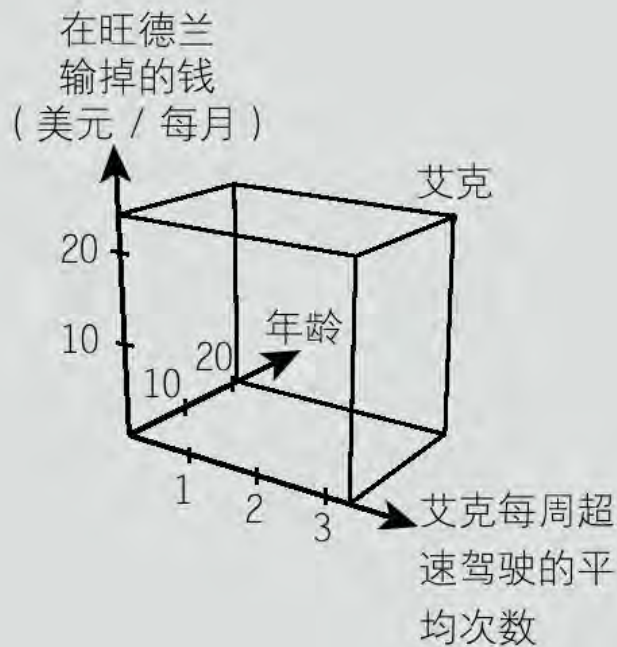


图1-2 三维艾克定位图。V形实线是三维定位图的坐标轴，标志为艾克的点对应的的是一个21岁、每月在旺德兰输掉24美元、每周超速驾驶平均3.3次的男孩。

可是，当描绘大多数人时，我们通常赋予他们不止1个甚至3个特点。阿西娜，艾克的妹妹，是一个酷爱读书的11岁小女孩，擅长数学，了解时事，还养猫头鹰做宠物。也许你想给这些也来个定位（尽管我不太确信是出于什么原因），那样的话，阿西娜就应该是五维空间里的一个点，对应的轴有年龄、每周所读书的数量、数学平均成绩、每天读报的分钟数以及所养猫头鹰的数量。但是，要我画出这样一幅图却很困难，

因为这需要一个五维空间。即便是电脑程序，也只能做得出三维图像。

不管怎么说，在抽象意义上存在着一个五维空间，它有一个5个数的集合，比如11、3、100、45、4，它表示的是：阿西娜11岁，平均每周读3本书，数学成绩满分，每天读报45分钟，现在养着4只猫头鹰。用这5个数字，我就把阿西娜描述了出来。如果你认识她，那么通过五维空间里的这个点，就能认出她。

以上3个人，每个人对应的维数就是我辨认他们所需的特征的数量：山姆，1个；艾克，3个；阿西娜，5个。当然，作为一个真实的人，仅凭这几条信息，远不足以了解他。

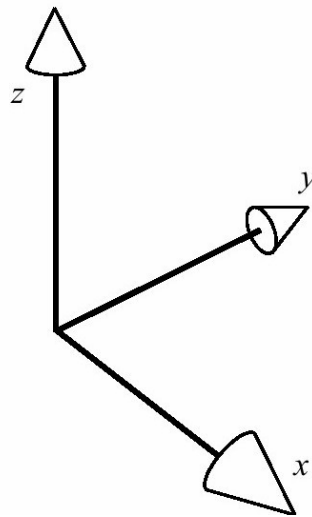


图1-3 三维空间里的三条坐标轴。

在后面的章节中，我们将用维度来探索空间而非人。这里的“空间”指的是，物质存在以及物理过程发生的区域。一个有着特定维数的空间，指的是需要特定数目的量来确定一个点的空间。在一维空间中，一个点的定位图只有一条x轴；在二维空间中，这个点的定位图有x轴和y轴；而在三维空间中，这个点的定位图就有x、y、z三个轴（见图1-3）。

在三维空间中，你只需要3个数字就可以知道自己的确切位置。你确定的数字可以是经度、纬度和海拔，也可以是长度、宽度和高度，或者你还有别的方法来选择你的3个数字。这都无关紧要，问题的关键在于，三维意味着你需要不多不少的3个数字；二维是，你需要2个数字；而多维是，你需要更多数字。

更多维度意味着，你可以在更多完全不同的方向上自由移动。一个在四维空间里的点，只需再加上一条轴——但还是难以描绘。不过想象它的存在并不困难，我们可以用语言和数学名词来尝试。

弦理论提出还有更多维度：它推想有6个或7个额外维度，也就是说，我们需要有6个或7个额外坐标轴来定位一个点。弦理论的最新成果显示，可能还远不止这些维度。在本书中，我会敞开思想，无论有多少额外维度，我都欣然接受。现在，要说出宇宙究竟包含多少维度，未免为时过早。我要描述的有关额外维度的许多概念都适用于任意数量的额外维度，只在很少几个场合有例外情况，届时，我会尽量解释清楚。

可是，要描绘一个物理空间，不仅仅要确定一些点，还需要明确一个度规（metric）。它确立测量的标度，即两点间的物理距离，也

就是轴线上的刻度。这就是说，仅仅知道几个点之间的距离是17远远不够，我们还要知道17代表的是17厘米、17公里，还是17光年。我们用度规来决定怎样测量距离，图上两点之间的距离对应的是其所代表的世界里的量。度规给出了一个测度标准，说明你选择以哪种测量单位来确定标度，就如在一张地图上，0.5厘米可能代表1公里，或者如在米制单位中，它给出了大家认同的1米的标杆。

度规

它确立测量的标度，即两点间的物理距离，它们是轴线上的刻度。但度规还不仅仅确定这些，它还会告诉我们空间是否弯曲或是卷曲。度规包含了有关空间形状的所有信息，弯曲空间的度规既指明长度，也指明角度。

但度规还不仅仅是确定这些，它还会告诉我们空间是否弯曲或是卷曲，就如同一个被吹起的气球表面。度规包含了有关空间形状的所有信息，卷曲空间的度规既指明长度，也指明角度。正如1厘米可以代表不同的长度，角度也可以对应不同的形状。之后，当我们探索弯曲空间与引力的关系时，我还会讲到这些。现在我们只说一点，即球体的表面与一张纸的平直表面是不同的，球面上的三角与纸面上的三角也不同，这种二维空间的差异可以从它们的度规看出来。

随着物理学的进步，度规里存储的信息量也在不断演变。爱因斯坦创立相对论时，认识到第四维度——时间，是与空间的其他三维密不可分的。时间也需要一个标度，于是爱因斯坦用四维时空（在三个空间维之上加入时间维）的度规构造了引力。

更新的研究成果显示，其他空间维度也有可能存在。那样的话，真正的时空度规将包含三个以上的空间维度。人们怎样描绘一个多维空间？那就是说出它有多少维度，以及那些维度的度规是什么。但是，在我们进一步探索度规和多维空间的度规之前，我们要再多想想“多维空间”这个术语的含义。

《巧克力工厂》旺卡梯的秘密

在挪威大名鼎鼎的作家罗尔德·达尔（Roald Dahl）的著作《查理和巧克力工厂》（*Charlie and the Chocolate Factory*）中，威利·旺卡（Willy Wonka）给客人介绍了他的“旺卡梯”，用他的话来说“电梯只能向上和向下，而旺卡梯却能够向前、向后、向左、向右、向旁边，无论是横向、竖向、斜向、侧向，只要是你能想到的方向，它都可以到达”。确实，他的装置可以向任何方向移动，只要不超出我们所认知的三维空间。这可真是个富有创意的好主意！

然而，“旺卡梯”并不真正能够向任何“你能想到的”方向移动。威利·旺卡也真够粗心的，他忘记了额外维度通道。额外维度是一个完全不同的方向，它们难以描述，可是通过一个比方，就很好理解了。

1884年，为了阐释额外维度这一观念，英国数学家埃德温·阿博特（Edwin A. Abbott）写了一本小说：《平面国》（*Flatland: A Romance of Many Dimensions*）。故事发生在一个虚构的二维宇宙——也就是书名里的平面国中，那里的居民都是二维生物（有着不同的几何形状）。阿博特要告诉我们的是，我们这个世界的人对四维概念充满了迷惑，正如平面国的居民感觉三维概念无比神秘，因为他们一生都生存在一个二维空间中，比如桌面上。

对我们来说，只要展开想象就能理解三维之外的世界，而对于平面国的居民来说，三维简直无法想象。那里的所有人都认为，宇宙显然只有他们所感知的两个维度，就像我们坚持三维观念一样，平面国的居民对二维自然也是深信不疑。

小说里的叙事者，正方形A（作者埃德温A的别名），被领进了三维世界中。在接受教育的第一阶段，他仍被限制在平面国，这时，让他观察一个三维球体垂直地穿过他所在的二维世界。因为受限于平面

国，当球体穿过正方形A所在的平面时，他看到的是大小不等的一摞盘子，先是慢慢变大，然后渐渐变小——其实，这就是球体的一个个切片（见图1-4）。

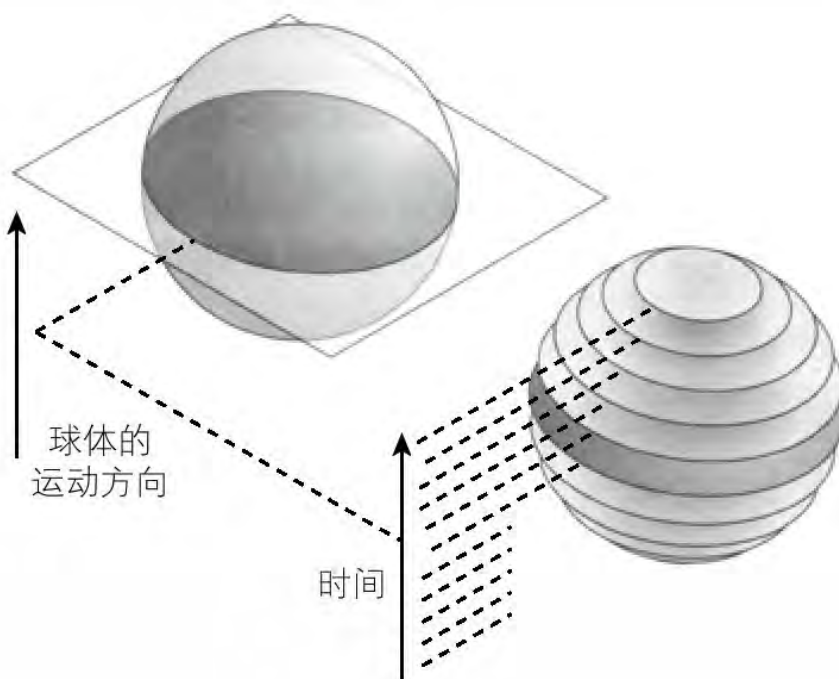


图1-4 球体穿过平面的图示。如果一个球体穿过一个平面，二维观察者看到的就是一个盘子，而观察者在一段时间内所看到的一系列盘子就形成了这个球体。

一开始，这令《平面国》中的二维叙事者颇感费解，他从未想到会有超出二维的东西，也就不能想象会有像球体一样的三维物体。直到正方形A被抬离平面国，进入一个三维世界，他才能真正想象一个球体。从这样一个新视角，他认识到球体就是他所看到的二维切片粘连在一起的结果。即使在二维世界里，正方形A也可以把他看到的盘子描绘成一个时间的函数，从而形成球体（见图1-4）。只有当他经历了第三维度的旅行之后，他才打开眼界，明白了球体和它的第三维。

通过这个类比，我们可以想象，当一个超球体（有4个空间维度的球体）穿过我们的宇宙时，它看起来就应该是一个三维球体随着时间的推移，先是慢慢增大，然后渐渐缩小。令人感到遗憾的是，我们无缘进入额外维度旅行，也就永远看不到一个完整、静态的超球体，但我们还是可以推想物体在不同维度里的样子，即使我们看不见那些维度。我们可以满怀信心地推断，一个穿越三维空间的超球体看起来就应是一系列三维球体。

再举一例，我们设想一下怎样构造超立方体——即立方体在三维空间里的延伸。一条一维直线连接两个点构成一条线段；在此线段上方再加一条一维线段，用另外两条线段将它们连接，就构成了一个二维正方形；以此方式继续，将另一个正方形置于这个上方，再在原正方形的每个边上以另外4个正方形连接，我们就能得到一个三维立方体（见图1-5）。

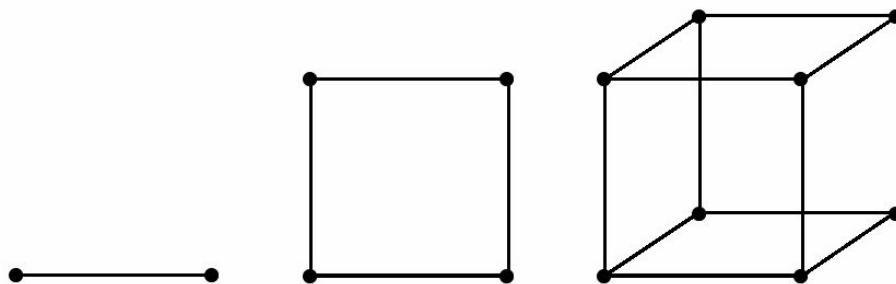


图1-5 三维立方体形成图。图中显示的是我们将低维度物体连接，形成高维度物体的方式。先连接两点形成线段，然后连接两条线段形成正方形，再连接两个正方形形成立方体，最后，连接两个立方体形成一个超立方体（在此没有图片，因为很难画出）。

依此类推，在四维空间中，我们可以得到一个超立方体；在五维空间也能得出某个东西，暂时还没有名称。即便我们三维中的凡人从未见识过这两种物体，我们也可以根据在低维空间中的方法作出推论：要形成一个超立方体（也叫超正方体），就是把一个立方体置

于另一立方体之上，然后加进另外6个立方体，并在原来两个立方体的每个面上进行连接。虽然这种构建很抽象，也很难描绘，但这并不影响超立方体的真实性。

在读高中的时候，我参加过一个数学夏令营。在那里，我看了电影《平面国》。电影结束的时候，叙事者徒劳地指向平面国居民根本看不到的第三维度，用一种愉快的英式口音说：“向上，不是向北。”遗憾的是，当我们试图指向第四维度，即一个通道时，我们面对的是同样的困境。在阿博特的小说中，即使第三维度存在，平面国的居民仍然无法看到，也无法在其中穿行；同样地，我们看不到额外维度，但并不代表它们不存在。因此，尽管我们还没有看到，也没有在这样的维度中穿行，贯穿本书的潜台词仍然是：“不是向北，而是沿着通道向前。”谁知道我们尚未看见的东西会是什么呢？

二生三

本章的后半部分，我们不再考虑三维之外的空间，而是讨论如何凭借有限的视觉能力，用二维图像来思考和绘制出三维图像。明白怎样将二维图像转化成三维现实，将有助于我们解释高维世界里的低维“图像”。在我们的思想随着额外维度一同卷曲之前，权且在这儿做一下热身运动。别忘了，我们在日常生活中一直在与维度打交道，所以其实这并不陌生。

通常，我们所看到的都是物体的表面，而表面只是其外缘。即使这个表面在三维空间里卷曲了，它也只有两个维度，因为在这个表面上，你只需两个数字就能定位一个点。因为它没有厚度，我们推定这个表面不属于三维。

当我们看照片、电影、电脑屏幕或书中的插图时，我们看到的总是二维而非三维的表象，但我们仍能推断出图像所体现的三维实体。

我们能用二维信息构筑三维，这要求我们在制作二维图像时，要忽略一些信息，同时还要保留足够的信息来重现原来物体的基本元素。现在，我们来回想一下，通常将一个三维物体降为二维，都使用什么方法——切片、投影或者全息图，有时干脆忽略一个维度。我们又怎样将其复原，推断出它们所表现的三维物体？



通常，最简单的办法就是切片。每一片都是二维的，但所有这些切片结合在一起就形成了一个真正的三维物体。例如，在熟食店买火腿时，你拿到的肯定是一摞二维火腿片，而不是整个三维火腿。^[4]将所有切片摞在一起，你就可以重构其完整的三维形状。

这本书是三维的，但其中的书页只有二维，所有二维书页集合在一起，就构成了这本书。^[5]我们可以用多种图示方法来显示这种书页的集合，如图1-6中，我们看到的书是平放的，在这个图中，我们再次玩起了维数游戏，因为每条线段代表的都是一页。只要我们明白那些线段代表的是二维书页，这个图示就很清楚了。以后描述高维世界的物体时，我们会使用同样的降维方法。

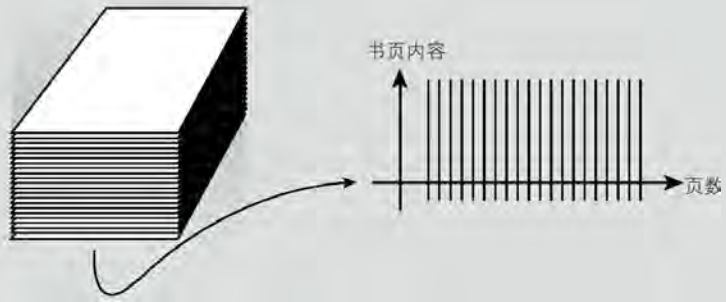


图1-6 由二维书页构成的三维图书。

切片只是以二维代替三维的众多方式之一，另外还有一种方式是投影，这是从几何学中借用的一个术语。为生成一个物体的低维图像，投影给出的是一个绝对处方——墙上的影子就是三维物体的二维投影（见图1-7）。从图中可以看到，我们（或是兔子）做投影时，会丢失许多信息。沿着墙左右和上下的方向，只需两个坐标就能定位影子上的点，但是被投影的物体实际上还有第三维，这个信息是投影所不能保留的。

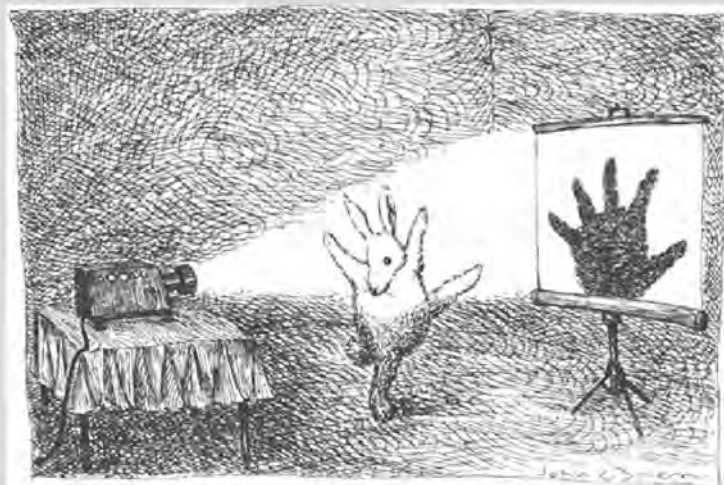


图1-7 兔子的二维投影。投影保留的信息要少于多维物体本身。

实现投影的最简单方法是忽略一个维度，例如，图1-8显示了一个三维立方体在二维平面上的投影，它的投影可以有多种形式，其中最简单的一个是正方形。

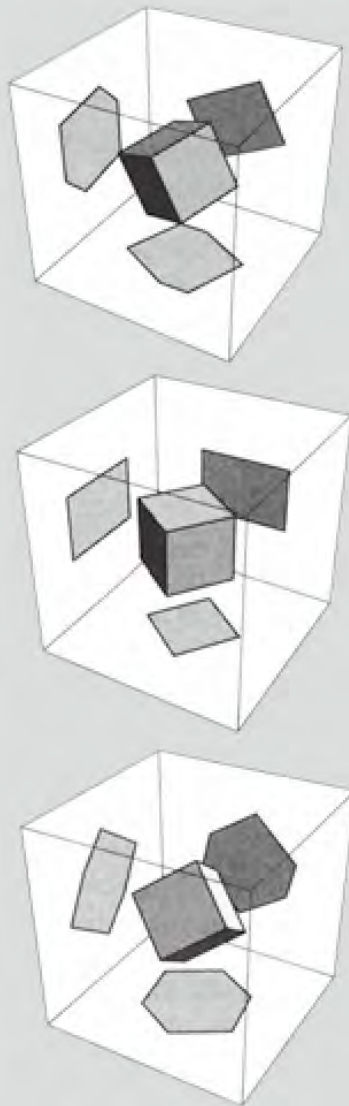


图1-8 一个立方体的投影。注意，投影可以是正方形，像中间那个图，也可以是其他多种形状。

再回到前面有关艾克和阿西娜的定位图，我们可以忽略艾克喜欢超速驾驶这一信息，把它变成一个二维图。我们也可以不管阿西娜所养的猫头鹰数量，这样，就可以绘制出一幅四维而非五维的图，而忽略阿西娜的猫头鹰就是一个投影方式。

投影忽略了原来三维物体的信息（见图1-7），不过，我们用投影制作二维图时，有时会加入一些信息，以帮助我们重获部分丢失的东西。这些附加信息可以是油画或全息图中的阴影、颜色，也可以是地形图中表示高度的数字，但有时什么标识都没有，这样一来，二维特写就不可能提供那么多的信息了。

如果不是我们双眼协同合作，帮助我们重构三维图像，那么我们看到的所有东西都将是投影。如果闭上一只眼睛，你就很难感知远近——一只眼睛只能构建一个三维现实的二维投影，所以我们需要两只眼睛来重建三维图像。

我一只眼睛近视，另一只眼睛远视，如果不戴眼镜，我就不能将两只眼睛看到的影像完美结合，而偏偏我又很少戴眼镜。尽管有人告诉过我，说我重建三维图像时会有困难，我却并没觉得有什么问题：在我看来，物体还是三维的，这是因为我可以凭借阴影和视角（以及我对这个世界的熟悉程度）来重建三维图像。

可是有一天在荒漠里，我和一个朋友正向远处的峭壁前进，朋友不停地告诉我说，我们可以直达峭壁。我就纳闷他怎么坚持要我们径直穿过一块大石头呢？我原以为那块大石头是峭壁直接突出形成的，会完全挡住我们的路。而实际上，它离我们很近，就在峭壁前面，根本没有与峭壁连在一起，因此也就不会挡路。之所以我会有这种误解，是因为我们是在正午靠近峭壁的，当时它没有影子，我也就没有办法以此构建第三维度，而只

有这个维度，才会让我知道远处的峭壁与大石头相隔的距离。直到那次失败，我才意识到阴影和视角的补偿作用。

绘画就是要求艺术家把他们所看到的物体简化成投影。中世纪艺术使用了投影的最简单方式，图1-9显示的是一幅镶嵌玻璃画，画上是—个城市的二维投影，这幅画上没有显示第三维度的任何信息，也没有任何标识或迹象来表明第三维度的存在。



图1-9 中世纪的一幅二维镶嵌玻璃画。

自中世纪起，画家们发展了投影的方法，来部分补偿绘画中的维度缺失。与这一方法对立的是20世纪立体派所使用的方法，立体主义油画（见图1-10，毕加索的《朵拉·玛尔的肖像》 [*Portrait of Dora Maar*]）能同时呈现多个投影，每个投影展现的是一个不同角度，以此来表现物体的三个维度。

但是，自文艺复兴以来，多数西方画家都用透视和阴影来形成第三维度的幻觉。绘画的一个基本技巧就是要将三维世界简化成二维图像，而且要让观赏者逆转这个步骤，重建本来的三维图像或物体。即

使并非所有三维信息都存在，因为我们适应了这种技法，仍然知道该怎样去解读图像。

艺术家甚至尝试过在二维平面上表现更多维度的物体，例如，萨尔瓦多·达利（Salvador Dali）的《耶稣受难图》（*Corpus Hypercubus*，见图1-11）将十字架表现为一个打开的超立方体。超立方体在四维空间里由8个相连的立方体构成，在图1-12中，展示了超立方体的几个投影。



图1-10 毕加索立体主义油画《朵拉·玛尔的肖像》。



图1-11 《耶稣受难图》

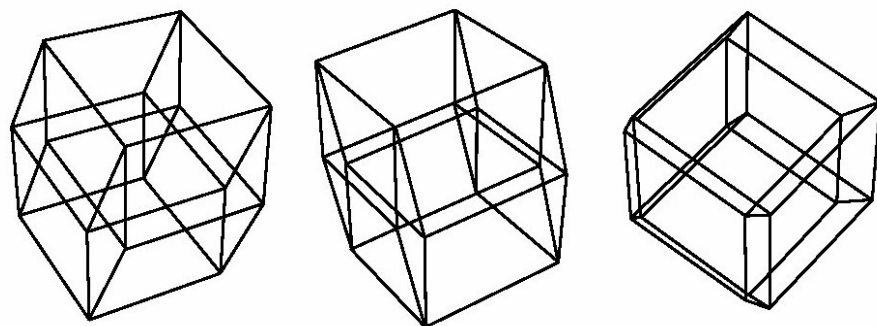


图1-12 超立方体投影图。

我曾在引言中举过一个物理学的例子：不粘锅里的准晶体，它看起来就像是一个高维晶体在我们三维世界里的投影。除了服务于艺术之外，投影当然也有其现实意义，医学中就有许多三维物体投影到二维空间的例子：X光片记录的就是一个二维投影；CAT（电脑辅助X线断层照相术）扫描将多个X光片结合起来，重建了一个更为详尽的三维图像。有了从不同角度拍摄的足够多的X光片，我们就可以把它们串连成一个完整的三维图像；另外还有MRI（核磁共振成像）扫描，它是通过切片来重构三维物体的。

全息照片是另外一种在二维表面记录三维信息的方式。尽管全息图像被记录在低维表面，但它实际上涵盖了原来三维空间的所有信息。也许你在钱包里就能看到这种技术：你信用卡上看起来像是三维的那个图像，就是一张全息照片。

全息图像记录了在不同地点的光的关系，这样，就能够重现一个完整的三维图像。这就好比立体声所使用的原理，它能让我们听出录音时不同乐器所处的位置。利用全息图像所存储的信息，眼睛就可以真实重现它所代表的三维物体。

这些方法显示的是，我们怎样从一个低维图像中获取更多信息，但你真正需要的也许是更少的信息。比如，有些东西在第三维度上非

常非常薄，在这个层面上，根本没有什么有趣的事会发生，尽管这张纸上的墨迹实际是三维的，但我们把它当作二维的也不会有什么损失。除非把它放在显微镜下，我们根本不会去想墨迹的厚度。一根电线看起来是一维的，但如果仔细观察，你还是能看到它有一个二维横截面，因此也该是三维的。

有效理论，忽略细枝末节

忽略另外一个看不见的维度没有什么不对，不仅仅是视觉效果，即使是物理作用，如果微小到难以察觉，也常常可以忽略。科学家们在阐述自己的理论或进行计算时，常常忽略（通常是无意识地）一些微小到不可察觉的物理过程。牛顿的运动定律在他能观测的距离和速度上是有效的，他不需要广义相对论的细节仍然作出了成功的预言；生物学家研究细胞时，也不需要了解中子里的夸克。

挑选相关信息，略去细枝末节，这是一种实用主义的做事方法，我们每个人每天都会这样做。这是一种应对冗余信息的办法，对于你所看到、听到、尝到、闻到或触摸到的任何东西，你都可以选择，是细细品味不放过任何一个细节，还是只需了解其“大概”，抓住主要特征？无论是欣赏油画、品味美酒、阅读哲学，或是安排旅游，你都会不由自主地将自己的想法按照兴趣归类——可能是大小、口味，也可能是观念，而当时你并未发现这些归类有什么相关。适当的时候，你会忽略一些细节，以便将精力集中在你感兴趣的问题上，而不至于被一些无关紧要的线索所迷惑。

这种摒弃细枝末节的过程应该并不陌生，因为它实际上是我们人类一直在做的。以纽约为例，身居这个繁华都市的纽约人都能够看到曼哈顿的细节和变化。对他们来说，闹市区更为繁华、古老，街道更为崎岖、狭窄；而城郊为了方便人们居住、建造了更多的房产，还有许多中心公园和

博物馆。尽管对外人来说，这些差别实际上是很模糊的，但在这个城市之内，它们却真实存在。

但想想远离纽约的人是怎么看的：对他们来说，纽约就是地图上的一个点，也许是一个重要的点、一个有着鲜明特征的点，但仅此而已。即便各不相同，可在别处看来，比如说中西部或是哈萨克斯坦，纽约人并无差别。当我提起这个比方时，住在闹市区（具体来说，是西村）的表弟大为不满，不愿将居住在闹市区和城郊的纽约人归为一类，这更证实了我的观点。但任何一个非纽约人都会告诉他，对并不生活在他们中间的外人来讲，其间的差别实在太小，真的无关紧要。

在物理学中，正式使用这种直觉并以相关的距离或能量来划分范畴已成为常规做法。物理学家接受这种做法，并为它取了名字——有效理论（effective theory）。

有效理论集中研究那些在相关距离内产生“效果”的粒子和力，我们不会用不可测量的、描述超高能行为的参数来描述粒子及其相互作用，只用那些与我们能探测的尺度相关的事物来构建我们的发现。在任何一个尺度上，有效理论都不会深入探究作为其基础的小尺度理论的细节，它只关注有望被测量或者观察到的事物。如果某个事物超出了你所在尺度的精度，那么你无须考虑其详细结构。这种做法并非科学诈骗，而是忽略冗余信息的一种方式，这是获得正确答案的一种“有效”方式。

有效理论

在物理学中，正式使用这种直觉并以相关的距离或能量来划分范畴已成为常规做法。物理学家接受这种做法，并为它取了名字——有效理论。如果某个事物超出了你所在尺度的精度，那么你无须考虑其详细结构。这种做法并非科学诈骗，而是忽略冗余信息的一种方式，这是获得正确答案的一种“有效”方式。

当高维细节超出我们的能力时，所有人，包括物理学家在内，都乐于回到三维世界。正如物理学家常常把一根电线当作一维事物对待一样，如果额外维度极其微小，那么高维细节则无关紧要，我们也常常以低维方式来描绘高维宇宙。额外维度小到无法看见，这样所有可能的高维理论，我们都可以通过这种低维描述来总结其可观察的效果。这个低维描述不受那些额外维度数量、大小和形状的影响，足以实现很多目的。

低维的数目不提供根本描述，但它们却是归纳发现和预言的简便方法。如果你确实了解一个理论的小距离细节（即微观结构），就可以利用它们导出发生在低能描述里的量，否则那些量就只能是等待实验来确定的未知数。

在接下来的章节中，我们将详细讲述这些观点，并探究微小、卷曲的额外维度的作用。我们将首先探讨的那些维度非常微小，小到根本不会产生任何影响；然后，当再次回到额外维度时，我们会探索庞大且无限延伸的维度，它们彻底改变了我们现在描绘的这一图像。



无论什么人，没有出路。

杰斐逊飞机乐队 (Jefferson Starship)

阿西娜“梦游”仙境

阿西娜突然醒了过来。前一天，为了得到有关维度的启发，她看了《爱丽丝梦游仙境》和《平面国》，可是，当天晚上她做了一个非常奇怪的梦，醒来后她才意识到，那是因为她同一天里看了这两本书。

阿西娜梦到自己变成了爱丽丝，滑进一个洞里，遇见了兔子。兔子把她推进一个陌生的世界，尽管阿西娜觉得它这样对待客人太粗鲁了，但她仍热切期盼着自己能在奇幻世界里游历一番。

然而，阿西娜势必是要失望的，因为那只兔子把她送进的是一维世界，一个奇怪却并不奇妙的一维世界。她环顾四周，或者应该说是环顾左右，却发现只能看到两个点，一个在左，一个在右（但是颜色很漂亮哦，她想）。

在一维世界里，所有人连同所有物品都是一维的，他们一同排列在这个维度上，就像是一根细线串成的长长珠链。但即使视觉范围有限，阿西娜仍知道，一维世界肯定不止她看到的这些，因为她能听到耳边人声鼎沸。“这是我见过的最荒谬的棋局！我一个子儿都不能动，连城堡都去不了！”从那尖厉的喊叫声中，阿西娜

听出躲在那个点后面的一定是红桃皇后。好在自己也是一维的，红桃皇后看不到她，不然又得承受红桃皇后的怒火。

但是阿西娜在这一维世界的舒适日子并没有维持多长时间，滑过一条沟后，她又回到了梦中的兔子洞。那里有一个电梯将她带入了一个假想的其他维度宇宙中，兔子当即宣布：“下一站，二维世界。”阿西娜并不喜欢“二维世界”这个名字，但她还是小心翼翼地走了进去。

其实，阿西娜大可不必那么犹豫，二维世界的所有东西与一维世界几乎没什么两样，但她还是发现了一样不同的东西——一个贴着标签的瓶子，上面写着“喝我”。因为实在厌倦了一维世界，阿西娜立即顺从地喝下。她“倏地”就变小了，随即，她看到了第二维度。这一维并不大，它是一个很小的、卷起来的圆圈。现在，她就像站在一根细长管子的表面上。一只渡渡鸟正沿着这个圆圈和自己赛跑，但它想停下来了。看到阿西娜很饿的样子，它善意地给了她一块蛋糕。

刚吃了一点渡渡鸟给的梦幻蛋糕，阿西娜就开始长大。只吃了几口（这一点她很确信，因为她仍感觉很饿），蛋糕就快没了，只剩下一点碎屑。她想，还好，至少还有点碎屑，可要使劲眯起眼睛才能看到。不止蛋糕从她的视线中消失了：当阿西娜回到她平常的大小时，整个第二维度都不见了。

她想：“二维世界实在是太离奇了，我最好还是回家吧。”她的归途同样充满了历险，这个我们留到以后再讲。

即便不知道3个空间维度为什么特别，我们仍可以问“它们哪里特别”。如果说宇宙最根本的内在时空包含更多维度，它怎么可能看起来只有3个维度？如果阿西娜在一个二维世界里，为什么有时她只看到1个维度？如果弦理论正确描述了自然，空间中确实存在9个维度（另加上时间维度），那么另外6个失踪的维度化作了什么？为什么我们看不到它们？它们会对我们的可见世界产生明显的影响吗？

最后的3个问题是本书的核心。然而，当务之急是，我们要确定是否有一种方法可以让额外维度隐藏起来，以至于让阿西娜的二维世界看起来只有一维，或者有着额外维度的宇宙会呈现出我们看到的三维空间结构。如果我们愿意接受世界还有额外维度的观点，那么无论它出自何种理论，对于为什么找不到它们存在的任何迹象，一定存在一个合理的解释。

本章我要讲到极端卷曲的维度，它们不像我们熟悉的三个维度那样无限延伸；相反，它们会很快将自己绕起来，像一个紧紧缠绕的线圈一样。沿着一个卷曲维度，任何两个物体都不会相隔太远，任何想远距离旅行的尝试最终都会成为一圈接一圈的环游，就像渡渡鸟跑圈儿一样。这些卷曲维度可能非常微小，我们根本注意不到它们的存在。事实上，我们明白，如果微小的卷曲维度真的存在，要探测它们还真是个难题。

为什么世界看起来只有三维

弦理论，最有希望将量子力学与引力统一起来的理论，给出了思考额外维度的具体理由。我们所了解的唯一和谐形式的弦理论负载着这些令人诧异的附属物。然而，虽然弦理论在物理学中的出现使额外维度备受瞩目，但额外维度观点的提出却由来已久。

早在20世纪早期，爱因斯坦的相对论就让人们想到可能存在额外维度。相对论描述了引力，却没有解释为什么我们感受到的引力会是如此。爱因斯坦的理论对空间维度的数量没有任何倾向，无论是三维、四维或十维都同样有效。那么，我们的世界为什么看起来只有三维呢？

1919年，紧随爱因斯坦广义相对论（完成于1915年）之后，波兰数学家西奥多·卡鲁扎在爱因斯坦的理论里看到了这种可能，并大胆地提出了第四维度——一个全新的看不见的空间维度^[6]。他提出额外维度可能与我们熟悉的三个维度不同，但他并没能明确指出具体的区别。卡鲁扎提出额外维度的目的是要将引力和电磁力统一起来，尽管这个失败的细节与我们不相干，但他大胆提出的额外维度却至关重要。

卡鲁扎的论文写于1919年，当时，一家学术期刊邀请爱因斯坦审稿，请他决定是否发表这篇文章。对于这一观点的价值，爱因斯坦犹豫不决，难下定论。直到两年之后，他才同意将论文发表，并最终承认了它的原创性。但爱因斯坦仍想知道这一维度是什么：它在哪里？它为什么不同？它会延伸多远？

这些问题是显然要问的，可能困扰你的也正是这些问题。但没有人回答爱因斯坦，直到1926年，瑞典数学家奥斯卡·克莱因才解答了这一问题。克莱因提出，额外维度会卷曲成一个圆圈，且极其微小，只有 10^{-33} 厘米，即十亿亿亿亿分之一厘米。这个极其微小的卷曲维度无处不在：空间的每一个点上都会有它自己的小圈圈，只有 10^{-33} 厘米。

这个微小的量代表普朗克长度，我们后面详细讲述引力时，这是一个非常重要的量。克莱因选择普朗克长度，是因为它是唯一自然出现在量子引力理论里的长度，而引力与空间形状又是密切相关的。现在你只需知道普朗克长度非常之小，小到不可测量——比我们可能探测到的任何东西都要小得多。它是如此之小，原子比它大 10^{24} 倍，质子比它大 10^{19} 倍，如此小的东西是很容易被忽略的。

日常生活中有很多物体在我们熟悉的三维中都是有一维小到不可察觉，墙上的涂料，或是从远处看的晾衣绳，都像是少于三维的东西：我们忽略了涂料的厚度及晾衣绳的粗细。在粗心的观察者看来，涂料好像只有二维，晾衣绳好像只有一维，尽管我们知道实际上这两者都有三维。要看清这类东西的三个维度只有靠近了仔细观察，或是借助足够精确的工具。如果我们扯着一根橡胶管穿过足球场，从直升机上往下看（见图2-1），橡胶管好像只有一维，但如果近距离观察，你就会发现橡胶管的二维表面及其包裹的三维容积。

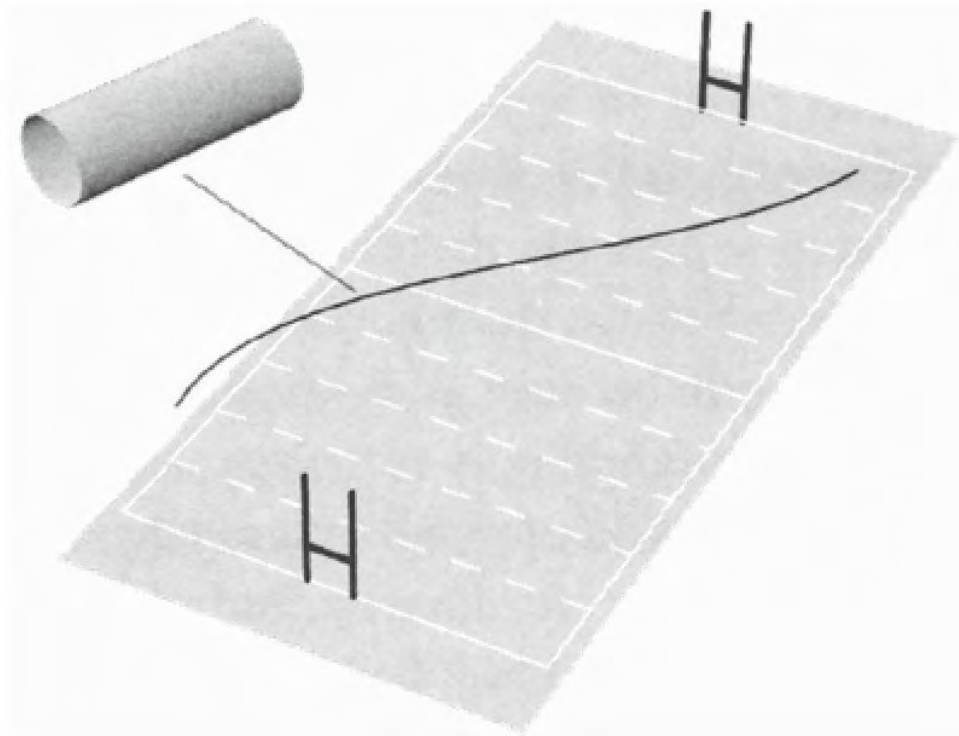


图2-1 在空中看到的穿过足球场的橡胶管。从空中看这根橡胶管，像是只有一维。但是走近了细看就会发现它的表面有二维，而其形成的容积有三维。



对克莱因来说，小到不可察觉的东西不是物体的粗细或厚度，而是整个维度本身。那么，小维度究竟代表什么意思呢？一个有着卷曲维度的宇宙对于居住在里面的人来说，看上去是什么样？同样地，这个问题的答案完全取决于卷曲维度的大小。我们以一些小生物为例，来看居住在这样一个世界里会是什么情形，这些生物可能比卷曲的额外维度要大，也可能比其小。因为要画出4个或是更多的维度实在是不可能的，因此，我们首先来看一个只有二维

的宇宙，其中包含一个卷曲的维度——紧紧卷起呈很小的形状（见图2-2）。



图2-2 一个维度卷曲时的二维宇宙。当一个维度卷曲时，一个二维的宇宙看起来就只有一维了。

再想想花园的橡胶管，我们可以把它看成由一长条橡胶片卷成的管子，它有一个圆形的横截面。这时，我们把这根橡胶管想象成整个宇宙（而不是宇宙里的一个物体）。如果宇宙的形状像是这根橡胶管，那么我们就会有长长的一维和卷曲且很小的另一维——这正是我们想要的。

对居住在花园橡胶管宇宙里的小生物，如一只小虫来说，宇宙看上去是二维的（在这一情景中，我们的小虫只能紧贴在橡胶管的表面上——二维宇宙是不包括内腔的，不然就成三维了）。小虫可以爬向两个方向：沿着橡胶管的长度向前或是绕着它转圈，就像只能在二维宇宙里转圈的那只渡渡鸟一样，绕着橡胶管爬行的小虫自一点开始，最终必然会爬到它开始的地方。因为

第二维度非常小，小虫爬不了多远就会回到原地。

如果居住在橡胶管上的一群小虫经受到外力，如电力或引力，这些力会将虫子吸引或推向橡胶管表面的任何方向。虫群可能会被打散——沿着橡胶管延伸的方向或者沿着它的圆周运动，而且会感受到出现在橡胶管上的任何作用力。只要有足够的能力看清如橡胶管直径般微小的距离，那么，其上的作用力和物体就会显示出橡胶管实际拥有的两个维度。

但是，如果小虫能观察其周围环境，它就能注意到这两个维度是非常不同的：沿着橡胶管长度的那一维非常大，甚至是无穷大；而另一维则非常小。在橡胶管圆周这一维度上，两只虫相隔永远不会太远，而想沿着这一方向长途旅行的虫子总会回到起点。一只喜欢动脑筋、有思想的虫子，会明白它的宇宙是二维的：一维延伸到很远；另一维则很小，并卷成一个圆圈。

但是，小虫的视角与我们假设在克莱因宇宙里的视角并不一样：在这个宇宙中，额外维度卷曲成极端微小的尺度，只有 10^{-33} 厘米。况且，我们还不如

小虫那么小，所以根本无法探知如此微小的维度，更别提在其中旅行了。

为完成我们的比喻，我们再来假想一个比虫子大一点儿的生物，居住在这样一个花园橡胶管宇宙里。它的感知能力比较弱，因此不能探知小的物体和结构。它观察这个世界的眼睛会忽视细小如橡胶管直径那样的细节，即便在对其有利的视角，这个大生物对另外一维仍是视而不见，它只能看到一维。如果某个生物的视力足够敏锐，能够探知如橡胶管粗细般微小的东西，那么它就能看到花园橡胶管宇宙有不止一个维度；如果它的视力不足以感知橡胶管的粗细，那么它所能看到的就只有一条线。

再者，物理作用也不会泄露额外维度的存在。花园橡胶管宇宙里的大生物会占满整个微小的第二维，因此小的生物永远也感觉不到还有这样一个维度。如果没有能力探查在额外维度上的结构和变化，如物质或能量的摇摆或波动，那么它们就永远无法感知额外维度的存在。第二维度上的一切变化都会被冲掉，这就像发生在原子结构尺度上的纸的厚度变化，那是你无论如何都注意不到的。

阿西娜梦到的二维世界，与这个花园橡胶管宇宙非常相似。因为她既有机会变得与二维世界的宽度一样大，也有机会变得比它小，她既能够从大于第二维度又能够从小于第二维度的视角观察同一宇宙。对于大的阿西娜来说，二维世界与一维世界好像都是一样的；只有变小了，阿西娜才能分辨其中的差别。同样的道理，如果花园橡胶管宇宙的另一空间维度小到看不见，其中的生物就不会知道它的存在。

现在，我们再回到卡鲁扎-克莱因宇宙，这里有我们所了解的三维，还有看不见的另一维，我们可以再次用图2-2来想象这种情形。理想的话，我应该画出这4个空间维度，但遗憾的是，这真的超出了我的

能力（即使将书翘起都做不到）。但是，因为构成我们空间的这3个无限维度本质都是一样的，我只需再画出一维来代表就行，这让我可以自由地使用其他维度来代表不可见的额外维度。这里显示的另一维度是卷曲的，与另外三维截然不同。

正如我们的二维花园橡胶管宇宙一样，四维的卡鲁扎-克莱因宇宙也有一个微小卷曲的维度，使其看上去比实际要少一维。我们无法了解额外的空间维度，除非我们能够找到其微观结构的证据，否则卡鲁扎-克莱因宇宙看上去就只有三维，如果这一维足够小，那么卷曲的额外维度将永远无从探知。后面，我们将探索它究竟有多小，但现在我们确信普朗克长度远远超出了我们的探测能力。

在生活和物理学中，我们只关注那些对实际产生作用的细节。如果你不能观察其细微结构，就可以假装看不到它。物理学中，这种忽略无关细节的做法，在前一章的有效理论里得到了最好的体现。在这一理论中，只有你能实际观察到的事情才是重要的。上面的例子，我们用的是三维有效理论，这样，额外维度的信息就被忽略掉了。

尽管卡鲁扎-克莱因宇宙的卷曲维度就在我们身边，但因为它实在太小了，以至于其中的任何变化都是不可感知的。正如纽约人之间自认为的差别对外人来说无关紧要一样，如果其细节变化只体现在微观尺度上，那么宇宙中额外维度的结构也是无关紧要的。就算从根本上讲确实存在许多额外维度，远比我们日常生活中认可的要多，但我们看到的東西仍然可以用我们观察到的维度来描述。极其微小的额外维度不会改变我们对世界的看法，甚至不会影响大多数物理计算。即使存在额外维度，如果我们不能看到或感受到它们，就可以将其忽略，这并不影响我们正确描述所看到的景象。可事情并不总是这样，以后我们会对这一简单景象进行修正，但这会与更多的假设相关。

图2-3显示的是一根橡胶管或其中一维卷成圆圈的宇宙，我们可以从中了解卷曲维度的另一个重要特点。我们集中来看无限维度上的任何一点可以发现，任一点上都有一个完整的卷曲的空间，即圆圈，橡胶管就是由这无数圆圈粘连在一起形成的，就如我们在第1章中讲到的切片一样。

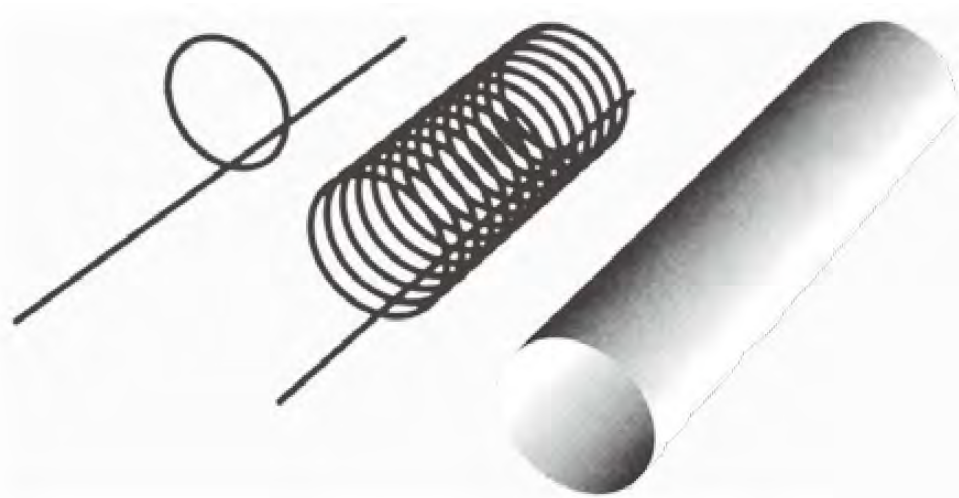


图2-3 一维卷成圆圈的两维宇宙。在两维宇宙里，如果一个维度卷起，那么，在其无限延伸的那一个空间维度上，每一个点上都有一个圆圈。

图2-4给出了不同的例子：有两个（而不是一个）无限延伸的维度，再加上卷成圆圈的另外一个维度。这种情况下，二维空间里的任一点上都有一个圆圈。如果有三个无限维度，那么三维空间里每一点上都会有卷曲维度。你可以把额外维度空间的每一点看作你身体里的细胞，因为每个细胞都携带了完整的DNA序列，同样，三维空间里的每一点上都寄居着一个完整的卷曲的圆圈。

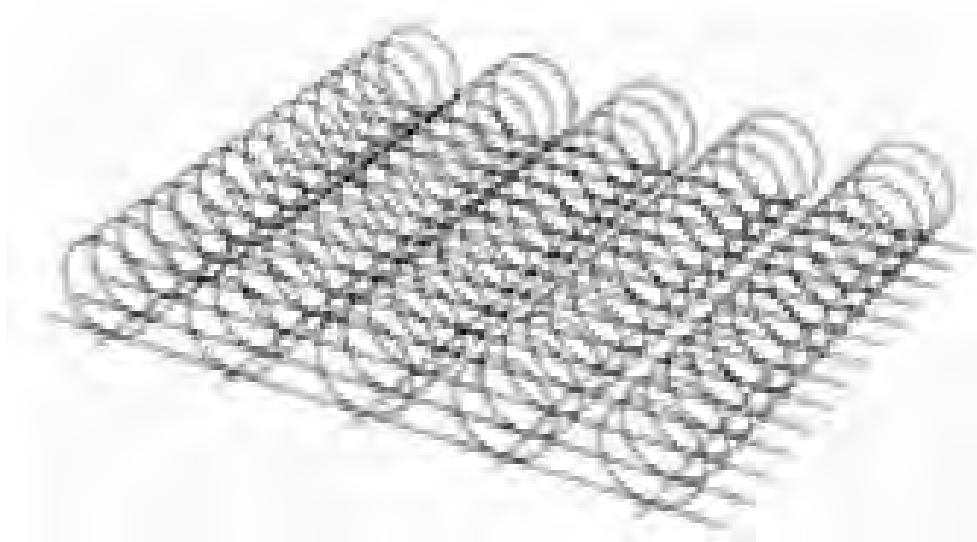


图2-4 一维卷曲的三维宇宙。在一个三维宇宙里，如果其中一维卷起，那么在这一平面的每个点上，都会有一个圆圈。

到目前为止，我们只探讨了一个额外维度，也就是卷成圆圈的维度。但即使卷曲维度表现为其他任何形状，我们所说的一切仍可以成立。我们选择一个环形，如炸面圈的形状，它的额外两维同时卷曲成圆圈（见图2-5），如果两个圆圈——绕着一个洞的圆和形成面圈身体的那个圆都足够小，那么卷曲的这两个维度都将永远不能被发现。



图2-5 两维同时卷曲的四维宇宙。如果四维中有两维都卷成炸面圈形状，那么空间的任一点上都有一个面圈。

但这只是一个例子。如果维度更多，会有大量可能的卷曲空间——有着卷曲维度的空间，因维度的具体卷曲方式不同而各不相同。有一类卷曲空间对弦理论至关重要，这就是卡拉比-丘流形。它得名于意大利数学家尤金尼奥·卡拉比（Eugenio Calabi），他首次提出这种特殊形状，而华裔哈佛数学家丘成桐证实了其在数学上的可能性。这些几何形状以一种非常特殊的方式将额外维度卷曲缠绕在一起，与所有的卷曲空间一样，这些维度卷曲得很小，却以一种更为复杂、更难绘制的形式缠绕。

无论卷曲维度采取何种形状，也无论它们有多少，在其无限维度的每个点上，都有一个包含所有卷曲维度的极小的卷曲空间。因此，如果弦理论学家是正确的，那么空间里的每个地方——你的鼻尖上、金星的北极点上、网球场上空你上次击球球杆划过的每个点上，都会有一个小到不可见的六维卡拉比-丘空间，多维几何在空间的每个点上无处不在。

与克莱因一样，弦理论学家也常常指出卷曲维度只有普朗克长度的大小，即 10^{-33} 厘米。这种尺度的卷曲空间会隐藏得很好，我们几乎肯定没有方法能探测到如此微小的东西。那么，普朗克长度的额外维度也就很可能不会留下任何痕迹。因此，即便我们生活在一个有着普朗克长度额外维度的宇宙里，我们依然只能认知熟悉的三个维度。宇宙中可能有很多这种微小的维度，但我们也许永远没有足够的能力去发现它们。

卡拉比-丘流形

它得名于意大利数学家尤金尼奥·卡拉比，他首次提出这种特殊形状，而华裔哈佛数学家丘成桐证实了其在数学上的可能性。这些几何形状以一种

非常特殊的方式将额外维度卷曲缠绕在一起，与所有的卷曲空间一样，这些维度卷曲得很小，却以一种更为复杂、更难画出的形式缠绕。

有额外维度的牛顿引力定律

图文并茂地解释为什么当额外维度卷曲至很小时会隐藏起来固然很好，但是检验物理定律符合这一直觉仍然至关重要。

牛顿于17世纪提出的万有引力定律告诉我们：引力的大小取决于两个有质量物体^[7]之间的距离。这就是我们熟知的平方反比定律，也就是说，引力强度会随着距离的增大而逐渐减弱，与距离的平方成反比。例如，如果你将两个物体之间的距离增加两倍，引力强度就会削弱到原来的1/4；如果相隔距离是原来的3倍，则引力会削弱到原来的1/9。万有引力平方反比定律是最为古老、最为重要的物理定律之一，正因如此，行星才会有其自己特定的椭圆轨道。任何可行的物理理论都必须遵循平方反比定律，否则必败无疑。

万有引力定律对距离的依赖，体现在牛顿的平方反比定律中，它与空间维数有着密切的关联。这是因为，维数决定了引力在空间里的发散速度。



我们来细想一下这种关联，这在我们以后探讨额外维度时至关重要。我们来设想一个供水系统，其中的水既可以直接流入一根橡胶管，也可以流向一个洒水装置。假设流过橡胶管和洒水装置的水量相同，而且都能浇灌花园里一定数量的花朵（见图2-6）。当水流过橡胶管时，橡胶管直接对准花朵，那么这株花就会得到所有的水，这与橡胶管根部到其对准花朵的龙头的距离是无关紧

要的，因为所有的水最终都会浇到花上，无论橡胶管有多长。

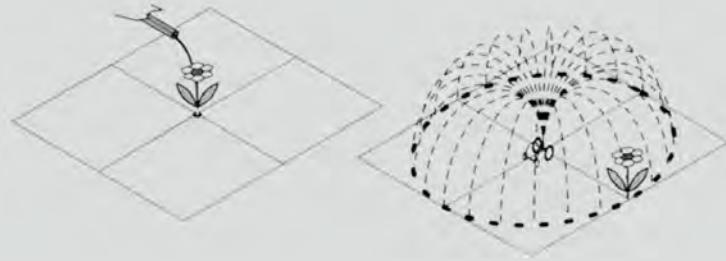


图2-6 两种不同的浇花方式。由一个把水洒向四周的洒水装置浇到花上的水量要少于由橡胶管直接浇到花上的水量。

但是，假设等量的水流进一个洒水装置，则可以同时浇灌许多花。就是说，洒水装置将水向四周喷洒，洒遍一定距离之内的花。现在水被喷洒至这个距离内的所有植物，那么原来的那株花就不能再得到所有水了。而且，花离水源越远，洒水装置需浇灌的植物就越多，水被喷洒的范围就越广（见图2-7）。这是因为，在3米的圆圈里比在1米的圆圈里可以种更多的花儿，水喷洒得越广，花儿离得越远，得到的水就越少。

同样的道理，任何一个在不止一个方向上被平均分配的东西，对于任何特定物体，无论是一株花，还是我们会看到的一个经受引力的物体，离得越远，

产生的影响就越小。引力与水一样，离得越远，分布得就越广。

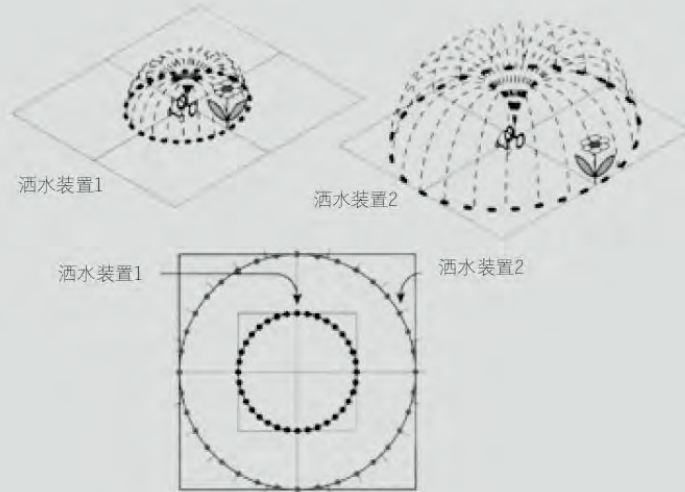


图2-7 用洒水装置浇花。当洒水装置将水洒向一个半径更大的圆圈时，水喷洒的空间更大，而花儿得到的水就更少。

从这个例子我们还可以看出，为什么维度的数量对水（或引力）的分布会产生如此强烈的影响：由一个二维洒水装置喷出的水，会随着距离的增大而喷洒得更广；而从单维橡胶管里流出来的水根本就没有分散。现在，我们再设想一个洒水装置将水以球形喷洒（这个洒水装置就像是蒲公英长出种子的绒球），而不仅仅是向四周喷洒。那么，随着距离的增大，水会分散得更快。

现在，我们将这一推理应用到引力，并得出在三维空间中引力与距离的精确关系。牛顿万有引力定律的成功需基于两个事实：引力在

各个方向上的作用都是相同的；空间有三个维度。现在我们设想一颗行星，它能吸引其影响范围内的所有物质，因为引力在各个方向上的作用都相同，行星作用于另一物体（如一颗卫星）的引力强度就取决于两者之间的距离，而非方向。

为形象地表示引力的强度，图2-8中左图显示了从一个行星核心延伸出的引力射线，就像是从洒水装置喷洒出的水。这些射线的密度决定了行星作用于邻近任何物质的引力强度：穿过一个物体的引力线越多，引力越大；而引力线越少，则引力越小。

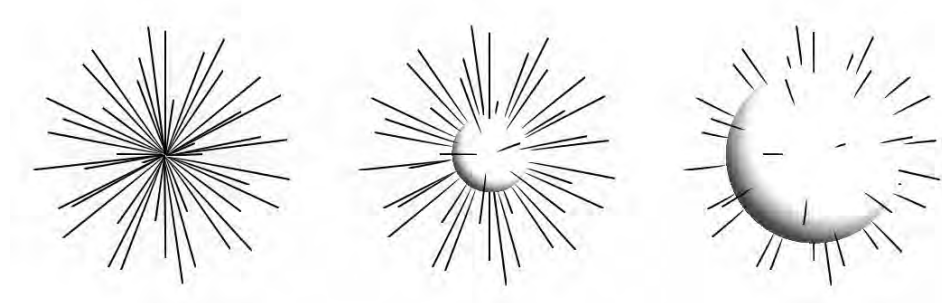


图2-8 一个大质量物体（如行星）发散出的引力线。穿过一个球面的引力线的数量是相同的，无论其半径大小。因此，离这个物体的中心越远，引力线就越分散，而引力也就越弱。

注意，穿过球面的引力线数量是相同的，而不论距离远近（见图2-8中间图和右图所示），但因为引力线分布于球面的每个点，距离越远，引力必然就越弱。精确的分散因子取决于定量度量一定距离上的引力线分布有多广。

穿过一个球面的引力线数量是一定的，无论它离其核心物质有多远。球体的表面积与其半径的平方成正比：表面积等于一个数乘以其半径的平方。分散于球面的引力线数量一定，因此引力必然随着半径的平方减弱，这种引力场的分散就是万有引力平方反比的来源。

牛顿定律与卷曲维度

现在我们知道了引力在三维空间里遵守平方反比定律，但要注意的是，这个论点似乎非常依赖一个既定事实：我们的空间有三个维度。假设只有两个维度，引力就只会以一个圆圈向外发散，那样，引力随距离减弱的速度就会慢得多。假设有不止三个维度，比如一个超球体，那么随着行星与其卫星之间距离的增大，其表面积增大的速度会更快，而引力也会因此迅速减弱。似乎只有三个维度才会产生这种与距离平方成反比的依赖关系，但如果真是这样，那么研究额外维度的物理学家们为什么也认可牛顿的引力平方反比定律呢？

了解卷曲维度如何解决这一潜在矛盾，是很有趣的。其基本逻辑是：引力线不能任意深入卷曲维度，因为卷曲维度的大小是有限的。尽管引力线最初是向所有维度发散的，但当其发散范围超出额外维度的大小时，它们就别无选择，只能沿着那些无限维度的方向延伸。

这仍可以用橡胶管的例子来说明。假设橡胶管的一端有一个盖子，水将通过盖子上的小孔进入水管（见图2-9）。正常情况是，流过小孔的水并非当即沿水管直流而下，而是先要充满水管的整个横截面。但是很显然，如果你正拿着管子的另一端在花园里浇水，那么水是如何进入水管的就根本无所谓了。虽然水刚进入水管时会向不止一个方向喷射，但很快就会撞到水管内壁，然后再流出。这时，水流看上去就只有一个方向。从根本上来讲，在微小的卷曲维度上的引力线就是如此。

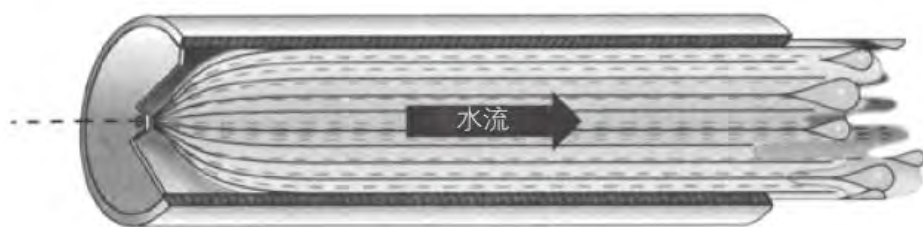


图2-9 流入橡胶管的水流。穿过小孔进入水管的水首先会向三个维度发散，然后再沿着水管水平方向的一维流动。

如前所述，我们仍可以想象一定数量的引力线由一个庞大的球体向外发散，在小于额外维度大小的距离内，引力线会均等地向所有方向发散，如果你能测量那个小尺度上的引力，那么就能测得高维里的引力。引力线的发散方式正如水穿过小孔进入橡胶管一样，会首先充满整个橡胶管的内部。

但是，在大于额外维度大小的距离上，引力线只能向着无限维度的方向延伸（见图2-10）；在微小的卷曲维度里，引力线在触碰到空间边缘之后，便不能再继续延伸，只能弯曲，而它们剩下的唯一出路就只有沿着仅有的一个方向延伸。因此，在大于额外维度大小的距离上，额外维度就仿佛根本不存在一般，而引力定律便会转而向牛顿的平方反比定律靠拢——就是我们现在看到的样子。这意味着，如果你只测量距离大于额外维度大小的两个物体之间的引力，那么即使是从数量的观点来看，你仍无法得知额外维度的存在；只有在卷曲空间的狭小区域内，引力与距离的依赖关系才能反映额外维度的存在。

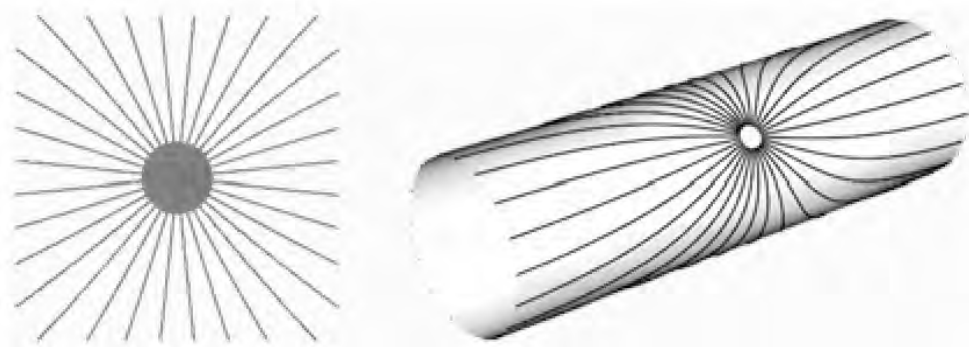


图2-10 当一个维度卷曲时，由一个庞大物体发散出的引力线。在短距离内，引力线会以辐射方式向外发散；在远距离上，它们只会沿着无限维度延伸。

额外维度必须小到不可见吗

现在我们可以肯定的是，如果额外维度足够小，它们就是不可见的，而且不会在我们能观察的距离尺度上产生可探测的影响。长期以来，弦理论学家们一直假设，额外维度的大小就是普朗克长度，但现在有人对此提出了质疑。

没有谁足够了解弦理论，能确定地说出额外维度究竟有多大。类似普朗克长度的大小是有可能的，任何小至不可探测的维度都有其道理，但普朗克长度实在太小了，即使远比它大的卷曲维度，我们也仍旧不可能察觉。由于至今我们还没能亲眼见到它们，因此，额外维度研究的一个重要问题就是：这些维度究竟有多大？

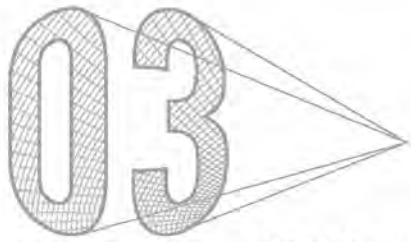
在本书中，我们将探讨以下问题：额外维度究竟有多大，这些维度对基本粒子是否会产生可辨别的影响，以及实验是怎样进行的。我们会发现，额外维度的存在会大大改变我们研究粒子物理所遵循的规则，而且其中一些改变会产生可观察的实验结果。

我们还将探讨一个更为激进的问题：额外维度是否必定是极小的？我们确实看不到微小的维度，但难道维度必须小到不可见吗？会不会有一个维度是无限延伸的，而我们却没有发现？如果真是这样，那么额外维度该与我们所见的维度截然不同。到目前为止，我才只举出了一些最为简单的可能性。以后我们会看到，即使与我们所熟知的三个无限维度截然不同，为什么我们仍不能排除无穷大额外维度这种近乎极端的可能性？

在第3章中我们将探讨另外一个问题，也许你也曾想过：为什么微小的额外维度不是局限在两堵“墙”之间的线段，而是卷曲的球呢？现在还没有人想到这种可能——但为什么不呢？原因在于，如果假设空间有终点，那就要知道，在那个终点发生了什么？会像过去图画里扁平地球所暗示的那样，事物到了宇宙末端会掉下去吗？或者它们会被反射回来，又或者它们根本到不了那里？要明确在终点究竟发

生了什么，我们首先要了解科学家所说的边界条件。如果空间有终点，那么终点在哪里？它又是怎样终结的？

膜——高维空间里的薄膜状物质给我们有“终点”的世界提供了必要的边界条件。正如我们将在第3章中看到的，膜能生成一个（或多个）不同的世界。



我们生活在一个膜宇宙上吗

WARPED PASSAGES

我会像胶一样地黏住你。黏住你，因为我已被你俘虏。

埃尔维斯·普雷斯利（猫王）

总是被困在路上的艾克

与好学的阿西娜不同，艾克不爱看书，他只喜欢玩游戏、搞机械和玩车。但是，艾克却很讨厌在波士顿开车，因为那里的司机驾车鲁莽，路标也不清楚，马路总在没完没了地施工，艾克总是被困在路上。尤其令人懊恼的是，前面也根本没什么车。虽然这样的路况非常诱人，可艾克一点儿办法都没有，毕竟他不是阿西娜养的猫头鹰，不能插翅飞过去。对被困在波士顿公路上的艾克来说，第三维毫无用处。

许久以来，几乎没有几个“自爱”的物理学家会认为额外维度值得思考。它纯属推测，又是那么怪异，没人能对它们说出个所以然来。但是最近这些年，额外维度的命运出现了转机：它不再是不受欢迎的不速之客，而成了人们追捧的、激发灵感的良友。这鹊起的声名要归功于“膜”，以及由这些引人入胜的结构所带来的众多可能的崭新理论。

1995年，当来自加州大学圣塔芭芭拉分校卡夫里理论物理研究所（KITP）的物理学家乔·波尔钦斯基（Joe Polchinski）确认膜是弦理论的基础时，整个物理界为之一震。但是，早在此之前，就有物理学家提出过薄膜状物质，例如p-膜（一些调皮的物理学家是这么叫它的），是一种只在某些维度无限延伸的物质，它是物理学家通过爱因斯坦的广义相对论推导出来的。

粒子物理学家也曾提出过一些把粒子束缚于膜表面的机制。但是，弦理论里的膜是第一种已知既能束缚粒子又能束缚力的膜，我们很快就会看到，正是这一部分使它们吸引了所有人的兴趣。就像在三维空间里被困在二维公路上的艾克一样，粒子和力也会被困在一些叫作膜的低维表面上，即便宇宙仍有许多等待探索的其他维度。如果弦理论正确描述了我们生存的世界，那么物理学家将别无选择，只能承认这种膜的确是可能存在的。

膜宇宙是一幅令人兴奋的新景象，它彻底改变了我们十几年来对引力、粒子物理学和宇宙学的理解。膜很可能在宇宙中存在，而且我们没有理由否认我们或许就生活在一个膜上。膜甚至有可能在决定我们宇宙的物理属性中发挥着重要作用，并能最终解释可见的现象。如果真是这样，那么膜以及额外维度必将永驻于此。

像切片一样薄

在第1章里，我们已看过一种思考二维世界（平面国）的办法：把它当成一个三维空间里的二维切片。在阿博特的小说里，正方形A走出二维平面国，进入三维世界游历了一番，然后认识到平面国只不过是一个更大的三维世界的切片而已。

回到平面国，正方形A提出他所见到的三维世界很可能也只是一个切片（这很符合逻辑）：一个更高维度空间的三维切片。当然，这里

的“切片”并不只是一个如纸片般的二维薄膜，而是这类事物的逻辑延伸，如果你喜欢，可以称其为一个假想的薄膜。你可以把正方形A猜想的三维切片当成四维空间里的一个三维切块。

但是，正方形A的三维“导游”很快就打消了他对三维切片的臆想。就像我们认识的绝大多数人一样，这个缺乏想象力的三维居民只相信他能看见的三维空间，根本不会去考虑第四维度。

膜

膜将一个数学概念引入到物理学中，它是一个特别的时空区间，只能在空间（可能是多维）的一个切片里延伸。无论哪种膜，它都是一个区域，维度少于包围它或与它相邻的整个高维空间。

膜将一个数学概念引入物理学中，它与一个世纪前《平面国》里的描述极为相似。现在，物理学家开始回到这个观点：包围我们的三维世界可能就是高维世界里的一个三维切片。膜是一个特别的时空区间，只能在空间（可能是多维）的一个切片里延伸。“薄膜”一词让我们选择了“膜”，这是因为，像膜一样，薄膜是裹在物质外面的一层，或者是穿过物质的一个夹层。有些膜是夹在空间里面的切片，而另一些则是裹住空间的切片，就像三明治的面包片一样。

无论哪种膜，它都是一个区域，比包围它或与它相邻的整个高维空间的维度少。注意，薄膜有两维，而膜的维度可能是任何数量。尽管最有可能引起我们兴趣的膜有3个空间维度，可是“膜”这个名词指代的却是所有这种类型的切片。有些膜有3个空间维度，而另外一些膜却可能有更多（或更少）的维度。我们会用3-膜来指有3个维度的膜，而4-膜指有4个维度的膜，依此类推。

边界膜和内嵌膜

在第2章里，我们解释了额外维度为什么不可见的原因：它们卷曲的形状可能太小，以至于未能显露出它们存在的任何证据。关键的一点在于，额外维度可能很小。但额外维度不可见的原因，没有一个是基于它们是卷曲的这一事实。

这就暗示了另外一种可能：也许维度不是卷曲的，而只是在一定距离内终结了。因为，如果维度消失得无影无踪，就会潜伏着危险，你一定不想某段宇宙会在其终点跌落下去，所以，这些有限的维度必定会有边界。边界让我们知道，它们在哪里终结以及是怎样终结的。那么，当粒子和能量到达这些边界时会发生什么？

答案是：它们会遇到一个膜。在高维世界里，膜有可能是被称作“体”的整个高维空间的边界。体与膜不同，它会向所有方向延伸。体会跨越每一维，既可能在膜上，也可能在膜外（见图3-1）。因此，体是庞大的，而相对来说，膜是扁平的（在某些维度上），就像一张煎饼。如果膜在某些方向上是体的边界，那么，体的某些维度可能与膜平行，而另一些维度则可能偏离它。如果膜是边界，那么脱离了膜的维度只能向一侧延伸。

为了了解终结于膜的有限维度的特征，我们可以设想一根长长的细管。管子里有三个维度：长的一维，短的二维。为了让与扁平膜的类比更直观，我们假设管子有方形的横截面，这样的话，无限长的管子里就有了4个无限延伸的方形横截面。如果管子自身就是宇宙，那么它是一个三维宇宙，其中有两维被各边的壁束缚，而另一维则延伸至无穷。

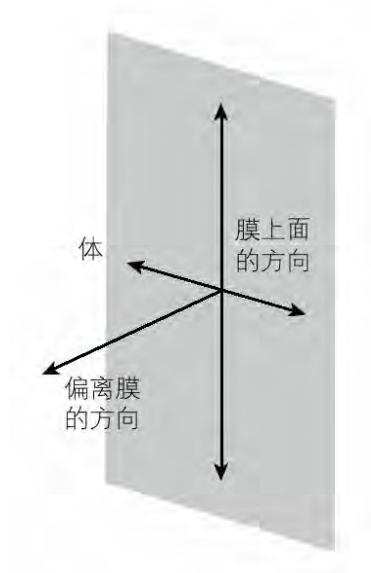


图3-1 膜与体。膜是一个低维表面，在膜上，既有方向顺其延伸，也有方向会离它远去，伸向更高维度的体。

我们知道，细长的管子如同第2章花园里的橡胶管一样，从远处看是一维的。但我们仍可以像以前一样问，这样一个方管宇宙（包括管壁及其内腔）对生活在其中的好奇生命来说，会是什么情形？

可能正如你猜想的那样，这要看那个生物的认知能力。一只可以在方管里自由行动的小飞虫，会感知到它是三维的。与二维花园橡胶管的例子不同的是，我们想象这只小飞虫可以在方管里自由飞翔，而不是只能紧贴在它的表面上。然而，如在花园橡胶管里一样，小飞虫会感知到有一条长长的维度，与另外两维完全不同。在这个方向上，小虫可以自由地飞翔，直至很远（假设我们的管子很长，甚至是无限长）；而在另外的两个方向上，小虫能飞越的距离很近，也就只是方管的宽度。

但是，除了各自所拥有的维数之外，花园橡胶管与方管宇宙还有一处差别：与之前的小虫不同，这个方管里的飞虫是在管子里面飞的。因此，飞虫有时会碰壁，它可以前后飞，也可以上下飞，但总是

会碰壁。而橡胶管上的小虫则不同，它永远不会遇到这样的边界；相反，它只会一圈接一圈地爬下去。

当飞虫撞到方管宇宙里的边界时，必须有一定的规范来约束它的行为。管子的壁就决定了它的行为：飞虫可能会撞到壁上而死去；也可能因管子是有弹性的而被反弹回来。如果管子真的是一个由膜作边界的宇宙，那么，这些二维的膜就能够决定，当粒子或携带能量的其他任何物体撞到边界时会发生什么。

当物体撞到边界的膜时，它们会反弹回来，如同台球从案边弹回或光从镜子上反射回来一样。这就是科学家所说的反射边界条件。如果物体从一个膜上弹回，则能量没有损耗：它没有被膜吸收或泄漏出去，没有任何物体超出膜外。边界膜就是“世界终点”。

在一个多维宇宙里，膜就发挥了以上方管宇宙中边界壁的作用。与管壁一样，这样的膜也比整个空间有更少的维度（边界肯定比其包围的物体维度少），面包的边界如此，空间的边界也是如此。我们房子的外墙也是一样：它比房间少一维，房间是三维的，而任何一面墙（当我们忽略其厚度时）都只会跨越两个维度。

这一节到目前为止，我们集中探讨了位于边界的膜，但膜并不总是位于体的边缘，它们可以存在于空间的任何地方，尤其是，膜可能位于空间内远离边界的某个地方。如果将边界膜看作长面包一端的那层薄壳，那么非边界膜就像长面包中间的一个切片。非边界膜与我们已经探讨过的那些膜一样，仍是低维物体，在它的每一侧都有一个高维的体空间。

下一节我们会看到，无论体或膜有几维，也无论膜位于空间内还是其边界，膜都能俘获在它上面的粒子和力，这使得它所占据的空间区域变得非常特别。

被膜俘获

尽管有许多空间是你能够到达的，但你却未必会探索所有的空间。也许还有一些地方是你一直渴望要去，却永远也不能成行的，比如外太空或是大海深处。虽然你从未去过这些地方，但是从理论上讲，你可以去。没有哪个物理定律会限制这种可能。

但是，假如你居住在一个黑洞里，那么你的旅行就将会受到严格的限制，甚至比沙特阿拉伯的女性受到的限制还要多。黑洞（直至其完全消亡）会把你困在里面，使你永远无法逃脱。



在我们熟悉的例子里，行动受限的东西有很多，对它们来说，有些空间区域是永远无法到达的。比如电线里的电荷，再比如算盘里的珠子都是位于三维世界的物质，却只能沿着一个维度运动。还有一些常见的事物也是被限制在二维表面上的：浴帘上的水珠，只能沿着浴帘的二维表面滑下（见图3-2）；被困于显微镜下玻璃片上的细菌，也只能体验二维的运动。另一个例子是山姆·劳埃德（Sam Loyd）的“十五数字推盘”游戏，这个烦人的游戏是：一个盛着15块塑料薄片的小塑料盘，里面的每个薄片上都有一个字母或是图案。你需要在塑料盘里水平移动它们，直到它们排列成正确的一句话或一幅图（见图3-3）。除非你作弊，否则那些字母或图

案不得离开塑料盘，不能在第三维度里移动。

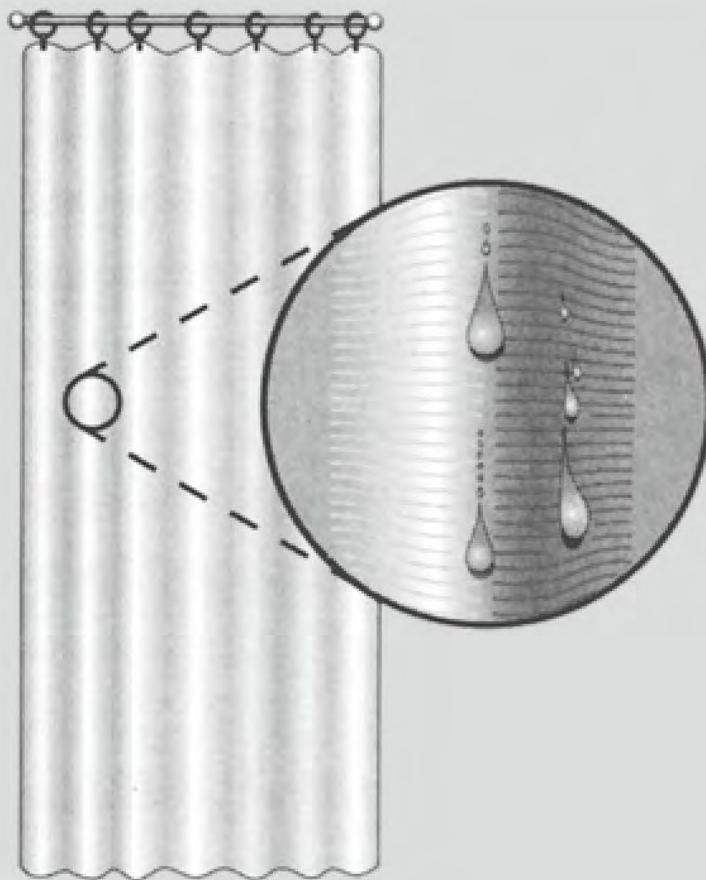


图3-2 浴帘上的水珠。在一个三维的房间里，水珠被挡在二维的浴帘上。



图3-3 山姆·劳埃德的“十五数字推盘”游戏。

膜，就像浴帘或劳埃德的十五数字推盘游戏一样，会把事物限制在一个低维的表面上。它们引入了这样一种可能：在一个有着更多维度的世界里，并非所有的物质都能够自由地在任何地方旅行。恰如浴帘上的水珠被束缚于一个二维的平面上一样，粒子和弦也可能被限制在一个三维膜上，而这个膜就处在高维世界中间。但与浴帘上的水珠不同的是，它们是真正地被束缚了；与十五数字推盘游戏也不同，膜不是随意挑

选的，它们可是高维世界的“职业选手”了。

被困在膜上的粒子被物理定律真正地束缚在了这些膜上，这些物质永远也不能进入延伸出膜外的额外维度。然而，并非所有的粒子都会被困在膜上，有些粒子可以自由地穿行于体中。但是，将有膜的理论和无膜的多维理论区别开来的，正是膜上的粒子——那些不能在所有维度上运动的粒子。

从理论上讲，膜和体的维度可以是任何数量，只是膜的维数永远要少于体的维数。膜的维数就是被膜束缚的粒子获准在其上面自由移动的维度的数量，尽管会有多种可能，但最让我们感兴趣的膜还是那些三维膜。我们不知道三维为什么看起来这么特别，但是有着3个空间维度的膜是与我们的世界密切相关的，因为它们可以沿着我们了解的3个空间维度延伸。这样的膜会出现在一个体中，这个体中可以有任何数量的维度，4个、5个或是更多，只要是超过3个。

即便宇宙确实有很多维度，但如果我们所熟悉的粒子和力被束缚在一个向3个维度延伸的膜上，那么它们仍会像在只有三维的世界里一样于膜上运动。如果光也被困于膜上，那么光线也只能沿着膜发散，在一个三维的膜上，光的表现就恰如它在一个真正的三维宇宙里一样。

而且，被限制在膜上的力，也只会对被困在同一个膜上的粒子发挥作用。构成我们这个世界的基本物质，如原子核和电子，以及这些物质之间的相互作用力，如电子力，都可能被限制在一个三维的膜上。被膜束缚的力只能沿着膜的方向扩散，而被膜束缚的粒子将产生交换，但只能沿着膜的维度运动。

因此，如果生存在这样一个三维的膜上，你将能够沿着它的维度自由行动，几乎与在我们眼下的三维世界里一模一样。其他维度将与膜相邻，但被困在这个膜上的物质，将永远也不能渗透到更高维度的体中。

虽然力和物质都可能被困在膜上，但膜宇宙的奇妙之处就在于，我们知道并非所有的东西都被限制在一个膜上。例如，引力永远也不会被限制在膜上，根据广义相对论，引力被交织在时空结构上。这就意味着，引力必定能穿越所有维度。如果它被限制在某个膜上，那我们就只能放弃广义相对论了。

好在事情并不是这样的。即便有膜存在，万有引力还是处处都能被感受到，无论在膜上还是离开膜。这点很重要，这意味着膜宇宙必须与体相互作用，即便只是通过引力。因为引力延伸到体中，而且所有东西都是通过引力而相互作用的，所以，膜宇宙总与额外维度相联系。膜宇宙并非孤立存在：

它们是一个更大的整体的组成部分，并与之相互作用。除了引力之外，在体宇宙中，还有可能存在其他的粒子和力。如果这些粒子真的存在，那么它们也会与被困在膜上的粒子相互作用，并把被膜束缚的粒子与更高维度的体宇宙联系起来。

除了我曾提到过的性质之外，后面我们将简要讲到的弦理论的膜还有一些特别属性：它们能够携带特定的电荷；受到外力时，会以特殊方式响应。但后面讲到膜时，我很少具体谈论这些性质。只要记住我们在本章里谈到的性质就足够了：膜是低维表面，它们能束缚住粒子和力，而且可以是更高维空间的边界。

神秘的膜宇宙

因为膜能俘获多数粒子和力，所以，我们生存的宇宙想必就构筑在一个三维的膜上，并漂浮在一个更高维度里。引力会延伸至额外维度中，而恒星、行星、人类以及我们能感受到的所有其他东西，都可能被限制在一个三维的膜上。如此说来，我们就生存在一个膜上，膜就是我们的家园。膜宇宙的概念正是基于这种假定（见图3-4）。

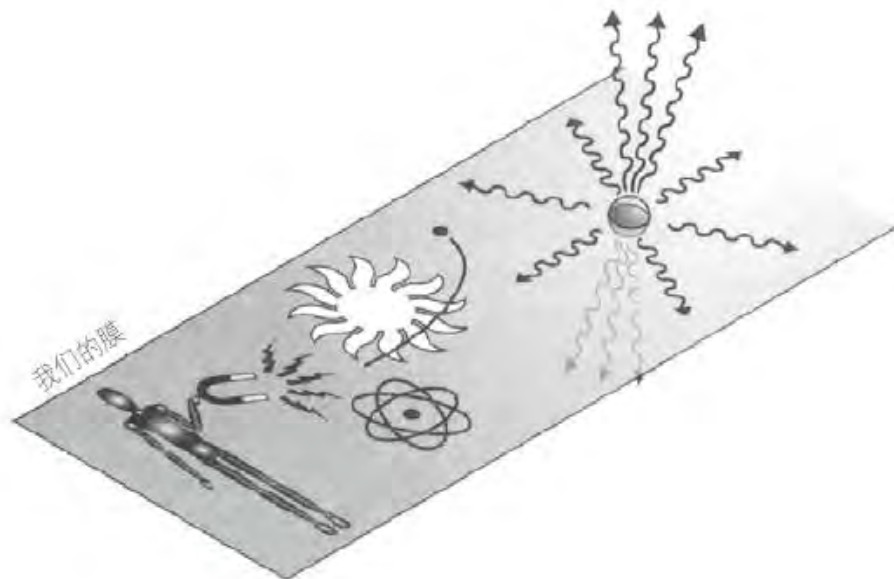


图3-4 我们生存的膜宇宙。我们可能生活在膜上，也就是说，组成我们的物质、光子以及其他标准模型粒子都可能是在膜上。但是引力是无处不在的，既在膜上，也在体中，就如图中扭动的曲线那样。

如果说，有一个膜悬浮在一个高维空间里，那么无可否认，就会有更多的膜也悬浮于其中。在膜宇宙的图景里，常常会涉及不止一个膜，我们还不清楚宇宙究竟会出现几个或是几种膜。“多重宇宙”这个名称与超过一个膜的理论相关联（见图3-5）。人们常常用这个词来描述这样一个宇宙：其组成部分要么没有相互作用，要么相互作用微弱。

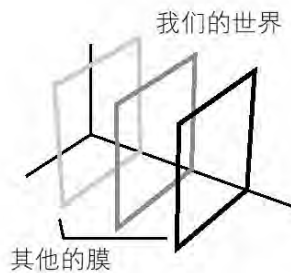


图3-5 多重宇宙。宇宙可以包含多重膜，它们之间或是只通过引力相互作用，或是根本互不相干，像这样的构成常被称作多重宇宙。

我觉得“多重宇宙”这个词有点奇怪：因为宇宙的定义本来就是许多部分联合在一起形成的一个整体。但是，也许会有一些不同的膜，它们相距遥远，无法联系；或是只能通过穿梭于它们之间的中介粒子产生微弱的联系。那样，在遥远膜上的粒子，感受到的力可能就会是完全不同的，而被膜束缚的粒子将永远不能与被束缚在另一张膜上的粒子产生直接的联系。因此，当不止一个膜所共同的力只有引力时，我有时就会把涵盖了两个膜的宇宙称作多重宇宙。

对于膜的思考让我们意识到：对于自己生存的这个空间，我们了解的实在是太少了。宇宙可能就是一个辉煌壮观的建筑，将所有位于其中的膜连接了起来。即使我们知道了它基本的组成部分，但在一个密布着多层膜的多重宇宙里，可想而知，总会有一些离奇的新景象，而且我们了解的、不了解的粒子在空间几何里的分布可能，也会有无数种。同样一副纸牌，会有多种不同的发牌法，这里存在着太多的可能。

其他的膜也许与我们的膜平行，也许包含平行的世界，但也许还会存在许多其他类型的膜宇宙。膜可以相互交叉，而粒子可能就被束缚在交叉线上。膜可以有不同的维数，它们可以弯曲，可以移动，也可以包裹未被发现的不可见维度。展开想象，你可以随心所欲地勾勒你的图景。宇宙中存在这样的几何图景也并非没有可能。

在一个有着内嵌膜的高维体宇宙里，有一些粒子可能会进入更高维度，也有一些粒子就只能待在束缚它的膜上。如果体将两个膜隔开，那么，有些粒子会在这个膜上，有些则会在另一个膜上，还有一些会夹在中间。有关粒子和力在不同膜与体之间是怎样分布的，理论上会有多种方式。尽管膜来自弦理论，但我们仍不明白为什么弦理论会独独钟情于某个特定的粒子和力的分布方式。膜宇宙给我们引入了一些全新的物理图景：

它们既描述了我们自以为了解的世界，也描述了在未知的其他膜上的其他未知世界。在一些看不见的维度上，它们与我们的世界遥遥相望。

在那些遥远的膜上，也许还存在着被束缚的新的力。我们不能直接与之相互作用的新粒子，兴许就在这些膜上传播。能够解释暗物质和暗能量（我们据其引力效应推测的一种身份不明的物质和能量）的其他物质，兴许就分布在不同的膜上，或既在体里，又在其他膜上。而当你从一个膜走到另一个膜时，引力甚至也可能对粒子产生不同的影响。

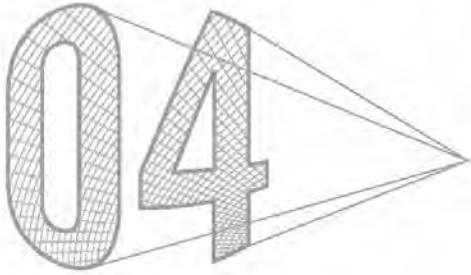
倘若其他膜上有生命，那么，由于他们被拘禁在完全不同的环境中，有着不同的感官，很可能体验着完全不同的力。我们的感觉会感知周围的化学反应、光和声音；而其他膜上的生物，倘使他们存在，也不大可能会与我们膜上的生命有太多相似的感觉，因为基本的力与粒子可能是不一样的。或者，其他的膜与我们的膜根本毫无相似之处，而我们能确定的唯一共同分享的力就是引力，而即使是引力，其作用也可能会发生变化。

膜宇宙的作用取决于膜的数量和类型，以及它们的位置。这对好奇心大的人将是个打击：被限制在遥远的膜上的粒子和力，对我们未必会产生强烈的影响。它们可能只决定在体中运动的东西，发出微弱的、可能永远不能到达我们的信号。因此，许多可以想象的膜宇宙即便真的存在也仍将难以探知。毕竟，引力是我们唯一明确知道的，由

我们膜上的物质与所有其他膜上的物质共同分享的相互作用，而引力又偏偏是一种极其微弱的力。如果没有直接证据，那么其他的膜就只能隐匿在理论和猜测的范畴里。

但是，我要讲的某些膜宇宙是可以产生可探测的信号的。这些可探知的膜宇宙，对我们世界的物理特性会产生深远的影响。尽管膜宇宙的泛滥在某种程度上让人心烦，可它却真的令人振奋。膜不仅有助于解决粒子物理里长期存在的问题，而且，如果我们足够幸运，且描述的某个图景正确的话，我们将很快在基本粒子物理实验里找到膜宇宙存在的证据。我们有可能真的生存在一个膜上——不出10年，我们就会真正弄清楚。

至于现在，我们还不知道，在众多的可能性里哪一个（如果有的话）会是对宇宙的正确描述。因此，我将保留所有选择，以免遗漏什么有趣的东西。无论最终证明哪种图景正确描述了我们的世界，我要讲的那些图景都会引进一些引人入胜的、全新的观点，而这都是前人从未想到过的。



理论物理世界的奇幻旅行

WARPED PASSAGES

她是模特，魅力四射。

发电站乐队 (Kraftwerk)

什么速度呀和新发明，以及第四维空间诸如此类

“嗨，阿西娜，你是在看电影《卡萨布兰卡》吗？”

“是的，不想和我一起看吗？这一幕可真棒！”

请你一定要牢记，

亲吻就是亲吻，

叹息就是叹息。

随着时光流逝，

爱情真谛永不变。

“等等，艾克，你不觉得最后这句有点奇怪吗？这里本来应该很浪漫的，可听起来倒像是在讲物理。”

“阿西娜，如果这你都觉得奇怪的话，那你最好还是先听听原版歌曲《任时光流逝》 (*As Time Goes By*) 的开始一段。”

我们生活的这个时代，

爱弄些缘由给人理解。

什么速度呀和新发明，

以及四维空间诸如此类。

我们还有些许烦恼，

有关爱因斯坦先生的理论。

“艾克，你真以为我会相信吗？我猜，下一步你就会说里克和伊尔萨逃进了第七空间！好了，就当我刚才什么都没说，还是坐下来，我们安安生生地看电影吧。”

20世纪早期，爱因斯坦提出了相对论；1931年，鲁迪·瓦利（Rudy Vallee）录制了由赫尔曼·赫普菲尔德（Herman Hupfeld）所作的歌曲《任时光流逝》，就是艾克叙述的那一版。但是，等到由杜利·威尔逊（Dooley Wilson）扮演的山姆一角演奏了《卡萨布兰卡》的曲调之后，那段被省略了的歌词以及时空科学在大众文化里就被遗忘了。尽管萨尔瓦多·卡鲁扎早在1919年^[8]就提出了宇宙中存在一个额外维度的观点，但是，直到最近，科学家才对这一观点给予了足够的重视。

我们已探讨了什么是维度，以及维度是怎样逃过我们的眼睛的。下一步，我们就要问：是什么重新激起了人们对额外维度的兴趣？为什么科学家要相信它们确有可能存在于真正的物质世界里？这要解释起来，话可就长了，它会涵盖20世纪一些最为重要的物理学成果。后面几章里，在开始描述可能的多重宇宙之前，我将首先回顾这些成果，并解释它们为什么是新近一些理论的先驱。我们将考察发生在20世纪早期的范式转变（量子力学、广义相对论）、当今粒子物理学的精髓（标准模型、对称性、对称破缺、等级问题）以及关于解决当前未解问题的新观点（超对称、弦理论、额外维度和膜）。

但是，在涉猎这些问题之前，本章将简要地介绍一下物质，目的是搭建一个物理平台。由于要理解我们的探索方向，得先熟悉当今物理学家所采用的推理类型，因此，我们还将探讨对于新近进展至关重要的理论研究方法。

最初，我觉得使用歌词“爱情真谛永不变”是一个明智的选择，但细想之后，我觉得这歌词也太物理了，我甚至怀疑是不是记忆在跟我开玩笑，因为这对记歌词来说是常有的事——即使你认为这首歌早已烙进了你的脑海里。可是，当我发现，这歌词实际比我想象的还要契合物理学时，就不由得我不惊讶了（而且觉得好玩）。以前我可从没想到过“时光流逝”就是第四维度！

物理见解常常也像这次发现一样：细小的线索往往会揭示一些不期然的联系。如果你足够幸运，你的发现甚至会超出你的预期，令你喜出望外，当然你得找对地方。在物理学中，一旦你发现了某种联系，即使只有一点细微的线索，你也会以你认为最好的方法去寻求其意义。那可能需要科学的推测，也可能需要为你认为可靠的理论导出数学结果。

下一节，我们将介绍现代物理学探寻这些线索所采用的方法：模型构建（这可是我的强项）以及基本高能物理的另一种方法——弦理论。弦理论学家试图从一个确定的理论得出对宇宙的预言；而模型构建者试图先找到解决特定物理问题的方法，然后由此出发创建理论。模型构建者和弦理论学家都意图寻找更为综合的理论，来解释更多的问题。他们旨在回答类似的问题，却以不同的途径接近它们。研究有时是科学的推测，如模型构建；有时是为已确信正确的最终理论推理出逻辑结果，如弦理论。很快我们会发现，额外维度的最新研究成功地结合了这两种方法。

自上而下，还是自下而上

尽管最初我喜欢数学和科学都是因为其彰显出的确定性，而今我却发现，那些未解的问题和不期然的联系，对我更有吸引力。量子力学、相对论和标准模型所包含的原理，拓展了人们的想象。但是，它

们却无法撼动当今物理学家正倾力关注的一些新奇的观念。鉴于现存观念的缺陷，我们知道是新理论出场的时候了。那些缺陷预示了新的物理现象，在我们完成更精确的实验时，它们必将出现。

粒子物理学家试图找到一些自然定律来解释基本粒子的运动规则。这些粒子以及它们所遵循的物理定律，就是物理学家所称理论的基本组成部分。理论是一整套的要素和原理，包括预测各要素之间相互作用的规则和方程。在本书中，当我说“理论”时，指的就是这层含义，而不是口语里说的“粗略猜想”的意思。从理想角度讲，物理学家当然希望找到一个能解释所有现象的理论，最好还是一个规则最少、基本组成要素最少的理论。有些物理学家的终极目标就是创建一个简单、凝练、统一的理论——一个可以用来预测所有粒子物理实验结果的理论。

寻求如此高度统一的理论，可谓志向远大，甚至有人说是胆大无畏。但在某种程度上，它反映了人们自古就有的对简洁理念的追求：在古希腊，柏拉图设想出完美的形式，如几何形状和理想状态，世间万物只能大致近似；亚里士多德也相信理想形式，但他认为只有观察才能揭示世间万物近似的理想形式；宗教也常常会想象出一个更为完美或更为统一的状态，它从现实中脱离出来，却又以某种形式与现实相联：伊甸园里人类堕落的故事，就设想了一个理想的极乐世界。尽管现代物理学的问题和方法与先哲们的不同，但物理学家同样是在追求一个简洁的宇宙，只不过表现为构成世界的基本成分，而不是哲学和宗教。

自上而下

有的理论学家采取一种“自上而下”的办法：他们由理论出发，首先相信理论是正确的，比如弦理论学家由弦理论出发，然后试图导出其结果，使其与我们观察到的纷繁世界联系起来。

然而，有个明显的障碍使我们无法找到一个能与现实世界相联的完美理论：我们周围找不到这种完美理论所应体现的简洁性。问题在于世界本身就很复杂。要将一个简洁、凝练的形式与一个复杂、真实的世界联系起来需要做大量工作。一个统一的理论，一方面要简洁凝练，另一方面又必须有足够的空间来容纳相应的观察和发现。我们倒乐意相信会有这么一个视点，从中看到的所有东西都是完美、可预见的，可世界却不像我们描述它的理论那般单纯、简洁和有序。

粒子物理学家以两种不同的方法——自上而下和自下而上，来缩短理论与现实之间的距离。

有的理论学家采取“自上而下”的办法：他们由理论出发，首先相信理论是正确的，比如弦理论学家由弦理论出发，然后试图导出其结果，使其与我们观察到的纷繁世界联系起来。而模型构建者选择的则是一条“自下而上”的道路：他们先找出观测到的基本粒子与其内在相互作用之间的联系，然后由此推导出其基本理论。他们在物理现象中寻找线索，创建模型，即一些样本理论。最终，这些理论可能正确，也可能错误。两种方法各有其优势和缺陷，而最佳的前进路线并不总是显而易见的。

两种科学方法之间的冲突很有意思，因为它反映了两种不同的科学研究方式，这种分化是科学界长期争议的新近体现。你是选择柏拉图的方法，还是亚里士多德的方法？柏拉图试图从一些基本真相里获得领悟，而亚里士多德则立足于经验观察。你是准备自上而下，还是自下而上？

自下而上

模型构建者选择的则是一条“自下而上”的道路：他们先找出观测到的基本粒子与其内在相互作用之间的联系，然后由此推导出其基本理论。他们在物理现象中寻找线索，创建模型，即一些样本理论。最终，这些理论可

能正确，也可能错误。两种方法各有其优势和缺陷，而最佳的前进路线并不总是显而易见的。

这种选择也可以称为“老年爱因斯坦Vs. 青年爱因斯坦”。爱因斯坦年轻时立足于实验与客观现实，即使他所谓的思想实验也都来自物质场景。在发展广义相对论时，他发现了数学的价值，之后，爱因斯坦改变了方法：他发现数学成果对其理论完成是至关重要的，这使得他在以后的事业中更多地使用理论方法。然而，向爱因斯坦看齐未必会解决问题。尽管他将数学成功地应用到了广义相对论中，但他后来为了统一理论而进行的数学探索却毫无结果。

从爱因斯坦的研究道路可以看出，科学真理有很多种，发现它们的道路也有很多种：其中一个基于观察，例如类星体和脉冲星的发现；另外一个基于抽象的原理和逻辑，例如，卡尔·史瓦西（Karl Schwarzschild）首先推导出黑洞是广义相对论的数学结果。最终，我们希望这些方式能够互相融合——如今，黑洞既能从观测数据的数学计算导出，也可以从纯理论导出，但在研究早期，我们基于两种真理所取得的进展却很少同步。对弦理论来说，它的原理和方程并不如广义相对论那般完善，因此由它推导出结果要困难得多。

弦理论一经崭露头角，粒子物理学界立即被彻底分裂。“弦革命”首次使粒子物理学界产生分化是在20世纪80年代中期，当时我还在读研究生。从那时起，就有一派物理学家决意要全身心地投入弦理论的优美数学王国中。

弦理论的根本前提是：自然界的基本物质是弦，而非粒子。我们现在观测到的周围的粒子只不过是弦的结果：它们是由振动的弦产生的不同振动方式，就像是从振动的小提琴弦上跳出的不同音符。弦理论之所以引人瞩目，是因为物理学家正在寻找一种能够有机融合量子力学和广义相对论的理论，能够作出可直达最为微观的可探测领域的预言。在许多人看来，弦理论似乎是最有希望的理论。

然而，另一派物理学家却决意要留在实验能够探索的相对低能的物理领域。当我在哈佛大学时，那里的物理学家，包括优秀的模型构建者霍华德·乔治（Howard Georgi）和谢尔登·格拉肖（Sheldon Glashow），以及众多天才的博士后及研究人员，全都是模型构建的忠实拥护者。

一开始，弦理论和模型构建两种对立观点的争论很热烈，两派都坚持认为自己才是探索真理的正确道路。模型构建者认为弦理论学家是在一个梦幻的数学领地里，而弦理论学家则认为模型构建者纯属在浪费时间，无视真理。

由于在哈佛大学有许多杰出的模型构建者，而我又喜欢模型构建的挑战，因此我初涉粒子物理学时，是站在这一阵营的。弦理论是一个光彩夺目的理论，它已得出了一些深刻的数学和物理见解，其中将很可能包含能最终描述自然的正确成分。但要找到弦理论与真实世界的联系却是令人却步的艰巨任务。问题在于，弦理论所定义的能量尺度比我们能以现有仪器探索到的能量要大1亿亿倍，即使粒子碰撞的能量增大10倍，我们也不知道会发生什么。

就我们现在了解的情况，弦理论与它描述世界的预言之间，还隔着一条巨大的理论鸿沟。弦理论方程所描述的物体实在是小到令人难以置信，而其具备的能量又高至超乎想象，即便是以我们可以想到的任何技术制造出任何探测仪器，都不大可能探测到它们。不仅从数学来讲要得出弦理论的结果和预言极具挑战，而且我们甚至都还不甚明了该怎样组织弦理论的基本要素，及确定该解决哪个数学问题。在一个充满了岔路的丛林里，是很容易走失的。

在我们能实际观测的距离内，弦理论能带来很多可能的预言，其预言的粒子立足于理论中尚未明了的基本成分构造。如果排除一些推想的假设，弦理论的世界可能比我们的可见世界包含更多的粒子、更

多的力以及更多的维度。我们需要知道，是什么使得额外的粒子、力以及维度与我们所见到的不同。我们还不知道会有什么物理特征倾向于某种构造而非另一种，甚至不清楚怎样找到弦理论与现实世界相符的任何表现。只有在非常幸运的情况下，我们才能萃取出所有正确的物理原理，使弦理论预言与我们的观察相匹配。

例如，弦理论中不可见的额外维度必须与我们看到的三维不同。弦理论的引力要比我们周围常见的引力更为复杂：与让苹果落下砸到牛顿脑袋的那种力不同，弦理论的引力要作用于6~7个额外维度。尽管弦理论非常新奇且炫目，但其诸如额外维度等一些令人迷惑的特征却模糊了它与可见宇宙的联系。是什么让那些额外维度与可见维度不同？为什么它们不全都一样？如果能发现自然为什么和怎么隐藏弦理论的额外维度，将是非凡的成就，无论使用什么探索方法都是值得的。

弦理论也想让自己“现实”起来，但到目前为止，所有尝试似乎都带着宇宙外科手术的意味：为了使其预言与现实世界相符，理论家不得不剔除所有不应具备的成分——去除粒子，将额外维度不动声色地隐藏起来。尽管由此得出的粒子与正确粒子相去不远，但你仍能断定它们并非完全正确。优雅的确是一个正确理论所应具备的优良特征，但只有在我们完全明了理论的所有含义时，才能判定它是否真的完美。弦理论起初是令人惊艳的，但弦理论学家最终必须面对这些根本的问题。

当我们在一片山区探险时，没有地图，就很难判断哪条路才是最终到达目的地的最近路线。在思想领域，也如同在地形复杂的山区一样，该走哪条路并不是从一开始就明晰的。即使弦理论最终能将所有已知的力和粒子有机地统一起来，我们仍不能断定，它包含的只是一个呈现了一定粒子、力和其相互作用的山峰呢，还是一个更为复

杂的、有着多种含义的广阔天地？如果一路坦途，路标明确，那么探索将轻而易举。但，事情却很少这样发展。

因此我要强调，超越标准模型向前的道路是模型构建。“模型”这个词很容易让人想起小时候建造的小型战舰或城堡，或者电脑里为了重建一种已知动态的数字模拟。例如，人口是怎样增长的、海水是怎样运动的。但粒子物理学的模型构建却不同于以上任何一种含义，它与时尚界或是杂志上所用“模特”的含义有些类似：无论是T台上的模特，还是物理学中的模型，展现的都是一种富于想象力的创作，它们会以各种各样的形式呈现出来，最美的必然会赢得最多的关注。

不用多说，相似仅此而已。粒子物理学的模型是对标准模型之外的另一些基本理论的推测。如果你认为一个统一的理论是山脉的顶峰，那么模型构建者就像是一个探路者，他试图找到连接山脚与山顶的路，一条能最终将所有观点都联系起来的路，而这基础就是已确立的物理理论。尽管所有模型构建者都承认弦理论的确很出色，且有可能最终被证明是正确的，但他们仍不能像弦理论家学那般确定地知道，如果最终登顶会发现什么理论。

在第7章我们会看到，标准模型是一个明确的物理理论，有一组固定的四维世界的粒子和力。超越标准模型的模型仍包含这些基本成分，并在已探明的能量水平上重现其结果，而且它们还包含了一些只在更小的距离内才能探测到的新力、新粒子和新相互作用。物理学家提出这些模型，是为了解决眼下的难题。而模型可能会为已知或假想粒子提出一些新的不同的行为，这些行为取决于模型假设所导出的一组新的方程。模型也可能提出新的空间场景，例如我们将以额外维度和膜来探索的那些场景。

即使我们完全通晓了一个理论及其含义，这个理论还可能以其他方式表现出来，对于我们生活的真实世界，它们可能会有不同的物理结果。比如，即使我们从理论上知道粒子和力是怎样相互作用的，仍需要知道在真实世界里存在哪些特定的粒子和力，模型使我们能抽样检验这些可能。

不同的假设和物理概念可以区分不同的理论，同样，理论原理所适用的距离和能量尺度也可以区分不同的理论。模型是直达这些不同特征核心的一种方法，它们让你探索一个理论的内在含义。如果你认为理论是指导你做蛋糕的一个大致说明，那么模型则是一个精确配方。理论会告诉你加糖，而模型则会明确说明加半杯糖还是两杯糖；理论会说葡萄干可加可不加，而模型则会给你一个更明白的指示：不要加。

模型构建者关注的是标准模型里那些未解决的问题，并试图使用已知的理论来应对其不足。模型构建的方法因直觉而激发起来，弦理论明确预言所在的能量远大于我们的观察。而模型构建者想了解全面的景象，以找到与我们的世界相关的部分。

现实地说，模型构建者承认不能立即导出一切东西。我们不是要导出弦理论的结果，而是要弄清楚基本物理理论的哪些要素能解释已知的观测结果，揭示实验结果之间的联系。模型的假设可能会是最终基本理论的组成部分，也可能在我们明白其更深层的理论基础之前就阐明某些新的关系。

物理总是努力以最少的假设来预言最多的物理量，但这并不意味着我们总能立即就明确找出最为根本的理论。常常是人们先有了进步，然后才从根本上明白其来龙去脉。例如，物理学家早就知道温度与压力的关系，并将其应用于热力学与发动机的设计，但直到很久以

后，人们才在一个更基本和微观的层面将其解释为大量原子和分子的无规则运动的结果。

因为模型关注物理“现象”（意即实验发现），因此与实验密切联系的模型建构者有时被称作“现象学家”。“现象学”可算不上是一个好措辞，因为它不能公正地涵盖数据分析，而数据分析在当今复杂的科学世界里与理论是密不可分的。模型构建者更多地是注重解释与数学分析，而不是像这个词在哲学里的含义所暗示的：简单的表象。

相反，最好的模型确实有其可贵之处：对于物理现象，它们能作出确定的预言，给实验者验证或反驳模型的断言提供一个途径。高能实验不仅仅是在寻找新的粒子，也是在测试模型，并寻找建立更好模型的线索。所有已创建的粒子物理学模型，都包括适用于可探测能量的新的物理原理及新的物理定律，因此，它能预见新的粒子及它们之间的可测试的关系。找到这些粒子并测量其性质，就能证实或排除一些思想。高能实验的目标就是帮助我们找到一些基本的物理定律，还有赋予它们解释能力的概念框架。

并非所有的模型最终都能被证明是正确的，但模型仍是研究可能性及构筑素材库的最好方法。如果弦理论正确，我们最终就可能得知某些模型是如何导出其结果的，就像热力学源于原子理论那样。但是，十多年来，两大阵营的对峙没有一丝结束的意思。阿尔比恩·劳伦斯（Albion Lawrence），一位来自布兰迪斯大学的年轻弦理论学家，在与我讨论这种分歧时，是这么评论的：“不幸的是，弦理论与模型构建是两个截然不同的课题，模型构建者与弦理论学家多年来互不交流。我总觉得，弦理论就像所有模型的祖父一样。”

弦理论学家与模型构建者都在探寻一条有迹可循的优雅路径，以将理论与可见世界联系起来。任何理论，只有其道路本身，乃至从山顶往下看

到的风景，都显现出它的优雅时，才真正是引人入胜和可能正确的。模型构建者自下而上，很可能要承受多次起点错误的风险，而由自下而上的弦理论学家同样也有风险：他们可能会发现自己正处在一个陡峭、孤立的悬崖边，远离营地，无法找到回去的路。

你也可以说我们正在寻找一种宇宙的语言：弦理论学家关注的是语法的内在逻辑，而模型构建者则专注于他们认为最有用的名词和词组。如果说粒子物理学家是在佛罗伦萨学习意大利语，那么，模型构建者则会懂得怎样寻找食宿，并学会问路所需要的最基本的词汇，但他们说出的话可能很好笑，而且可能永远也不能完全领会“炼狱”的含义；相反，弦理论学家可能立志于掌握意大利文学的精妙之处，但他们可能不等学会怎样点餐，就将面临被饿死的危险。

好在情况现在已有了改观。最近，理论研究和低能物理现象两者都支持了对方的进展，而许多人现在也已开始同时考虑弦理论和以实验为指导的物理学了。在我的研究中，我还是继续沿袭模型构建的道路，但是，也会纳入弦理论的观点。我认为，最终很有可能会通过结合两种途径的最佳方法而取得进展。

阿尔比恩说：“两者之间的界限不再那么鲜明了，这很大程度上得益于额外维度的研究，人们开始互相交流。”两派不再是那么泾渭分明，也有了更多共同的立场。目标和观念重新开始融合，无论是在学术上还是社交上，模型构建者和弦理论学家都有了强烈的重合。

我将描述的额外维度理论最美的一面就是，两个阵营的观点开始融合起来。弦理论的额外维度也许有些痴人说梦，但兴许它们最终是一个机会，可以帮助我们找到解决老问题的新方法。我们当然可以问：额外维度在哪儿？为什么我们没见到它们？但我们也可以这样问：这些看不见的维度对我们的世界是否有重要意义？它们可能会帮助我们揭示一些与观测现象相关的内在联系。模型构建者喜欢尝试

把一些概念联系起来，比如额外维度与可见的三个维度、粒子质量之间的关系。而且，如果幸运的话，基于额外维度模型的见解有可能会成功解决弦理论面临的一大难题：它的不可实验性。模型构建者已使用了由弦理论得出的理论要素来解决粒子物理学中的问题，而那些模型，包括那些有额外维度的模型，将会产生可以检验的结果。

我们后面研究额外维度模型时会看到，模型构建的方法正与弦理论并肩探索着粒子物理学、宇宙演变、引力和弦理论，并产生了重要的新见解。有了弦理论的“语法知识”和模型构建者的“词汇”，两者就可以开始携手编撰一本合乎逻辑的“短语手册”了。

进入物质内核

我们最终将探讨的观点会包含整个宇宙，但这些观点的根源在于粒子物理学和弦理论——志在描述物质的最小组分的理论。因此，在开始我们的旅程，进入这些理论性极强的领域之前，让我们先来看看物质，深入其最微小的部分进行一番简单的游历。在探索原子的旅途中，我会做你的向导，请记住物质的基本建构成分以及不同物理理论所讨论的物质的尺寸，它们可以作为地标，然后，你可以根据它们识别方向，并分辨每个物理领域所关注的基本成分。

物理学中，大多数理论的基本前提是：物质是由一些基本粒子构成的。层层剥离、逐步深入之后，最终你总能发现基本粒子。这是粒子物理学家研究的宇宙，在这里，粒子是最小的元素。弦理论将这一假设再推进一步，设想那些最小的微粒是一些基本的振动弦。但即使是弦理论学家，他们也相信物质是由一些微粒——其核心不可再分的实体构成的。

相信所有东西都是由粒子组成的，这或许有些困难。确实，对我们的肉眼来说，这很不明显。这是由于我们的感官能力太过粗糙，无

论何时何地，都不能直接探知如原子一般微小的东西。但是，即使我们不能直接看到它们，基本粒子仍是物质的基本组成部分。就如在你的电脑或电视屏幕上呈现的虽然是连续的图像，但这些图像实际上是由一些小点组成的。物质是由原子构成的，而原子又是由更小的基本粒子构成的。我们周围的物质看上去连续完整，但其实质并非如此。

物理学家要探查物质内部，并推导出其基本成分，首先需要技术上的进步，创造出灵敏的测量仪器。但是，每当他们开发出更为精确的技术工具，就会出现一些更细微的结构、更为基本的成分。而每次物理学家得到能够探索更小尺度的工具时，都会发现还有更基本的组成成分——亚结构，构成前面已知结构元素的成分。

粒子物理学家的目标就是发现物质的最基本成分以及这些基本成分所遵循的基本物理定律。我们研究小尺度，是因为基本粒子就在这些尺度上相互作用，而且很容易解析出基本的力。在大尺度上，基本成分被联结成化合物，很难解析出基本物理定律，因此也就比较模糊。小尺度的奇妙之处就在于新的原理和新的联系只在这里适用。

物质不仅仅只是一个俄罗斯套娃，里面一个套一个都是一样的复制品，而更小的距离会揭示出真正新奇的现象。直到17世纪，在威廉·哈维等科学家将人体解剖开、看到人体内部之前，即便是人体的运行，例如心脏和血液循环，都是被错误理解的。而最近的实验对物质也做了同样的事：探索更小的距离，新的世界在那里通过更基本的物理定律运行着。正如血液循环对所有的人体活动都很重要一样，基本物理定律在大尺度上对我们同样有着重要的影响。

现在我们知道所有物质都是由原子构成的，它们通过化学过程联合在一起



形成分子。

原子非常小，大约是一埃，即一亿分之一厘米。但原子不是最基本的：它还包括一个位于中心的、带正电的原子核，原子核周围环绕着带负电的电子（见图4-1）。原子核比原子要小得多，只占原子大小的十万分之一。而带正电的原子核本身也是一个复合物：它由带正电的质子和中性（不带电）的中子组成，两者合称为核子，它比原子核小不了多少。

这是20世纪60年代以前科学家们所持有的有关物质的图景，也很有可能是你在学校里学到的蓝本。

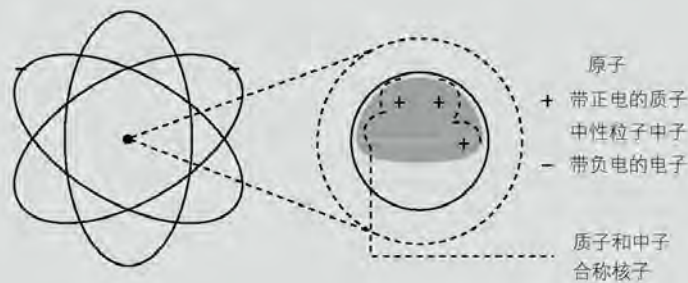


图4-1 原子构成图。原子包括一个微小的原子核和围绕原子核旋转的电子，而原子核又包括带正电的质子和不带电的中子。

尽管以后我们会看到，量子力学给出了一幅比你能画出的任何图形都更为有趣的、有关电子运行轨道的图画，可

原子的这一模板还是正确的。但是现在我们知道，质子和中子也不是最基本的粒子。与我在引言里提到的伽莫夫的话相反，质子和中子还包含亚结构，一种更为基本的组成成分，叫作夸克。质子包含两个上夸克和一个下夸克，而中子包含两个下夸克和一个上夸克（见图4-2），这些夸克通过一种叫作强力的原子核力束缚在一起。而原子的另一组成成分电子——却不同，就我们现在所知，它是基本的粒子：电子不能被分成更小的微粒，里面不再含有亚结构。

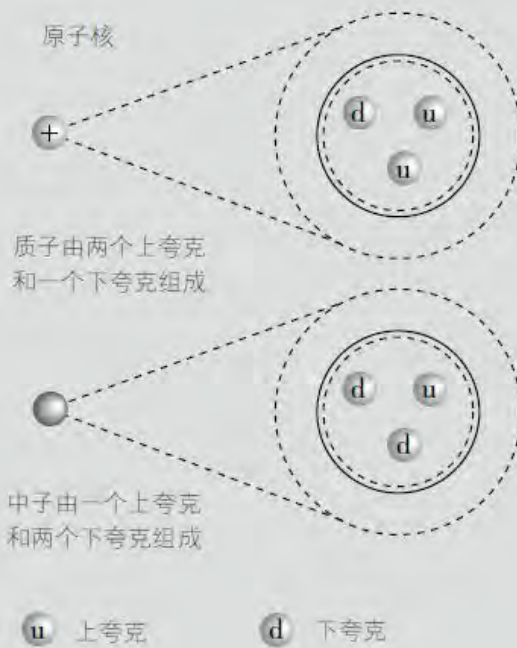


图4-2 原子核构成图。质子和中子是由更为基本的夸克通过强力束缚在一起形成的。

物理学家、诺贝尔奖获得者斯蒂芬·温伯格（Stephen Weinberg）发明了“标准模型”一词，用它来称呼已确立的粒子物理学理论，它描述了物质的基本组成成分——电子、上夸克、下夸克之间的相互作用以及我们很快就将谈到的其他基本粒子。标准模型还描述了基本粒子4种相互作用力中的3种——电磁力、弱力和强力（引力通常被省略）。

虽然几百年前我们就知道了引力和电磁力，但直到20世纪后半叶，还无人知道后面这两种不太熟悉的力：弱力和强力作用于基本粒子，在核反应过程中至关重要。例如，它们将夸克束缚在一起，使原子核衰变。

如果愿意，我们还可以把引力也包括在标准模型里，但我们通常并不这样做，因为在与粒子物理相关的距离尺度上，引力实在是太微弱了，远不足以在实验所能达到的能量水平上产生任何影响。我们关于引力的通常概念，在极度高能和极小尺度上失去了作用。这对弦理论很重要，但并不出现在可测量的距离尺度上。研究基本粒子时，引力只在标准模型的某些延伸里才有意义，例如，在我们后面将探讨的额外维度模型里。而在所有其他有关基本粒子的预言里，我们都可以忽略引力。

既然已进入到基本粒子世界，那我们就四处看看，参观一下邻居们的领地。上夸克、下夸克和电子位于物质核心，可是，现在我们知道，还有另外一种更重的夸克以及其他一些更重的像电子一样的微粒，它们在通常的物质里是找不到的。

例如，电子的质量只是质子质量的0.5%，而有一种叫作 μ 子的粒子，与电子有精确相等的电荷，其质量却比电子质量大200倍；另一种叫作 t 子的粒子，也有等量的电荷，其质量还要再大出10倍。而且，在过去的近40年里，高能对撞机实验还发现了更重的粒子。为了产生它

们，物理学家需要大量高度集中的能量，当今高能粒子对撞机实现了这种需要。

我知道这节是我们进入物质内核的游览，但我刚才所说的这些粒子并不处于物质世界的稳定物体中。尽管所有已知物质都由基本粒子构成，但更重的基本粒子并不是物质的组成部分：在你的鞋带里、桌面上、火星上或是你已知的任何其他现实物体上，你都找不到它们；它们是由当今高能对撞实验产生的，是紧随宇宙大爆炸之后早期宇宙的组成部分。

但不管怎样，这些重粒子是标准模型的基本成分。与我们熟悉的粒子一样，它们也通过相同的力互相作用，而且很可能会帮助我们更深入地了解物质的最基本物理定律。如图4-3和图4-4所示，我列出了标准模型的粒子，中微子和传递力的规范玻色子也包括在内，在第7章详细讨论标准模型的元素时再细细讲述。

第一代	上夸克 3 MeV	下夸克 7 MeV	电子中微子 ~0	电子 0.5 MeV
第二代	粲夸克 1.2 GeV	奇夸克 120 MeV	μ 子中微子 ~0	μ 子 106 MeV
第三代	顶夸克 174 GeV	底夸克 4.3 GeV	t子中微子 ~0	t子 1.8 GeV

图4-3 标准模型里物质的粒子以及它们的质量。同一列里的粒子电荷相等，质量却不同。

	电磁力	弱力	强力
传递力的规范玻色子	光子 无质量	弱规范玻色子 W^+ 80 GeV Z 91 GeV	胶子 无质量

图4-4 标准模型里传递力的规范玻色子、它们的质量以及所传递的力。

没人知道为何存在标准模型的重粒子，有关它们的目的、它们在终极基本理论中发挥的作用，以及为何它们的质量与我们更为熟悉的物质组成成分的质量相差如此之大，诸如此类的问题是标准模型面临的主要未解之谜。而这还只是标准模型未能解决的众多谜题中的几个而已，例如，为什么只有4种力，而没有其他力？是否还有其他力我们未能探测到？为什么引力相比其他已知力如此微弱？

标准模型还留下了一个更为理论性的问题，这也是弦理论希望能够解答的：我们怎样将量子力学和引力在所有距离尺度上协调起来？这一问题不同于其他，它与当今可见现象无关，而是一个关乎粒子物理学内在局限的问题。

两种类型的未解问题——有关可见现象和纯理论现象的，给了我们足够的理由去超越标准模型。尽管标准模型强大而成功，可我们仍然相信，必定还有更为基本的结构等待我们去发现，而对更为基本原理的求索也必将有所回报。就如作曲家史蒂夫·里奇（Steve Reich）在《纽约时报》上的精辟言论一样（他为自己的一个作品打的一个比方）：

先只有原子，然后有质子和中子，然后又有夸克，而现在我们开始谈论弦理论。似乎每隔20、30、40、50年，就会有一扇新的大门打开，向我们展现又一层次的世界。

当前和将来的粒子对撞机实验不再寻找标准模型的成分，因为它们都已被发现了。标准模型根据它们的相互作用，已很好地组织了这些粒子，它们的全家福已经很清楚，而实验家们要找的是更为有趣的粒子。现在的理论模型包括了标准模型的成分，但为了解决标准模

型的未解问题又加进了一些新的元素。我们希望，现在和将来的实验能够提供线索，让我们识别它们，发现物质的真正本质。

尽管我们已经在实验和理论上了解了更根本的理论应具有的特征，但在高能实验（探索更小距离）提供答案之前，我们仍不太可能知道哪个才是对自然的正确描述。正如我们在后面看到的，理论线索表明，未来10年的实验几乎肯定会发现新的东西。它可能不是弦理论的确切证据，因为这很难找到，但我们会看到它们将是一些奇异的事物，如新的时空关系或新的却仍未发现的额外维度，它们会是在弦理论以及其他的粒子物理理论里发挥重要作用的新现象。即便我们有着丰富的想象力，这些实验仍有可能揭示我们从未想过的事物。它们究竟是什么？我和我的同事们都在好奇地等待着。

宇宙的想象力远比我们丰富

我们已知道刚才所介绍的物质结构是20世纪物理学研究重要进展的结果，这些伟大的进展对于我们将提出的有关世界的更为普适的理论至关重要，而它们本身就是伟大的成就。

从下一章开始，我们将回顾这些进展。理论在观察和克服前人理论的缺陷基础上发展起来，通过了解那些非凡的早期成果，你可以更好地认识新近成果的作用。图4-5指出了我们所探讨的理论的内部联系。我们会看到每种理论是怎样通过总结旧理论的经验建立起来的，而新理论又是怎样填补了只有在完善了旧理论之后才能发现的空白的。

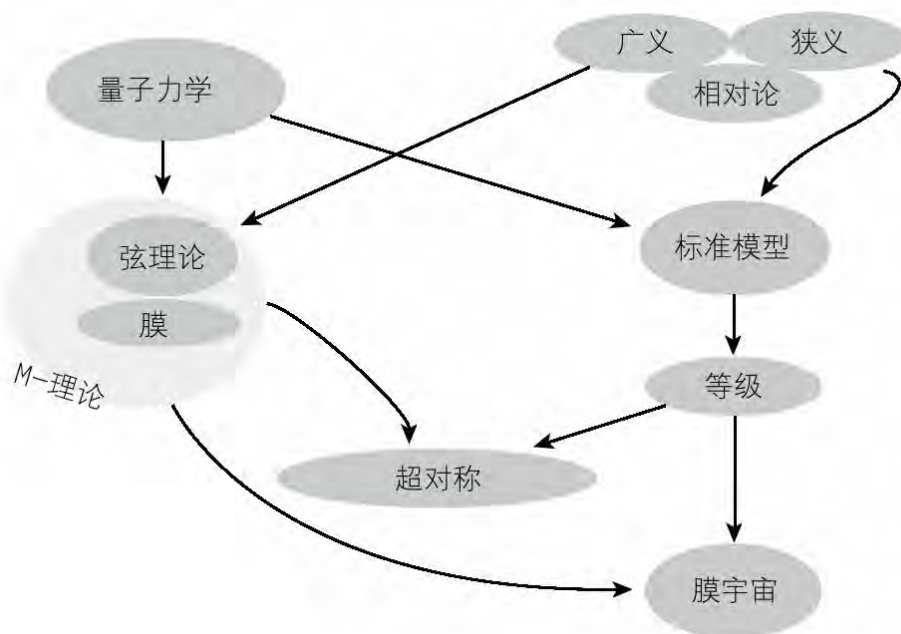


图4-5 我们将探讨的物理领域及其联系。

我们首先介绍的是20世纪早期两个革命性的观念：相对论和量子力学，通过它们，我们知道了宇宙的形状、它所包含的物质以及原子的构成和结构。然后，我们介绍了粒子物理学的标准模型，它发展于20世纪六七十年代，预言了我们刚才所看到的基本粒子的相互作用。我们还将探讨粒子物理学中一些最为重要的原理和概念：对称、对称破缺以及物理量对尺度的依赖，通过它们我们了解到物质的最基本成分如何生成了我们看到的结构。

标准模型虽然很成功，却留下了一些没有解决的基本问题。这些问题是如此根本，解决它们会让我们对于这个世界的建构基础产生新的认识。在第10章，我们将展现标准模型里最为有趣也最为神秘的一面：基本粒子质量的来源。我们会看到，如果要解释已知粒子的质量以及引力的微弱问题，我们几乎肯定需要一个比标准模型更为深刻的物理理论。

额外维度模型探讨粒子物理学的这些问题，却还是利用了弦理论的观点。在讨论过粒子物理学基础之后，我们将介绍弦理论的根本动机和概念。我们不会由弦理论直接得出模型，但弦理论包含了一些我们在创建额外维度模型时所使用的元素。

这个回顾涉及许多立场，因为额外维度研究将粒子物理学的两条主线——模型构建和弦理论的许多理论成果结合在了一起。在一定程度上了解这些领域里最为有趣的新成果，会有助于你更好地理解构建额外维度模型的根本动机和方法。

但是，如果你想跳过这些知识，那么在成果回顾的每一章结尾，我都会列出一些重要概念，以便以后在我们回到额外维度模型构建时参考。这些“探索大揭秘”列出了本章的概要、总结，以便在你想跳过某一章，或者想专注于某个资料时，可以用于以后查阅。我也偶尔会提到不在“探索大揭秘”里的内容，但这些“探索大揭秘”回顾的是一些关键概念，对本书后半部分的主要结果至关重要。

从第17章开始，我们将探索充满了额外维度的膜宇宙。这些理论提出，组成我们宇宙的物质被限制在一个膜上。膜宇宙观点提供了关于广义相对论、粒子物理学和弦理论的新见解。我将呈现的这些不同的膜宇宙会作出不同的假设，解释不同的现象。同样，在每章结尾的“探索大揭秘”里，我会总结每个模型的特征。我们还不知道哪种观点（如果有的话）会正确地描述自然。但我们完全可以相信，我们最终会发现膜是整个宇宙的一部分，而我们连同其他平行宇宙都被限制在了膜上。

我在这项研究中学到了一样东西：宇宙的想象力远比人类要丰富。有时，它们的秉性是如此出人意料，我们只是碰巧才能发现它们。能遇到这样的惊喜真是太神奇了，我们的已知物理定律也因此会产生令人吃惊的结果。

现在，我们就开始探索那是些什么定律。



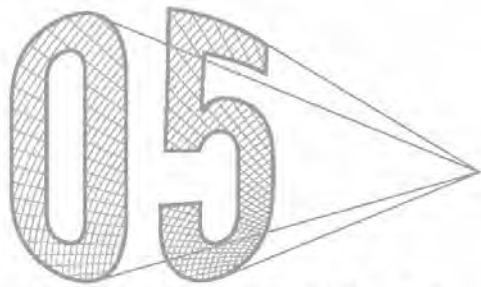
WARPED

第二部分

20 世纪两大革命性宇宙观

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



相对论：弯曲的时空

WARPED PASSAGES

引力定律是非常、非常严格的，只不过是我們为了自己的利益，一直在歪曲它們。

比利·布拉格 (Billy Bragg)

艾克的保时捷和GPS

伊卡洛斯·拉什莫尔三世（艾克）迫不及待地要向迪特尔展示他的新保时捷，但是，虽说这新车很值得他骄傲，可更令他兴奋的是他最近自行设计安装的全球定位系统（GPS）。

艾克想让他的朋友迪特尔佩服一下，于是，他说服迪特尔一起驾车到当地的赛车场。他们上了车，艾克将目的地输入程序，两人就出发了。可让艾克大失所望的是，他们最终停错了地方——GPS系统根本不像他预计的那么管用。迪特尔的第一反应是，艾克一定犯了什么可笑的错误，比如说把米和英尺弄混了。但艾克坚信自己不会犯这么低级的错误，他向迪特尔保证，问题肯定不在这儿。

第二天，艾克和迪特尔摆弄了几下。可令人沮丧的是，他们在兜风时发现GPS比以前表现得更糟了。艾克和迪特尔继续找问题，终于，在困惑了一周之后，迪特尔忽然灵感闪现：他快速地计算了一下，吃惊地发现，如果不考虑广义相对论，艾克的GPS系统累积的误差每天会达到16公里。艾克没想到他的保时捷会快到需要考虑相

对性计算，但迪特尔解释说，车本身没有那么快，可GPS信号是以光速传播的。迪特尔修改了软件，将GPS信号必须穿过的引力场的变化也计算在内，这次，艾克的GPS系统终于能像市面上买到的一样正常运行了。

艾克和迪特尔长舒一口气，开始计划他们的下一次兜风。

20世纪初，英国物理学家开尔文爵士说：“现在，物理学家已经没什么新东西可以发现了，剩下的就是让计算越来越精确。”开尔文爵士的错误从此可是出了名：就在他讲过这番话后不久，相对论和量子力学彻底改变了物理学，而且出现了当今人们研究的各个不同的物理领域。但是，开尔文爵士的另一句更为深刻的话却千真万确，“科学财富的积累是遵循复利定律的”，以它来形容这些革命性的进步尤为恰当。

本章将探讨引力科学，看它如何从牛顿定律的伟大成就演进到爱因斯坦引力论的革命性进步。牛顿定律是经典的物理定律，几个世纪以来科学家们一直使用它计算机械运动，包括由引力促进的运动。牛顿定律非常伟大，它使我们对于各种运动的预言异常精确，精确到足以将人类送上月球，将卫星送入轨道，让欧洲的高速列车在转弯时不致脱轨，激发人们在发现天王星轨道异常时去寻找第八颗行星——海王星；但是可惜它不足以让GPS系统精确无误。

我们现在应用的GPS系统，正是因为爱因斯坦的广义相对论才使精确度达到了1米之内，这实在令人难以置信；更让人不可思议的是，目前为了确定火星上积雪深度的变化，我们使用的是来自绕火星运转的宇宙飞船的激光扫描数据，如果将广义相对论计算在内，产生的数值精确度可在10厘米之内。当然，在相对论创建初期，任何人——甚至包括爱因斯坦自己都没有料到，这样一个抽象的理论会得到如此实际的应用。

本章将探索爱因斯坦的相对论，这个在众多领域得到广泛应用的、异常精确的理论。我们将首先简要回顾牛顿的引力论，它在我们日常见到的能量和速度中还是颇为有效的；然后，我们将转到它失效的极限，即极高速（接近光速）和极大质量或极高能量的区域。在这些极限区域里，牛顿引力被爱因斯坦的广义相对论所超越。有了爱因斯坦的广义相对论，空间（以及时空）由一个静止的舞台演变成了一个能够移动和弯曲的、有着自己丰富生命的动态实体。我们将探讨这一理论，介绍它的起源，以及让科学家们相信它的实验验证。

牛顿的万有引力王国

引力就是让我们站在地面上的力，也是你把球抛向空中时使它加速返回的那种力。16世纪末，伽利略证明，这种加速对于地球表面的所有物体（无论其质量大小）都是同等的。

但这种加速取决于物体离地球核心有多远。通常来讲，引力强度取决于两物质之间的距离——两物体相隔越远，其间的引力就越弱。而且如果产生引力吸引的不是地球，而是其他物体，那么引力强度就要取决于那个物体的质量。

牛顿创建了万有引力定律，此定律总结了引力对质量和距离的依赖关系。



牛顿定律认为，两个物体之间的引力与它们各自的质量成正比，这两个物体可以是任何东西：如地球和月球，太阳与木星，篮球和足球，或你喜欢的任何东西。物体质量越大，两者之间的引力作用越强。

牛顿的万有引力定律还指出，引力取决于两个物体之间的距离：两个物体之间的引力与其距离的平方成反比。这种平方反比定律正是源于那个著名的苹果，因为苹果就在地球表面，牛顿能够推算出苹果由于地球引力而产生的加速度，然后将它与月球被吸引产生的加速度相比。月球离地球中心的距离，要比地球表面离中心的距离远60倍，因此，月亮由于地球引力作用而得到的加速度是苹果的加速度的 $\frac{1}{3600}$ （3600是60的平方），这种引力的减弱符合离地球中心的距离的平方。

然而，即使知道了引力强度取决于质量和距离，我们仍需要一个信息才能确定总的引力强度。这一缺失的信息是一个数值，被称作牛顿万有引力常数，也就是计算经典引力时的一个因子，引力作用与它成正比。引力非常微弱，这在牛顿常数的极小值上也得到了反映。

地球引力或太阳与行星之间的引力，看起来似乎非常大，这只不过是因为地球、太阳和行星的质量非常庞大。因为牛顿常数很小，在基本粒子之间的引力也就极其微弱。引力为何极其微弱，这本身就是一个极大的谜题。

尽管牛顿定律正确，可直到20年以后的1687年，他才将它发表，他试图找到理由来证明自己理论里的一个关键假设：地球施加引力的方式好像是将其庞大的质量全部集中在地球核心上。而就在牛顿致力于发展微积分以解决这一问题时，约翰尼斯·开普勒经测量发现，行星的运行轨道是椭圆的，埃德蒙·哈雷、克里斯托弗·雷恩（Christo

pher Wren)、罗伯特·胡克以及牛顿自己通过分析行星的运动，都取得了重大进展，从而确立了万有引力定律。

这些人对行星运动问题都作出了重大贡献，而牛顿则因平方反比定律集盛誉于一身。这是因为，牛顿最终证明了只有在平方反比定律成立时，作为一个核心力量（太阳引力）的结果，才会产生椭圆轨道；而他还以微积分证实，球体的质量似乎真的是集中在其核心。但是，牛顿还是对其他人的重要贡献表示了感谢：“如果说我看得更远些，那是因为我站在巨人的肩膀上。”（但也有传闻说，牛顿之所以这样说，是因为他很反感身材矮小的胡克。）

在高中物理课上，我们学到了牛顿定律，并对这种有趣系统的活动做了计算。可我的老师鲍梅尔先生却告诉我们，刚学的万有引力定律是错误的。我仍记得当时有多么生气：为什么明知它错误还要教给我们？以我在高中时代对这世界的观点，科学的最大优点就在于它总是正确可信的，并能作出准确和现实的预言。

兴许鲍梅尔先生是为了追求轰动效果吧，他显然将对于万有引力定律的结论简单化了。牛顿定律没有错，只不过它是一种近似，一种在大多数场合效果很好的近似。对于大量的参数而言（如速度、距离、质量等等），它非常正确地预言了引力。更为精确的基本理论是相对论，而只有当你在处理极高速度、极大质量或极大能量时，相对论才会作出可以观测的不同预言。牛顿定律对一个小球运动的预言精确得令人钦佩，这是因为它不受以上条件的约束。而使用相对论去预言一个小球的运动，则纯属自找麻烦。事实上，爱因斯坦自己最初以为狭义相对论不过是牛顿定律的改进，而没想到会是彻底颠覆科学范式。当然，他也过分低估了他工作的终极意义。

狭义相对论，颠覆时空观念

物理定律最为合理的一点就是，对于任何人它都是适用的，如果人们在不同国度，或坐在以不同速度行驶的列车上，或正乘飞机旅行，却体验到了不同的物理定律，那么肯定没有人会因此责备我们质疑物理定律有效性和实用性的做法。物理定律应是基本的，而且对所有观察者都成立。计算的任何差异都应因环境而异，而不是因物理定律所导致。事实上，如果一个普适的物理定律要依赖于某个特定事物，倒是非常奇怪的。你要测量的特定量也许会依赖于你的参照系，但制约这些量的定律是不会改变的。爱因斯坦对狭义相对论的阐释便证实了这一点。

爱因斯坦对引力的研究最终被称作“相对的理论”，实际上，这确实有点儿讽刺意味。促使狭义相对论和广义相对论产生的根本原因是，物理定律对所有人都是适用的，无论他处于何种参照系。事实上，爱因斯坦倒是喜欢“不变性理论”这一术语。一个记者曾建议他重新考虑它的名称，在给这个记者的回信里，他承认“相对论”这个词的确很不幸，但那时，这个名称早已深深烙在了他的大脑里。

爱因斯坦最早关于参照系和相对性的见解来自他对电磁学的思考。电磁学理论自19世纪就已被人们所熟悉，其基础是麦克斯韦定律，该理论描述了电场、磁场的活动和电磁波，给出了正确结果，但最初所有人都以“以太”的运动来错误地解释预言。

以太

以太是一种假想的不可见物质，电磁波被当作它的振动。爱因斯坦意识到，如果有一个以太，那就也应该有一个更有利的观察角度或参照系，这是因为以太在这个参照系中是静止的。

以太是一种假想的不可见物质，电磁波被当作它的振动。爱因斯坦意识到，如果有一个以太，那就也应有一个更有利的观察角度或参照系，这是因为以太在这个参照系中是静止的。他的理由是，同样的物理定律对匀速^[9]运动的人都应是适用的，匀速可以是相对于彼此的

运动，也可以是相对于静止的人——即在一个被物理学称为惯性系的参照系里。为了要求所有物理定律（包括电磁学理论）对所有惯性参照系内的观察者都成立，爱因斯坦放弃了以太的观点，而最终创立了狭义相对论。

爱因斯坦的狭义相对论彻底改变了传统的时空观念，是一次巨大的理论飞跃。物理学家、科学史学家彼特·加里森（Peter Galison）提到，将爱因斯坦推上正确轨道的，不仅仅是以太理论，还有爱因斯坦当时的工作。加里森之所以这样说的理由是，生长于德国，而当时在瑞士伯尔尼一家专利局工作的爱因斯坦，对时间和时间协调一定有着很强的观念。所有去过欧洲旅行的人都知道，那些国家，如瑞士和德国，都非常看重精准，这让在那里旅行的乘客有着非常愉快的体验——他们完全不用担心火车会延误。在1902—1905年间，爱因斯坦在这家专利局工作，那时乘火车旅行已变得越来越频繁，而协调时间则成了尖端的新技术。20世纪早期，爱因斯坦很可能想到了现实世界的问题，例如该怎样协调在这一车站与下一车站之间的时间。

当然，爱因斯坦没必要为了解决现实的时间协调问题而发明相对论，（对于已经习惯了火车常常误点的美国人来说，协调时间无论怎样听上去都还有点新奇^[10]）。但协调时间引出了有趣的问题：时间协调对相对行驶的火车来说，可不是一件容易的事。如果我要与一个人对表，而这个人正在一列行驶的火车上，那我就需要考虑信号在我们之间传递所延误的时间，因为光速是一定的，而与静止坐在我身边的一个人对表和与远处的一个人对表又是不同的。^[11]

将爱因斯坦引向狭义相对论的一个关键发现是，我们必须重建时间概念。根据爱因斯坦的理论，时间和空间不应再被人们孤立地来看待，虽然它们不是一回事——时间和空间显然是不同的，但你测量的量却依赖于你的旅行速度，狭义相对论就是这一见解的结果。虽然很

离奇，但人们可以由两个假设得出爱因斯坦狭义相对论的所有新奇结果。要陈述它们，我们需要懂得惯性系（参照系的一个特定类型）的含义。首先，我们选择匀速（速度和方向）运动的任一参照系，一个静止的就很好。惯性参照系就是指相对于第一个参照系以固定速度移动的参照系，比如一个以恒定速度从旁边跑过或驾车驶过的人。

爱因斯坦的假设指出：

- 物理定律在所有惯性系都是不变的；
- 光速在所有惯性系都是不变的。

这两个假设告诉我们：牛顿定律是不完善的。一旦接受了爱因斯坦的假设，我们就别无选择，只能以符合这些规则的、更新的物理定律来取代牛顿定律。随后的狭义相对论导致了所有你可能听说过的令人瞩目的结果，如时间膨胀、观察者对同时性的依赖、移动物体的洛仑兹收缩。当在以相对光速很慢的速度运行的物体上应用时，新的定律看上去与旧的经典定律非常相像，但当在高速（以光速或接近光速）运行的物体上应用时，牛顿定律和狭义相对论的解释是有很明显的差异的。例如，在牛顿力学中，速度只是简单相加。在高速路上一辆迎面向你驶来的车，驶近你的速度就是你们两车速度之和。同样的道理，如果你在行驶的火车上，一个人从站台上向你扔一个球，那么那个球的速度应该是球本身的速度加上火车的速度（我以前有一个学生维泰克·斯奇巴可以证实这一点：有人向正在靠近的列车投了一个球，车上的维泰克恰巧被球打中，他差点没被击昏过去）。

根据牛顿定律，你在一列行驶的火车上，所看到的迎面向你射来的一束光的速度，应该是光速与火车行驶的速度之和。但如果像爱因斯坦第二条假设所说的那样，光速不变，这就不对了。如果光速总保持不变，那么，你在行驶的火车上时一束向你射来的光的速度，与

你静静地站在地面上时一束向你射来的光的速度都是一样的。虽然这与你的直觉有些相悖，因为你的直觉都是来自日常生活中所遇见的低速，但光速的确是不变的，而且，在狭义相对论里，速度不像在牛顿定律中那样只是简单地相加，相反，速度的相加需要遵循由爱因斯坦假设所得出的相对论公式。

狭义相对论的许多含义并不符合我们习惯的时间和空间观念。狭义相对论与以前牛顿力学对时间和空间的看法是不同的，正因如此，才产生了许多有悖于直觉的结果。时空的测量要取决于速度，而且在相对于彼此运动的系统里融合在一起。但是，尽管它们让人惊讶，可一旦你接受了这两个假设，那么不同的时空观念就是其必然结果。



我们来看为什么会这样：假设有两艘相同的船，有着相同的桅杆，一艘船停泊在岸边，另一艘正在驶离。再假设两位船长在第一艘船启航时已对好了表。

现在假设两位船长要做一件非常奇怪的事：两人同时决定在他们各自的船上测量时间。他们在船的桅杆顶部和底部各放置一面镜子，然后将光由底部的镜子照向顶部的镜子，以此测量光在两面镜子之间往返的次数。当然，从现实角度来讲，这确实有点荒谬，因为光往返的频率实在是太快了，根本无法计量。但是，请耐住性子，就让我们假设两位船长的计数非常快，因为我要用这

个有点缺乏真实性的例子来说明，时间在行驶的船上被拉长了。

如果两位船长都知道光往返一次要花多少时间，那么，用光往返一次的时间乘以光在两面镜子之间往返的次数，就能算出时间的长度。但是，现在假设停泊船只的船长不用他自己静止的镜子钟，而是以行驶船只上光在桅杆两端镜子之间往返的次数来测量时间。

从行驶船只船长的角度来看，光只是上下穿梭；而从停泊船只船长的角度来看，光就必须旅行得更远一些（才能走完行驶船只走过的距离，见图5-1）。但光速是不变的——这正是有悖于直觉的地方，无论是射向停泊船只的桅杆顶部，还是射向行驶船只的桅杆顶部，光速都是一样的。那么，行驶的镜子钟则必须“滴答”得慢一些，以此来弥补行驶船只上光旅行的更长的距离。

这个与直觉相悖的结论（行驶船只与静止船只上的钟必须以不同的速率“滴答”）所遵循的事实是：在一个移动参照系里的光速和静止参照系里的光速是相同的。尽管以这种方法测量时间很可笑，但无论以何种方式测量，同样的结论——移动的钟表走得要慢一些

都会成立。如果船长戴着表，他们会观察到同样的事实（再次提请注意，对通常速度来说，这种效应非常微小）。

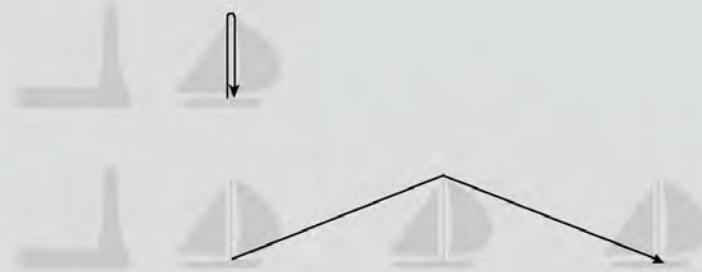


图5-1 光束在静止船只上和行驶船只上从桅杆顶部反射下来的路径。在静止的观察者（在岸边停泊的船上或是在灯塔里）看来，行驶船只上光束走过的路径会更长。

尽管上面的例子不真实，但所描述的现象真正产生了可测量的结果。例如，狭义相对论使快速运动的物体经历了不同的时间，这种现象叫作时间膨胀。

物理学家在研究对撞机或大气产生的基本粒子时，测量了时间膨胀。这些粒子以相对论速度（接近光速穿行），例如，称作 μ 子的基本粒子与电子有相同的电荷，但它更重，而且会衰变（即它会转变为其他更轻的粒子）。

μ 子的寿命，即在其衰变之前的时间，只有2微秒。如果一个运动的 μ 子与一个静止的 μ 子有着同样长的寿命，那么在它消失之前，就只能穿行600米。但 μ 子却成功穿过了大气，在对撞机里，直达大型探测仪的边缘。因为它接近光速的高速运动，使其寿命在我们看来要长许多。在大气中， μ 子穿行的距离比在基于牛顿定律的宇宙里穿行的

距离至少要远10倍。我们能看到 μ 子，这一事实本身就证实了时间膨胀（及狭义相对论）产生了真切的物理效应。

狭义相对论之所以重要，既因为它与经典物理学有着巨大差异，也因为它是广义相对论和量子场论发展的基础，而这两者在最近的研究进展中都发挥了重要作用。因为我在后面探讨粒子物理学和额外维度模型时，不会用到狭义相对论预言，所以，我只好遏制自己的欲望，不再去细讲狭义相对论出神入化的结果，例如为什么同时性要取决于观察者是否运动，以及运动物体的大小与它们静止时有什么不同。相反，我将细述另一重大进展，即广义相对论，这在我们以后讨论弦理论和额外维度时是非常关键的。

等效原理，广义相对论的开始

1905年，爱因斯坦完成了他的狭义相对论。1907年，当他打算写一篇论文总结他关于这一课题的最新研究时发现，自己已经在质疑这一理论是否适用于所有场合了。他注意到两处重要的疏忽，其中一处是，物理定律只在某些特定的惯性参照系里才是同样的，这些参照系以固定速度相对于彼此运动。

在狭义相对论里，这些惯性系占据了一个优势地位，该理论没有考虑任何正在加速的参照系。开车时，如果你踩了油门，那么，你所在的参照系将不再是狭义相对论所适用的特定参照系。这就是狭义相对论的“狭义”之处：“狭义”的惯性系只是所有可能参照系的一小部分。对那些相信没有特定参照系的人来说，该理论只突出惯性参照系问题就表现了出来。

爱因斯坦的第二处疑虑与引力有关。虽然他指出了物体在某些场合是怎样回应引力的，但他起初并没有想出确定引力场的公式。在某

些简单场合，引力场定律已被熟知，但爱因斯坦还不能够对所有可能的物质分布导出引力场。

1905—1915年，经过一番艰苦卓绝的探索，爱因斯坦解答了这一问题，其结果便是广义相对论。

新的理论围绕着等效原理展开。等效原理指出，加速度的作用与引力的作用是不可区分的。所有物理定律在一个正在加速的观察者和一个处在引力场中静止的观察者看来都应该是相同的：引力场使所有在静止参照系内的物体加速，与原来的观察者的加速度相等，只不过方向相反。换句话说，你没有办法辨别你是正在稳定地加速还是静止地站在引力场中。根据等效原理，没有办法能够将这两种场合区分开来，观察者也永远无从知道自己是处于哪种场合。

等效原理源自惯性质量和引力质量的相等。原则上讲，这本应是两个互不相同的量。惯性质量决定一个物体如何对外力作出回应，即如果施以外力，物体会如何加速。惯性质量的作用在牛顿第二定律里也有总结： $F=ma$ ，即如果你对一个质量为 m 的物体施以作用力 F ，产生的加速度则为 a 。牛顿第二定律告诉我们，同样的力如果施加在一个更大质量的物体上，则会产生更小的加速度。这在经验上是我们非常熟悉的：如果你以同样的力去推一个小板凳和一架钢琴，小板凳肯定会被推出去更远，且滑得更快。注意，这一定律适用于任何作用力，如电磁力。在任何与地球引力无关的场合，它都是适用的。

等效原理

等效原理指出，加速度的作用与引力的作用是不可区分的。等效原理源自惯性质量和引力质量的相等。原则上讲，这本应是两个互不相同的量。

而另一方面，引力质量遵循引力定律，并决定引力强度。正如我们所看到的，牛顿引力的强度与相互吸引的两个物体的质量成正比，这些质量即引力质量。引力质量与惯性质量同时遵循牛顿第二定律，结果殊途同归，也因此，我们给了它们同样的名称：质量。但从原则

上讲，它们应是不同的，我们应称其中一个为“质量”，而另一个为“重量”。好在我们不必这么做。

两种质量相同这一神秘事实有着深刻的含义，它促使爱因斯坦去辨认、去开发。

引力定律指出引力强度与质量成正比；而牛顿定律又告诉我们用这种（或其他）力会产生多大的加速度。因为引力强度与质量成正比，而这质量还决定其加速度，那么这两个定律合在一起就告诉我们，即便根据公式 $F=ma$ ，力要取决于质量，而由引力导致的加速度则完全可以不考虑正在加速的质量。

对于距一个物体同等距离的所有物体或人来说，它们所经受的引力加速度必然是一样的。这正是伽利略所证实的断言，从比萨斜塔上坠落的两个物体，地球引力对其产生的加速度都是相同的，与它们的质量无关。加速度的大小与被加速的物体质量无关，这一事实是引力所独有的特征，因为除引力外，所有其他力都要依赖于质量。因为引力定律和牛顿定律以同样的方式依赖质量，所以计算加速度时，质量就被抵消了，因此，加速度也就与质量无关了。

这一推理相对比较简单，但其含义却颇为深刻。因为在同样不变的引力场中，所有的物体都有着相同的加速度，如果这唯一的加速度能够被抵消，那么引力的证据也就被同时抵消了。一个自由落体恰恰正是这样：它的加速度恰好抵消了重力。

等效原理告诉我们，如果你和周围的物体一同自由落下，那么，你根本意识不到还有一个引力场，你的加速度会抵消引力场对静止时的你产生的重力加速度。从在轨道中运行的宇宙飞船图片里，我们已熟悉了这种失重状态。在宇宙飞船中，宇航员和他周围的物体都感受不到重力。



教科书里常常用一幅图来说明引力失效的后果（从一个自由落体的观察者角度来看，见图5-2）。图中画的是，一个人在一部自由下落的电梯里抛掷小球。在图中，你看到的是人和球同时下落，电梯里的人看见球总是在离电梯地面的相同高度，他看不到球的下落。

在物理课本中所呈现的自由下落的电梯，好像是这世界上再自然不过的一件事：里面的观察者静静地看着并不下落的球，镇定自若，丝毫不担心自己的个人安危。而在电影里电梯缆绳若被割断，演员坠向地面，我们看到的则是一幅惊悚的场景，这形成了鲜明的对比。为什么会有如此截然不同的反应？如果所有东西都在自由下落，原也没有什么可惊恐的：这情形就和所有东西都是静止的没有什么区别，只不过是在一个零引力的环境里。但是，如果像在电影里那样，一个人在自由下落，而下面的土地却仍是坚挺不动的，那他被吓呆便是理所当然的了。如果一个人在一个自由下落的电梯里，而下面等待这降落的是坚实的地面，那么，在他的自由落体运动终结时，他肯定会意识到引力是多么好（见图5-2最后一幅所示）。

爱因斯坦的结论看上去是如此地令人惊讶和离奇，这是因为，我们生长在地球上，脚下就是静静的大地，这让我们的直觉产生了谬误。当地球引力让我们静止于地面时，我们能注意到引力的作用，是因为我们不会沿着重力的方向被直接吸进地核中去。在地球上，我们已习惯了这种使物体坠落的引力，但“坠落”真正的意思是“相对于我们的坠落”。如果我们与一只下落的球同时坠落，就像我们在自由下落的电梯里那样，球也并不会比我们下降得更快，因此我们也就看不到它的下落。

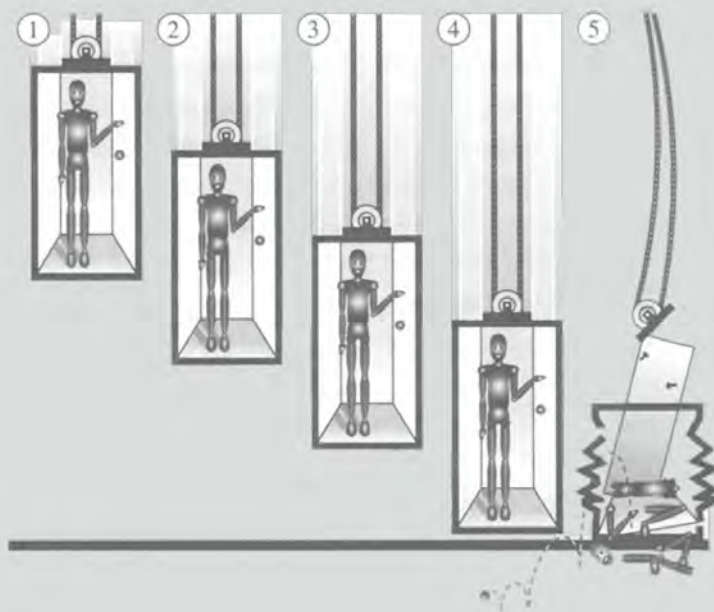


图5-2 一个人在自由下落的电梯里抛掷小球。在电梯里，观察者虽然松开了球，但他看到的球并不下落。但是，当自由落下的电梯撞到静止的地面时，里面的观察者可就不会太高兴了。

在这一自由落体的参照系里，所有的物理定律都会与你和周遭的一切都静止时所遵循的物理定律完全一致。自由落下的观察者观察到的运动，应以狭义相对论里的方程来描述，即对在惯性的、不加速的参照系里的观察者适用同样的方程。1907年，爱因斯坦在回顾相对论的一篇论文里解释了引力场的存在为何是相对的：“因为对于一个从屋顶自由坠落的观察者来说，至少在其邻近的周围，是不存在引力场的。”这就是爱因斯坦的重要见解。

自由下落的观察者的运动方程也就是惯性系里观察者的运动方程，自由下落的观察者感觉不到重力——只有那些非自由下落的物体才会受重力影响。

在日常生活中，我们通常并不会遇上自由下落的东西或人。当自由落体发生时，常常都是惊悚、危险的，但正如在参观爱尔兰的莫赫尔悬崖时，一个爱尔兰人对物理学家拉斐尔·布索（Raphael Bouso）所说的：“杀死你的并非下落，而是你下落终止时的碰撞！”有一次，我在攀岩中摔折了几根骨头，被迫缺席了我组织的一场会议。这时，就有很多人来取笑我，说我是在验证引力理论，而我可以信心十足地宣布，引力加速度与预言是完全一致的。

两大实验，验证广义相对论

广义相对论还有很多内容，很快我们会讨论到其他部分——它们耗费了科学家更长的时间。可是，仅等效原理就解释了广义相对论的许多结果。一旦爱因斯坦认识到在加速参照系里引力可以被排除，他便可以通过假设一个加速系统等同于一个有引力的系统而算出引力的作用。这使得他能够计算一些有趣系统里的引力作用，而别人则以这些系统来检验他的结论。现在，我们来探讨几个最为重要的实验验证。

第一个是光的引力红移，红移使我们探测的光波频率低于其发出的频率。你可能遇到过类似的效应，比如，当一辆摩托车从你身边呼啸而过时，你听到的声音先是升高，而后降低。

有很多方法可以帮助我们理解引力红移的起因，但最简单的可能还是打比方。设想你向空中抛一个球，球上升的速度会因重力的作用减慢下来，但是，尽管球速减慢了，球的能量并未损失，它已转化为势能，然后，当球下落时，作为动能被释放出来。

同样的推理也适用于光的粒子——光子。就如一个球被抛向空中失去了冲力一样，当光子由引力场逃逸时也会失去动力。这意味着光子和球一样，在挣脱引力场时失去了动能但获得了势能，但是光子不会像球那样慢下来，因为光速是不变的。抢先说一下，下一章我们会看到量子力学的一个结果，光子的频率下降时，其能量就会下降。一个穿过变化的引力场的光子所经历的正是这样：为了降低其能量，光子必须降低频率。这种频率的降低就是引力的红移。

反过来说，一个正朝着引力源方向运动的光子，其频率就会增加。1965年，加拿大物理学家罗伯特·庞德（Robert Pound）和他的学生格伦·雷布卡（Glen Rebka）通过研究由伽马射线辐射源放射出的伽马射线测量了这种效应。他们把这个伽马射线辐射装置放在哈佛大学杰弗逊物理实验室的“塔”顶上，现在我就在这里工作（尽管它是整座建筑的一部分，但杰弗逊实验室高耸的阁楼连同它下面的几层被统称为“塔”）。

塔顶和塔底的引力场会有轻微的差异，因为塔顶离地球中心会稍远一些。测量这种差别最好是能找个高塔，那样会使伽马射线的发射（塔顶部）和探测（地下室）之间的高度差达到最大。但即使塔只有三层，加一个尖顶，再加上塔尖上往下俯瞰的窗户，一共22.5米高，庞德和雷布卡还是测量出了被放射和吸收的光子之间的频率差异，精

确得令人难以置信：一千万亿分之五而已。他们因此确认，广义相对论对引力红移的预言可达到1%的精度。

等效原理第二个实验观察结果是光的偏折。引力既能吸引能量也能吸引质量，毕竟，著名的 $E=mc^2$ 公式意味着能量与质量是密切相关的：如果质量经受引力，那么能量也一样。太阳的引力影响质量，同样也会影响光的轨迹。爱因斯坦的理论预言，光在太阳的影响下会产生偏折，并产生一定的偏折。这一预言在1919年的日全食时得到了验证。

英国科学家阿瑟·爱丁顿（Arthur Eddington）组织了两支探险队，一支到西非海岸的普林西比岛，一支到巴西的索布拉尔——那是观测日食的最佳地点。他们的目的是拍摄太阳周围的恒星，以检验看似离太阳很近的恒星相对于它们通常的位置而言是否会产生偏移。如果恒星看上去真的移动了，那就意味着光是沿着弯曲的轨迹运行的（科学家们需要在日全食时进行测量，这样就不会因为太阳过强的光芒掩盖了黯淡的星光）。可以肯定的是，恒星恰好出现在那个“错误的”位置上，测到的正确的偏折角度提供了支持爱因斯坦广义相对论强有力的证据。

令人难以置信的是，光的偏折已被明确确立并理解，现在已成了人们的探索工具，用于探索宇宙的质量分布，并在一些已燃烧殆尽、不再发光的小恒星里寻找暗物质，这些物质就像是黑夜里的黑猫，是很难被发现的。观测它们唯一的方式就是通过引力作用。

引力透镜是天文学家了解暗物质的一种方式。暗物质与所有其他东西一样，通过引力相互作用。尽管燃尽的恒星自身已不发光，但在它们身后（从我们的角度来看）可能会有亮的物质，它们发出的光是我们能够看到的。如果光的路途中没有暗恒星，那么光就会直射过来，但是，一颗明亮恒星发出的光如果要经过暗恒星，就会发生偏折。由左侧经过的光线会比从右侧经过的光线更明显地折向相反方向；而从顶部经过的光线会比从底部经过的光线更明显地向对侧偏折。这样，在暗物质后面的一个明亮星体就会产生一个多重影像，这

种作用就被称作引力透镜。图5-3描绘出的是，当恒星发出的光线因中间经过的大质量物体而向不同方向偏折时，所产生的多重影像。

引力透镜

引力透镜是天文学家了解暗物质的一种方式。暗物质与所有其他东西一样，通过引力相互作用。尽管燃尽的恒星自身已不发光，但在它们身后（从我们的角度来看）可能会有亮的物质，它们发出的光是我们能够看到的。

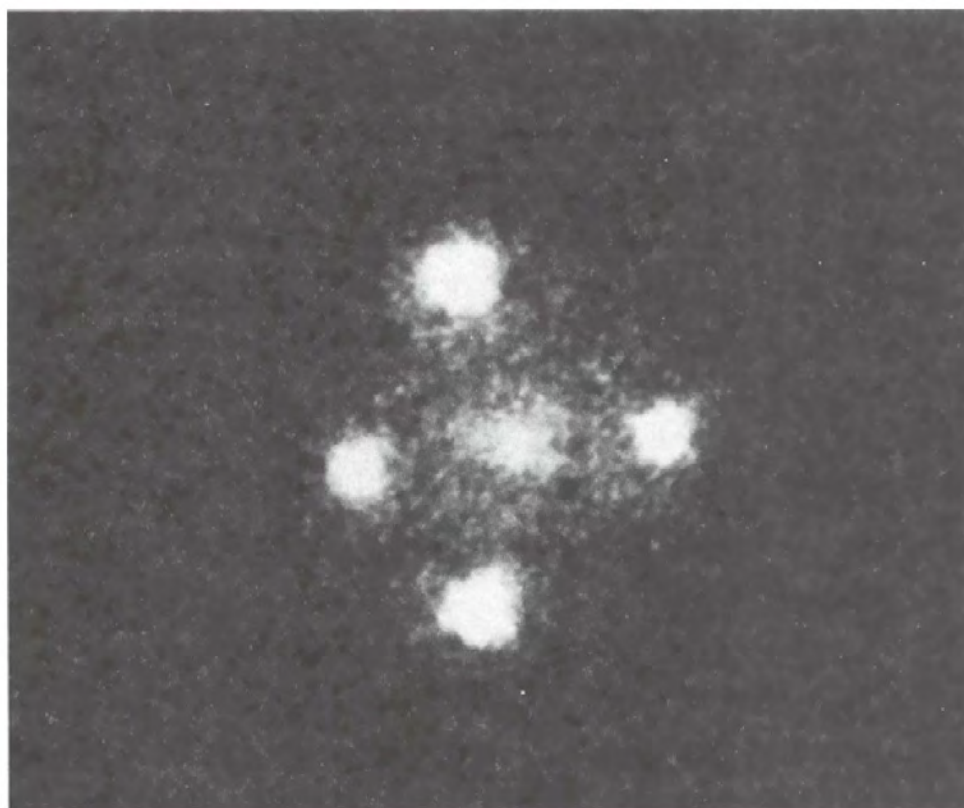


图5-3 类星体的多重影像。一个遥远、明亮的类星体在经过中间一个大质量星云时，光线会折向不同方向，从而形成“爱因斯坦十字”的多重影像。

一切源于宇宙的弯曲

根据等效原理，引力与恒定加速度是不可区分的。很高兴你已明白这一点，因为我必须承认，我将其简化了。毕竟，这两者并非完全

不可区分，那么，区别在哪儿呢？如果引力等于加速度，那么，在地球相对的两个半球上的人就不可能同时落向地球了。地球毕竟不可能同时向两个方向加速。比如，处在中国和美国两个不同方向上的人感受到的地球引力，不可能被看作一个加速度。

解决这一矛盾的方法是等效原理所主张的——引力只能够由同一地点的加速度取代。在空间的不同位置上，可以根据等效原理取代引力的加速度通常也应处在不同方向上。关于中国和美国关系问题的答案就是：与美国引力相等的加速度和能产生中国引力的那个加速度的方向是不同的。

这一重要见解使爱因斯坦对引力论又进行了一次完整的、全新的阐释。爱因斯坦不再将引力看作直接作用于一个物体的力，相反，他把它描述成是时空几何的弯曲，反映了在不同地点抵消引力所要求的不同加速度。时空不再是一个事件发生的空白背景，而是一个主动的参与者。有了爱因斯坦的广义相对论，引力就被理解成时空的弯曲结构，反过来，它又由物质和能量所决定。现在，我们来探讨时空弯曲的概念，爱因斯坦的革命性理论正是立足于此。

弯曲的时空

数学理论必须是内部连贯一致的，但不同于科学理论的是，它不必对外部现实作出回应。当然，数学家们的灵感常常来自对周围世界的观察，立方体、自然数等数学对象也确实能在现实世界里找到对应。但是，数学家们会将这些关于熟悉概念的假设推广为物理现实不那么确定的对象，如超正方体（在四维空间的超立方体）和四元数（一个新奇的数字体系）等。

公元前3世纪，欧几里得写下了有关几何学的5个基本公理，由这些公理产生了一个优美的逻辑结构，在中学时代你可能对此有过简单

的涉猎。但后来数学家们发现第五公理（即著名的平行线公理）很麻烦。这一公理说的是，过线外一点，有且只有一条直线与原直线平行。

欧几里得提出他的公理后两千年来，数学家们一直争论不休，这第五公理究竟是独立于其他四个而存在，还是只是另外四个公理的逻辑延伸？会不会有一种几何体系除了最后一个外其他都成立？如果不存在这样的几何体系，那么第五公理就不是独立的，因此也就无关紧要了。

直到19世纪，数学家们才给第五公理找到了恰当的位置。伟大的德国数学家卡尔·弗雷德里克·高斯（Carl Friedrich Gauss）发现，第五公理正如欧几里得断言的那样，可以被另一公理取代。他继续研究并取代了它，由此发现了另外的几何体系，并证明第五公理是独立的。在此基础之上，非欧几何诞生了。

俄国数学家尼古拉·伊万诺维奇·罗巴切夫斯基（Nikolai Ivanovich Lobachevsky）也发展了非欧几何。可是，当他将自己的研究寄给高斯时却失望地发现，这位老数学家早在50年前就想到了同样的观点，只不过这位德国人因为怕被同事嘲笑，一直隐藏着自己的研究，所以，无论是罗巴切夫斯基还是任何其他人都不会知道高斯的结果。

高斯原本不必这么担忧的，很显然，欧几里得的第五公理并非总能成立，因为我们都知道还会有其他可能。例如，虽然经线在赤道是平行的，但它们在南、北极点会相交。球体上的几何就是非欧几何的一个例子。如果古人是在球体上写字，而不是在羊皮纸上，那么，这点对他们也会是显而易见的。

但是，非欧几何还有许多例子，与球体不同，它们是不能在三维世界里直观地实现的。高斯、罗巴切夫斯基和匈牙利数学家雅诺什·

鲍耶 (János Bolyai) [12] 等人最初的非欧几何，研究的正是这样一些根本不可描绘的几何理论，也难怪他们花了这么久时间才发现它们。

有几个例子可以说明是什么使得弯曲几何与这页上面的平面几何有所不同：图5-4显示了三个二维平面。第一个是球体表面，有稳定的正弯曲；第二个是平面的一部分，零弯曲；第三个是双曲抛物面，有稳定的负弯曲，马鞍的形状、两座山峰之间的地形和一片品客薯片的形状都是负弯曲表面的例子。



图5-4 正弯曲、零弯曲和负弯曲的表面。

很多测试方法可以让我们知道，某个特定的几何空间属于这三种可能形状的哪一种。例如，你可以在这三种形状的表面画一个三角形：在平坦表面上，三角形的内角之和总是 180° 。而如果三角形的顶点在北极，另外两点在赤道上，间距 $1/4$ 赤道长度，这样一个在球体表面的三角形会是什么样？这个三角形的每个角都是 90° 直角，因此，它的内角之和是 270° 。这在一个平坦表面上是永远也不会发生的，但在一个正弯曲的表面上，因为表面突出，三角形的内角之和必然大于 180° 。

同样道理，画在双曲抛物面上的三角形内角之和必然小于 180° ，这是其负弯曲的反映。这不太容易见到：在靠近马鞍两边顶部的位置各画一点，另一点向下，沿着双曲抛物面延伸至底部，如果你骑在马上，就是你一只脚踩的地方。最后这一角比起它在平面上要小，因此三个角加起来小于 180° 。

一旦内在一致地确立了非欧几何，即它的假设不会有任何矛盾和悖论，德国数学家乔治·伯恩哈德·黎曼（Georg Bernhard Riemann）便研究出了一个丰富的体系来描述它们。一张纸不能被卷成一个球体，但可以被卷成一个圆筒；如果不碎裂，不向后折叠，你无法将马鞍摊平。在高斯研究的基础上，黎曼创建了又一数学形式，将这些事实囊括在内。1854年，针对如何通过它们的内禀性质来刻画所有几何，他发现了一个一般性的解决方法。他的研究奠定了现代数学领域微分几何的基础，这一数学分支专门研究曲面和几何。

因为从现在起，我将几乎总是把空间和时间放在一起考虑，所以我们会发现时空的概念比空间更为有用。时空要比空间多一维：除了“上下”“左右”“前后”之外，它还包括时间。1908年，数学家赫尔曼·闵可夫斯基（Hermann Minkowski）使用几何概念发展了这一时空结构的绝对观念。爱因斯坦研究时空所用的时间和空间坐标要依赖于一个参照系，而闵可夫斯基则明确指出一个独立于观察者的时空结构，任何一个给定的物理情境都可以由它表现出来。

在本书其余部分，除非特别说明，我说维度时都指时空维度数。例如，我们看周围时，看到的就是我以后所指的维宇宙。偶尔我会把时间单列出来，说“三维加一维”的宇宙，或三维空间。但请记住，所有这些术语指的都是同一场景——即有三个空间维度和一个时间维度。

时空结构是一个很重要的概念，它简明地表现了一个时空几何的特征，这一时空几何与由特定能量和物质分布所产生的引力场密切呼应。但起初，爱因斯坦并不喜欢这一观念，要对自己已经解释过的物理再作重新阐述，在他看来，这未免太过分了。但他最终还是认识到，这一时空结构对于完整、概括地描述引力和计算引力场是非常重要的。从记录来看，闵可夫斯基在最初认识爱因斯坦时对他也没有什

么好印象：根据以往爱因斯坦学生时代在他微积分课上的表现，闵可夫斯基给爱因斯坦的评价是“一条懒狗”。

拒绝非欧几何的不仅仅是爱因斯坦一个，他的朋友、瑞士数学家马塞尔·格罗斯曼（Marcel Grossmann）也觉得它太复杂了，劝爱因斯坦不要用它。但他们最终一致认为，解释引力唯一可行的方法就是用非欧几何来表现时空结构，直到此时，爱因斯坦才得以诠释和计算弯曲的时空，它与引力相等，而这也成为完成广义相对论的关键。在格罗斯曼承认失误之后，他们两人携手，共同致力于研究错综复杂的微分几何，以简化他们高度复杂的早期尝试，来明确诠释引力理论。终于，他们完成了广义相对论并对引力本身有了一个更为深入的理解。

爱因斯坦的广义相对论

广义相对论提出了一个全新的引力概念：现在我们知道，引力（这种力让我们站在地球上，也让我们的星系和宇宙联系在一起）并非直接作用于物体，而只是时空几何弯曲的结果。这一概念使爱因斯坦有关时空一体的概念形成了一个合理的理论结果。

广义相对论利用了惯性质量和引力质量的深层联系，单纯以时空几何诠释了引力的作用。任何形式的物质和能量分布都会弯曲时空，时空里的弯曲路径决定着引力的运动，而宇宙里的物质和能量又引起时空本身膨胀、波动或是收缩。

在平坦空间里，两点之间的最近距离，即大地线（地球椭球面上连接两点的最短程曲线），是一条直线。在弯曲空间里，我们仍可以把大地线定义为两点间的最短距离，但这条路径看上去未必是直的。比如，飞机沿着地球大圆飞过的航线就是大地线（一个大圆是绕过地

球最粗部分的任一圆，如赤道或一条经线）。虽然这些路径不是直的，但除非你凿穿地球，它们就是地球上两点间的最短距离。

在弯曲的四维时空里，我们仍然可以定义大地线。对于被时间隔开的两件事，大地线就是时空中把一件事与另一件事联系起来的自然路径。爱因斯坦意识到，自由落体就是沿时空的大地线运动，因为这是阻力最小的路径。他由此得出结论：在没有外力的情况下，下落物体会沿着大地线落下，就像一个人在自由下落的电梯里那样，他感觉不到重力，也看不到小球的下落。

但是，即使事物是在沿着大地线穿越时空，且没有外力，但引力还是有其明显作用。我们已经看到了，引力与加速度在同一地方是等效的这一重要见解，导致爱因斯坦发展了一种有关引力的全新的思考方式。他推断出，因为由引力导致的加速度对于同一地方的所有质量相同的物体都是相同的，所以，引力必然是时空本身的一种属性。

这就是为什么自由落体在不同地方意味着不同的事情，而只有在当地，引力才可以由唯一的一个加速度所取代。我和对应的一个在中国的人，即使我们两人都是在当地一个爱因斯坦意义的电梯里，我们仍是在以不同的方向下落。自由落体的方向并非处处都是一样的，这是因为受到了弯曲时空的影响——一个单一的加速度不可能抵消地球上所有地方的引力作用。在弯曲的时空里，不同观察者的大地线是各不相同的，因此，引力在全球范围里都会产生可见的作用。

广义相对论要比牛顿引力深入得多，因为无论能量和质量以何种形式分布，它们所引致的相对性引力场都能够被计算出来。而且，时空几何与引力作用内在联系的发现，使爱因斯坦得以弥补他最初对引力诠释的一大空白。尽管当时物理学家都知道物体会对引力作出何种反应，但他们都不知道引力是什么。现在，他们知道了引力场是由物质和能量引起的时空的弯曲，这一弯曲遍布宇宙各处，或者，还如我

们很快将会看到的，延伸至一个可能包含膜的更高维度的时空里。所有这些情形中更为复杂的引力作用，都会体现在时空表面的褶皱和弯曲里。



图画能最好地描绘物质和能量怎样使时空结构弯曲，以产生引力场。图5-5所示的是一个位于空间里的球体。球体周围的空间被扭曲了：这个球使空间表面陷落，陷落的深度反映出球的质量和能量。另一个球从旁边滚过，它会滚向中心的陷落处，也就是质量集中的地方。根据广义相对论，时空就是以这样一种类似的方式弯曲。同时，从旁边滚过的另一球会加速滚向这球的中心。在这种情况下，其结果正好符合牛顿定律的预言。但它们对运动的诠释和计算却极为不同。根据广义相对论，小球是随着时空表面的波动，由此表现出由引力场导致的运动。

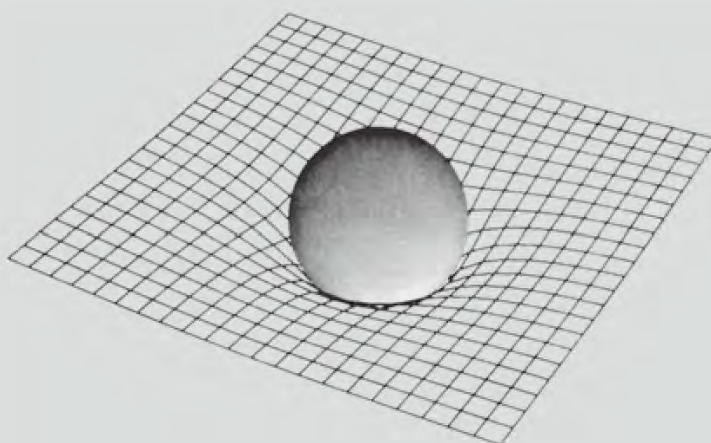


图5-5 一个位于空间的球体。一个有质量的物体使周围空间产生弯曲，由此产生了引力场。

图5-5可能会让人产生误解，因此我要提醒大家注意几件事：

第一，图中球周围的空间是两维的，而实际上，被弯曲的是整个三维空间和整个四维时空，时间也会弯曲，因为从狭义相对论和广义相对论的视角看，它也是一维。例如，狭义相对论告诉我们，在不同地方时钟会以不同速率“滴答”，这就是时间的弯曲。第二，在弯曲几何里，绕着第一个球滚动的第二个球同样会影响时空的几何形状，我们假设它的质量相比第一个大球太小，因而忽略这一微弱作用。第三，使时空产生弯曲的物体可以是任何维度的。以后，在这类图画里，我们会以一个膜来代替这个球体的作用。

但是，在所有情况下，物质决定时空如何弯曲，而时空决定物质如何运动；在没有外力的情况下，弯曲时空确定了大地线，物体将沿着它行进；引力是体现在时空的几何形状里的。爱因斯坦花费了近10年时间推出这一时空和引力的精确关系，并将引力场的作用也涵盖其中——毕竟，引力场也有能量，因此也会使时空弯曲。爱因斯坦的付出是英勇的、伟大的。

爱因斯坦在他的著名方程里明确了怎样找到宇宙的引力场。尽管他最为著名的方程是 $E=mc^2$ ，但物理学家称作“爱因斯坦方程”的却是他确定引力场的方程。这一方程的背后是艰苦卓绝的付出，它告诉我们怎样通过一个已知的物质分布确定一个时空度规。你计算的度规决

定了时空几何的形状，它会告诉你怎样把一些与任意尺度相关联的数转化成确定几何的物理距离和形状。

随着广义相对论的最终建立，物理学家可以确定引力场并计算它的作用。与使用以前的引力公式一样，物理学家使用这些方程算出物质在一个已知引力场中是怎样运动的。例如，把地球或太阳等庞大球体的质量与位置代入方程，就可以计算出著名的牛顿的万有引力。在这个特殊例子中，结果没有改变，但意义大不相同。物质和能量弯曲了时空，而这种弯曲产生了引力。但广义相对论更大的优势在于：它的物质和能量分布包含了所有形式的能量——包括引力场自身的能量。即使在某些场合，引力本身是最为主要的能量，广义相对论仍然非常有用。

爱因斯坦方程适用于任何形式的能量分布，因此改变了宇宙学家，即宇宙历史学家的看法。现在，科学家如果知道了宇宙的物质和能量分布，就能计算出宇宙的演变。在一个空白空间里，空间是完全平坦的，没有任何褶皱和波纹——没有任何弯曲。然而，如果宇宙充满了能量和物质，那么它们就会弯曲时空，演化出各种有趣的、可能的宇宙结构和行为。

我们生活的宇宙绝非静止不变：正如我们很快会看到的：我们还可能生活在一个弯曲的五维宇宙里。值得庆幸的是，广义相对论告诉了我们怎样计算它们的结果。就像我们二维几何的例子中有正、负、零弯曲一样，四维时空几何结构也有正、负、零弯曲，它们是物质和能量恰当分布的结果。以后我们探讨宇宙学和额外维度里的膜时，由物质和能量引起的时空弯曲非常重要。这些物质和能量既存在于我们的可见宇宙中，也存在于膜上和体里。我们会看到三种类型的时空弯曲（正、零、负）在更高的维度上同样能够实现。

广义相对论的许多结果是牛顿引力所计算不出的，除了其他众多的优点之外，广义相对论还消除了牛顿引力让人心烦的超距作用（action-at-a-distance）：物体一出现或运动，它所产生的引力作用便处处都能感觉到。而用广义相对论，我们知道时空先弯曲，然后才有引力。这一过程并非立即产生的，它需要时间。引力波是以光速穿行的，只有在一个信号花费一定时间、到达某一地点并使时空产生弯曲之后，引力才会在这一定位置产生作用。这怎么也不可能比光到达那里发生得更早，因为光速是我们已知的最快速度。比如，你接收到一个无线电信号或是手机信号的时间，怎么也不会早过一束光到达你那里的时间。

而且，物理学家还能以爱因斯坦方程探索其他类型的引力场。可以用广义相对论描述和研究黑洞，当质量高度集中于一个很小的体积时，就形成了这种吸引人的、谜一样的物质。在黑洞里，空间几何被极度弯曲，以至于进入黑洞的任何东西都被困在里面，即使光也不能逃逸。尽管在广义相对论创建之后，德国数学家、天文学家卡尔·史瓦西几乎立刻便算出了爱因斯坦方程的解，从而发现了黑洞^[13]，但直到20世纪60年代，物理学家才真正开始重视这一观点——黑洞有可能是我们宇宙里真实的东西。如今，黑洞已在天文物理领域被广泛接受，事实上，似乎在每个星系的中心，包括我们自己的星系，都存在一个超大质量黑洞。而且，如果存在隐藏的宇宙维度，那就一定存在更高维度的黑洞，如果它们很大，就像天文学家观察到的四维黑洞一模一样。

奇迹总发生在最不可能的地方

让我们总结一下GPS系统的故事：如果我们计算的位置要精确至1米之内，那么，对时间的测量必须精确至 $1/10^{13}$ 。要达到这一精度，唯一可能的办法是使用原子钟。

但是，即使我们有了完美的时钟，时间膨胀还会使它们以 $1/10^{10}$ 的速度减慢。这一误差如果得不到纠正，到我们理想的GPS系统就会放大1 000倍。我们还必须考虑引力蓝移，这也是广义相对论的作用，由光子穿越变化的引力场引起，它所产生的误差至少也是这么大。这一作用加上其他的广义相对论偏差，累积产生的误差，如果被忽略，最终会达到每天误差16公里。艾克（以及现下的GPS系统）必须纠正这些相对论作用的误差。

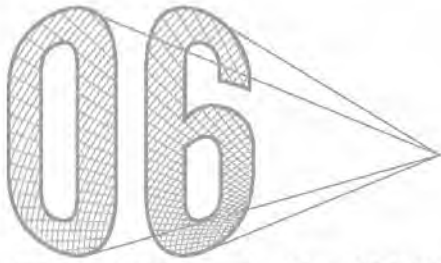
尽管相对论现在已得到了验证，还产生了现代实用仪器必须考虑在内的效应，但在我看来，最初竟有人听信爱因斯坦，也实在令人惊讶。他没有一点儿名气，因为找不到更好的工作才在伯尔尼一家专利局做小职员。他就是在这样一个最不可能的地方，提出了一个有悖于当时所有物理学家信念的理论。

哈佛大学的科学史学家杰拉尔德·霍尔顿（Gerald Holton）告诉我，德国物理学家马克斯·普朗克是爱因斯坦的第一个支持者，如果不是普朗克当即认识到爱因斯坦研究的伟大成就，它可能要等更长的时间才会被认识和接受。普朗克之后，紧接着是其他一些著名的物理学家对它逐渐了解、认识、给予关注，很快，全世界都关注了起来。

-
- 光速是不变的，它不依赖于观察者的速度。
 - 相对论改变了我们有关时间和空间的观念，它告诉我们可以把它们看作单一的时空结构。
 - 狭义相对论列出了能量、动量（即一个物体遇到外力时会作出何种反应）和质量的关系。例如， $E=mc^2$ ，其中E是能量， m 是质量， c 是光速。
 - 质量和能量使时空产生弯曲，由此，你就可以把弯曲的时空当作引力场的来源。



-imension
探索大揭秘



量子力学：上帝不掷骰子

WARPED PASSAGES

也许你会问自己，我对了？……还是我错了？

传声头乐队 (Talking Heads)

她是从哪个门进来的？

艾克在纳闷，不知到底是因为阿西娜让他看了太多的电影，还是因为迪特尔对他讲了太多的物理知识，不管什么原因吧，总之艾克昨晚梦到了一个量子侦探。那个侦探穿着军大衣，一脸木然。在梦中，侦探说：“除了名字之外，我对她一无所知，而她就站在我面前，但从看到伊莱克特拉^[14]的第一眼起，我就知道她是个大麻烦。我问她从哪儿来，她拒绝回答。这屋子有两个门，她肯定是从其中某一个进来的。但伊莱克特拉沙哑着嗓子低声说：

‘先生，别问了。我是不会告诉你从哪个门进来的。’”

“看到她在发抖，我努力想让她镇静下来。但当我走近她时，伊莱克特拉却疯了似地后退一大步，她恳求我不要再靠近了。见她那么不安，我只好离她远远的。跟不确定的人打交道我并不陌生，但这次我却束手无策了，看来这会儿要跟不确定耗上了。”

量子力学尽管有悖于直觉，却从根本上改变了科学家对世界的看法。大多数现代科学都起源于量子力学：统计力学、粒子物理学、化

学、宇宙学、分子生物学、进化生物学、地质学（通过放射性元素推定年代），所有这些无一不是这一进步的结果，它们要么因量子力学而创立，要么因量子力学而得到改进。如果没有半导体和现代电子，则根本不可能有电脑、DVD播放机、数码相机等现代世界的诸多便利设施，它们的发展全都仰仗于量子现象。

我不太确信在大学里初学量子力学时，是否真正理解了它的奇妙。我学会了它的基本原理，也能够在不同场合应用它。但是直到多年以后我开始讲授量子力学的知识并仔细研究量子力学的逻辑时，才发现它竟是这么神奇。尽管我们现在只是把量子力学当作物理学课程的一部分，但它真的令人称奇。

科学应该怎样发展，量子力学的故事可谓一个非常完美的范例。量子力学早期是以模型构建理论开始的——不等任何人形成任何理论，它就开始探索一些令人费解的现象，实验和理论同时蓬勃迅速地发展着。物理学家创立了量子理论，来解释经典物理学不能解释的实验结果；反过来，量子理论又提出了更进一步的实验来验证假设。

科学家们花了很长时间才厘清这些实验观察的所有含义——量子力学的意义实在是太激进了，以至于大多数科学家并不能立即领悟它们。在接受量子力学的假设之前，科学家们只能暂时搁置心里的迷惑，因为它们与熟悉的传统观念有着太大的差异。甚至像马克斯·普朗克、欧文·薛定谔（Erwin Schrödinger）和阿尔伯特·爱因斯坦这样的理论先驱，也从未真正地接受量子力学的思考方式。“上帝不掷骰子。”爱因斯坦以此表达了他的反对。这句话也因此成了名言。大多数科学家最终的确接受了这一真理（正如我们现在理解的这样），但这种接受并非一蹴而就。

20世纪早期科学进步的激进特征弥漫在现代文化里：当时，文学和艺术的根本特征以及人们对心理学的理解都发生了翻天覆地的变

化。尽管有人将这些进步归因于第一次世界大战带来的动荡和破坏，但艺术家如瓦西里·康丁斯基（Wassily Kandinsky）却说：“原子能够穿透一切的事实说明任何事情都可以改变。因此，在艺术上，什么都有可能发生。”康丁斯基这样描述了他对核原子的反应：“原子模型的坍塌给我心灵的影响无异于整个世界的坍塌——最坚固的城墙顷刻间轰然倒塌；如果一块石头出现在我面前，在空气里融化、消失，我一点儿都不会感到惊诧。”

康丁斯基的反应是有点极端。量子力学非常激进，以至于当被应用于非科学场合时，它还是很容易被用过头。我发现最大的麻烦是，不确定性原理常常被滥用，人们往往似是而非地误用它，把它作为不准确的借口。本章我们将看到，不确定性原理实际上是对可观测量的一个极为精确的表述。当然，这一表述有着令人惊讶的内涵。

现在，我们来介绍量子力学和它的基本原理，这些原理使它有别于之前的经典物理。我们将讲到的离奇的新概念包括量子化、波函数、波粒二象性和不确定性原理，本章概括了这些主要观点，并简要回顾了它们的发展历史。

没有比量子力学更离奇的了

粒子物理学家西德尼·科尔曼（Sidney Coleman）曾说过，如果说成千上万的哲学家花了几千年的时间寻找世界上最为奇异的东西，那么，他们再也找不到比量子力学更为离奇的事物了！量子力学之所以难以理解，是因为它的结论是那么的有悖于常理，又是那么的出人意料。它的基本原理不仅有悖于以前所有已知物理的基本前提，也有悖于我们自己的经验。

量子力学看上去如此奇怪，其中一个原因是，我们的生理机能根本就不能够让我们观察到物质和光的量子本性。通常只有在一埃的尺

度上，也就是原子的尺度上，量子作用才会有意义。如果没有特殊仪器，我们根本不可能观察到那么小的尺度，即使是具有高分辨率的电视或电脑显示器的像素通常都小得我们根本看不见。

量子力学如此奇怪的另一个原因是，我们看到的只是大量原子的集合，这么多原子，足以让经典物理掩盖量子的作用。通常，我们看到的光也是大量光量子的集合。尽管眼睛的感光器官非常灵敏，足以捕捉到光的最小单位——单个量子，但是通常因我们的眼睛要处理这么多的量子，以至于任何可能的量子效应都被更明显的经典行为淹没了。

如果说量子力学让人难以解释，那也有充足的理由。量子力学意义深远，它足以容纳经典物理的预言，但反过来并非如此。在许多情况下，例如，当涉及大物体时，量子力学与牛顿经典力学的预言是一致的，但没有一个尺度可以让经典力学作出量子预言，因此，当我们试图用熟悉的经典术语和概念去理解量子力学时，我们必然是要遇到麻烦的。试图用传统的观念去描述量子作用，就好像是要以100个英文单词为限去翻译法语，以这么有限的英文词汇，常常只能含糊地解释某些概念和词语，而有一些则根本无法表达。

量子力学的先驱之一，丹麦物理学家尼尔斯·玻尔（Niels Bohr）意识到，要描述原子内部的活动，人类的语言远远不够。他在回顾这一主题时，讲了他的模型是如何“像图画一样，本能地闯进了他的脑海”。正如物理学家沃纳·海森伯（Werner Heisenberg）解释的：“我们只记得，通常的语言完全失去了作用，在物理学领域里，语言已毫无意义。”

因此，我不会尝试用经典模型去描述量子现象，相反，我要描述一些使量子力学与以前的经典理论如此不同的、重要的基本假设和现象，我们会逐个回忆那些对量子力学及其发展作出贡献的重要发现和见解。尽管这一讨论大致遵循着它们的历史顺序，但我真正的目的是要把量子力学内在的许多新观点、新概念逐一地介绍给大家。

起源：一切就像“卡路里灾难”那样

量子力学的发展分为几个阶段。一开始，它只是一些恰好与现象相符的、随意的假设，虽然没人知道它们为什么相符。这些突发奇想没有任何基本物理理论的支持，但的确给出了正确答案。我们现在所称的旧量子论包含的就是这些猜想，理论的基本假定是：像能量和动量这样的物理量不可能是任意值，而只可能限于一组离散的、量子化的数值。

量子力学由粗陋的旧量子论发展而来，它肯定了我们很快就会遇到的神秘的量子化猜想。更重要的是，量子力学给出了一个明确的程序，预言了量子力学体系会怎样随时间演变，这大大增加了理论的威力。但量子力学最初也只是摸索，因为没人知道后面的发展会怎样，开始的时候，量子化假设就只有这点内容。

旧量子论开始于1900年，德国物理学家马克斯·普朗克提出，光如砖块一样，也只能以某些量子单位配给。根据普朗克的假设，任何特定频率的光所包含的总能量只能是那个基本能量单位与其特定频率的乘积。这个基本单位等于一个常量乘以频率 f ，这个常量就是我们现在所称的普朗克常数 h ，那么有着特定频率 f 的光的能量就可能是 hf 、 $2hf$ 、 $3hf$ ，等等，但根据普朗克的假设，不可能介于两者之间。光与砖不同，因为砖的量化是我们随意假定的，它不是最根本的——砖还可以一劈两半，而光必然会有一个不可再分的既定频率的最小能量单位，它是一个永远不会出现中间值的能量。

普朗克最初提出这一颇具先见之明的非凡建议，是为了解决一个理论难题，称作黑体紫外光^[15]灾难。黑体^[16]如一块煤，是一种能吸收所有辐射然后再把它辐射出来的物体，它所放射的光的总量和其他能量取决于它的温度。一个黑体的物理属性完全体现在它的温度上。

但是，关于黑体辐射的光，经典预言是有问题的：经典计算预测的高频率辐射所放射的能量，要远远大于物理学家的观察和记录。测量显示，黑体的辐射并非由不同频率平等奉献，高频率反而比低频率贡献更少，只有低频率才可以放射出显著的能量，因而，放射性物体是“红热”（red-hot）而不是“蓝热”（blue-hot）。但经典物理学却预言了大量的低频率辐射。事实上，放射的总能量中，据传统推理所预测的辐射是有限的，传统物理学面临着紫外光灾难问题。

解决这一两难境地的权宜之计是，假设黑体辐射的能量全部是由低于一个上限值的频率所贡献的。普朗克对这一可能不以为然，他偏重于另一假设：光是量子化的，而这一假设显然同样的随意。普朗克的理由如下。

如果任一频率的辐射包含的都是一个基本辐射单位的整数倍，那么，因为基本能量单位太大，高频率的辐射根本就不可能发出。因为光量子单位所包含的能量与频率成正比，高频率辐射即便是只有一个单位，也仍旧会含有大量的能量，如果频率太高，一个量子所含有的最小能量值则会因为太大而不能放出。黑体只能放射那些低频率的量子，因此普朗克的假设就杜绝了过高频率的辐射。



有一种比喻可以帮助我们理解普朗克的逻辑。你可能在聚餐时遇到过这样的人：该点甜点了，他们却拒绝。因为害怕进食太多会增肥的食品，所以即使嘴上很馋，他们也很少给自己点甜食。如果服务生说甜点其实很小，那他们可能就会点上一份，但如果份量大，再切成一块块的蛋糕或冰激凌、布丁什么的，他们就干脆不要了。

面对这种情形，会有两种人：艾克属于第一种，他很有节制，而且确实不喜欢甜食。如果甜点太大，艾克根本就不会去吃它。而我则属于另外一种，阿西娜也一样，这种人觉得甜点太大了，因此自己就不点了，但又做不到像艾克那样拒绝美食，然后又会忍不住诱惑，就从每人的盘子里分一点。就这样，即使阿西娜拒绝给自己点甜食，但最终吃下去的却一点儿也不少。如果阿西娜与很多人一起吃饭，那么，从每个人的盘子里她都尝一点，最终她就会遭受不幸的“卡路里灾难”。

根据传统理论，黑体就像是阿西娜一样，它会发出任何频率的小量的光，使用经典推理方式的理论学家因此就会预测到“紫外光灾难”。为避开这一困境，普朗克提出的黑体就像是那种绝对饮食有度的人，就像艾克一样，他绝对不会去吃一点甜食。

黑体的表现会严格遵守普朗克的量子规则，它只会以量子能量单位发出一定频率的光，能量值等于常数 h 乘以频率 f 。如果频率太高，能量量子会变得太大，那个频率的光就不能发出，因此，黑体发出的辐射大部分都是低频率的，高频率会被自动排除。在量子理论里，一个黑体不会发出大量的高频率辐射，因此，它发出的辐射要远远小于经典理论的预言。如果一个物体发出辐射，我们称它的辐射模式为光谱（见图6-1），即物体在每一频率每一温度会释放多少能量。某些物体的光谱，比如恒星，就与黑体光谱类似，我们在许多不同温度下测量过这样的黑体光谱，它们都符合普朗克的假设，图6-1显示，黑体所发出的辐射都是低频率的，到了高频率，辐射则关闭了。

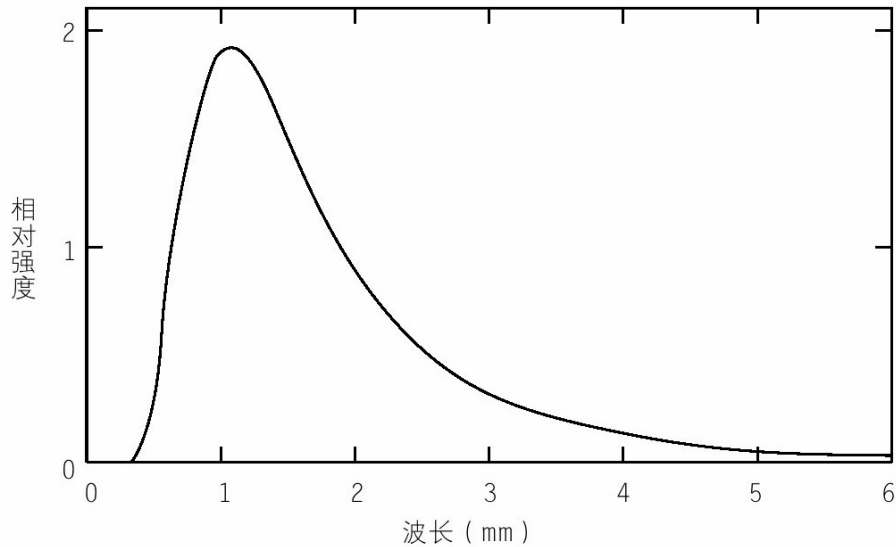


图6-1 宇宙里一个微波背景的黑体光谱。黑体光谱显示了辐射物体在固定的温度值下所发出的所有频率的光的总量，要注意的是，光谱里摒除了高频率的光。

20世纪80年代以来，实验宇宙学最为杰出的一个成就就是，黑体光谱测量越来越精确，它记录了宇宙产生的辐射。最初，宇宙是一个火热的、大密度的、包含高温辐射的火球，此后，宇宙膨胀，辐射大大冷却下来。这是因为随着宇宙的膨胀，辐射的波长也会增长，而波长越长则对应频率越低，频率越低则能量越低，能量越低则温度越低。宇宙现在所包含的辐射就像是一个绝对温度只有2.7K的黑体所产生的辐射——这比其初始状态凉快多了。

卫星最近测量了这一宇宙微波背景辐射（见图6-1），它看上去与一个绝对温度为2.7K的黑体光谱完全一致，这一测量显示的偏差要小于万分之一。事实上，这一遗留的辐射是迄今所测得的最为精确的黑体光谱。

1921年，当普朗克被问及他是怎么想到光是量子化的这一奇异假设时，他回答道：“这纯属走投无路才做的假设。6年来，我一直在研究黑体理论，我知道这是一个根本性问题，而且我知道答案，我

必须不惜一切代价找到一个理论性的解释……”^[17]对普朗克来说，光量子化是一种工具，一个对正确的黑体光谱进行胡乱组装的机器。以他的观点，量子化不一定是光本身的属性，但可能是辐射光的某种原子属性。尽管普朗克的臆测迈出了理解光量子化的第一步，但他自己却从未完全领会它。

5年以后，也就是1905年，爱因斯坦对量子理论作出了重大贡献——他确定光量子是真实的东西，而不仅仅是抽象的数学概念。那一年是爱因斯坦非常忙碌的一年，他提出了相对论、通过研究物质的统计性质帮助证实了原子和分子的存在、给出了量子论的有效解释——所有这些都是他在瑞士伯尔尼专利局工作期间完成的。

爱因斯坦利用光量子假说解释了被称作光电效应的现象，由此提高了光量子假说的可信度。实验者把同一频率的辐射照向物质，辐射会放出电子。实验显示，若用更多的光（这意味着携带更大的能量）集中射向物体，并不能改变所放出电子的动能（运动能量）最大值。这有悖于我们的直觉：参与能量越多，所产生的电子也必然会有越大的动能。由此，电子动能的限度就成了一个谜，为什么电子不能吸收更多的能量？爱因斯坦的解释如下。

辐射是由单个的光量子组成的，只有某些特定量子会将其能量贡献给特定的电子，供给某个电子的光就像是一枚导弹，而不像枪林弹雨。因为只有一个光量子能击发出电子，因此再多的参与量都不能改变被发出电子的能量。加大参与量子的数量，只会使光发射出更多电子，但不能影响单个电子的能量最大值。

爱因斯坦以固定大小的能量包来解释光电效应，这些固定大小的能量包即光的量子单位。经过这样的解释，被发出电子的动能最大值总是不变的属性就变得容易理解了——一个电子所拥有的最大动能是固定的，就是它接收到的光量子的能量减去将电子从原子中击发出来所需要的能量。利用这一逻辑，爱因斯坦导出了光量子的能量。他发

现，它们的能量取决于参与光的频率，这与普朗克预测的完全一致。对爱因斯坦来说，这就是光量子真实存在的确切证据。他的解释给出了一幅有关光量子的非常具体的景象：一个光量子击中一个电子，电子由此被击发出来。正是这一成就（而非相对论）使爱因斯坦获得了1921年的诺贝尔物理学奖。

但奇怪的是，尽管爱因斯坦承认光量子单位的存在，却不愿接受这些量子是无质量的粒子，即它们虽然携带能量和动量，却没有质量。关于光量子粒子特征的第一个有说服力的证据，是1923年的康普顿散射（Compton Scattering）测量。在此实验中，一个光量子击中一个电子，然后产生散射（见图6-2）。通常，你可以通过测量粒子碰撞后的偏转角度来测得粒子的能量：如果光子是没有质量的粒子，当它们与其他粒子如电子碰撞时，就会表现出一种特定的活动方式。而测量显示，光量子的表现确实就像是一个没有质量的粒子在与电子相互作用，由此我们便得出了一个无可更改的结论：光量子实际是一些粒子。我们现在称这些粒子为光子。



图6-2 康普顿散射。在康普顿散射里，一个光子（ λ ）击中一个静止的电子（ e^- ）而产生散射，并表现为不同的能量和动量。

爱因斯坦拒绝接受光量子理论，这确实令人费解，因为这一理论恰是在他帮助之下创立起来的。但相比普朗克的反应，他的表现又不那么令人吃惊了：对爱因斯坦量子化的建议，普朗克干脆不信。尽管普朗克和其他几个人对爱因斯坦的许多成就都赞赏有加，但这次他们

的热情却有所保留。普朗克甚至说过，多少有点儿贬低的意思：“他的猜测并未言中，比如他的光量子假说也错失了目标，但这对他并不能造成多少影响，因为即使在一个最为确定的学科里，想要引进一个全新的观念而又不冒任何风险根本是不可能的。”不要搞错，爱因斯坦的光量子假说可是正中目标，普朗克的言论只是反映了爱因斯坦见解的颠覆性特征，以及科学家们最初不情愿接受的事实。

发现原子

量子化与旧量子论的故事并未在研究过光之后就结束，结果发现，原来所有的物质都是由基本量子构成的。尼尔斯·玻尔是第二个站到量子假说队列里的人，在他的研究中，他将量子用于一个早已确立的粒子——电子。

玻尔对量子力学的兴趣，部分源于他当时正试图弄清楚原子的神秘性质。19世纪有关原子的概念模糊得令人难以置信：许多科学家根本不相信原子的存在，只不过把它当作一种有用的启发式教学工具，而这又没有现实基础。即便有些科学家真的相信它存在，却又把它们与分子混淆。而我们现在知道，分子是由原子组成的。

直到20世纪初，原子的真正属性和构成才被人们广泛接受，问题的部分原因是，希腊词汇“原子”的本意是“不可分割的东西”，因此人类对原子的最初定位就是一个不可改变、不可分割的事物。但到19世纪，随着物理学家对原子行为有了更多的了解，他们开始意识到这一观点肯定是错的。

到19世纪末，人们已正确测量了原子的一些属性，如放射性和光谱线（spectral lines）——光被发出和吸收的特定频率，而这两种现象都显示原子是能够改变的。在此基础之上，1897年，J. J. 汤姆逊

(J. J. Thomson) 发现了电子，并提出电子是原子的组成部分，这意味着原子一定是可以再分的。

到20世纪初，汤姆逊综合当时的原子观察，合成了他的“梅子布丁”模型 (plum pudding)，这个名称源于一种把水果丁分散地嵌在蛋糕里的英式甜点。他提出，有一种带正电的成分遍布整个原子（蛋糕部分），而带负电的电子（水果）则镶嵌于其中。

1910年，新西兰人欧内斯特·卢瑟福 (Ernest Rutherford) 证明了这一模型是错误的。当时，汉斯·盖革 (Hans Geiger) 和他的一个研究生欧内斯特·马斯登 (Ernest Marsden) 进行了由卢瑟福提议的实验。他们发现了一个坚实、紧密的原子核，远比原子小。

氦²²²（一种镭盐放射性衰变所产生的气体）放出了我们现在称为氦核的阿尔法粒子。物理学家通过发射出原子中的阿尔法粒子并记录阿尔法粒子散射的角度，发现了原子核的存在。只有存在一个坚实、紧密的原子核时，才会产生他们记录的这种急剧的散射，一个分散的、弥漫于整个原子的正电荷根本不可能将粒子打得如此分散。用卢瑟福的话来说：“这可能是我一生中见过的最不可思议的事了：就好像你朝一张卫生纸发射了一枚长38厘米的炮弹，而它竟然反弹回来并击中了你。这真是太令人难以置信了！”

卢瑟福的结果推翻了“梅子布丁”模型，他的发现意味着正电荷不是遍布于整个原子的，相反，它被局限于一个极小的内核中。而且肯定有一个坚实的核心组成部分——原子核。根据这幅图像，原子的组成就是：电子沿轨道围绕着中心一个小原子核旋转。

2002年夏天，我参加了弦理论的年会，碰巧那年会议在剑桥大学的卡文迪什实验室召开。许多重要的量子力学先驱，包括两名领军人物卢瑟福和汤姆逊都在这里完成了他们的重要研究。装饰走廊的是早年那些令人兴奋的回忆，游览走廊的工夫，我听说了许多有趣的故事。

例如，中子的发现者詹姆斯·查德威克 (James Chadwick)，他学物理只是因为不好意思承认自己在等待录取时排错了队；汤姆逊当实验室主任时很年轻，有封贺信是这样写的：“请原谅我的失误，没有致函祝您像

一个教授那么成功、愉快，因为您当选的消息对我是天大的惊喜，以至于都忘了要这么做。”（物理学家也并不总是大度有礼。）

但是，尽管20世纪早期在卡文迪什和别的地方，一幅清晰的原子图像已经建立，但其表现出来的组成成分却几乎击溃了物理学家最根本的信念。卢瑟福的实验提出，原子是由一个原子核与沿轨道绕其运行的电子组成的。这幅图像，尽管简单，却有着不幸的缺陷：它必然是错误的。

经典电磁学理论预言，如果电子沿圆形轨道运行，那么它们会通过发射光子辐射能量（经典的说法是电磁波辐射）。因此，光子会带走能量，留下能量更小的电子，它旋转的圆形轨道会越来越小，就这样螺旋着绕向中心。实际上，经典电磁学理论预言原子不可能是稳定的，而且不到一纳秒就会坍缩。原子稳定的电子轨道完全是个谜：为什么电子没有丢失能量，并螺旋地向中心的原子核落下去？

要解释原子的电子轨道，需要与经典论证彻底决裂。为找到通向这一必然结论的合理逻辑，探索者们发现了经典物理学的缺陷，只有发展量子力学才能弥补这一缺陷。尼尔斯·玻尔提出的就是这样一个革命性的建议，他将普朗克的量子观念延伸到了电子，而这也是旧量子论的基本组成部分。

原子的电子轨道

玻尔确定电子不可能在任何旧轨道上运行：电子轨道的半径必然满足他提出的公式。他发现这些轨道全凭幸运和天才的猜想。他认为，电子的行为必须像波一样，这就意味着它们在绕原子核运动时会上下振动。通常来讲，一个特定波长的波每经过一段特定距离就会上下振动一次，这个距离就是波长。那么绕着一个圆运动的波同样也有

其相应的波长。在这种情况下，波长就确立了波在绕着原子核转动时上下振动一次的弧度。

在固定半径轨道上运行的电子不可能有任意波长，而只能有一个特定的波长，以使它能以固定的次数上下振动。这就隐含了确定允许波长的法则：波在确定电子轨道的圆上旋转时，上下振动的次数一定是一个整数（见图6-3）。

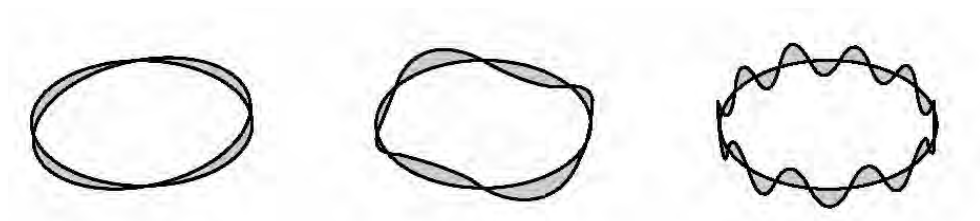


图6-3 根据玻尔量子假说，一个电子波的可能模式。

尽管玻尔的提议非常激进，而其意义又非常含糊，但他的猜想却似乎真的起到了作用。如果正确，它就保证了稳定的电子轨道：只允许存在某些特定的电子轨道，不允许出现中间的轨道；如果没有外来作用使电子从一个轨道跳向另一轨道，那么电子就没有办法向中心的原子核移动。你可以把一个有着特定电子轨道的玻尔原子想象成一座多层建筑，在这个建筑里，你只能在二层、四层、六层等双数楼层里活动，因为你不可能踏足中间的楼层如三层和五层——你永远只能被限制在你所在的双数楼层上，没有办法下到底楼走出去。

玻尔的波真是一个天才的设想。他没有说自己明白它的含义，他作出这一猜想只是为了给稳定的电子轨道找到解释。但他建议里的量子特征使它可以得到验证，尤其是，玻尔的假说正确预言了原子光谱。光谱线给出了一个未电离原子发出和吸收的光的频率，未电离原子指的是一个中性原子，它的所有电子携带的净电荷为零。物理学家注意到，光谱线显示的是一个条纹状的图案，而不是一种连续的分布

（即所有频率的光都有）。但没人知道这是什么原因，也没人知道为什么恰恰会是他们见到的这些频率。玻尔用他的光子假说解释了为什么光子的发出和吸收恰好是这些测到的频率。

尽管电子轨道对孤立的原子是稳定的，但当有一个恰当频率（根据普朗克的理论，是恰当能量）的光子释放或是带走能量时，是能够改变电子轨道的。

玻尔使用经典物理学的推理方式计算了遵守他的量子假说的电子的能量，从这些能量中，他预言了由只含有一个电子的氢原子所发出和吸收的光子的能量，即频率。玻尔的预言是正确的，这些正确预言大大增加了他的量子假说的可信度，也正因如此，爱因斯坦和其他科学家才相信玻尔一定是对的。能发射和吸收因而也能改变电子轨道的量子包，借前面多层建筑的比喻，可以被比作放在窗户外面的绳子的长度。如果每条绳子的长度恰好是从你所待的楼层到另一双数楼层所需的长度，而那层楼的窗户又是开着的，那么绳子就为转换楼层提供了工具。但这种转换只能在双数楼层之间。同样的道理，光谱线只能表现出某些特定的值，即占据容许轨道的电子之间的能量差。

即便玻尔没有为他的量子化条件提供任何解释，但他显然是正确的。人们测量了许多光谱线，而他的假设可以用来再现它们。如果说这种一致只是巧合，那也太神奇了。终于，量子力学证明了他的假设。

粒子的使命恐惧

尽管量子假说非常重要，但只有在法国物理学家路易·德布罗意（Louis de Broglie）和奥地利籍物理学家欧文·薛定谔及德裔物理学家马克斯·玻恩（Max Born）取得一定的进展后，量子力学中粒子与波的联系才渐渐变得明朗起来。

德布罗意提出了一个伟大的建议，他转变了普朗克的量子假说，由此跨出了关键的第一步：旧量子论的自由漫步终于迈向了真正的量子力学理论道路。普朗克把量子与辐射波联系起来，而德布罗意（像玻尔一样）假想粒子也可以像波一样活动。德布罗意的假说意味着粒子也可以表现出像波一样属性，而这些波又是由粒子的动量所决定的。（对低速运动来说，动量即质量乘以速度；对所有速度，动量告诉我们物体会如何回应外力。尽管在相对论速率里，动量是一个更为复杂的质量与速度的函数，但它适用于高速的动量推广，也指示了相对论速率的物体会怎样回应外力。）

德布罗意假定，动量为 P 的粒子关联着一个其波长与动量成反比的波，即动量越小，波长越长；波长还与普朗克常数 h 成正比。^[18]德布罗意建议的背后含义是，一个振动激烈的波（即有很小的波长），相比一个振动迟缓（有较大的波长）的波携带着更大的动量；较小的波长意味着更快的振动，德布罗意为它赋予了更大的动量。

如果你觉得这种粒子——波的存在很令人迷惑，那是因为它的确如此。当德布罗意第一次提出他的波时，没人知道它们有什么用。马克斯·玻恩提出了一个惊人的解释：波是一个位置函数，它的平方给出了在空间的任一位置找到一个粒子的概率，^[19]他称其为波函数。马克斯·玻恩认为，粒子不可能固定，而只能以概率来描述，这与经典的设想有着天壤之别。它意味着，你不可能知道粒子的确切位置，只能确定在某些地方找到它的概率。

但即便量子力学的波描绘的只是概率，量子力学还是能预言波随时间的确切演化方式：给定任一时间的值，你就可以确定它在以后时间里的值。薛定谔创立了波动方程，它显示了与量子力学的粒子相关的波的演变。但找到一个粒子的概率有什么意义？这是一个令人迷惑的观点，毕竟没有分数粒子这样的东西。在当时，用波来描述粒子是

（在某种意义上，现在依然是）量子力学最令人吃惊的观点，尤其是人们都知道粒子的活动就像是一个台球，而不是像波。这样看来粒子说与波动说似乎是不可调和的。

要解决这一明显的矛盾，你需要了解这样一个事实：用一个粒子永远无法探测到粒子的波动特征。当你探测单个粒子时，只能探测到它处在某个特定的位置上。为了能勾画出完整的波，你需要许多相同的电子，或多次重复一个实验。即使每个电子都关联着一个波，一个电子也只能测得一个数字。但如果你能准备大量的相同电子，就会发现在每个位置上的电子的分布与量子力学给定的电子的概率波成正比。

单个电子的波函数告诉我们，有着同一波函数的许多同样的电子会如何表现。任何单个电子只能出现在一个位置上，但如果有许多相同的电子，它们就会呈现像波一样的位置分布图。波函数显示的就是一个电子出现在那些位置上的概率。

这好比是人群的身高分布，每个人都有各自的身高，但分布图显示的则是个体具有某特定身高的可能性。同样的道理，即便一个电子的表现像是粒子，但是许多电子就会形成一个由波来描述的位置分布图；不同的是，单个电子也仍然关联一个波。

图6-4显示的是一个电子的概率函数，这个波给出了在某一特定位置上找到一个电子的相对概率。我画的曲线在空间的每一点（或沿一条线的每个点，因为纸是平面，我只能画出空间的一个维度）上都有具体的值。如果我能多次复制同一个电子，那么我就能得到电子位置的一系列测量结果，进而就会发现我所测量的这个电子出现在某个特定位置上的次数与这一概率函数成正比：测量值越大，电子越有可能出现在那个位置上；测量值越小，电子越不可能在那里出现。波反映了许多电子的累积效应。

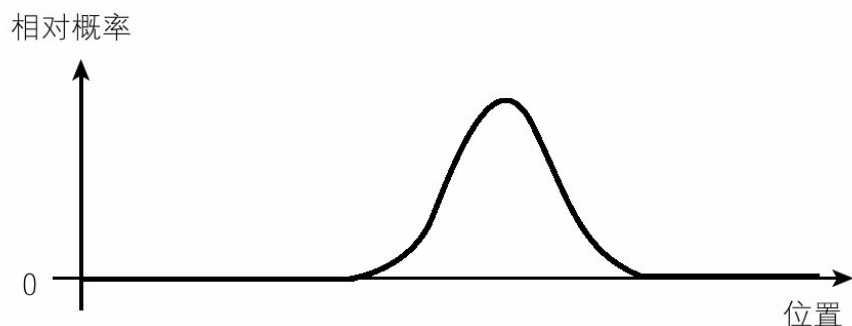


图6-4 一个电子的概率函数。

即使你用了许多电子来勾画一个波，但是量子力学的特殊之处在于，单个电子也可以用波来描述。这就意味着，你永远也不可能确定地预测到这个电子的所有事情：如果你测得了它的位置，你会发现它是处在一个特定的点上的；但在你作出这个测量之前，你只能预测这个电子有一定的概率会在那个地方出现，而不能确切地说出它最终会出现在哪儿。

这一波粒二象性由著名的双缝实验得出，篇首故事里伊莱特克拉的不明来历指的就是双缝实验。以前，双缝实验只是物理学家的思想实验，用以说明电子波函数的含义和结果。直到1961年，德国物理学家克劳斯·琼森（Claus Jonsson）才在实验室里完成了这一实验。实验的组成是：一个电子发射器发出电子，穿过有两条平行狭缝的隔板（见图6-5），电子穿过狭缝，射到隔板后面的一个屏幕上，由屏幕记录下来。

这个实验本意是要效仿19世纪中期的一个相似实验，那个实验显示了光的波动特征。当时，英国医生、物理学家以及埃及古物学家托马斯·杨（Thomas Young）^[20]，将一束单色光穿过两个狭缝打在屏幕上，以观察光在屏幕上形成的波状图案。那一实验显示，光的表现就

像波一样。用电子来做同样的实验的关键是，看你将怎样观察电子的波动特征。

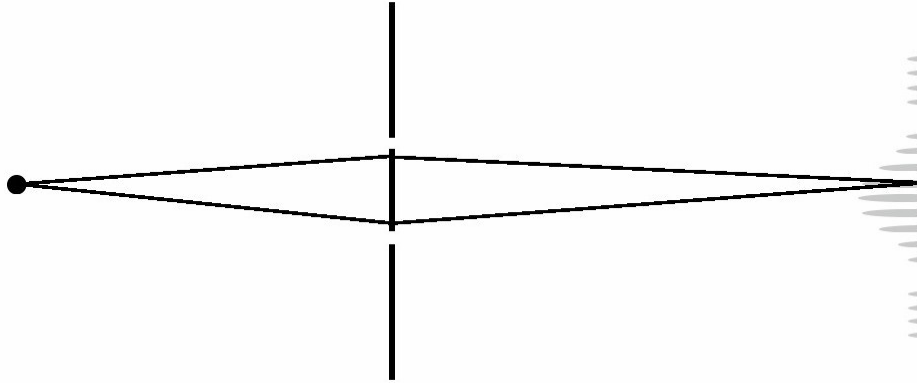


图6-5 精心设计的电子双缝干扰实验。电子可以穿过两个狭缝中的任一个，最终射到屏幕上，记录在屏幕上的波状图案就是两条路径相互干扰所产生的结果。

事实上，如果你用电子做双缝实验，那么便会看到与托马斯·杨在光里所看到的同样的现象：狭缝后的屏幕上出现的是波动图案（见图6-6）。在光的例子里，我们理解波是由于干扰引起的：一些光通过这条狭缝，而另一些光通过另一条狭缝，由此记录下的波状图案反映的就是两者之间的干扰。那么，波状图案对电子又意味着什么？

屏幕上的波状图案告诉了我们一个有悖于直觉的事实：我们应该会想到每个电子都通过了两个狭缝。对于单个电子，你不可能了解它所有的事情，任何一个电子都有可能穿过其中的一个狭缝。即便每个电子的位置在它们到达屏幕时都被记录了下来，但没人知道一个特定的电子究竟是从哪个狭缝穿过的。

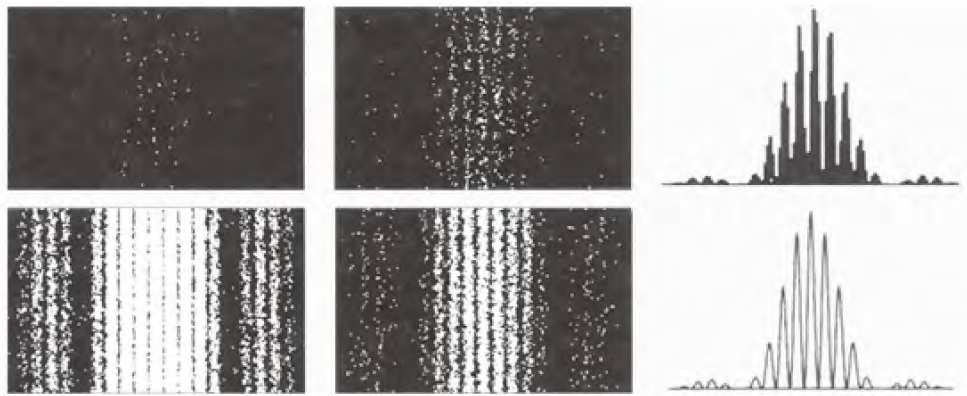


图6-6 双缝实验记录的干扰图案。左边的四块镶板由左上起顺时针依次显示的是：50、500、5 000、50 000个电子射过狭缝时所看到的图案。右边曲线是将电子数量的分布（上面的曲线）与通过两条狭缝可能形成的一个波状图案相比较，它们几乎是完全相同的，这表明电子的波函数真的像波一样。

量子力学告诉我们，一个粒子从其起点到终点可以采取任何可能的路径，粒子的波函数就反映了这一事实，这是量子力学的众多新奇特征之一。与经典物理学不同，量子力学并不为一个粒子明确指定运动轨迹。

但当我们已经知道了电子是粒子，那么双缝实验又是怎样表明单个电子会像波一样活动的？毕竟，没有像半个电子这样的东西，被记录下的任一电子都有它特定的位置。事情究竟是怎样进行的？答案就是我之前说过的：只有在记录许多电子时，你才能看到波的图案。每个单独的电子都是一个粒子，它打在屏幕上只能有一个位置，但许多电子被射在屏幕上的累积效果就是一个经典的波状图案，反映了有两条电子途径干扰的事实。在图6-6中，你就可以看到。

波函数给出的是一个电子可能击在屏幕上某个特定位置的概率。一个电子可能会出现在任何地方，但你在某个特定位置找到它的概率是一定的，这个概率由波函数在某点的值决定。许多电子一起形成一个波，假设电子通过两个狭缝，你就会得到同样的波。

1970年，日本的外村彰（Akira Tonomura）、意大利的佩尔基奥基奥·默里（Piergiorgio Merli）、朱利奥·波济（Giulio Pozzi）和奇安弗兰科·米西罗利（Gianfranco Missiroli），在真正的实验里清晰地观察到了这些，他们令电子一个个地穿过狭缝。当越来越多的电子被打到屏幕上时，一个波的图案就形成了。

你可能会感到奇怪，像波粒二象性这么戏剧性的东西，为什么直到20世纪才被人们注意到？例如，为什么人们没有早一点注意到光虽然像波，但实际是由一些独立的小东西——光子组成的？

答案在于，没人（除非超人）能看到单个光子^[21]，因此，量子力学的作用就不容易察觉。通常的光看起来也不像是由量子组成的，我们看到的都是形成可见光的光子束。而大量的光子在一起表现得就像一个经典的波。要观察光的量子特征，你需要一个很弱的光子源，或一个精心设计的系统。如果有太多的光子，你就不能区分单个光子的效果。在经典光里，因为包含许多光子，再加进一个光子，也根本不足以造成任何差别。你的灯泡表现的都是经典现象，即使它多发出一个光子，你也根本不会注意到。只有在精心设计的系统里，你才会观察到量子现象的细节。

如果你不相信这最后一个光子实际是无足轻重的，那就想想你投票时的感觉吧。成千上万的人都参与了投票，当你知道你这一票根本不足以对结果造成任何影响时，你还愿意费时费力地去投这一票吗？单个人的投票结果一般会被淹没在人群中。即便选举结果由许多单个选票的累积作用所决定，但单张选票却很少能改变结果。但有一个著名的例外：佛罗里达，那是一个“摇摆州”，当重复投票时有可能产生不同结果。让我们把这对比再推进一步，你可能也看到了，只有在量子力学上——还有在佛罗里达，它的表现就像一个州量子——重复测量还真的会产生不同的结果。

海森伯的不确定性原理

物质的波特征有许多含义有悖于常理，现在我们从电子的不确定性转向海森伯的不确定性原理，这也是物理学家和人们茶余饭后津津乐道的话题。

德国物理学家沃纳·海森伯是量子力学的先驱之一。在他的自传里，他讲述了在慕尼黑神学培训学院总部驻扎期间，自己如何萌生了有关原子和量子力学这一革命性观点的。1919年，他被派驻在那里对抗巴伐利亚。枪击事件平息以后，他坐在学院的屋顶上看柏拉图的《对话录》，尤其是《蒂迈欧》。柏拉图的教义使海森伯相信：“为了阐释物质世界，我们需要了解其最小的成分。”

海森伯憎恨年轻时代周围世界的喧闹与躁动，他希望能回到“普鲁士时代，那时人们的生活原则是：个人抱负要服从国家事业，人们生活简朴、正直廉洁、英勇无畏、克己守时”。但海森伯的不确定性原理却无可否认地改变了人们的世界观。也许是海森伯生活的动荡年代将他带上了一条革命的科学道路，却非政治道路。无论是怎样一种状况，我还是觉得多少有点好笑，不确定性原理的作者竟是有着这样一种矛盾性格的人。

不确定性原理告诉我们，某些成对的物理量不可能同时被准确测量到。这与经典物理学截然不同，经典物理学认为，至少在理论上，只要你愿意，就可以准确地测量一个物理系统的所有特征，如位置和动量。

对这些特定的物理量，先测量哪一个，是至关重要的。例如，如果你先测量位置，然后再测量动量（这个量既给出了速度也给出了方向），得到的结果会和你先测量动量再测位置是不同的。而在经典物理学中不会有这样的情形。当然，我们习惯的情形也都不会这样。只有在量子力学里，测量顺序才会对结果有影响。而且，根据不确定性原理，对测量顺序会产生影响的两个量的不确定性相乘总会大于一个

基本常数，即普朗克常数 h ，它的值是 6.582×10^{-25} GeV秒。^[22]如果你执意要非常准确地知道位置，那么你就不可能同样准确地知道动量，反之亦然。无论你的测量工具有多精密，也无论你尝试过多少次，你永远不可能同时将两个量都进行精确的测量。

普朗克常数在不确定性原理中的出现，有着重大的意义，因为它是一个只在量子力学中出现的量。回想一下，根据量子力学，一个有着特定频率的粒子的能量等于普朗克常数乘以频率。如果经典力学统领世界，那么，普朗克常数的值等于零，而且根本不会有基本量子。但以真正的量子力学来描述世界，普朗克常数则是一个非零的固定值，而且这个数值告诉了我们不确定性的存在。

从理论上讲，任何单一的量都可以精确地被测量到。有时，物理学家用波函数坍缩来明确指出某种东西已被精确地测量，因而有确定的值。“坍缩”一词描绘的是波函数的形状，它不再扩展，而是在某一特定位置上有了一个非零的值，因为这以后测得其他值的概率是零。在这种情况下（一个量已被精确测量），不确定性原理告诉我们，在测量了一个量之后，对不确定性原理中的另一个与之成对的量，你会一无所知。对另一个量的值，你的不确定性是无限的。当然，如果你先测量第二个量，那么第一个量就是你事先所不知道的。无论哪种方式，对一个量你知道得越精确，另一个量的测量就必然越不精确。

在本书中，我们不去探究不确定性原理的详细推导过程，但对其来源，我们还是简单了解一下，这些对后面的内容并非至关重要，所以，如果你不感兴趣，大可直接跳到下一部分。但也许，你对不确定性的基本推理，还想知道得更多一点。

在这一推导中，我们集中来看时间-能量不确定性，因为这个较容易理解和解释。时间-能量不确定性原理把能量（根据普朗克假设，也

指频率)的不确定性与表征系统变化率的时间间隔联系起来。也就是说,能量不确定性与表示系统变化的典型时间的乘积总会大于普朗克常数 h 。

当你打开电灯听到附近收音机的静电噪声时,时间-能量的不确定性就出现了。打开开关,会引发各种频率的无线电波,这是因为通过电线的电量突然发生变化,因此能量(由此引起频率)也会发生很大变化,你的无线电由此把它当作静电干扰。



为理解不确定性原理的来源,我们来看另一个截然不同的例子:滴水的水龙头。^[23]我要说的是,要确定水龙头滴水的准确速度,你需要长时间的测量,这就非常类似于不确定性原理的断言。一个水龙头和流过的水量,涉及许多原子,这是一个非常复杂的系统,不可能显示出可观察的量子力学作用,因为它们已被经典过程所淹没。但有一点是不变的,即越是要精确地测量频率,就需要越长的时间来测量,而这正是不确定性原理的核心。因为在一个精心设计的量子力学系统里,能量和频率是互相关联的,因此量子力学的一个系统会将这种相互依赖的关系再延伸一下。这样,频率不确定性与测量时间的长短(就如我们马上就要看到的)之间的关系,就转化为能量和时间之间的不确定性关系。

假设水的滴落速度是大约每秒1次。如果你的秒表精确度是1秒，即你至多会让它在1秒内停下，那你测得的速率会怎样？如果你等上1秒钟，看到落下1滴水，你就会认为可以由此得出结论：水龙头的滴水速度是每秒1次。但如果你的秒表只能在1秒内停下，你的观察就不能明确地告诉你，测量水龙头滴水花了多长时间。如果你的秒表滴答了1下，那么，时间有可能是稍稍多于1秒，甚至也有可能是接近2秒，那么在1秒和2秒之间，水龙头滴水的准确时间究竟是几秒呢？如果没有更为精确的秒表或是更长的测量时间，你就找不到更为满意的答案。用这个秒表测量，你只能得出这样的结论：水的滴落速度可能介于1秒1次或者2秒1次之间。如果你说水的滴落速度是每秒1次，那么从根本来讲，测量误差可能会达到100%。也就是说，你只有两种可能，要么对，要么错，你出错的可能是一半对一半。

但是，假设在做这一测量时，你等了10秒钟，那么在秒表滴答了10次之后，测得水滴了10次。用这只精确度只有1秒的表，你只能推测出10滴水滴落所花的时间大约在10~11秒之间。你的测量仍旧是水的滴落速度：大约是每秒

1次，而这次的误差就只有10%了，因为你等了10秒钟，测量的频率达到1/10秒之内。注意，你测量所花的时间（10秒）与频率的不确定性（10%或0.1）的乘积大约是1。再注意，在上一例里，你的频率测量误差可能更大（100%），但是所花时间更短（1秒），频率的不确定性与测量所花时间的乘积，大约也是1。

你可以沿着这一思路继续下去：如果要进行100秒的测量，那么，所测得的滴水频率可能精确到1/100；如果测量水的滴落长达1 000秒，那么测量的频率可能精确至1/1 000。在所有这些情况中，测量的时间间隔与测量频率的精确度的乘积总是1。^[24]

时间-能量不确定性原理的核心在于，越想测得更为准确的频率，所需花费的时间就越长。要使频率的测量越来越精确，就必须测量更长的时间。时间与频率的不确定性乘积大约总是1。^[25]

现在让我们完成这一简单的不确定性原理的推导，如果你有一个足够简单的量子力学系统，如一个光子，它的能量就等于普朗克常数 h 乘以频率。对这样一个物体，你测量能量的时间间隔与所测得的能量的误差的乘积总会超过 h 。只要你愿意，可以尽可能准确地测量能量，但这个实验就相应需要更长的时间。这就是与我们刚才所推导的同样的不确定性原理，唯一不同的就是多转了一道弯，增加了能量与频率的量子关系。

两个重要的能量值及其在不确定性原理下的意义

至此，我们已基本完成了量子力学基本内容的介绍。在这一节及下一节，我们将回顾量子力学的另外两个元素，在以后的章节中，我们会用到它们。

在这一节中，我们会讲述不确定性原理与狭义相对论的一个重要应用，不会再涉及任何新的物理原理。它探讨了两个重要能量与具有那些能量的粒子所敏感的物理过程的最小尺度之间的关系——粒子物理学家一直在用这些关系。后面一节介绍自旋、玻色子和费米子，这些概念还会出现在下一章有关粒子物理的标准模型里，而且，在后面我们探讨超对称时也会出现。

根据位置与动量的不确定性原理，位置与动量的不确定性乘积必然会超过普朗克常数。它告诉我们，任何东西——无论一束光、一个粒子或任何其他物质，或所能想象的任何系统，只要它对发生在小尺度的物理过程敏感，就必然涉及大范围的动量（因为动量必然是极不确定的），尤其是，所有对这些物理过程敏感的物体，一定包含极高的动量。根据狭义相对论，动量高，能量就高。这两个事实结合在一起就告诉我们，探索小距离尺度的唯一方式就是使用高能量。

另一种解释是，我们需要高能量来探索小距离，是因为只有波函数在很小范围内变化的粒子才会受小距离物理过程的影响。就如维米尔（Vermeer）不可能用一个5厘米宽的画笔完成他的作品，而你用模糊的视力不可能看到精妙的细节一样，如果粒子的波函数不是只在很小的范围内变化，它们也不可能对小距离内的物理过程产生反应。但根据德布罗意的观点，粒子的波函数涉及的波长很小，它的动量就很大，粒子-波的波长与其动量成反比。因此，根据德布罗意的观点，我

们也可以得出这样的结论：要敏锐地测知小距离物理，就需要高动量，因而也需要高能量。

这对粒子物理学有着极为重要的意义，只有高能粒子才能感知小距离物理过程的作用。我用两个具体例子来说明那些能量会有多高。

粒子物理学家通常用电子伏的倍数来表示能量，我们将电子伏缩写为eV，就念它的字母音。1电子伏是克服一个电位能差移动一个电子所需要的能量，就好比是由一个1伏的很弱的电池所提供的能量。我在文中还会用到与它相关的单位：吉电子伏或GeV和太电子伏或TeV，1 GeV等于10亿eV；1 TeV等于1万亿eV（或1 000 GeV）。

粒子物理学家还发现，用这些单位不仅可以测量能量，也可以用来表示质量。这是因为，狭义相对论通过光速将质量、动量、能量联系起来，而光速是一个常量 $c=299\,792\,458$ 米/秒，因此，我们可以用光速将能量转换为质量或动量。例如，爱因斯坦著名的公式是 $E=mc^2$ ，这意味着与一个特定的能量相关的是一个既定的质量。因为所有人都知道转换的系数是 c^2 ，我们就可以把它包括在内，用eV来表示质量，这样一来，质子的质量用这一单位表示就是10亿eV，即1 GeV。

这种转换就好比我们常在日常生活中告诉别人“火车站离这有10分钟的路”一样，我们心里假设了一个特定的转换系数。如果是步行10分钟，这个距离有可能是1公里；如果是以高速路上开车的速度，那10分钟的路程有可能是16公里。你和你的对话人之间会有一个公认的转换系数。

这些狭义相对论关系，连同不确定性原理，就确定了带有特定能量或质量的一个波或粒子所能经历和探测的物理过程的最小空间范围。现在，我们就将这些关系应用到粒子物理学中两个非常重要的能量上。在以后的章节里，它们会常常出现（见图6-7）。

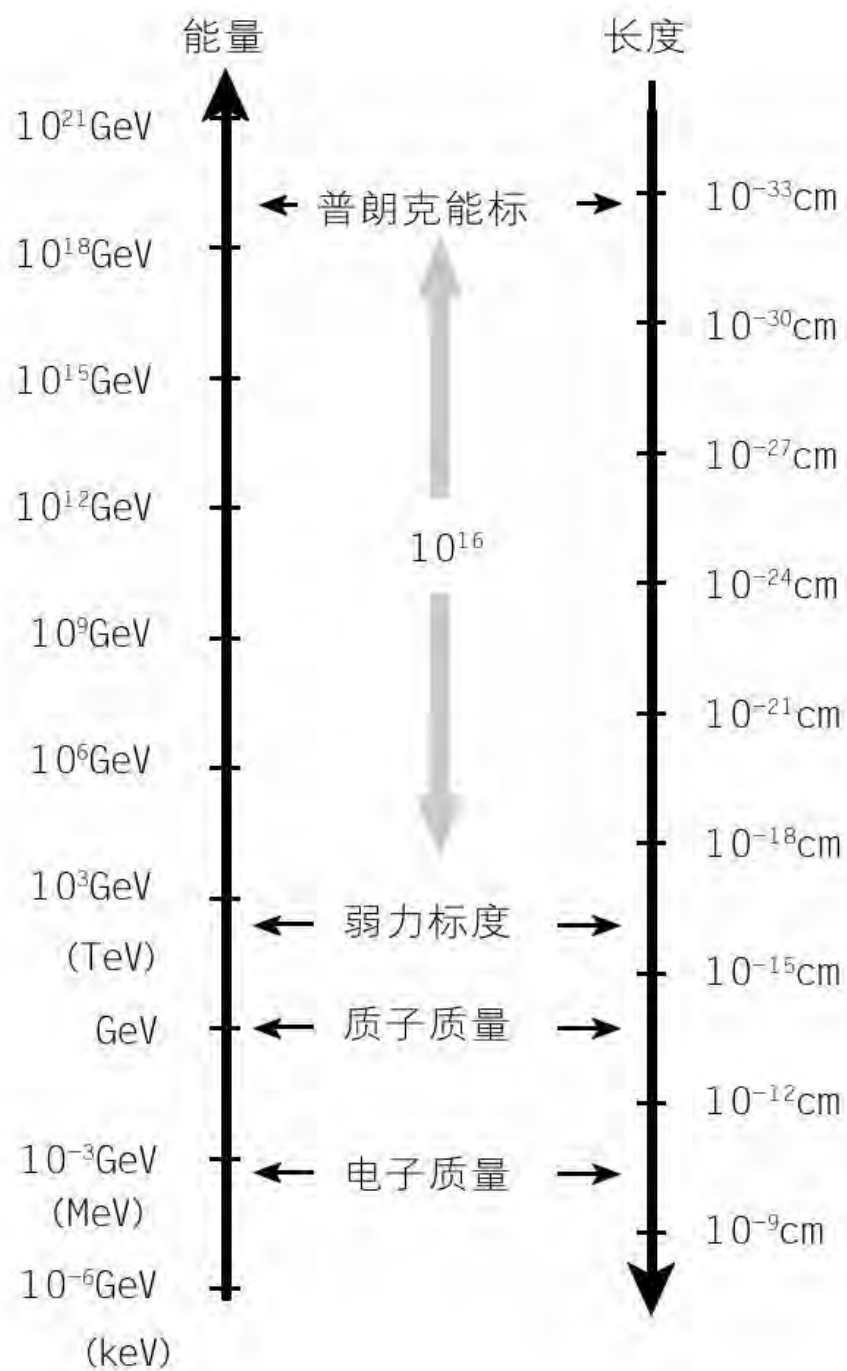


图6-7 粒子物理学中一些重要的长度和能量标度。更大的能量（根据狭义相对论和不确定性原理）对应的是更短的距离——一个更为活跃的波对于发生在更小距离上的相互作用很敏感。引力作用与普朗克能量成反比，普朗克能量很大则意味着引力作用很微弱；

弱力级能量（通过 $E=mc^2$ ）确立了弱力规范玻色子质量的范围。弱力长度是指弱力规范玻色子传递弱力的距离范围。

第一个能量称作弱力能标，是250 GeV。这一能量的物理过程确定了弱力和基本粒子的主要性质，更重要的是确定了基本粒子是如何获得质量的。物理学家（包括我自己）都期望我们在探索这一能量时能发现一些新的、由未知物理理论所预言的作用，并能更多地了解物质的基本结构。幸运的是，实验很快就会探索到弱力能标，而且应该很快就能告诉我们想知道的。

有时，我也会用弱力质量，它通过光速与弱力能标相联。以更为传统的质量单位表示，弱力质量标度是 10^{-21} 克，但就如我刚才解释的，粒子物理学家共同认可用GeV来表示质量。

与之相应的弱力长度是 10^{-16} 厘米，或一亿亿分之一厘米，这就是弱力的范围，即粒子能相互影响的最大距离。

因为不确定性告诉我们，探知小距离必须用高能量，弱力长度也就是能量为250 GeV的粒子所能探知的最小长度，这是物理过程所能影响的最小范围。如果想以这一能量探测更小的距离，那么距离不确定性必须小于 10^{-16} 厘米，而这将打破距离-动量的不确定性关系。

现如今正在使用的费米实验室加速器和在日内瓦CERN建成的LHC要探索的就是直达这一尺度的物理过程，我将讨论的许多模型都将在这一能量水平产生可见的结果。

第二个重要的能量是普朗克能标—— M ，即 10^{19} Pl GeV。这一能量对所有引力理论都非常重要，例如，牛顿引力定律的引力常数与普朗克能标的平方成反比。两个质量之间的引力作用很小，是因为普朗克能标很大。而且，普朗克能标是经典引力理论所能适用的最大能量。超出普朗克能标，重要的就是量子引力理论，它将量子力学和引

力统一在一起。在后面的章节，当我们谈到弦理论的时候，会看到在旧的弦理论模型里，弦的张力很可能取决于普朗克能标。

量子力学和不确定性原理告诉我们，当粒子达到这一能量时，它们就能探知在微小如普朗克长度^[26]的物理过程，这个微小长度只有 10^{-33} 厘米。这是一个极端微小的距离，远小于我们能测得的任何长度。但要描述发生在这样微小尺度的物理过程，就需要量子引力理论，而这一理论可能就是弦理论。因此，普朗克长度与普朗克能标都将在后面的章节反复出现。

玻色子和费米子

玻色子

玻色子的命名是为了纪念印度物理学家萨地扬德拉·玻色（Satyendra Nath Bose），被称作玻色子的粒子都是整数自旋——量子力学的自旋与自转无关，玻色子的自旋可以是0、1、2等。

量子力学在粒子世界做了一个重要的划分，所有粒子被分为玻色子和费米子。这些粒子可以是基本粒子，如电子和夸克；也可以是复合体，如质子或原子核。任何粒子要么是玻色子，要么是费米子。

一个粒子究竟是属于玻色子还是费米子要取决于它们的内禀自旋，这一名称会让人产生联想，但粒子的“旋转”与实际的空间运动并无联系。如果一个粒子有自旋，是指它们的内部作用就像是在旋转，尽管实际并非如此。

例如，电子和电磁场的相互作用依赖于电子的经典旋转——它在空间的实际自转，但这种相互作用还依赖于电子的自旋。经典的旋转是由现实空间里的实际活动^[27]引起的，而自旋则不同，它是一个粒子的属性：是固定的，而且有着不变的、特定的值。例如，光子是一个

玻色子，它的自旋是1，这就是光子的属性，就如同光子以光速行进一样，是一个基本事实。

在量子力学里，自旋也被量子化。量子自旋值可以是0（即没有自旋），也可以是1或2，或其他任何整数的自旋，我称它们是自旋-0，自旋-1，自旋-2，如此等等。玻色子的命名是为了纪念印度物理学家萨地扬德拉·玻色（Satyendra Nath Bose），被称作玻色子的粒子都是整数自旋——量子力学的自旋与自转无关，玻色子的自旋可以是0、1、2等。

费米子自旋的量子化是量子力学出现之前人们无论如何都想象不到的。费米子得名于意大利物理学家恩里科·费米（Enrico Fermi），它的自旋是半整数，即 $1/2$ 或 $3/2$ 。一个自旋为1的粒子转一圈才能回到原来的结构，而自旋为 $1/2$ 的粒子旋转两圈才能回到原来的结构。尽管费米子自旋为半整数倍令人感到奇怪，但质子、中子和电子等都是自旋为 $1/2$ 的费米子。从根本上讲，所有我们熟悉的物质都是由自旋为 $1/2$ 的粒子构成的。

大多数基本粒子都是费米子，这一性质决定了我们周围物质的许多属性。尤其是泡利不相容原理，它阐明了两个同一类型的费米子永远不会出现在同一位置。正是这一不相容原理，形成了原子化学结构的理论基础。因为有着相同自旋的电子不可能出现在同一位置，它们必须占据不同的轨道。

正因如此，我才以一座有着不同楼层的建筑来做比喻。根据泡利不相容原理，不同楼层代表许多电子在围绕原子核旋转时，电子所占据的不同的量子化电子轨道。正因为这一原理，你才不会将手插进桌子，也不会掉进地球中心。桌子和你的手都有着坚实的物质结构，是因为不相容原理使物质形成了它们各自的原子、分子和晶体结构。你手上的电子与桌子的电子是相同的，所以当手拍击桌子时，电

子无处可去，两种相同的费米子不可能同时出现在同一位置上，因此物质才不会坍缩。

费米子

大多数基本粒子都是费米子，这一性质决定了我们周围物质的许多属性。尤其是泡利不相容原理，它阐明了两个同一类型的费米子永远不会出现在同一位置。正是这一不相容原理，形成了原子化学结构的理论基础。因为有着相同自旋的电子不可能出现在同一位置，它们必须占据不同的轨道。

而玻色子的表现与费米子恰恰相反，它们会出现在同一位置上。玻色子就像鳄鱼——喜欢一个摞一个地叠在一起。如果你向本来就有光的地方再射去一束光，它的表现可完全不同于你空手拍桌子。光是由玻色子——光子构成的，光与光可以互相渗透，两束光完全可以射向同一位置。事实上，激光所依赖的事实正在于此：玻色子可以占据同一领地，因此允许激光生成强烈的、协调一致的光束。超流体和超导体也都是由玻色子组成的。

玻色子属性的一个极端例子是玻色-爱因斯坦凝聚：许多相同的粒子表现得像一个粒子一样——这是费米子永远无法做到的（它们不能出现在相同位置上）。玻色-爱因斯坦凝聚的实现，正是因为组成它们的玻色子不同于费米子，可以具备完全相同的性质。2001年，埃里克·康奈尔（Eric Cornell）、沃尔夫冈·克特勒（Wolfgang Ketterle）和卡尔·韦曼（Carl Wieman）因为发现玻色-爱因斯坦凝聚而获得诺贝尔物理学奖。

在之后的章节，我将不再讨论费米子和玻色子行为的具体性质。我会用到的唯一内容是，基本粒子有其内禀自旋，可以表现为向一个或另一方向旋转；而且，所有粒子的特征都可以刻画为费米子或玻色子。

-
- 量子力学告诉我们，物质和光由一些离散的单位“量子”组成。例如，光看上去是连续的，但实际是由被称作光子的离散量子组成的。量子是粒子物理学的基础。粒子物理学的标准模型解释了已知的物质和力，它告诉我们，所有物质和力归根到底都可以用粒子和它们的相互作用来解释。
 - 量子力学还告诉我们，任何粒子都有一个相关的波，即粒子的波函数。这个波的平方就是在某个特定位置找到一个粒子的概率，为了方便起见，我称它为概率波，即波函数的平方，概率波的值会直接给出概率。当我们后面探讨引力子的时候，会出现这样的波，引力子即传递引力的粒子；当我们探讨卡鲁扎-克莱因模式时，概率波同样很重要，卡鲁扎-克莱因模式是沿着额外维度，即在垂直于常见维度的方向上具有动量的粒子。
 - 量子力学与经典力学的一个显著差别在于：量子力学告诉我们不可能确定一个粒子的路线——你永远无法知道一个粒子从起点到终点的确切路径。这告诉我们，在粒子传递力时，我们必须将它可能采取的所有路径都考虑在内。因为量子路径可能包括所有互相作用的粒子，因此，量子力学作用可能影响质量和相互作用的强度。
 - 量子力学将所有的粒子分为玻色子和费米子。两种不同类型粒子的存在是标准模型结构的关键，对标准模型的延伸超对称也至关重要。量子力学的不确定性原理以及其狭义相对论的关系告诉我



-imension
探索大揭秘

们，我们可以用物理常数将粒子的质量、能量和动量，与具有那个能量的粒子所能经历的力或相互作用的最小区域尺度联系起来。

- 两种最常用的关系包括两种能量，即弱力能标和普朗克能标。弱力能标是250 GeV（吉电子伏），而普朗克能标要大得多——1 000亿亿GeV。只有在小于 10^{-17} 厘米范围之内，力才会对一个携有弱力能标的粒子产生可测量的影响。这一长度范围非常微小，但对原子核的物理过程和粒子获得质量的构造是非常重要的。
 - 尽管弱力长度已经非常微小，但普朗克长度还要远远小于它，只有 10^{-33} 厘米，这是力对携有普朗克能量的粒子能够影响的区域范围。普朗克能标决定了引力的强度，粒子必须具备这一能量，引力才会成为强力。
-



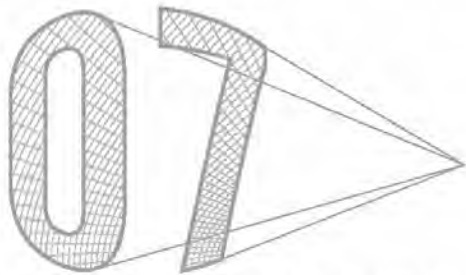
WARPED

第三部分

探秘基本粒子物理

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



物质的已知最基本结构

WARPED PASSAGES

你从不孤单，你从不会被疏远！即便你独自守候，只要呼唤陪伴，必有人到你身边誓死保护你！……当了“喷气机”，你就永远是“喷气机”！

里夫《西区故事》

豌豆公主的秘密

在所有故事里，最让阿西娜感到迷惑的是安徒生的《豌豆公主》。

故事讲的是一位王子欲寻觅一位真正的公主结婚，却一直没有找到。在徒然地寻找了几个星期后，有一天，在暴风雨中，碰巧有位公主路过王宫，请求躲避风雨。由此，不知不觉中，这位狼狈的公主就成了王后验证真正公主的试验对象。

王后准备了一张床，上面铺了无数层床垫，然后又铺了无数层羽绒被。在这厚厚的床垫和被褥底下，她放了一粒小小的豌豆。当晚，她把这位公主请进了这一精心准备的客房。次日一早，这位公主（她最终证明自己是真正的公主）抱怨，昨晚一夜都没合眼。整个晚上，她在床上翻来覆去，发现身上被硌得青一块紫一块——之所以这么不舒服，全都是因为那粒豌豆！王后和王子这回相信，这位公主有真正的王室血统，别人又怎么可能如此娇嫩呢？

阿西娜把这个故事寻思了一遍又一遍。她觉得，如果一个人只是安安静静地躺在床上，即便像公主那般娇嫩、敏感，要发现厚厚床垫下的一粒豌豆也未免实在太

离谱了。经过好几天的琢磨，阿西娜终于想出一个说得通的解释，她立即跑去告诉哥哥艾克。

通常的看法是，公主证明自己是真正的王室血统，是因为她的皮肤娇嫩，能够敏感地察觉到哪怕是埋在厚厚的床垫和被褥下的那粒小小的豌豆。但阿西娜不相信，提出了另外的解释。

阿西娜的解释是：王后走后，留下公主一个人在房间里。这下子，公主可撒了欢儿，她抛开平日的高贵矜持，什么礼仪，什么端庄，统统让位于活泼好动的孩子天性。公主一个人在屋子里，床上床下地跑啊、跳啊，可算闹了个够。终于把自己折腾累了，她想睡觉了。她一个高蹦上床，然后又“咚”地倒了下去，这么一来，那床垫被压得厉害，下面的豌豆便突起来，一下子把公主身上硌得青一块紫一块。阿西娜觉得这公主还是蛮令人佩服的，可更让她满意的，是自己对这个故事的新阐释。

在原子里发现亚结构，就如同公主发现豌豆一样，是一项非凡的成就。被称作“夸克”的粒子，是质子的组成成分，它在质子里所占的体积与质子相比，就如同一粒豌豆在厚厚的床垫下一样。一个1立方厘米的豌豆在一个2米×1米×1/2米的床垫里所占的比例是床垫体积的一百万分之一，夸克在质子里所占的比例与此没什么差别；而物理学家发现夸克的方式与阿西娜设计的公主活泼好动的发现方式也有些类似。安安静静的公主怎么也不可能发现被层层盖住的豌豆。同样地，如果不是将高能粒子打进质子去探索它的内部，物理学家怎么也不会发现夸克。

在这一章里，你可以像公主一样纵情跳跃，跳进粒子物理学的标准模型里。标准模型是描述物质的已知基本成分以及它们之间相互作用力的理论，^[28]它是一项辉煌的成就，展现了许多令人称奇、令人兴奋的成果。你不必记住所有细节，因为后面讲到它们时，我会重提所有这些粒子的名称或它们相互作用的性质，但标准模型是我很快将讲到的许多离奇的、额外维度理论的基础，而且，随着你对这些激动人心的最新成果的了解，会加深你对标准模型及其主要观点的认知，且

让你对物质的基本结构以及当今物理学家思考世界的方式产生更深入的了解。

电子与无处不在的电磁场

列宁在他的哲学著作《唯物主义和经验批判主义》里用了一个有关电子的比喻，他说“电子是不可穷尽的”，这里指我们用以解释它的那些层出不穷的理论观点。事实上，自从量子力学改变了我们的观点之后，我们对电子的理解如今已与20世纪初大为不同。

可是，从物理意义上来讲，与列宁的话正好相反：电子是可以穷尽的。就目前的发现，电子是基本的、不可再分的。电子，因为不存在“不可穷尽的”结构，对粒子物理学家来说，是标准模型里最容易描述的粒子。电子是稳定的，又没有更小的组成成分，因此，我们只要列出它的几个属性，包括质量和电荷，就可以完整地描述它的特点。电子会向电池的正极流动；移动的电子还会对磁力作出回应：电子在穿过磁场时，路径会发生弯曲。这两个现象都是电子带负电荷的结果，它使电子对电和磁产生回应。

电子

电子会向电池的正极流动；移动的电子还会对磁力作出回应：电子在穿过磁场时，路径会发生弯曲。这两个现象都是电子带负电荷的结果，它使电子对电和磁产生了回应。

19世纪以前，人们以为电和磁是两种互不相干的力，但在1819年，丹麦物理学家、哲学家汉斯·奥斯泰德（Hans Oersted）发现，移动的电流会产生磁场。由这一发现，他推想描述电和磁应该有一个统一的理论：它们一定是一枚硬币的两面。指南针会对闪电作出反应，这就印证了奥斯泰德的结论。

如今我们还在使用的经典电磁场理论创立于19世纪，它所立足的观察是：电和磁是相互联系的。“场”的观念对此理论至关重要，物理学家将弥漫于空间的量称为“场”，例如，引力场在任一点的值表示那一点引力的作用有多大。任何类型的场都是这样：场在任一位置的值，都表示场在那里的强度。

19世纪后半叶，英国化学家、物理学家迈克尔·法拉第提出了电磁场的概念。由于家境贫寒，他14岁就不得不辍学，但他却能作出这番具有革命性影响的物理研究，不能不令人称奇。那时他是一名订书匠学徒，他的师傅鼓励他读那些书并教育他，这无论对他还是对物理学史，都可谓幸事。

法拉第的观点是：电荷会在空间的所有地方形成电磁场，而电磁场又反过来作用于其他物体，无论这些物体在哪儿。但电磁场对带电物体的作用强度要依赖于它们的位置。场的值越大，其作用越强；值越小，则作用越弱。

把铁屑撒向磁铁周围，你就会看到磁场存在的证据：这些微粒会根据磁场的强度和方向自行排列。把两块磁铁靠近，你也能感觉到磁场的存在，在它们触碰之前，你就能感觉到它们的相互吸引或相互排斥。每块磁铁都会对弥漫于它们周围的场产生回应。

有一天，就在我即将结束攀岩时，我深刻体会到了电磁场的无处不在。当时，我和一个同伴正在科罗拉多靠近博尔德的一个山脊上，同伴是攀岩新手，却有着丰富的徒步旅行经验。一场电子风暴正悄然临近！因为不想让他紧张，我只是催促他动作快点，并没有告诉他绳子就要断裂，而那一刻他的头发也竖了起来。等安全到达山下后，我们愉快地回顾这次历险：这是一次难忘的攀登经历，总的来说非常幸运。同伴告诉我说，他当然知道我们当时所处的险境，因为我的头发也明显地竖了起来。电磁场不只存在于一个地方——就在我们的周围，它无处不在。

19世纪以前，还没有人以场的形式描述过电磁，人们习惯性地将这种力描述成超距作用。在上中学的时候，你可能学到过这种表达方式，它描绘了一个带电物体怎样对另一带电物体立即产生吸引或排斥，无论它在何处。这看起来似乎并不神秘，因为我们早就习以为常。可一个地方的东西会立即影响远处的另一个东西，这应该是非常奇特的，那这种作用是怎么传递的呢？

场和超距作用虽然听起来只是语义上的差别，但实际上是两个极为不同的概念。根据电磁场的解释，一个电荷对空间的其他区域不会立即产生影响，场需要时间来调整。一个移动的电荷会在周围形成一个场，然后（但会很快地）渗透到空间中去，只有在光（由电磁场形成）经过一定时间抵达后，物体才会感受到远处电荷的活动。因此，电磁场的变化不会超过光速。在空间任一点，只有当远处的电荷经过一定时间到达那一点后，场才能作出回应。

尽管法拉第的电磁场观点非常重要，但其意义还仅停留在启示作用上，没有很大的数学意义。也许是因为他所受的教育有限，数学并不是法拉第的强项。而另一位英国物理学家，詹姆斯·克拉克·麦克斯韦则融合了法拉第的观点创立了经典电磁场理论。麦克斯韦是一位杰出的科学家，他兴趣广泛：光与色、椭圆数学、热动力学、土星环、用蜜糖罐测量纬度以及猫头朝下坠落怎样做到四脚落地，同时还保持其斜向冲势的问题^[29]等都在他的涉猎范围之内。

麦克斯韦对物理学最大的贡献是描述了电磁场的方程组，运用这一方程组可以从电荷和电流的分布推导出电磁场的值。^[30]由此，他推想到电磁波的存在，这些电磁波可以是各种各样的电磁辐射，如电脑、电视、微波炉以及现代的许多便利设备都会产生电磁辐射。但麦克斯韦犯了一个错误，他与其他同时代的物理学家一样，把场的观点太过物质化了，以为场是由以太的振动产生的——正如我们看到

的，爱因斯坦否定了这一观点。但爱因斯坦还是将狭义相对论的起源归功于麦克斯韦：麦克斯韦的电磁场理论使他想到光速是不变的，由此引出了他的不朽之作。

光子，力的传递者

麦克斯韦的经典电磁场理论作出了许多成功的预言，但由于它的出现要早于量子力学，因此，它并没有包括量子作用。如今，物理学家以粒子物理学来研究电磁场力，粒子物理学的电磁场理论，除了包括麦克斯韦久经研究和验证的经典理论预言之外，还兼容了量子力学的预言。因此，相比经典电磁学，它是一个更为全面、更为精确的电磁场理论。事实上，电磁场的量子理论作出了精确的预言，经验证，其精确度令人难以置信地达到了一百万分之一。^[31]

光子

量子电磁场理论将电磁力归因于被称作光子的粒子交换，光子就是我们上一章讨论的光的量子。其运作方式是，一个电子发出一个光子，光子再传向另一电子，将电磁力传递给它后消失。通过它们的交换，光子就传递了一个力。

量子电磁场理论将电磁力归因于被称作光子的粒子交换，光子就是我们上一章讨论的光的量子。其运作方式是，一个电子发出一个光子，光子再传向另一电子，将电磁力传递给它后消失。通过它们的交换，光子就传递了一个力。这就好像是一封秘函，将信息从一地传递到另一地后，就立即被销毁了。

我们知道电磁力有时相吸，有时相斥：当携带异性电荷的物体相互作用时，它们相吸；当携带同性电荷，无论是同正还是同负时，都相斥。你可以把光子传递一种相斥的力看作两个冰球运动员在来回地传递一个球：运动员每次接到球后，会迅速带球从另一队员身

边滑开，穿过冰场；而相吸的力，则更像是两个新手在抛飞盘。两个滑冰运动员会越离越远；而初学抛飞盘的两个人，则会越靠越近。

光子是我们将遇到的第一种规范玻色子。规范玻色子是一种基本粒子，它的任务就是传递某种特定的力。“规范”（gauge）一词听上去挺吓人，19世纪末物理学家选用它，是因为它指的是两条路轨间的标准距离，而这比较切合物理情境。100年前，这个词远比现在常用。弱力玻色子和胶子是规范玻色子的另外两个例子，它们分别传递弱作用力和强作用力。

20世纪20年代末至40年代期间，英国物理学家保尔·狄拉克（Paul Dirac）与美国物理学家理查德·费曼和朱利安·施温格（Julian Schwinger），以及朝永振一郎（Sin-Itiro Tomonaga）发展了光子的量子力学理论，他们称这一新发展的量子理论分支为量子电动力学（QED）。量子电动力学包括了经典电磁理论的所有预言以及物理过程中的粒子（量子）作用，即由交换或生成量子化粒子所引起的相互作用。

QED预言了光子的交换将怎样产生电磁力，例如，图7-1示例的过程：两个电子进入相互作用区域，交换一个光子，然后受其所传递的电磁场力的影响，最终出现在其他路线（如运动速度和方向）。

场理论为图示里的每一部分都标定了数值，这样我们就能用它作出量子预言。这幅图就是费曼图的一个例子，费曼图以图画的形式描述了量子场论的相互作用（这些图以理查德·费曼的名字命名，费曼也为自己的创意感到自豪，他甚至在自己的小房车上都涂上了这些图标）。



图7-1 费曼图例。右边的费曼图有多种释义，其中一个如左图所显示（由下往上看），两个电子进入作用区域，交换一个光子，然后两个电子离开。这一图示还可以解释为电子-正电子湮灭。

但是，并非所有QED过程中的光子都会湮灭。除了一些虚幻的介子或是内部粒子^[32]之外，如那些引起电磁作用，然后瞬间湮灭的光子，还有一些真正的、外来的光子。它们是进入作用区域，然后离开的粒子。这些粒子有时会产生路线偏移，有时会转化为其他粒子。无论是哪种情形，这些进入或离开的粒子都是真实的物理粒子。

量子场论

量子场论是我们研究粒子的工具^[33]，它的基础便是能够产生或毁灭这些粒子的永恒的、无处不在的对象。这些对象即量子场论的“场”，这个名字受到了电磁场的启发。与经典的电磁场类似，量子场也遍布于时空。但量子场却发挥着不同的作用，它们会产生或吸收基本粒子。根据量子场论，粒子可以在任何地点、任何时间产生和毁灭。

量子场论

量子场论是我们研究粒子的工具，它的基础便是能够产生或毁灭这些粒子的永恒的、无处不在的对象。这些对象即量子场论的“场”，这个名字受到了电磁场的启发。与经典的电磁场类似，量子场也遍布于时空。但量子场却发挥着不同的作用，它们会产生或吸收基本粒子。根据量子场论，粒子可以在任何地点、任何时间产生和毁灭。

例如，一个电子或光子可以在空间中任何地方出现或消失。通过粒子的产生与毁灭，量子过程允许宇宙的带电粒子数量发生变化，每一个粒子的产生与毁灭都有其特定的场。在量子场论中，不仅电磁力，其他所有力和相互作用都是以场来描述的。场能产生新粒子，也能让已存在的粒子消失。

根据量子场论，粒子可被想象成量子场的激发态。一个没有任何粒子的状态是真空，它所含的场是不变的；而有粒子存在的状态所包含的场会发生相应的隆起和弯曲：场得到一个隆起就产生一个粒子；当它将这隆起吸收，场再归于稳定，粒子便被摧毁了。

产生电子和光子的场必须是无处不在的，这样才能保证在时空中的任一点，所有的相互作用都会发生。这一点很重要，因为相互作用只发生在局部，也就是说，只有同一地点的粒子才有可能参与相互作用。超距作用更像是天方夜谭，粒子没有超感觉（ESP）——它们必须通过直接接触才能发生相互作用。

当然，电磁场里的相互作用发生在并非直接接触的远距离粒子之间，但这是因为有能与相互作用的带电粒子直接接触的光子或其他粒子的帮助。在这种情况下，电荷看似立即对彼此产生了影响，这只是因为光速非常快。实际上，相互作用只发生在局部过程中。光子先是遇到一个带电粒子，然后再遇到另外一个。因此，就在带电粒子的那一确切位置，场生成，并毁灭了光子。

反粒子与正电子，从科幻进入现实

量子场论还告诉我们，每一个粒子都必然存在一个与它相反的对应物，称作反粒子。汤姆·斯托帕德（Tom Stoppard）在他的剧作《谍变》（*Hapgood*）中这样描述反粒子：“当一个粒子遇到它的反粒子时，它们会彼此消融，会转化成能量爆炸，这你能理解。”所有的

科幻迷都知道反粒子——它们是用来制造毁灭宇宙武器的东西，也是电影《星际迷航》中“进取”号航空母舰的驱动力。

反粒子

量子场理论还告诉我们，每一个粒子都必然存在一个与它相反的对对应物，称作反粒子。反粒子是粒子物理世界观里的真实组成部分。

这几种应用都是虚构的，但反粒子并非虚构。反粒子是粒子物理世界观里的真实组成部分。在场论与标准模型里，它们同粒子一样重要。实际上，反粒子就和粒子一样，只不过是具有相反的电荷。

保尔·狄拉克在创立量子场论描述电子时，第一次遇到反粒子。他发现，与量子力学和狭义相对论都协调的量子场论必须包括反粒子。他并非刻意地将它们添加进来，在他纳入了狭义相对论时，理论就自然出现了。反粒子是相对性量子场论的必然产物。

为什么说反粒子会随狭义相对论出现，这里大致解释一下：带电粒子能在空间里前后移动，那么根据狭义相对论，我们就可能天真地以为，这些粒子在时间中也能够前后移动。但就我们所知，无论粒子还是我们知道的其他任何东西，都不能在时间里向后流动。事实是带相反电荷的反粒子取代了逆时间倒流的粒子。反粒子产生了倒流粒子所能产生的效应，这样，即便没有逆流，量子场论的预言也能与狭义相对论互相协调。

假想一部电影的播放是带负电荷的大量电子从一点流向另一点。现在，想象倒着放映这部电影，那么负电荷会向后流动，（就电荷来说）这就等于正电荷在向前流动。正电子（电子的带正电荷的反粒子）就形成一个向前流动的正电子流，它们的表现就像是逆时间倒流的电子。

量子场论告诉我们，如果存在任何一种类型的带电粒子，如电子，就必然存在一个对应的带相反电荷的反粒子。例如，电子的电荷是-1，而正电子的电荷则是+1。除了电荷之外，反粒子与这个电子完全一样。质子的电荷也是+1，但质子要比电子重2 000倍，因此不可能成为它的反粒子。

正如斯托帕德说过的：当两者发生接触时，反粒子和粒子确实会互相消融，因为粒子与其反粒子的电荷相加总是为零。所以，当粒子遇上其反粒子时，它们会彼此湮灭。粒子和反粒子在一起，它们不带电荷，因此，我们由爱因斯坦的公式 $E=mc^2$ 可知，所有的质量都会转化为能量。

另一方面，如果有足够的能量，也可以让粒子与反粒子之间互相转换。在粒子加速器里，既可以毁灭粒子，也可以生成粒子。物理学家用粒子加速器来进行重粒子研究的实验，重粒子质量太大，在我们常见的物质里找不到。在对撞机里，粒子和反粒子相撞，彼此湮灭，由此会突发大量的能量，从而生成新的正反粒子对。

物质——尤其是原子，是由粒子而非反粒子构成的，因此，像正电子等反粒子，通常在自然界中并不存在。但在对撞机里，在宇宙的高温区域，甚至在医院里（人们用正电子发射断层扫描[PET]来发现癌症），它们都可以暂时生成。

格里·加布里埃尔斯（Gerry Gabrielse）是我在哈佛大学物理系的同事，他一直在杰弗逊实验室的地下室里制造反粒子。多亏格里及其他人的贡献，我们才精确地知道，尽管电荷相反，但反粒子确实和它们对应的粒子一样，有相同的质量和引力作用。只是反粒子很少，也不足以构成危害。科幻迷们尽可以放心，反粒子对建筑物的危害，远不及不断兴建新的实验楼和办公楼所带来的危害严重，因为这些建筑的兴建，总是伴随着大规模的环境破坏和噪声污染。

电子、正电子和光子是最简单、最容易得到的粒子，电磁力和电子成为物理学家最先理解的标准模型成分，并非出于偶然。但电子、正电子和光子不是仅有的粒子，而电磁力也不是仅有的力。

弱力

尽管在日常生活中我们可能注意不到弱作用力，因为它们非常微弱，但在许多核过程中，它至关重要。许多形式的核衰变都要归因于弱力；事实上，中子本身的衰变也要归因于弱力，核过程改变了原子核的结构，而通过这些过程，原子核里的中子数发生变化，释放出大量能量，这些能量可以用于核武器或核弹，也可以用于其他目的。

在图4-3和图4-4中，我列出了已知的粒子和除引力之外的力。^[34]图中未列出引力，是因为它与其他力有着本质的区别，必须区别对待。弱力和强力，尽管名字平凡，但其性质却非常有趣。后面两部分，我们就来看看它们究竟是怎样的。

弱力和中微子

尽管在日常生活中我们可能注意不到弱作用力，因为它们非常微弱，但在许多核过程中，它至关重要。许多形式的核衰变都要归因于弱力，比如钾-40的衰变（就在我们的地球上，它的衰变非常缓慢——平均大约要100万年，足以持续给地核供热）；事实上，中子本身的衰变也要归因于弱力，核过程改变了原子核的结构，而通过这些过程，原子核里的中子数发生变化，释放出大量能量，这些能量可以用于核武器或核弹，也可以用于其他目的。

弱力在重元素的生成中发挥了重要作用，重元素是在剧烈的超新星爆发时产生的。弱力与恒星（包括太阳）发光也有紧密的关系：它会激发连锁反应，使氢转化为氦。由弱力引发的核过程，有助于使宇宙组成不断演变。根据我们的核物理知识可以算出，大约10%的宇宙

史前氢被用作了恒星的核燃料（令人高兴的是，剩下的90%保证了宇宙短期内无须依赖于外来能源）。

尽管弱力非常重要，但科学家意识到它的存在却是不久之前的事。1862年，当时最受推崇的一位科学家威廉·汤姆逊（William Thomson，后来的开尔文爵士^[35]），因为不知道核过程是由弱力引起的（这很正常，因为那时弱力还未被发现），只是大致地估算了太阳和地球的年龄。汤姆逊的依据只有唯一的已知光源——白炽光，他推测太阳得到的能量使它的生命不可能超过3 000万年。

达尔文却不喜欢这一结果，通过估算位于英格兰南部的威尔德峡谷侵蚀形成所需要的年限，他得出的地球最小年龄更接近正确的结果。达尔文认为3亿年这个数字可能性更大，因为依据自然选择，地球才能有足够的时间出现如此大量的物种。

但所有人——包括达尔文自己，都以为汤姆逊是对的，因为他是那么一位负有盛名的物理学家。达尔文如此信赖汤姆逊的推导和声誉，以至于在他最新版的《物种起源》里，略掉了自己的估算时间。直到卢瑟福发现辐射的重要意义后^[36]，达尔文关于地球年龄的观点所遭受的质疑才得以澄清。不过，最终人们确定地球和太阳的年龄大约是45亿年，这远远地超过了汤姆逊和达尔文的估算。

20世纪60年代，美国物理学家谢尔登·格拉肖和斯蒂芬·温伯格（Steven Weinberg），以及巴基斯坦物理学家阿卜杜斯·萨拉姆（Abdus Salam），虽然都在独立地（且并非一定和谐地）工作，却共同创立了弱电统一理论。这一理论解释了弱力并提供了有关电磁力的新见解。^[37]

根据弱电统一理论，就像光子的交换传递电磁力一样，称作弱规范玻色子的粒子交换产生了弱力作用。弱规范玻色子有三种：带电的

两种， W^+ 和 W^- （ W 代表弱力，+号和-号是规范玻色子的电荷）；中性的
一种，称作 Z （因为其电荷为零）。

如光子的交换一样，弱规范玻色子的交换所产生的力也可以是相吸或相斥的，这取决于粒子的弱荷。弱荷在弱力中所发挥的作用，就如电荷在电磁力中所发挥的作用。只有携带弱荷的粒子才会经受弱力作用，而它们携带的特定弱荷会决定它们所经受的相互作用的强度和类型。

但是，在电磁力和弱力之间也有几处重要的差别，其中最令人吃惊的是弱力要区分左右，或者如物理学家所说的“宇称不守恒”定律（violates parity symmetry）。宇称不守恒定律意味着粒子与其镜像的表现互不相同。20世纪50年代，美籍华裔物理学家杨振宁和李政道提出了宇称不守恒定律，1957年，该定律被另一位华裔物理学家吴健雄所验证。杨振宁和李政道获得了当年的诺贝尔物理学奖。可奇怪的是，吴健雄这位唯一在标准模型构建中作出贡献的女性，却没有因其重大发现而获得诺贝尔奖。

有些宇称不守恒现象我们应该是熟悉的，例如，我们的心脏都在身体的左边。如果当初我们换种方式进化，让人的心脏长在右边，那我们也会像现在一样，认为所有的人体特征就应该是这样子的。心脏在左边还是右边，应该不会影响根本的生物进程。

1957年，吴健雄进行了这一实验，但在这之前的许多年里，物理定律（尽管不一定是物理对象）“显然”不会有任何手征偏好，毕竟没有理由必须这样。当然，引力和电磁力以及许多其他的相互作用力都没有这种区别。可弱力，自然界里一种基本的力，却要区分左右。尽管令人惊讶，可弱力确实打破了宇称守恒定律。

一种力为什么会选择这种旋向而放弃另一种呢？答案就在费米子自旋里。就如要旋紧螺丝，你必须顺时针旋转而不能逆时针旋转一样，粒子也有它的旋向性，以表明其自旋方向（见图7-2）。许多粒子，如电子和质子，会选择一种自旋方向：要么向左，要么向右。“手征”一词来源于希腊语的“cheir”，意思是“手”，在这它指的是两种可能的自旋方向。粒子可以是左旋的，也可以是右旋的。就像我们的指纹，有的向左旋，有的向右旋。

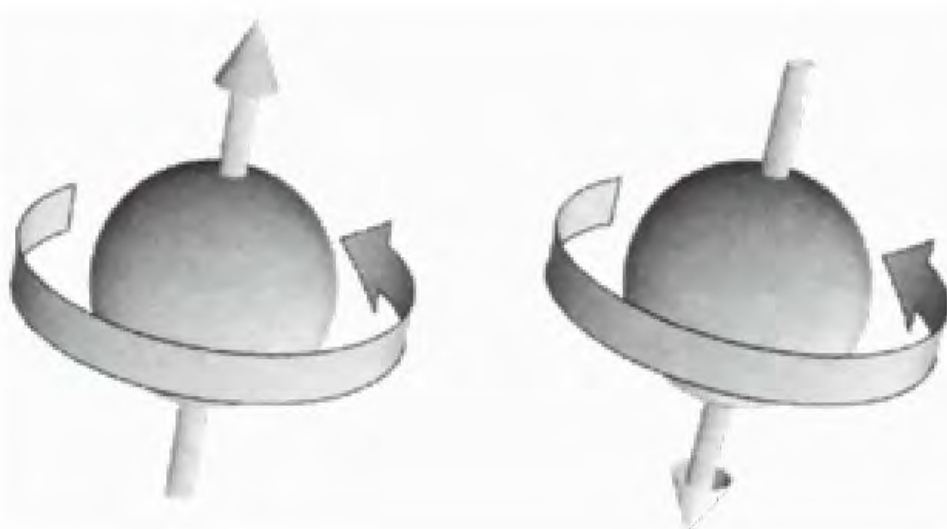


图7-2 夸克和轻子可以是左旋的，也可以是右旋的。

弱力打破了宇称守恒定律，是因为它会以不同方式作用于左旋粒子和右旋粒子，只有左旋粒子才经受弱力。例如，左旋电子会经受弱力，而右旋的则不会。实验清晰地显示了这些——这就是世界运行的方式，但为什么会这样，却没有一个直观的、力学上的解释。

你能想象一种力只作用于你的左手而不作用于你的右手吗？我所能说的就是，宇称不守恒定律虽然令人瞠目，却是久经验证的弱相互作用的属性，它是标准模型最引人入胜的特征之一。例如，中子衰变时出现的电子总是左旋的。弱相互作用打破了宇称性，因此，当我列

出全部的基本粒子和作用于它们的力时（见图7-6），就需要分别列出左旋粒子和右旋粒子。

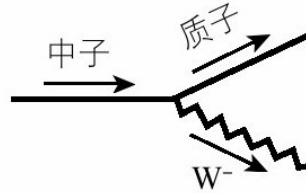


图7-3 中子与弱规范玻色子相互作用图。与一个W⁻规范玻色子的相互作用将一个中子转换成一个质子（中子里包含的下夸克转换成质子里的上夸克）。

虽然宇称不守恒似乎非常奇特，但它还不是弱力唯一的新奇属性。第二个同样重要的属性是：弱力能真正将一种粒子类型转化为另一种（却仍旧维持原来的电磁荷总量）。例如，当一个中子与一个弱规范玻色子相互作用时，就可能出现一个质子（见图7-3）。这与光子的相互作用完全不同，光子永远不会改变任何类型的带电粒子的净数量（即粒子数减去反粒子数），如电子数减去正电子数。为了对比，图7-4列举了一个光子与一个电子的相互作用，一个电子进入，而后变成另一个电子（同时附上了我们以前用过的简图）。一个带电弱规范玻色子与中子和质子的相互作用，使一个孤立的中子衰变成了一个完全不同的粒子。

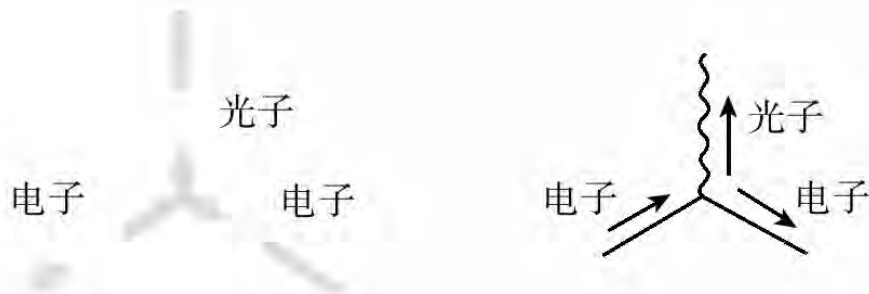


图7-4 光子-电子相互作用图。曲线指的是光子，如左边简图所示，电子进入，与光子相互作用，然后沿相互作用的顶点方向离

开。

但是，因为中子与质子有着不同的质量，且携带着不同的电荷，所以为了维持电荷、能量和动量守恒，中子衰变成质子必须加上其他粒子。结果是，当中子衰变时，它不仅产生了一个质子，还产生了一个电子和一个中微子^[38]，这一过程叫作贝塔衰变（beta decay，见图7-5）。

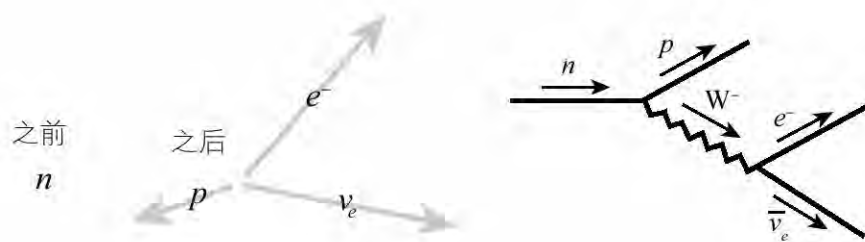


图7-5 贝塔衰变。贝塔衰变中，一个中子通过弱力作用衰变成一个质子、一个电子和一个反中微子。右图为这一过程的费曼图：一个中子转化成一个质子和一个虚拟的 W^- 规范玻色子， W^- 规范玻色子接着转化成一个电子和一个反中微子。

第一次观察到贝塔衰变时，没人知道还有中微子，它只通过弱力相互作用，而不通过电磁力。粒子探测器只能发现带电粒子或放出能量的粒子，因为中微子没有电荷且不会衰变，所以探测器找不到它，也就没人知道它的存在。

中微子

第一次观察到贝塔衰变时，没人知道还有中微子，它只通过弱力相互作用，而不通过电磁力。粒子探测器只能发现带电粒子或放出能量的粒子，因为中微子没有电荷且不会衰变，所以探测器找不到它，也就没人知道它的存在。

但如果没有中微子，贝塔衰变似乎就不能维持能量守恒。能量守恒是一个基本的物理定律，即能量既不会产生，也不会消亡——它

只会从一个地方传递到另一个地方。贝塔衰变不能维持能量守恒的假设引起了轩然大波，但许多负有盛名的物理学家，因为不知道中微子的存在，宁可作出这一激进的（也是错误的）论断。

1930年，沃尔夫冈·泡利指出了一条“亡命的出路”：一个新的中性粒子，^[39]这为持怀疑态度的人拯救科学开辟了一条道路。他的观点是，中子衰变时，中微子偷偷挟走了一部分能量。三年后，恩里克·费米将这一“小小的”中性粒子命名为“中微子”，并给出了坚实的理论基础。但中微子的主张看起来是那么的缺乏基础，以至于当时著名的科学期刊《自然》杂志拒绝刊登费米的论文，因为“它所包含的假设太过遥远，读者根本不会感兴趣”。

但泡利与费米的观点是对的，现在，物理学家都一致认可中微子的存在。^[40]事实上，我们现在知道中微子一直在我们身边“流淌”，在太阳的核过程中它与光子一同被释放出来，每秒钟都有上亿的中微子穿过我们，但它们的作用非常微弱，我们从未注意到。我们确定存在的中微子都是左旋的；右旋中微子要么不存在，要么很重——重到根本不可能生成，或者作用太过微弱。无论是哪种情形，总之，至今在对撞机里都没有生成右旋中微子，我们也从未见过它们。

因为我们对于左旋中微子远比右旋中微子更为确定，所以我在图7-6中分列左旋和右旋粒子时，只列出了左旋中微子。

现在，我们知道了弱力只作用于左旋粒子，而且能改变粒子类型。但为了更好地理解弱力，我们需要一个理论来预测传递力的弱规范玻色子的相互作用。物理学家最初发觉构建这一理论并非易事，他们需要取得重大的理论进展，然后才能真正理解弱力及其作用。

	夸克：经受强力				轻子		
第一代	上夸克 (左) 3 MeV	下夸克 (左) 7 MeV	上夸克 (右) 3 MeV	下夸克 (右) 7 MeV	电子中微子 (左) ~0	电子 (左) 0.5 MeV	电子 (右) 0.5 MeV
第二代	粲夸克 (左) 1.2 GeV	奇夸克 (左) 120 MeV	粲夸克 (右) 1.2 GeV	奇夸克 (右) 120 MeV	μ 子中微子 (左) ~0	μ 子 (左) 106 MeV	μ 子 (右) 106 MeV
第三代	顶夸克 (左) 174 GeV	底夸克 (左) 4.3 GeV	顶夸克 (右) 174 GeV	底夸克 (右) 4.3 GeV	t子中微子 (左) ~0	t子 (左) 1.8 GeV	t子 (右) 1.8 GeV
	左旋夸克：经受弱力				左旋轻子：经受弱力		

图7-6 标准模型里的三代。左旋与右旋夸克和轻子都分别列出。每一列包含的粒子都含有相同的电荷（不同味的粒子类型），弱力能够将第一列的元素转化为第二列，也能把第五列的元素转化为第六列。夸克经受强力，而轻子不经受强力。

问题出在弱力的最后一个奇异特征上：它在很小的距离内陡然下降，这个距离只有 10^{-16} 厘米。这使得它与引力和电磁力的作用极为不同，正如我们在第2章里看到的，那两个力的强度与距离的平方成反比。尽管引力和电磁力都会随着距离的增长而变弱，但它们不会像弱力这样突兀地下降。光子会在很远的距离里传递电磁力，而弱力的表现为什么会有这么大的差别？

显然，物理学家需要找到一种新的作用类型来解释像贝塔衰变这样的核过程，但尚不明了这种新的作用究竟会是什么。在格拉肖、温伯格和萨拉姆创立弱力理论之前，费米插入了一个理论，包括一个涉及光子、中子、电子和中微子这4种粒子的新型的相互作用。这一费米相互作用直接产生了贝塔衰变，而没有激起一个作为中介的弱规范玻色子。换句话说，这一相互作用允许质子直接转化成它的衰变产物——中子、电子和中微子。

然而，即便当时所有人都很清楚，费米的理论不可能成为对所有能量都起作用的正确理论。尽管它的预言对低能量下的情况是正确的，但对高能量下的情况却明显地错了。在高能量之下，粒子的相互作用要强烈得多。事实上，如果你错误地以为可以将费米理论应用于高能粒子，那么你会得出根本不合情理的预言。比如，粒子相互作用的概率会超过1，而这是根本不可能的，因为任何事情的发生频率都不可能超过永远。

虽然在解释低能和距离足够的粒子的相互作用时，基于费米相互作用的理论非常有效，但物理学家发现，如果要了解高能量现象，他们还需要对像贝塔衰变这样的过程有一个更为根本的解释。弱规范玻色子传递力的理论似乎在高能量领域更有效，但没人能解释弱力的作用距离为什么会这么小。

物理学家最终发现，这一短小的作用距离是弱规范玻色子质量非零的结果。在粒子物理学中，不确定性原理与狭义相对论暗含的关系会产生显著的作用。我在第6章末尾讨论过带特定能量（如弱力能量或普朗克能量）的粒子受外力影响的最小距离。由于能量与质量之间的狭义相对论关系 $E=mc^2$ ，有质量的粒子，比如弱规范玻色子，会自动地将质量与距离之间的类似关系包含进来。

尤其是，一个既定质量的粒子交换所传递的力，当质量很小时，距离稍长一点，力就会消失（这一距离也与普朗克常数成正比，与光速成反比^[41]）。这一质量与距离的关系告诉我们，弱规范玻色子的质量约为100 GeV，只能在 10^{-16} 厘米的范围内自动地将弱力传递给粒子，超出这一距离，由粒子传递的力就将变得极其微弱，我们根本探测不到。

弱规范玻色子的非零质量对弱力理论的成功是极为关键的，弱力只在很短的距离内发挥作用，而且非常微弱，距离稍长一点，几乎就

同不存在一样。从这一角度来看，弱规范玻色子与光子和引力子是不同的，后两者都是无质量的。因为光子与引力子——传递引力的粒子，携带能量和动量却没有质量，所以，它们能穿越更远的距离来传递力。

粒子没有质量，这一概念虽然听起来很特别，但从粒子物理学角度来看不足为奇。粒子没有质量，说明它们能以光速行驶（毕竟，光是由无质量的光子组成的），而且能量和动量总会遵守特定关系：能量与动量成正比。

然而，弱力的传递者却有质量，从粒子物理学角度来看，有质量的规范玻色子——而非无质量的——确实非常奇怪。铺平弱力理论道路的关键，就是要理解弱规范玻色子质量的来源，它使得弱力对距离的依赖关系与电磁力对距离的依赖关系非常不同。在第10章，我们将探讨使弱规范玻色子产生质量的机制——希格斯机制。当今粒子物理学面对的一个最大困惑就是要寻求一个根本理论——一个使粒子获得质量的精确模型（在第12章我们会详细讲述）。额外维度的魅力之一就是，它们有可能帮助解开这一谜题。

夸克和强力

我的一个物理学家朋友曾这样对我妹妹解释他所研究的“强力”（strong force）：“之所以叫作强力，是因为它的作用非常强大。”虽然我妹妹并不觉得这一解释足够明白，但强力确实名副其实。它是一种极度强大的力量，可以将质子的组成要素紧紧地束缚在一起。在正常情况下，这些要素是不可能分开的。强力与本书的后半部分并无直接关联，但为完整起见，在此我还是讲一些基本事实。

强力

强力由量子色动力学理论所描述，是标准模型里我们能够以规范玻色子交换来解释的最后一个力。它也是在20世纪才被发现的，强规范玻色子称作胶子，因为它们传递的力就像“胶”一样，能把作用的粒子紧紧地黏合在一起。

强力由量子色动力学（QCD）理论所描述，是标准模型里我们能够以规范玻色子交换来解释的最后一个力。它也是在20世纪才被发现的，强规范玻色子被称作胶子，因为它们传递的力就像“胶”一样，能把作用的粒子紧紧地黏合在一起。

20世纪50~60年代，物理学家相继发现了许多粒子，他们给每一种粒子都分别以希腊字母命名，例如， π 、 θ 和 Δ ，这些粒子被统称为强子——来源于希腊语的“hadros”，意为“胖的，重的”。

确实，强子比电子质量要大很多，几乎可以与质子质量相比，而质子质量要比电子大2 000倍。强子的庞大质量一直是个谜，直到20世纪60年代，物理学家默里·盖尔曼（Murray Gell-Mann）提出：许多强子并非基本粒子，它们本身是由被称作夸克的粒子所构成的。

盖尔曼的“夸克”一词来自詹姆斯·乔伊斯（James Joyce）的小说《芬尼根的守灵夜》（*Finnegans Wake*）里的诗句：“冲马克叫三声夸克，他一定没有从吼叫声中听出什么，他所得到的都一定都偏离中心。”而就我的推论来说，这与物理学里的夸克实在没有什么相干，只除了两件事：它们有三个；它们一样令人费解。^[42]

盖尔曼提出有三种夸克^[43]，现在它们被称作上夸克、下夸克和奇夸克，并对应于可能束缚在一起的众多夸克组合，其中会有大量强子的存在。如果他的观点正确，强子应正好符合预言的类型。就如以前新物理原理提出时常有的情况一样，盖尔曼在提出他的建议时，并不真正相信夸克的存在。不过，他的提议是非常大胆的，因为科学家当时只发现了几个被预言的强子。因此，当隐匿的强子相继被发现时，

夸克假设得到了证实。对盖尔曼来说，这真是一个伟大的胜利。由此，他获得了1969年的诺贝尔物理学奖。

即便物理学家一致同意强子是由夸克组成的，但直到夸克假设提出9年之后，强子物理学才以“强力”的形式得到解释。强力是最后一个被理解的力，这确实令人奇怪，其原因很大一部分是由于它的力量太过强大了。现在我们知道，强力如此强大，以至于经受强力的粒子如夸克，总是被紧紧地束缚在一起而难以分离——经受强力的粒子不会没有约束地自由闲逛，因此对它们的研究就变得很困难。

每种夸克类型里又有三类，物理学家饶有趣味地用颜色来标识它们，有时称它们为红、绿、蓝。而这些色夸克总会与其他夸克或反夸克束缚在一起，形成色中性组合（color-neutral combinations）。

在这些组合里，夸克与反夸克的强力“色荷”会互相抵消的情形，这就好比白色光中各种颜色互相抵消的情形。^[44]色中性组合有两种：稳定的强子结构要么包含一个夸克和一个反夸克，两者形成一对；要么包含三个自行紧紧束缚在一起的夸克（没有反夸克）。例如，在 π 子里，是一对夸克与反夸克；而在质子和中子里，则是三个被束缚在一起的夸克。

在强子里，夸克间的强力色荷互相抵消，这与在原子里质子电荷与带负电的电子电荷互相抵消类似，但你可以任意使一个原子电离，而分离如质子和中子等强子却非常困难。它们被强力胶子紧紧地束缚在一起——胶子更恰当的名称应该是“疯狂胶子”^[45]，因为它们的束缚状态实在难以打破。

夸克

夸克可以被看作一个大的、糊状质子里的坚硬的一点——它嵌在质子或中子里，就像一粒豌豆藏在床垫里。正如在床上弹跳的公主被豌豆硌青了一样，实验者会向其中射进一个高能电子，它发出一个光子，而光子从夸克

上直接弹开，这看上去与光子从一个大的、蓬松物体里弹开大不相同。就像卢瑟福的阿尔法粒子从一个坚硬的原子核上弹开会与从一个分散的正电荷上弹开看起来大不相同一样。

现在我们就回到夸克的发现上，这就与阿西娜修正解释的比喻完全一样了。质子和中子包含的都是三个夸克的组合，其中，与强力相关的电荷互相抵消了。质子包含两个上夸克和一个下夸克——不同的夸克类型有不同的电荷。因为上夸克的电荷是 $+2/3$ ，而下夸克电荷为 $-1/3$ ，所以质子的电荷是 $+1$ ；而一个中子有一个上夸克和两个下夸克，所以它的电荷是 0 （ $-1/3$ ， $-1/3$ 和 $+2/3$ 的和）。

夸克可以被看作一个大的、糊状质子里的坚硬的一点——它嵌在质子或中子里，就像一粒豌豆藏在床垫里。正如在床上弹跳的公主被豌豆硌青了一样，实验者会向其中射进一个高能电子，它发出一个光子，而光子从夸克上直接弹开，这看上去与光子从一个大的、蓬松物体里弹开大不相同。就像卢瑟福的阿尔法粒子从一个坚硬的原子核上弹开会与从一个分散的正电荷上弹开看起来大不相同一样。

在斯坦福线形加速器中心（SLAC）进行的弗里德曼·肯德尔·泰勒的深度非弹性散射实验，通过追踪记录这一作用而发现了夸克。实验显示了电子从质子上散射的表现，由此提供了夸克真实存在的第一手实验证据。由于这一发现，杰瑞·弗里德曼（Jerry Friedman）、亨利·肯德尔（Henry Kendall）（他们是我在麻省理工学院的同事）以及理查德·泰勒（Richard Taylor）共同获得了1990年的诺贝尔物理学奖。

夸克在高速碰撞里产生时，还没有被束缚形成强子，但这并不意味着它们是独立的——它们总会有其他夸克和胶子的陪伴，使其在强力之下最终呈现为中性。夸克从不会是自由的、没有“随从”的物体，许多相互强烈作用的粒子会护卫着它。粒子实验捕捉到的也不会

是一个独立的夸克，而是一系列由夸克、胶子组成的，差不多向同一方向运动的粒子。

这一队朝着同一方向运动的、由夸克和胶子组成的粒子合称为“喷射流”（jets）。一个高能的喷射流一旦形成，就会像一条绳子一样永不消失。当你截断一条绳子，你只会得到两条绳子。同样地，当相互作用将喷射流断开时，这些片段只会形成新的喷射流：它们永远不会分离成单个的、独立的夸克和胶子。估计斯蒂芬·桑德海姆（Stephen Sondheim）在为《西区故事》（*West Side Story*）写《喷射机帮之歌》（*Jet Song*）的歌词时，一定不会想到高能粒子对撞机，但他的歌词却令人钦佩地如此符合强烈作用的粒子喷射流：高能的、强烈作用的粒子总在一起，“你从不会孤立……总有人保护你”。

已知的基本粒子

本章描述了四种已知力的三种：电磁力、弱力和强力，另一种力——引力非常微弱，不会对粒子物理学的预言产生可实验观察的影响。

但我们还没有完成标准模型粒子的介绍。我们依据它们的电荷以及旋向来分辨它们，正如我们前面描述过的，左旋和右旋粒子会带有不同的弱荷。

粒子物理学将这些粒子分作两类：它们要么是夸克，要么是轻子。夸克是经受强力的基本费米子；轻子是不经受强力的费米子，电子和中微子都属于轻子。“轻子”一词源于希腊语“leptos”，意为“小的”或“精细的”，指电子的微小质量。

奇怪的是，除了电子和上夸克、下夸克这些对原子结构至关重要的基本粒子之外，还有另外一些粒子，它们虽然很重，却与我们介绍过的粒子有着相同的电荷。所有的最轻的稳定夸克与轻子都有它们相对应的更重的翻版，没有人知道它们为什么会存在，也没人知道它们存在的目的是什么。

μ 子是首先在宇宙射线中被发现的粒子。当第一次意识到，它只不过是电子的一个更重的形式（要重200倍）时，物理学家拉比（I. I. Rabi）问道：“谁要它来的？”尽管 μ 子像电子一样带负电，但它比电子重，而且可以衰变成电子。也就是说， μ 子是不稳定的（见图7-7），很快会转化成电子（及两个中微子）。就我们所知，它在地球上没什么用，是无关紧要的。那它为什么存在？这是标准模型里的众多谜题之一，我们希望科学进步最终能够解答。

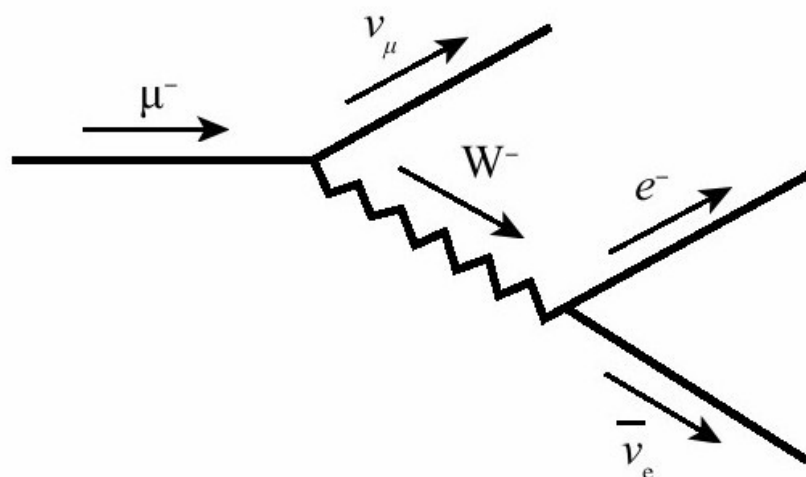


图7-7 在 μ 子衰变中， μ 子转化成了一个 μ 子中微子和一个虚拟的 W^- 规范玻色子， W^- 再接着转化成电子和一个电子中微子。

事实上，标准模型的所有粒子每组都有三个携带相同电荷的复本（见图7-6）。每套复本被称作一代或是一族。

第一代粒子包括一个左旋和一个右旋的电子，一个左旋和一个右旋的上夸克，一个左旋和一个右旋的下夸克，以及一个左旋的中微子。第一代粒子包含了原子所有的稳定组成部分，即所有物质都是由第一代稳定粒子组成的。

第二代和第三代粒子会产生衰变，且不会在通常的已知物质中出现，它们并非第一代粒子的绝对翻版：虽然与第一代的对应粒子电荷相同，却更重。只有在高能粒子对撞机里才能找到它们，但科学家至今不清楚它们存在的目的。第二代包含一个左旋和一个右旋的 μ 子，一个左旋和一个右旋的粲夸克，一个左旋和一个右旋的奇夸克，以及一个稳定的左旋 μ 子中微子。第三代包含一个左旋和一个右旋的 t 子，一个左旋和一个右旋的顶夸克，一个左旋和一个右旋的底夸克，一个左旋 t 子中微子。有着相同电荷的一组粒子的同样复本，即每一代里的一个成员，都被称作这一粒子类型的“一味”。

由图7-6可以看出，尽管在盖尔曼首次提出时，只有三味已知夸克，但现在我们已发现了六味：上类型三味，下类型三味——每一代里都有一味。除了上夸克本身以外，上类型里还有另外两味携带相同电荷的夸克——粲夸克和顶夸克；同样地，下类型夸克里也有三味——下夸克、奇夸克和底夸克。 μ 子和 t 子是电子的较重翻版，属于“轻子”。

物理学家一直在努力，想弄清楚为什么存在三代粒子以及粒子为什么具有特定质量。这是标准模型面对的主要问题，也是激发当前研究的动机。与他人一起，我的整个职业生涯都在致力于此项研究，而至今，仍在孜孜以求地探寻着这些问题的答案。

重味夸克相对于它们的轻味明显要重许多，尽管在1977年我们就发现了第二重的夸克——底夸克，但最重的顶夸克却直到1995年才显

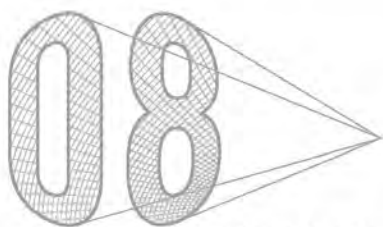
露出其行踪。下一章的主题是两个粒子实验，其中包括引人注目的顶夸克的发现实验。

- 标准模型里包含了引力之外的作用力以及经受这些力的粒子。除了众所周知的电磁力之外，原子核里还有两种作用力：强力和弱力。
- 弱力提出了标准模型仍未解答的最主要谜题：另外两种力都是通过无质量的粒子来传递的，而传递弱力的规范玻色子却有质量。
- 除了传递力的粒子之外，标准模型里还包含经受这些力的粒子，它们分作两类：经受强力的夸克和不经受强力的轻子。
- 在物质里发现的轻夸克和轻子（上夸克、下夸克和电子）并非仅有的已知粒子。另外还存在较重的夸克和轻子：上、下夸克和电子都有它们对应的两个较重版本。
- 这些重粒子是不稳定的，它们会衰变成较轻的夸克和轻子。粒子加速器实验能生成它们，并显示出这些较重粒子与那些熟悉的、稳定的轻粒子所经受的是同样的力。
- 每一组包含的一个带电轻子、一个上类型夸克、一个下类型夸克的粒子被称作一代，一共有三代。每代里的同类型粒子相继加重，这些粒子种类被称作“味”，有三味上类型夸克、三味下类型夸克、三味带电轻子、三味中微子。
- 在后面的章节，我不会用到特定类型的夸克和轻子的细节和名称，但我们需要知道这些“味”和“代”，因为它们强烈约束着粒子的



属性。它们会给我们提供重要的线索，让我们知道标准模型之外的物理会受到什么约束。

- 在这些约束中，最重要的是，带有相同电荷的夸克和轻子的不同“味”很少会互相转换，我们排除能轻易换“味”的粒子理论。后面我们会看到，这对超对称破缺模型及标准模型的其他发展提出了很大挑战。
-



高能粒子加速器的惊人新发现

WARPED PASSAGES

哪怕历经千难万险，我一定会找到你。

金发女郎乐队 (Blondie)

他的猎物迟早会出现

在梦中，艾克再一次遇到了那个量子侦探。这次，侦探非常清楚地知道自己要找什么，而且非常清楚地知道它在哪里，他所要做的就是等待，如果判断正确，他的猎物迟早会出现。

找到重粒子并非易事，但要揭示标准模型的基本结构并最终解释宇宙的物理构成，那么我们必须这么做！我们对粒子物理学的了解，大部分来自高能粒子加速器实验，它们首先将快速运动的粒子束加速，然后使它们与其他物质相碰撞。

在粒子对撞机里，被加速的粒子束与被加速的反粒子束相撞，在它们相撞的极小区域内会产生大量能量，这些能量接着被转化成在自然界里难以被找到的重粒子。宇宙大爆炸时，宇宙温度极高，所有粒子大部分都存在。但自那以后，唯一能找到已知最重粒子的地方，就只有在高能粒子对撞机里了。从理论上讲，只要有足够的能

量，即由爱因斯坦 $E=mc^2$ 所给定的能量，对撞机就能够生成所有的粒子与反粒子对。

但是，高能物理的目的不仅仅是寻找新的粒子。高能粒子对撞机实验会告诉我们以其他任何方式观察不到的自然的基本定律——这些定律作用的领域太过微观，以至于我们根本不能直接看见。高能实验是探索发生在微小距离尺度的相互作用的唯一方式。

本章讲的两个对撞机实验，对证实标准模型的预言以及限定物理理论的有效范围都非常重要。这两个实验本身都令人难忘，但它们还使我们体会到：物理学家将来探索新现象（如额外维度）时，会面临什么样的困难。

顶夸克，众里寻他千百度

顶夸克

虽然顶夸克不是已知物质原子的组成部分，但没有它，标准模型就不完善，因此，自20世纪70年代开始，物理学家就坚信它的存在。可是直到1995年，人们才探测到它。顶夸克显然很重——远远重于其他所有已探测到的夸克。经过20年的追寻，人们终于追踪到了它的踪迹，这才发现它的质量几乎是质子质量的200倍。

顶夸克的探寻过程，充分展现了在对撞机里寻找粒子的困难。那时，对撞机的能量还不足以制造它，而实验者发挥了聪明才智，千方百计地应对了这一挑战。虽然顶夸克不是已知物质原子的组成部分，但没有它，标准模型就不完善，因此，自20世纪70年代开始，物理学家就坚信它的存在。可是直到1995年，人们才探测到它。

那时，对于顶夸克，物理学家已徒然地寻找了多年。标准模型里第二重的粒子——底夸克，已在1977年被发现，它的质量是质子的5倍。虽然当时物理学家都以为顶夸克马上就要现身了，而且实验者们

也争先恐后地要找到它，并宣布这一盛事，但令人惊讶的是，一次次的实验却均告以失败，人们以40倍、60倍甚至是100倍于产生质子所需要的能量进行对撞实验，还是未能找到它。这说明，顶夸克显然很重——远远重于其他所有已探测到的夸克。经过20年的追寻，人们终于追踪到了它的踪迹，这才发现它的质量几乎是质子质量的200倍。

因为顶夸克如此之重，狭义相对论隐含的关系告诉我们，只有在极度高能的对撞机里才能生成它。高能量总是需要一个大型的加速器，这在技术上难以实现，而且耗资巨大。

最终生成顶夸克的对撞机是费米实验室的Tevatron，它位于伊利诺伊州离芝加哥将近50公里外的巴达维亚草原。这一对撞机最初设计的能量，距离生成一个顶夸克所需的能量还差得很远，但工程师和物理学家们进行了多次改进，大大地提高了它的性能。Tevatron凝聚了多次改进的成果，1995年，它的运行能量大大增加，并进行了多次对撞，这是最初的机器根本无法做到的。

Tevatron现在仍然在运行，坐落于费米实验室。该实验室是一个加速器实验中心，为纪念物理学家恩里科·费米而得名，于1972年投入使用。

我第一次到费米实验室时感到非常有趣，在那里有野生的玉米、鹅，更奇特的是，甚至还有野牛，但除此之外，这里便再无奇趣可言，甚至可以说是呆板、枯燥的。电影《反斗智多星》(*Wayne's World*)就在费米实验室南9公里的奥罗拉取景，如果你熟悉这部电影，就知道费米实验室的周边环境了。好在物理学家善于鼓舞人心，总会让你高兴起来。

Tevatron的得名是因为它加速的质子和反质子能量都要达到1 TeV（太电子伏），等于1 000 GeV（吉电子伏），这是迄今为止所有加速器所能达到的最高能量。Tevatron产生的高能质子和反质子束在环形机里循环，每隔3.5微秒在两个对撞点发生对撞。

粒子束和反粒子束路径交叉的地方会发生有趣的物理过程。因此，在两个对撞点上有两个独立的实验团队，并分别安装了探测器，其中一个叫CDF（费米实验室对撞机探测器），另一个叫Do（这是放置探测器的质子与反质子对撞点的名称）。这两个实验旨在广泛地搜寻新的粒子和物理过程，但在20世纪90年代早期，物理学家梦寐以求的就是找到顶夸克——两个实验团队都想最先发现它。

多数重粒子是不稳定的，很快就会发生衰变。衰变发生时，实验要寻找的就是粒子衰变的产物，而不是粒子本身。例如，顶夸克会衰变成一个底夸克和一个W子（传递弱力的带电规范玻色子）；而W子也会衰变成轻子或夸克。因此，寻找顶夸克的实验要找的就是底夸克协同其他夸克或轻子的组合。

然而，粒子出现的时候可不会挂着姓名标签，因此，探测器只能通过它们的特别属性来辨认，如它们的电荷或参与的相互作用，而且，这些属性分别需要探测器里不同的部分来记录。分置于CDF和Do的两个探测器也都分成了几个部分，每一部分记录不同的特征。

有一部分叫作追踪仪，它探测的是从原子中电离的电子在其轨迹里所留下的带电粒子；另一部分叫热量计，可以用于测量粒子通过时所释放的能量；探测器里还有其他一些组成部分，分别用来辨识有着其他特别属性的粒子，如底夸克，它在衰变之前的寿命比其他大多数粒子都长。

探测器一经捕捉到信号，就会通过大量电线和放大器来传输这一信号，并记录结果。但是，并非所有的探测结果都值得记录，当质子和反质子相撞时，很少会生成有趣的粒子，如顶夸克或底夸克；更多时候，碰撞只能产生轻夸克和胶子，甚至有些时候根本不会产生任何有价值的东西。事实上，在费米实验室里，为产生一个顶夸克所进行的对撞，有千百亿次，但它们根本不含顶夸克。

面对如此大量的无用数据，没有一个计算机系统足够强大，以至于能够找到有意义的对撞。因此，实验总会包含一些感应器，这种装置包含一些软、硬件设施，只允许记录一些可能有价值的结果，其作用就像是夜总会雇用的保安，专门驱逐那些捣乱的人。在CDF和Do两个实验站里的感应器，将实验所需筛选的对撞数量缩减到了万里挑一——这仍是一项巨大的挑战，但相比从上百亿次里选择已经容易多了。

一旦信息被记录下来，物理学家就会尽力去解读，并重建任何有意义的对撞里会出现的粒子。因为总是会有很多次对撞、产生很多的粒子，而信息却是有限的，所以重建一次对撞的结果是一项艰巨的任务，它也拓展了人们的聪明才智。或许在未来的几年里，数据分析就会有更进一步的发展。

在1994年之前，CDF的几个工作小组都发现，有几次对撞很像顶夸克（见图8-1），但他们并不确信。尽管CDF小组不敢肯定那年他们已发现了顶夸克，但1995年Do和CDF两个实验站都确认了这个发现。我在Do里的朋友达伦·伍德（Darren Wood）向我描述了最后编委会里的紧张气氛——他们要在会上完成数据分析和报告实验结果的论文。会议进行了一整夜，他们累了时，就趴在桌子上小睡一会儿。然后，一切继续。

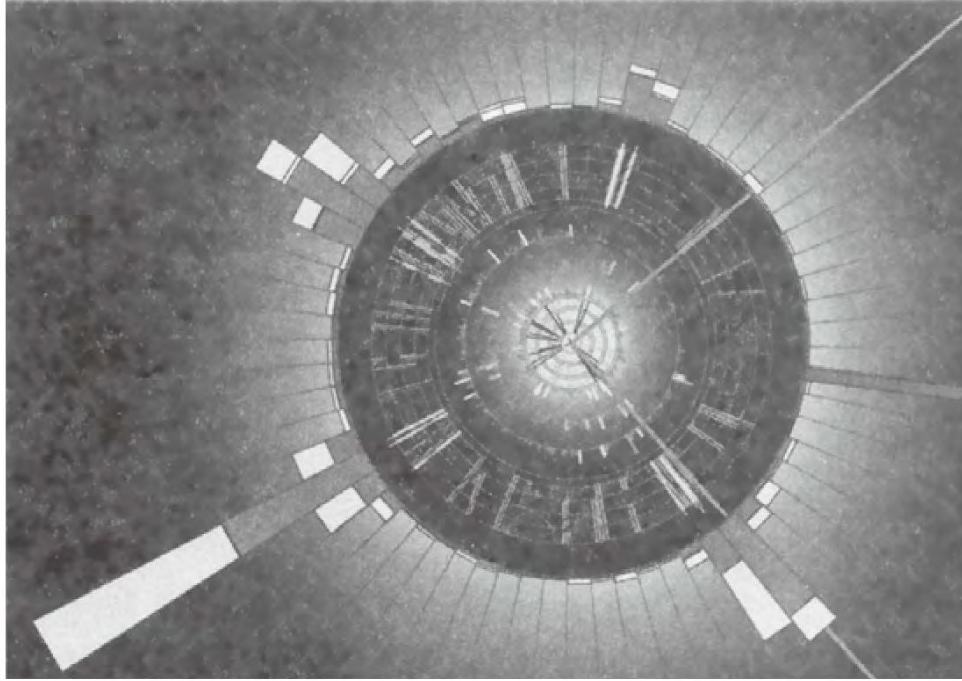


图8-1 由Do记录的一次顶夸克探测实验的结果。该实验探测顶夸克以及同时产生的反顶夸克的衰变产物。右上角的直线是一个 μ 子，它直达探测仪的外围；4个长方形的组块是4个“喷射流”；向右的直线是中微子消失的能量。

Do和CDF两个实验团队为顶夸克的发现共同作出了贡献：一个从未出现过的粒子终于被制造出来，这一新发现的粒子与其他已确定的粒子一起，加入到了标准模型的队列里。现在，我们已发现了许多顶夸克，因此已确切地了解了顶夸克的质量和属性。我们甚至担心，将来高能粒子对撞机会产生太多的顶夸克，以致顶夸克反倒会成为混淆和干扰其他粒子发现的东西。

新的理论肯定会被发现。很快我们将看到，为什么说标准模型的未解问题告诉我们，只要对撞机的能量比现在再高一点，就有可能出现新的粒子和物理过程。LHC实验就是要寻找超出标准模型的结构证据。如果这些实验成功，回报将是丰厚的——它是对所有物质基本结构的更好的理解，要完成这一艰巨的任务，既需要高能的、多粒子的碰撞，也需要新观点。

标准模型的精确验证

现在，让我们从伊利诺伊平原转移到多山的瑞士——欧洲核子研究中心CERN的驻地。验证标准模型预言的实验很多，但最为壮观的还要数1989—2000年间在CERN的大型电子-正电子对撞机（LEP）里进行的实验。

CERN的中心驻地设在欧洲，其主入口离法国边境非常近，分隔两国的边防警卫亭就在大门外，许多居住在法国的CERN员工每天要穿越边境线两次。穿越边境时，他们很少遇到麻烦——除非他们的车不符合赫尔维蒂标准（Helvetic Standards），那样，瑞士就不允许进入了；再有一种危险的情况就是，科学家心思太过专注于工作而犯规。这是一个同事曾经历过的：过边境时，他还在想着黑洞的问题，警卫请他停车检查，而他长驱直入，根本没有停下。

更引人注目的是，CERN与费米实验室的周边环境大不相同。CERN位于欧洲最高峰勃朗峰脚下，毗邻美丽的汝拉山脉（Jura mountains），离霞慕尼山谷（Chamonix）很近。这是一个美丽的山谷，两边山上冰雪覆盖，一直能延伸到路上（虽然由于全球气候变暖，现在路上已少见冰雪）。尽管城里的冬天常常阴云密布，但在CERN，幸运的物理学家们在很多日子里都沐浴着阳光，皮肤黝黑地度过冬季，因为他们在附近的山上就能滑雪、溜冰或徒步旅行。

CERN创建于第二次世界大战之后，诞生于新兴的国际大合作的氛围中。最初的12个成员国包括：联邦德国、比利时、丹麦、法国、希腊、意大利、挪威、荷兰、英国、瑞典、瑞士和南斯拉夫（于1961年退出）。随后，奥地利、西班牙、葡萄牙、芬兰、波兰、匈牙利、捷克斯洛伐克共和国以及保加利亚也加入进来。参与CERN活动的观察员

国包括印度、以色列、日本、俄罗斯联邦、土耳其和美国。CERN是一个真正的跨国团体。

与Tevatron一样，CERN也有着许多傲人的成就。1984年，卡罗·卢比亚（Carlo Rubbia）和西蒙·范德梅尔（Simon Von DerMeer）被授予诺贝尔物理学奖，因为他们设计了最初的CERN对撞机，并发现了弱力规范玻色子。这一成就，打破了美国在粒子发现上的垄断。CERN还是万维网WWW、HTML（超文本标记语言）和http（超文本传输协议）的诞生地，创造这些的英国人蒂姆·伯纳斯-李（Tim Berners-Lee）正是CERN的员工。他创建Web是为了让分散于各国的实验者能够通过链接立即得到信息，而且这样能使计算机共享数据。当然，网络的影响力已远远地超出了CERN——科学研究得到如此广泛的现实应用是人们不曾预见的。

几年后，CERN就会成为一条纽带，联结起一些最为振奋人心的物理成果。LHC就坐落于此，它现在的能量达到了Tevatron能量的7倍。由LHC所作出的发现必然是全新的、质的飞跃，LHC要寻找并且很有可能找到标准模型的未知的物理基础，它们或是证实、或是排除我在本书中所描述的这些模型。尽管LHC位于瑞士，但它是真正的国际大合作，所LHC实验正在世界各地进行着。

20世纪90年代，在CERN的物理学家和工程师们建造了令人难以置信的LEP（大型电子-正电子对撞机）。它是一个Z玻色子“工厂”，生产了上百万的Z玻色子——传递弱力的三种规范玻色子之一。通过研究上百万的Z玻色子，LEP的实验者（以及在SLAC——加利福尼亚的斯坦福线形加速器中心的人们）能对Z玻色子的属性作出详尽的测量，以空前的精确度验证标准模型的预言。具体描述这些测量会离题太远，但我会很快让你明白它们的精度是多么令人惊讶！

验证标准模型的基本假设非常简单，标准模型预言了弱规范玻色子的质量以及基本粒子的衰变和相互作用，我们可以通过检查所有这些量的相互关系是否符合理论的预言来验证弱相互作用理论是否内在一致。如果有新的理论，其中的新粒子和新的相互作用在弱力能标上有显著作用，那就会出现改变弱相互作用预言的新因素，使它们与标准模型的值有所不同。

因此，标准模型以外的模型所作出的有关Z玻色子属性的预言会稍不同于标准模型的预言。20世纪90年代早期，为使预言得到验证，所有人都在使用其他模型来预言Z玻色子的属性，但所使用的方法都烦琐得令人难以置信。这些方法让人很难领会，而概括它的文件页数多到我根本不愿携带。当时，我在加州大学伯克利分校做博士后，1992年夏天，在我参加费米实验室的一个暑期项目时，我突然间得出结论：不同的物理量之间的关系是不可能烦琐到需要这么长的篇幅来概括的。

我和当时在费米实验室的博士后米奇·戈登（Mitch Golden）一起创立了一种更为简洁的方法来阐释弱相互作用的实验结果。我们证实，只需在标准模型里添加3个能概括所有非标准模型的贡献量，就能系统地囊括新的（虽然尚未看见的）重粒子效应。我花了几周时间厘清它，终于在一个周末的紧张工作之后得到了一个答案。这一发现让我感到无比欣慰，Z-工厂让测量的所有过程都可以有机地联系起来，我和米奇都觉得我们描绘了一幅更加清晰的图像，将理论与测量联系起来，效果是令人满意的。但作出发现的并不只有我们，SLAC的迈克尔·佩斯金（Michael Peskin）和他的博士后竹内（Takeuchi）几乎同时也做了类似的工作，随后，又有许多人也跟上了我们的脚步。

但真正的成功故事却是关于标准模型的LEP验证，这一验证精确得令人难以置信。我们不讲细节，只通过两件事来看看它们的精确是多么令人钦佩。

第一件事是有关发现正电子和电子对撞的确切能量，实验者需要知道这一能量来确定Z玻色子质量的确切值。他们必须将影响能量值的所有可能考虑在内，但即使把所有想到的可能都计算进来，他们仍注意到，当在某些特定时间进行测量时，能量似乎还是有轻微的涨落。究竟是什么引起了这一变化？结果是日内瓦湖的潮汐，这太不可思议了！湖的水位随着潮汐和那年的大雨有涨有落，这就影响到周围的地形，而电子和正电子在对撞机里穿行的距离就产生了细微的变化。一旦将潮汐作用也考虑在内，Z玻色子质量会随着时间变化的假象便被排除了。

第二件事是每隔一段时间，电子和正电子的位置似乎就会稍稍发生偏移，这表明对撞机里的磁场发生了变化。现场的一位工作人员发现，这一变化与TGV的通过密切相关。TGV是穿行于日内瓦和巴黎之间的高速列车。显然，一些与法国直流电相关的电力尖峰轻微地扰乱了加速器。一位在CERN工作的巴黎物理学家阿兰·布隆代尔（Alain Blondel）给我讲述了这一故事最有趣的部分：实验者们有绝好的机会来证实这一假设，因为TGV的许多员工都是法国人，难免会罢工。因此，罢工日时实验者们便享受到了一个“无尖峰日”。

-
- 研究粒子物理学最重要的实验工具是高能粒子加速器，高能对撞机是使粒子碰撞到一起的粒子加速器。如果能量足够大，对撞机就能产生出因为质量太大而不能在我们周围存在的粒子。
 - Tevatron是当今运行的最为高能的对撞机。
 - 位于瑞士的LHC，能量是Tevatron的7倍，可以检验粒子物理学的许多模型。



-imension
探索大揭秘



啦，啦啦啦啦，啦啦啦啦，啦啦啦啦啦啦啦啦啦。

头脑简单乐团 (Simple Minds)

谁才是闯祸那一只

阿西娜打开笼子，让她的3只猫头鹰出来放放风。但不幸的是，艾克那天也把汽车的折叠车顶放了下来。这可倒好，好奇的猫头鹰就飞了进去。一只最为捣蛋的猫头鹰把车里面踩得乱七八糟，还划破了一道小口。

看到这片狼藉，艾克大发雷霆，他冲进阿西娜的房间，警告她看好那些猫头鹰，以后不许再闯祸。阿西娜抗议说，她的猫头鹰一向都很乖，她只要看住最捣乱的那只就行了。但这会儿，3只猫头鹰已然都飞回了笼子，艾克和阿西娜谁也分不清究竟是哪一只闯了祸。

标准模型运行得很好，但这只是因为理论中的夸克、轻子和弱规范玻色子——在带弱荷的物质间传递弱力的带电的W子和Z子，都有质量。当然，基本粒子的质量对宇宙里的所有东西都是至关重要的：如果物质真的没了质量，就不可能形成真实、喜人的事物，而宇宙的生命与结构也自然是无稽之谈。但在有关力的最简洁的理论里，弱玻色子和其他基本粒子似乎都应该是无质量且以光速行驶的。

你可能会觉得奇怪，既然力的理论会偏好零质量，那么为什么不是任何质量都可以呢？那是因为力的最基本的量子场论不能容忍这一点，它明确禁止标准模型基本粒子质量的任何非零值。标准模型的一大成果就是，它找到了解决这一问题的方法并形成了一个理论，其中的粒子具有我们观测所确定的质量。

下一章，我们就将探索让粒子获得质量的机制，即著名的希格斯机制。但在本章，我们会探讨一个重要的话题——对称。对称和对称破缺帮助确定了宇宙如何从一个混沌状态发展成我们现在所见的复杂结构。希格斯机制与对称，尤其是与对称破缺密切相关。若想理解基本粒子是如何获得质量的，首先要熟悉这些重要的概念。

化繁为简，完美的对称

对多数物理学家来说，对称是一个神圣的字眼。我们可能会想到其他许多社团也都非常看重对称，像基督教的十字架、犹太教的大烛台、佛教的达摩轮、伊斯兰教的新月以及印度教的曼荼罗，它们都呈现出了对称的形态（见图9-1）。说某样东西对称，就是说你可以以某种方式操作它，比如，让它旋转、用镜子将它反射或将它的组件对调，而操作后，新的结构与原来的相比不会有所改变。例如，将大烛台上两根相同的蜡烛对调，你不会看出有任何差别，而从镜子里照出的十字架与它本身也完全相同。

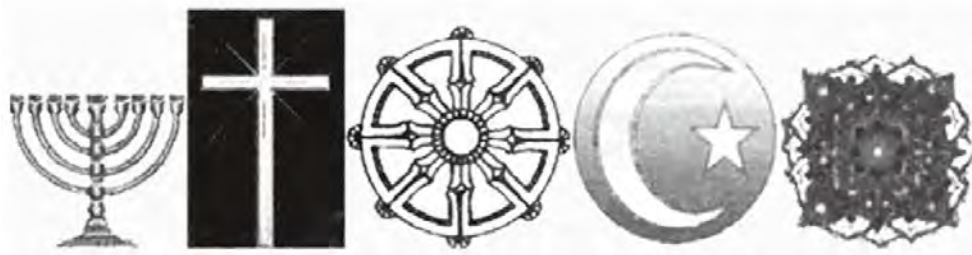


图9-1 宗教中的对称。犹太教的大烛台、基督教的十字架、佛教的达摩轮、伊斯兰教的新月、印度教的曼荼罗都呈现出了对称性。

无论数学、物理还是现实世界，当存在对称时，我们就可以做一些转换，而不改变其原始形态。如果有人在你背转身时，对调了一个系统的组成，或让它旋转，或用镜子将它反射，但当你转过身再看它时却未发现任何差别，那么这个系统就是对称的。



对称通常是指一种静态属性：比如，一个十字架的对称与时间无关。但物理学家常喜欢将对称描述为假想的对称变换——即我们可以在一个系统中应用这个操作，而不会改变它的可观察属性。比如，我不说大烛台的蜡烛都是一样的，而是说如果将两支蜡烛对调，大烛台仍不会呈现出其他变化。我并不一定为了证明对称而真的去调换蜡烛，但是假设我真正地调换了，我也看不出任何改变。为了方便起见，有时我将以这种方式来描述对称。

我们熟悉的不仅是科学和宗教象征中的对称，世俗世界的艺术也是如此。在许多绘画、雕塑、建筑、音乐、舞蹈和诗歌中都体现了对称。这一方面，伊斯兰艺术也许表现得最为淋漓尽致，在其建筑和装饰艺术上，无论细节还是整体都广泛地使用对称，见过泰姬陵的人都可证实这一点：无论从哪一角度望去，这一建筑都是一样的；再加上陵前水池中的倒影，就像有两座泰姬陵互相辉映；即使是树木的种植，都保持了这

一建筑的对称。我参观的时候，看到一个导游正在指点人们看一些对称的地方，就请他告诉我其他的对称，结果为了看全所有的对称，我要从一些非常奇特的角度来观赏这一建筑，甚至爬到旁边的碎石堆上。

在日常应用中，人们常常将对称与完美等同起来。当然，对称之所以吸引人，是因为它很规律和整齐。对称还有助于我们的学习，因为无论是空间还是时间上的重复，都会在我们大脑里形成难以磨灭的印象。正是由于大脑对于对称的天生知觉及纯粹的审美意识，我们才会让自己的周围布满对称。

但对称不仅仅出现在艺术和建筑里，在自然界没有任何人为的干涉，仍旧存在对称。因此，在物理学中，你常会遇到对称。物理的目的是将一些彼此独立的量联系起来，以依据观察作出预言。在这一意义上，对称是自然的产物。当物理系统存在对称时，相比没有对称，你就可以依据更少的观察来描述这一系统。比如，如果两个物体有相同的属性，在测量了其中一个的行为之后，我就得知了决定另一物体行为的物理定律——两个物体相同，它们的行为也必然相同。

在物理学中，一个系统中对称变换的存在意味着存在某些既定的程序，可以重新安排这一系统，而不会对其物理属性留下可测量的改变。^[46]例如，如果一个系统具有旋转和平移对称，这是空间对称中人们熟知的两个例子，那么物理定律在任何方向、任何位置都同等适用。旋转对称和平移对称告诉我们，当你挥棒击球时，你面朝什么方向、站在什么位置都无关紧要，只要你用同样大的力，棒球就会以同样的方式飞出。即使你变换方向，或是在另一房间或另一位置做重复的测量，所有实验都会产生同样的结果。

在物理定律中，对称的重要性是怎么强调都不为过的。许多物理定律，例如麦克斯韦的电动力学定律、爱因斯坦的相对论都深深地依赖于对称，通过利用各种各样的对称，我们常常可以简化以理论作出物理预言的任务。例如，对行星在椭圆轨道的运动、宇宙的引力场（或多或少地属于旋转对称）、电磁场中粒子的表现以及其他许多物理量的预言，只要我们将对称考虑在内，其运算就要简单得多。

物理世界的对称并不明显，但即便不那么明显或只是理论工具，对称也常常能大大简化物理定律的形式。我们很快要讨论的力的量子理论，也不例外。

内部对称，与空间无关

物理学家通常将对称分为不同的几类，我们最熟悉的可能要数空间对称——即在外部世界能够移动或旋转物体的对称变换。这种对称，包括我刚提到的旋转和平移对称，都告诉我们，无论一个系统指向什么方向或位于什么位置，物理定律都是普适的。

内部对称

内部对称变换以不可察觉的方式交换或混合不同的对象。事实上，我已经给出了一个内部对称的例子——大烛台上蜡烛的对调，内部对称认为两支蜡烛是相同的，这一陈述针对的是蜡烛，而非空间。

现在，我要说的是另外一种对称——内部对称。空间对称告诉我们，物理学同等对待所有方向和位置；而内部对称则告诉我们，物理学以相同方式作用于性质不同但效应不可分辨的对象。内部对称变换以不可察觉的方式交换或混合不同的对象。事实上，我已经给出了一个内部对称的例子——大烛台上蜡烛的对调，内部对称认为两支蜡烛是相同的，这一陈述针对的是蜡烛，而非空间。

但是，传统的大烛台既有空间对称又有内部对称：蜡烛是相同的，意味着存在内部对称；而将烛台绕其中心旋转 180° ，蜡烛仍保持不变，这意味着它还存在空间对称。但即便不存在空间对称，内部对称仍可存在，例如，即使马赛克镶嵌形成的树叶形状并不规则，但你仍可以任意地对调其中两片相同的绿色马赛克。

内部对称的另一例子是，你可以任意地调换两个相同的红色玻璃球。如果你两只手里都抓着一个相同的玻璃球，那么哪个球在哪只手中就都无所谓了。即使你在两个球上标明了“1”和“2”，我是否已调换了两只球，你仍无法知道。注意，玻璃球的例子不像大烛台和马赛克的例子那样，与空间位置密切相关：内部对称关注的是物体本身，而非它们在空间的位置。

粒子物理学处理的是关于不同的粒子类型的内部对称，它有点抽象。这种对称将粒子以及生成粒子的场都看作可以对调的，正如当你滚动两个相同的玻璃球或是将它们弹到墙上时，它们的表现会完全相同，有着相同电荷和质量的两种粒子类型也应遵循相同的物理定律。描述这一现象的对称，就是味对称（flavor symmetry）。

在第7章我们看到，“味”是有着相同电荷的3种不同的粒子类型，每一代里都有一味。例如，电子和 μ 子是两味带电轻子，这意味着它们有着相同的电荷。如果在我们生活的世界里，电子和 μ 子的质量也相同，那么这两个就完全可以对调了，如此一来，就会有一种味对称。据此，电子和 μ 子在其他所有粒子和力面前就会有完全相同的表现。

在我们的世界里， μ 子比电子要重，因此味对称并不完全贴切，但是对于某些物理预言，质量的不同并不十分重要，因此，如 μ 子和轻子等有着相同电荷的轻粒子的味对称，对计算还是非常有用的。有时，利用即使并不完美的对称也能帮助我们算出足够准确的结果，粒子间的质量差异通常很小（相比能量和大质量物体而言），对预言不会形成可测量的差别。

但现在对我们最为重要的对称类型，是与力的理论相关的对称，它是精确的对称。这一对称也是粒子间的内部对称，但它比我们刚才说的味对称还要抽象一点。下面的例子与这一特定类型的内部对称更为相似：回想你上中学时，物理课上、电影院里或美术课上见过的三盏射灯，通常是一红、一绿、一蓝，它们同时发光会产生白光。如果将三盏灯的位置调换，新的布置仍旧会产生白光。光束从哪一位置射出并不重要，因为我们关注的是它的最终结果——在这种情况下，白光进行内部对称变换（对调不同的光）不会产生任何可见的影响。

现在我们看到，这一对称与力的对称非常相似。因为在这两种情形里，你都不能看到所有的状况，即只能看到混合的灯光，而不能看到所有变化，因此，灯光的放置就会呈现对称。如果你能看到灯光，就会知道它们已经做了调换。正如我们早先提到的，色与力的类似，就是描述强力的量子色动力学中“色”字的由来。

1927年，物理学家弗里茨·伦敦（Fritz London）和赫尔曼·外尔（Hermann Weyl）经实验证实：**最简洁的量子场论对于力的内部对称的描述非常类似于以上例子中射灯的对称。**力与对称的联系非常微妙，因此，在课本之外你根本见不到它。即便不能完全理解这一联系，也并不影响你跟上有关质量问题的讨论——包括随后几章的希格斯机制和等级问题——因此，如果愿意，现在你可以直接跳到下一章。但如果你有兴趣知道内部对称在力的理论和希格斯机制里发挥了什么样的作用，就请继续看下去。

对称和力，排除虚假偏振

电磁力、弱力和强力都要涉及内部对称（引力与空间和时间对称相关，因此必须分开来讨论），如果没有内部对称，量子力学有关力的理论就会成为无迹可循的一团乱麻。为理解这些对称，我们首先要探讨规范玻色子的极化。

可能你熟悉光的极化（或偏振）概念，例如，偏光太阳镜只允许垂直偏振的光通过，摒除水平偏振的光，这样的结果是光不会那么刺眼。在这种情况下，偏振就是与光相关的电磁波振动的独立方向。

量子力学将波与每一个光子都联系起来，每个光子都会有不同的偏振，但是并非所有可以想见的偏振都可以被接受。光子向任一个方向行驶时，波的振动方向只能垂直于它的运动方向。波的这一表现就

像是海浪一样，海浪也只是在竖直方向上振动，因此，当海浪经过时，我们会看到浮标或小船在水面上上下下浮动。

与光子相关的波会以垂直于其运动方向的任何方向振动（见图9-2）。实际上，存在无数这样的方向：想象垂直于运动直线的一个圆，你会看到波可以沿着圆的任一半径方向振动，而由圆心辐射出来的方向是无数的。

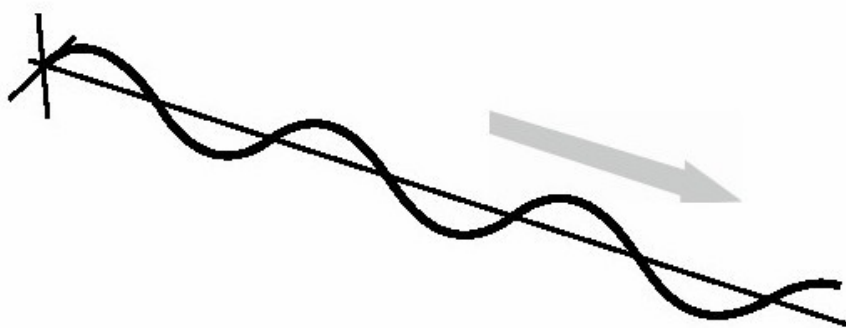


图9-2 一个横波相对于它的运动方向垂直地振动。在这一例子中，波在向右运动，而振动是在上下的方向上。

但是，在物理中描述这些振动时，我们只需两个独立的垂直振动就可以了。以物理学术语来说，它们叫作横向偏振（transverse polarizations）。这就好比你用x轴和y轴来标志一个圆，从圆的中心，无论你画哪一条线，它总会在某一点上与圆相交——有特定的x、y的值，因此，你只需两个坐标就能使这一点与其他所有点区分开来。同样地，尽管会有无数的方向垂直于一个波的运动方向，但所有这些偏振光的方向都可以归于一个集合，用两个互相垂直的方向就可以全部描述。

重要的是，理论上还可能存在第三种偏振方向，即波会沿着它的运动方向振动（假设它存在，我们就只能称它为纵向偏振）。

例如，声波就是这样传播的，但不存在这样的光子偏振。三种可以想见的独立的偏振方向，自然界中只存在两种。光子永远也不会沿着它的运动方向和时间的方向振动：它只会垂直于其运动方向振动。

即便我们单从独立的理论思考中看不到纵向偏振的谬误，但量子场论也会助我们将它排除。如果物理学家错误地将三种偏振方向考虑进来，使用这样一个力的理论进行计算，那么这一理论对其属性的预言根本没有意义。例如，它所预言的规范玻色子相互作用的频率会高得荒谬。事实上，它预言的相互作用频率会超过永远——即在时间上大于100%。任何理论，作出这样不合情理的预言显然是错误的，因此自然界和量子场论都明确表示：这种非垂直方向的偏振是不存在的。

不幸的是，物理学家创建的关于力的最简单理论包含了这一虚假的偏振。这并不奇怪，因为要适用于所有光子的理论不可能将某个沿着特定方向的特定光子的信息都包含在内。而没有这一信息，狭义相对论就不能识别所有的方向。一个理论若要保持狭义相对论对称（包括旋转对称），则需要三个（而非两个）方向来描述光子振动的所有方向，在这样的描述里，光子会在空间的所有方向上振动。

但我们知道，事实并非如此。对任何特定的光子，它的运动方向是选定的，而在另一方向的振动是不被允许的。但因为所有光子都有各自不同的运动方向，你可不想为每个光子都构造一个不同的理论，我们需要的是无论光子向哪个方向运动都适用的理论。尽管你可以尝试创造一个不包括虚假偏振的理论，但是保留旋转对称而以其他方式消除那个虚假的极化方向，岂非更为简单而明晰？物理学家的目的就是简化自己的工作，他们发现，把虚假的纵向偏振包含在内，运用一个额外要素来将好的、有物理意义的预言与虚假的分开，这样的量子场论最为有效。



内部对称就在这里进入了理论图景。内部对称在力的理论中所发挥的作用，就是为了排除不需要的偏振所引起的矛盾，让我们不致损失狭义相对论的对称。无论是独立的理论思考还是实验发现，都告诉我们运动方向上的偏振并不存在，内部对称是滤除这一虚假偏振最简单的方法。

它们把偏振分为好的和坏的两类，即符合对称的一类和不符合对称的一类。它的作用方式有点太过专业，不太好解释，但我会打一个比方，让你对它有个大致的了解。

假设你有一台制作衬衫的机器，可以制作长袖和短袖两种衬衣。但不知什么原因，机器的制造者却忘了加进一个控件来保证左衣袖和右衣袖的长度是一样的。有一半时间你作出的衬衣是好的——两只都是长袖或两只都是短袖，可另一半时间你做的衬衣全是残次品：一只袖长，一只袖短。不幸的是，你只有一台机器。

你有两种选择：要么把机器扔掉，一件衬衣都做不成；要么留着这台机器，做一半好衬衣、一半残次品。但你也不会很困惑，因为需要扔掉哪些衬衣

很明显，只有左右对称的才能穿。即便仍然用这台机器制作衬衣，只保留左右对称的那些，这样你的穿戴仍会很得体。

与力相关的内部对称完成的就是类似的使命，它提供了一个有用的标记，将那些在理论上我们可能观察到的量（我们想要保留的偏振）和那些不应出现的量（沿着运动方向的虚假偏振）区分开来。这就像计算机里的垃圾邮件过滤器，它要找的是垃圾邮件的明显特征，并将它与有用邮件区分开来。同理，内部对称“过滤器”会将保持对称的物理过程与那些虚假的过程区分开来，这使清除像垃圾邮件一样的偏振更为简单。如果它们出现，就会打破内部对称。

这一对称方式与我们前面讨论的三色射灯的例子非常相似，在那个例子里，我们观察到的只是三种颜色一起形成的白光，而不能看到单独的灯光的颜色。同样地，在有关力的理论里，只有一些特定的粒子组合符合内部对称，也只有这些组合才会出现在物理世界。

与力相关的内部对称，摒除了所有涉及虚假偏振的过程——即沿着运动方向的振动（这在自然界里并不真实存在）。就如不符合左右衣袖对称的衬衫很容易被挑出扔掉一样，不符合内部对称的虚假偏振也会被自动删除，不会干扰计算。一个规定了正确的内部对称的理论，就会排除可能出现的虚假偏振。

电磁力、强力、弱力都是通过规范玻色子来传递的：电磁力通过光子，弱力通过弱规范玻色子，强力通过胶子。每一类型的规范玻色子都有一个相关的波，在理论上，它可以在任何方向上振动；但实际上，它只会在垂直方向上振动。因此，这三种力，每种力都需要它自己特定的对称来排除传递力的规范玻色子的虚假偏振。因此，与电磁

力相关的有一对称，与弱力相关的有一对称，而与强力相关的又是另一对称。

在力的理论里，内部对称似乎非常复杂，但要形成一个有用的力的量子场论，并作出正确预言，这是物理学家所知道的最简单的方法。要区分真实的和虚假的偏振，只有通过内部对称。

我们刚刚探讨的内部对称对力的理论是非常关键的，它们也是希格斯机制的基础，这一机制告诉我们标准模型的基本粒子怎样获得质量。下一章，我们将不再需要内部对称的细节，但我们会看到对称（及对称破缺）是标准模型的基本组成部分。

规范玻色子、粒子和对称

到现在为止，我们只讨论了规范玻色子的对称作用，但是，与力相关的对称变换却不仅仅对规范玻色子发挥作用。规范玻色子会与经受它所传递的力的粒子发生相互作用：光子与带电磁力电荷的粒子相互作用；弱规范玻色子与带弱荷的粒子相互作用；胶子与夸克相互作用。

由于这些相互作用，只有将规范玻色子及与它们相互作用的粒子同时改变时，才能保持内部对称。我们来看一个比方：如果旋转只作用于一些物体而不作用于其他物体，那么就不能算是对称变换。如果你只旋转奥利奥饼干的上层，而不旋转其他部分，那么你只会把它分开；你只有将整个奥利奥饼干同时旋转，才不会看出改变。

同样的原因，只改变传递力的规范玻色子而不改变经受力的粒子，这种转换将无法维持对称。排除胶子虚假偏振的内部对称，也要求夸克和胶子一样可以对调。事实上，调换夸克与调换规范玻色子的对称变换都是同一对称变换。保持对称的唯一办法就是把两者混在一

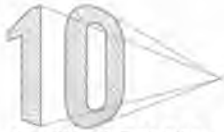
起，就如要保持奥利奥饼干不变一样，唯一的办法就是将它同时整个旋转。

本书中最为吸引我们的力是弱力，与弱力相关的内部对称将3种弱规范玻色子同等对待，它还将一些粒子对看作对等的，如电子和中微子、上夸克与下夸克等。这一弱力对称变换会将三种弱规范玻色子对调，也会把这些粒子对对调。至于胶子和夸克，只有当所有东西都一起对调时，才有可能保持对称。

-
- 对称告诉我们两个不同的结构在什么情况下会有相同的表现。在粒子物理学中，对称的用途是禁止某些相互作用：那些不能维持对称的相互作用。
 - 对于力的理论，对称非常重要，因为力的最简单有效的理论包括了与每种力相关的对称。这些对称排除了我们不想要的粒子，也排除了这一理论可能作出的关于高能粒子的错误预言。



-imension
探索大揭秘



基本粒子的质量来源：有趣的自发对称破缺和希格斯机制
UNIVERSITY OF CHINA PRESS

某一天早晨，锁链终将会打破。

艾瑞莎·富兰克林 (Aretha Franklin)

警察可不会看到保时捷的风驰电掣

更严格的限速法令使得远距离开车成了艾克的噩梦，他多么渴望能够随心所欲地放开速度飞驰啊，可是几乎大约每跑不到一公里他就会被警察逮到。对那些常见的、大众化的车，警察是从不愿费心去管的，他们只会骚扰像他这种开着鲜亮的、涡轮发动机车的人。

为了表示顺从，艾克只开很短的距离，因为这样，他就可以完全避开警察了。在离他出发点方圆不到一公里的区域内，警察从不干涉，而他就能肆意地开快了。尽管出了他周围的社区，没人见识过保时捷的风驰电掣，但在他家周围，那可是有目共睹的。对称非常重要，但宇宙通常并不能呈现完美的对称，稍微有点不完美的对称使世界变得更为有趣（且更有条理）。对我来说，物理研究最有趣的一面就是寻找一些联系，让对称在一个不对称的世界里变得有意义起来。

当对称并不完美时，物理学家称其为对称破缺。尽管对称破缺常常很有趣，但从审美角度来看，它可能不那么吸引人。它可能会损失（或降低）系统或理论潜在的美感。泰姬陵虽那么讲究对称，但它的对称也并非绝对完美：原计划在对面要再建造一座陵墓，以保持一个完美的四面旋转对称，但建造者的后人因为节俭还是放弃了这一计划，而是在原来的基础上加了一个偏离中心的陵墓，这就稍稍损失了其本来应有的美感。但幸运的是，对于颇具审美情趣的物理学家来

说，对称破缺甚至比绝对的对称还要完美和有趣。绝对对称往往是枯燥的，如果蒙娜丽莎展现的是一个对称的微笑，那肯定不会有现在的艺术效果。

对称破缺

当对称并不完美时，物理学家称其为对称破缺。对于颇具审美情趣的物理学家来说，对称破缺甚至比绝对的对称还要完美和有趣。尽管物理学家非常看重和崇尚对称，我们仍不得不在对称的理论和不对称的世界间寻找一种联系。

在物理学中，就如在艺术中一样，简洁并不一定是终极目标。现实生活 and 世界很少是完美的，你能说出的所有对称几乎都是破缺的。尽管物理学家非常看重和崇尚对称，我们仍不得不在对称的理论和不对称的世界间寻找一种联系。虽然最好的理论崇尚对称理论的典雅，但同时也会兼容必要的对称破缺，以作出符合我们现实世界的预言。我们的目的是让理论更为丰富，有时甚至是更为完美，而不致损害其原本的优雅。希格斯机制的概念，就是这样一个成熟的、优雅的理论观点，它所依赖的现象正是自发对称破缺（这是我们下节讨论的内容）。苏格兰物理学家彼得·希格斯提出的这一机制，使标准模型的粒子——夸克、轻子和弱规范玻色子获得了质量。

如果没有希格斯机制，则所有的基本粒子都不能具有质量；而粒子具有质量的标准模型，如果没有希格斯机制，在高能领域就会作出不合情理的预言。希格斯机制的神奇之处就在于，它既然让你得到了蛋糕，就可以让你享用它：粒子获得了质量，但在高能情况下，当有质量的粒子遇到问题时，它们的表现又像是没有质量一样。我们会看到，希格斯机制既允许粒子具有质量，又使它们能够在有限的范围内自由行驶，就像艾克的车一样，驶出不到一公里他会被警察拦住，但在有限的距离内，他可以自由驰骋。这就解决了高能问题。

尽管希格斯机制是量子力学最完善的观点之一，给出了所有基本粒子质量的基础，但它还是有点抽象。由于这一原因，除了专家，大多数人对它并不了解。但即使不明白希格斯机制的细节，也不会影响你理解我后面讨论的观点（如果你愿意，可以跳到后面的“探索大揭秘”部分），但本章确实给我们深入理解粒子物理学和支撑当今粒子物理学理论发展的思想（如自发对称破缺）提供了一个机会。

作为一个额外收获，对希格斯机制的更多了解还会让你知道一个关于电磁场的神奇见解，这是在20世纪60年代，人们正确理解了弱力和希格斯机制后才发现的。以后，当我们探索额外维度模型时，对于希格斯机制的理解会让那些新近观点的潜在优势变得有意义起来。

没有永远的对称

在描述希格斯机制之前，我们首先要看一下自发对称破缺，这是一种特殊的对称破缺，也是希格斯机制的核心。自发对称破缺在宇宙中许多我们已知的属性里发挥了重要作用，且在我们即将探索的所有事情中都可能发挥作用。自发对称破缺不仅在物理中比比皆是，而且在我们日常生活中也普遍存在。自发对称破缺是物理定律仍旧维持的对称，但现实世界的事物排列却做不到这样。当一个系统不能维持它本应呈现的对称时，就会出现自发对称破缺。或许最好的解释方法还是举几个例子。



首先设想一张圆形餐桌，桌边围坐了几个人，桌上的杯子都摆在两人中间。那么，人们该用哪个杯子呢？是右边的，还是左边的？这不好说。按礼仪，我该用右边的。但礼仪规范也是随

意确定的，实际上左边和右边的杯子作用都一样。

可是，一旦有人选择了杯子，对称就被打破了。选择的契机并不一定是系统的组成部分，在这个例子中，它是其他因素——口渴。但如果一个人自动选择了用左边的杯子喝水，那么，旁边的人也会用左边的，结果是，桌上的所有人都用左边的杯子喝水。

在有人拿起杯子之前，一直存在着对称，但直到那一刻，对称便被自动打破了。没有哪条物理定律规定你必须选择左还是右，但你必须作出选择。此后，左右便不再相同，因为对称不复存在，你也不能再将两个杯子交换。

再看另一个例子，一支铅笔立在一个圆的中心。有那么短暂的一刻，它垂直地立着，这时所有的方向对它都是相同的，存在着一种旋转对称。但铅笔不会总这么立着：它肯定会自动倒向某个方向，一旦铅笔倒下，原来的旋转对称就会被打破。

请注意，决定倾倒方向的并非物理定律本身。无论铅笔倒向何方，决定铅笔倒向的物理定律都是一样的。打破对

称的是铅笔，是系统的状态。铅笔不可能同时倒向所有方向，它肯定会选择某个方向倒下。

一座墙如果无限高、无限长，那么沿着它的任一方向看，它应该处处都是一样的。但是，实际的墙总会有边界，只有离它很近，且边界超出了你的视野时，你看到的墙才是对称的。墙的终点会告诉你，墙并非处处相同。但是，如果你愿意靠近它，只看到很小的范围，那么这对称似乎就能维持。或许你愿意简要地思考一下这个例子，它说明：虽然从一个距离尺度来看，对称产生了破缺，但在另一距离尺度，对称仍然保持——这一概念的意义很快将显现出来。

世界上，你能想出的所有对称几乎都不能维持。例如，一个空的空间会呈现出多种对称，如旋转对称或平移守恒，这意味着所有的方向和位置都是相同的。但实际上空间却不是空的：里面充斥着各种各样的结构，如恒星和太阳系，它们都占据一定的位置并向特定的方向运动，这就不再维持原有的潜在对称。它们会出现在任何地方，可又不能出现在所有地方。潜在对称肯定会被打破，尽管在描述世界的物理定律中它们仍旧很清晰。

与弱力相关的对称也会出现自发破缺。我将在本章的其余部分解释我们是怎样知道这一点的，并讨论它的一些作用。我们会看到，弱力的自发对称破缺是解释粒子质量的唯一方法，同时又避免了其他任

何候选理论所不能避免的有关高能粒子的错误预言。希格斯机制既承认要满足与弱力相关的内部对称，又允许它在必要的时候产生破缺。

弱力谜题

弱力有一种奇特的属性。电磁力能够穿越很远的距离——每次打开收音机时，你便能体会到这一点——而弱力却不同，它只在极近的距离范围内对物质产生作用。两个粒子的距离只有在一亿亿分之一厘米的范围之内，才会通过弱力相互影响。对于早期研究量子场论和QED的物理学家来说，这一有限的范围是一个谜。QED指出，力似乎都应和人们已熟知的电磁力一样，可以传到离荷源任意远的距离，但为什么弱力不能在任意距离的粒子间传递，而只传递给那些附近的粒子？

结合了量子力学原理与狭义相对论原理的量子场论规定，如果低能粒子只在短距离内传递力，那么它们必须具有质量；且粒子越重，粒子的作用范围则越小。正如第6章所讲述的，这是不确定性原理与狭义相对论的结果。不确定性原理告诉我们，需要高动量粒子来探索或影响小距离的物理过程；而狭义相对论则将动量与质量联系起来。尽管这一陈述只是讲了其性质，但量子场论却给出了这一关系的精确值。它告诉我们一个有质量的粒子会行驶多远：质量越小，行程越远。

因此，根据量子场论，弱力的微小作用范围只意味着一件事：传递弱力的规范玻色子质量一定非零。但是，前几章里我所描述的力的理论只适用于像光子等规范玻色子，它们能在远距离上传递力，并且质量为零。根据最初有关力的理论，非零质量的存在就很奇怪，而且出现了问题——当规范玻色子有质量时，理论作出的高能预言是无意义的。例如，理论预言：能量极大、有质量的规范玻色子的相互作用

会过于强大——事实上，这些粒子的相互作用频率会超过100%。这一天真的理论显然是错误的。

而且，弱规范玻色子、夸克和轻子（所有这些都非零质量）的质量不能够维持内部对称，而内部对称，正如我们在上一章看到的，是力的理论的一个关键因素。物理学家想要构筑一个包含质量粒子的理论，显然需要一个新的观点。

物理学家证实，要使一个理论避免作出有关高能的、有质量的规范玻色子的无意义预言，唯一的办法是通过希格斯机制的程序使弱力对称自发破缺，以下就是解释。

你可能还记得，在上一章里，我们想以包含内部对称来排除规范玻色子的虚假极化的一个原因就是：没有这一对称，理论会作出同样的不合情理的预言。最简单的没有对称的理论会预言：高能规范玻色子，无论有无质量，与其他规范玻色子的作用要频繁得多。

通过禁止导致不正确预言且在自然界中不存在的极化，力的理论成功地排除了高能粒子的不良表现。虚假极化是关于高能散射问题预言的根源，因此，对称只允许保留那些实际存在并符合对称的物理极化。对称既排除了理论不存在的极化，又排除了它可能导致的不正确预言。

可能当时我并未明确说明，但所述观点只对无质量规范玻色子有效。与光子不同，规范玻色子是非零质量的，弱规范玻色子的速度没有光速快，这就给工作迎头一击。

无质量的规范玻色子在自然界中只有两种极化，而有质量的规范玻色子却有三种极化方向。有一种方法可以帮助理解这一差别：无质量规范玻色子总是以光速行驶，这就告诉我们，它们永远不会静止，

因此它们总能明显表现出其运动方向，你也总能将垂直极化与其他的沿其运动方向的极化区分开来，结果就是，无质量规范玻色子只存在两个垂直方向上的极化。

而有质量的规范玻色子就不同了。就如我们熟悉的所有物体一样，它们可以静止下来。但当有质量规范玻色子不运动的时候，我们就无从辨别它的运动方向。对一个静止的、有质量的规范玻色子，所有三个方向都是相同的，而如果三个方向相同，那么所有三个可能的极化方向在自然界中都将存在，而它们也确实存在。

尽管你可能觉得上面的逻辑很神秘、难以理解，但请放心，实验者们已经观察到了第三种极化的效应，并证实了它的存在。第三种极化叫作纵向极化。当一个有质量的规范玻色子运动时，纵向极化就是沿其运动方向的波的振动，例如声波的振动方向就是如此。

这种极化不存在于像光子等无质量的规范玻色子中，但是，第三种极化却是像弱规范玻色子等有质量规范玻色子的一种真实的自然属性。因此，第三极化必须成为弱规范玻色子理论的一个组成部分。

因为第三种极化是弱规范玻色子在高能下作用频率明显超高的根源，因此，它的存在就带来了矛盾。我们已经知道，要消除不良高能表现需要一个对称，但这一对称在排除不正确预言的同时，也排除了第三种极化；而这一极化对有质量规范玻色子和对描述它的理论都是必不可少的。尽管内部对称会排除高能表现的错误预言，但其代价却是高昂的：这一对称还摒弃了质量！就如给孩子洗完澡后，将洗澡水连同孩子一起倒掉了。

乍看起来，障碍似乎不可逾越，因为有质量规范玻色子理论的要求似乎完全是自相矛盾的：一方面，上一章描述的内部对称不应保持，否则有三个极化方向的有质量规范玻色子就会被禁止；另一方

面，如果没有内部对称来排除另一极化方向，当规范玻色子具有高能量时，力的理论会作出不正确的预言。如果我们希望排除不良的高能表现，还是需要一个对称来排除单个有质量规范玻色子的第三种极化。

解决这一明显矛盾并找到描述有质量规范玻色子的正确量子场论的关键是，要找到高能和低能的有质量规范玻色子的差别。在没有内部对称的理论里，似乎只有关于高能规范玻色子的预言才会出现问题，关于低能有质量规范玻色子的预言还是符合情理（且正确）的。

这两个事实结合在一起就暗含了一个深刻的意义：**要避免有问题的高能预言，内部对称是必须的**——上一章的内容仍然适用。如果有质量规范玻色子的能量很低（相比爱因斯坦方程 $E=mc^2$ 与质量相关的能量），则不应再保持对称，对称应被排除。这样，规范玻色子就能具有质量，第三种极化也可以参与低能量的相互作用，而质量使得作用有所不同。

1964年，彼得·希格斯等人发现了力的理论是如何通过我们刚才讲述的方式来兼容有质量的规范玻色子的：**高能时保留对称；低能时排除对称。**

希格斯机制依赖的基础是自发对称破缺，它打破了弱相互作用的内部对称，但这只是在低能的时候。这就保证了另一种极化在低能时会出现，这也是理论所需要的。但另一种极化不会参与高能过程，因此也就不会出现不合情理的高能相互作用。

现在，我们就来探讨体现希格斯机制弱力对称自发破缺的模型。有了希格斯机制的这一典型例子，我们会看到标准模型的基本粒子是怎样获得质量的。

希格斯场，上帝粒子的源泉

希格斯机制

希格斯机制包含了一个物理学家称为希格斯场的场。希格斯场就是希格斯粒子的源泉。

希格斯机制包含了一个物理学家称为希格斯场的场。正如我们看到的，量子场论的场能在空间任何位置产生粒子，每一种场都会产生它自己特定类型的粒子，如电子场是电子的源泉；同样，希格斯场就是希格斯粒子的源泉。

就如重夸克和轻子一样，希格斯粒子也很重，因而在通常物质里找不到它们。但不同于重夸克和轻子的是，即便在高能加速器的实验里，也没人见到希格斯场产生过希格斯粒子。这并不意味着希格斯粒子不存在，而只是因为希格斯粒子太重，迄今为止，我们实验能探索的能量还不足以将它们生产出来。物理学家预计，如果希格斯粒子存在，在几年的时间里，当LHC开始运行时，我们就能生成它们。

但我们仍然确信希格斯机制对我们的世界是有用的，因为，这是已知的使标准模型粒子具有质量的唯一方法，它也是前几节所提到问题的唯一解决办法。不幸的是，因为还没有人发现希格斯粒子，我们仍不能确切地知道希格斯场究竟是怎样的。

希格斯粒子的性质是粒子物理学中最热门的极具争议性的话题。在这一节里，我将给出几个最简单的模型——包含不同粒子和力的可能理论——它们展示了希格斯机制是如何发挥作用的。无论真正的希格斯场论最终会怎样，它都要遵循希格斯机制——自发地打破弱力对称，给基本粒子以质量。下面这一模型就给出了一种方式。

在这一模型里，有一对场同时经受弱力。这在以后会非常有用，我们要考虑两个希格斯场，因为携带弱荷，它们必须受弱力支配。希

格斯机制的术语并不很严格，“希格斯”有时同时指两个场，有时指单个场（有时还指我们渴望找到的希格斯粒子）。这里，我会将这些可能区分开来，用希格斯1和希格斯2来代指单个的场。

希格斯1和希格斯2都有可能生成粒子，但即使没有粒子存在，它们也可能取非零值。到现在为止，我们还没有在量子场里遇到这样的非零值。迄今为止，除了电磁场以外，我们只探讨过会产生和销毁粒子的量子场，可没有粒子时，它们的值是零。但是，就如经典电磁场一样，量子场也可以是非零值。根据希格斯机制，其中的一个希格斯场会采取非零值，这一非零值就是粒子质量的根本来源。

当一个场取非零值时，理解它的最好办法是将它想象成一个空间——呈现了场所带的电荷，却不包含任何实际粒子。你应该想到场所携带的电荷是无处不在的，因为场本身就是一个抽象的物体，而这一概念就更为抽象。但当场取非零值时，其结果却是具体的：现实世界中就存在着非零值的场所携带的电荷。

尤其是，一个非零的希格斯场产生的弱荷贯穿于整个宇宙，这就好像是一个携带弱荷的非零希格斯场将弱荷涂满了整个空间。希格斯场的非零值意味着，即便不存在任何粒子，希格斯1（或希格斯2）所携带的弱荷仍将无处不在。当其中一个希格斯场取非零值时，真空（没有粒子存在的空间状态）本身也携带弱荷。

就如会与所有其他弱荷作用一样，弱规范玻色子也会与真空里的这些弱荷相互作用。而弥漫于真空的这些弱荷会妨碍规范玻色子在远距离上传递力，它们想穿越的距离越远，所遭遇的“喷涂”就越多（因为弱荷实际上是在三维空间弥漫，或许将它想象成是一场大雾更为恰当）。

希格斯场所发挥的作用就像是篇首故事里的交警，它将弱力的作用限制在很短距离之内。如果试图将弱力传递给远处的粒子，传递弱力的规范玻色子必然要遭遇希格斯场，它会挡住它们的去路，并将它们困住。就像艾克一样，他只能在方圆不到一公里的范围内自由行驶，弱规范玻色子也只能在很小的范围内无障碍地行动，这一范围大约只有一亿亿分之一厘米。弱规范玻色子和艾克一样，都是在小范围内可以任意驰骋，到了远距离就会受到阻碍。

真空的弱荷分布非常稀薄，因此，在近距离内非零希格斯场及相关弱荷的踪迹也比较罕见。夸克、轻子和弱规范玻色子在近距离内就可以自由地行驶，就好像真空中不存在弱荷一样。因此，弱规范玻色子能在短距离内传递力，似乎两个希格斯场的值都是零。

但在远距离上，粒子行驶得较远，遭遇得弱荷越多。究竟遭遇多少，要看弱荷的密度，而弱荷密度又依赖于非零希格斯场的值。远距离行驶（及弱力的传递）可由不得低能弱规范玻色子自己选择，真空的弱荷会拦住它长途旅行的去路。

关于弱规范玻色子，这正是我们需要弄明白的。量子场论指出，能在近距离内自由行驶却极少会行驶到远距离的粒子，它的质量非零。弱规范玻色子受阻的行程告诉我们，它们的表现就像是具有质量，因为有质量的规范玻色子不会行驶得很远。弥漫于空间的弱荷妨碍了弱规范玻色子的行驶，使得它们的表现恰好与实验相符。

弱荷在真空里的密度大致符合间距为一亿亿分之一厘米。在这一密度里，弱规范玻色子（带电的W及中性的Z）的质量测量值大约是100 GeV。

希格斯机制的成就远不止于此，它还解释了夸克和轻子的质量。夸克和轻子是标准模型里构成物质的基本粒子，它们获得质量的方式

与弱规范玻色子非常相似。夸克和轻子与密布于空间的希格斯场相互作用，因此也受到了宇宙中弱荷的阻挡。与弱规范玻色子一样，因为夸克和轻子被时空中到处弥漫的希格斯弱荷弹回，从而获得了质量。如果没有希格斯场，这些粒子的质量也会是零，但非零的希格斯场和真空的弱荷再一次干涉了运动，使粒子具有了质量。希格斯机制对夸克和轻子获得质量也是必要的。

可能你会觉得，以希格斯机制来解释质量的来源太过牵强，但根据量子场论，这是规范玻色子获得质量的唯一合理方式。希格斯机制的魅力在于，它既赋予了弱规范玻色子质量，又恰巧完成了本章一开始我所提出的任务：有了希格斯机制，弱力对称似乎就能既在短距离内（根据量子力学和狭义相对论，这就等于高能量）保持，又在远距离上（等于低能量）破缺。它自发地打破了弱力对称，而这一自发破缺正是解决规范玻色子质量问题的根源。下一节，我将解释这一更高深的话题（如果愿意，你尽可以跳到下一章）。

弱力对称的自发破缺

我们已经看到，与弱力相关的内部对称变换能够对调弱力作用下的所有东西，因为对称变换会作用于所有与弱规范玻色子相互作用的东西。因此，与弱力相关的内部对称必然也作用于希格斯1和希格斯2场或它们将产生的希格斯1和希格斯2粒子，把它们都同等看待，就如对待同样经受弱力的上夸克和下夸克一样，把它们看作可交换的粒子。

如果两个希格斯场的值都是零，它们就会是同等的、可以交换的，这就完全保持了与弱力相关的对称。但当其中的一个希格斯场取非零值时，希格斯场便自动打破了弱力对称。如果一个场为零而另一场非零，那么希格斯1和希格斯2赖以交换的弱电对称就会被打破。

就如第一个选择了左边或右边杯子的人打破了圆桌上的左右对称一样，一个希格斯场取非零值，就打破了将两个希格斯场对调的弱力对称。对称的打破是自发的，因为打破它的是真空——系统的实际状态，在这一例里是非零值的场。但物理定律没有改变，它仍保持对称。

我们来看一幅图，这有助于理解一个非零的场如何打破弱力对称。图10-1是有两个轴的图，分别标以 x 和 y 。两个希格斯场等价，就像两个没有点的坐标轴的等价一样。如果将图旋转，使两轴变换位置，图看上去还是同样的。这是通常的旋转对称的结果。

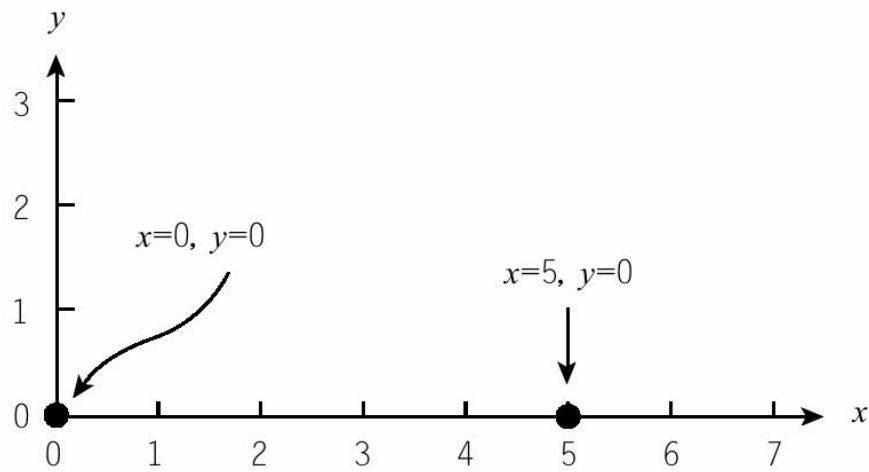


图10-1 两个轴的坐标图。当选取一点 $x=0, y=0$ 时，旋转对称仍旧保持；但如果选取的点是 $x=5, y=0$ 时，旋转对称则被打破。

请注意，如果我在 $x=0, y=0$ 的位置标一个点，这一旋转对称仍能完全保持。但如果标一个非零坐标值的点，如 $x=5, y=0$ ，则不再保持旋转对称。两个坐标轴不再对等，因为这一点的 x 值，而非 y 值，不等于零。

希格斯机制自发打破弱力对称的方式与此类似。如果两个希格斯场都是零，则保持对称；如果一个为零，而另一个非零，弱力对称则自发破缺。

弱规范玻色子的质量说明了弱力对称发生自发破缺的精确能量。那个能量是250 GeV，即弱力能标，非常接近弱规范玻色子 W^- 、 W^+ 和 Z 的质量。当粒子的能量大于250 GeV时，发生的相互作用就好像对称仍旧保持；但当能量小于250 GeV时，则对称被打破，弱规范玻色子的行为就像有了质量。有了非零希格斯场的正确值，弱力对称就在那个能量被打破，而弱规范玻色子则正好获得正确的质量。

作用于弱规范玻色子的对称变换同样也作用于夸克和轻子，而夸克和轻子没有质量，那么这些变换就会使事物保持原状。这意味着弱力对称只有在夸克和轻子没有质量时才会保持。因为弱力对称在高能量时是必不可少的；而对称的自发破缺，不仅弱规范玻色子获得质量需要，夸克和轻子要获得质量，对称破缺也是必需的。希格斯机制是标准模型的所有这些基本粒子获得质量的唯一途径。

希格斯机制的作用在于保证了所有将它包含在内的理论，即使弱规范玻色子（以及夸克和轻子）具有质量，也能对高能表现作出正确预言。具体来说就是，对高能弱规范玻色子（能量超过250 GeV的那些）似乎保持了对称，因此不会作出不正确的预言。在高能量时，与弱力相关的内部对称仍旧滤除了可能引起过高作用频率的弱规范玻色子的问题极化，但在低能量上，质量对产生弱力的小距离相互作用是必不可少的，因而弱力对称就被打破了。

这就是希格斯机制非常重要的原因，所有赋予这些质量的其他理论都没有这种性质。其他方法要么在低能出现错误，得到错误的质量；要么在高能出现错误，作出错误的相互作用预言。

意外收获：电磁力为何如此特别

标准模型还有一个更为成功的特点我没有解释，尽管随后的几章都离不开希格斯场，但与希格斯机制的这一独特方面却并无联系，可因为它是那么的神奇和出人意料，所以还是值得一提。

希格斯机制不仅仅给我们解释了弱力，令人吃惊的是，它还给出了一个新的见解，让我们知道电磁力为什么特别。20世纪60年代以前，没人会想到电磁场还有新的东西需要学习，因为一个世纪前，人们就把它彻底领会了。但到了60年代，由格拉肖、温伯格和萨拉姆提出的弱电统一理论显示，当宇宙在高温、高能之下开始它的演变时，有三种弱规范玻色子，还有一个有着不同作用强度的独立的中性玻色子，而如今无处不在的重要的光子那时却不在此之列。弱电统一理论的作者从数学和物理两方面的线索都推导了这四种弱规范玻色子的性质，在此，我不再细述。

令人瞩目的是，光子原来没什么特别的。事实上，现在我们谈论的光子是原来四个规范玻色子中的两个混合的产物。之所以将光子单列出来，是因为它是参与弱电统一作用的唯一规范玻色子，弱电统一作用对真空的弱荷无动于衷。光子最为独特的特征就是，它会不受羁绊地自由穿越充满弱荷的真空，因此也就没有质量。

与W子和Z子不同，光子的行驶不会受到希格斯场非零值的阻挡。这是因为，尽管真空带有弱荷，却不带电荷。光子传递的是电磁力，它只与带电物体相互作用，因此，光子能在很大的范围内传递力，而不会受到真空的干涉。由此，光子就成了唯一无质量的规范玻色子，即便在非零的希格斯场里也能保持无质量。

这一情形非常类似于艾克所必须遵守的速度限令（虽然无可否认比喻的这部分有点牵强）。限速令让那些普通的车自由驰骋，而不惩罚它们。光子就像普通的“中性车”一样，总能不受约束地自由穿行。

谁会想到这个？多年来，物理学家以为他们完全了解了光子，然而其来源却只有由一个更为复杂的、融合了弱力和电磁力的统一的新理论才能作出解释。这一理论通常被称作弱电统一理论，其相关的对称是弱电对称。

弱电统一理论和希格斯机制是粒子物理的主要成就，弱规范玻色子的质量及光子的意义都在这一框架内得到了完美的解释。更为重要的是，它让我们理解了夸克和轻子质量的来源。刚才我们遇到的这些概念虽然非常抽象，却能完美地解释宇宙的许多特征。

两难境地

希格斯机制非常有效，它赋予夸克、轻子和弱规范玻色子质量，却又不会作出荒谬的高能预言，而且还解释了光子的根源。然而，希格斯粒子还有一个根本属性，物理学家至今未能完全理解。

弱电对称必须在250 GeV的能量上被打破，才能使粒子获得质量。实验显示，能量超过250 GeV的粒子似乎没有质量；而能量低于250 GeV的粒子好像有质量。但只有当希格斯粒子（有时也称作希格斯玻色子）本身也是这一质量时（再次说明，使用 $E=mc^2$ 进行转换），弱电对称才会在250 GeV的能量上破缺。如果希格斯粒子质量过大，弱电统一理论就不会有效。如果希格斯粒子质量更大，对称破缺就会发生在更高的能量，而且弱规范玻色子也会更重——这便与实验结果产生了矛盾。

但是，希格斯粒子过轻也会带来重大的理论问题，我会在第12章解释原因。由量子力学得出的计算表明，希格斯粒子会很重，但物理学家还弄不明白为什么希格斯粒子的质量反而会这么小。这个矛盾非常关键，它会激发出新的粒子物理学观点和后面我们将探讨的额外维度模型。

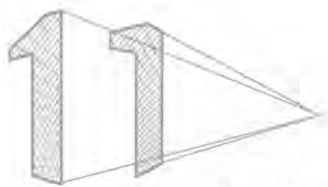
即便不了解希格斯粒子的确切性质以及它为什么这么轻，但其对质量的要求告诉我们，在瑞士CERN投入运行的LHC，一定会发现一个或多个关键的新粒子。无论是什么让弱电对称破缺，但它一定具备大约弱力级的质量。我们期待LHC会发现它究竟是什么，如果真能如愿，这一至关重要的发现会大大增进我们对于物质基本结构的了解。它还可以告诉我们，所有解释希格斯粒子的假说（如果有的话），究竟哪个才是正确的。

但是，在我们研究这些假说之前，我们先要看标准模型的一个可能，它的提出纯粹是出于对一个简洁的自然兴趣。下一章，我们将探讨虚粒子、力对距离的依赖以及一个引人入胜的话题：大统一理论。

-
- 尽管对称对高能粒子的正确预言意义重大，但夸克、轻子和弱规范玻色子的质量却告诉我们，弱力对称必须被打破。
 - 因为我们仍需要防备错误预言，因此弱力对称在高能时还必须保持。所以只有在低能时，弱力对称才被打破。
 - 当所有物理定律保持对称，而实际的物理系统并不保持时，就产生了自发对称破缺。高能时保持、低能时打破的对称就是自发破缺的对称，弱力对称就是自发破缺的。
 - 弱力对称自发破缺的过程叫作希格斯机制。要打破这一对称，必须要有一个能量大约是250 GeV的带质量粒子（记住，狭义相对论通过 $E=mc^2$ 将能量与质量联系起来）。



-imension
探索大揭秘



大统一理论：不同尺度和能量的相互融合

WARPED PASSAGES

某一日，希望你会成为我们的一员，世界将会更简单。

约翰·列侬（John Lennon）

力的影响力取决于你在哪儿

阿西娜常常觉得，她总是最后一个才知道发生了什么有趣的事。艾克开车的冒险经历，都快过去一个月了，阿西娜才听说，而且还不是艾克直接告诉她的——她是从朋友那儿听说的；朋友是听迪特尔的表哥的同学说的；同学是听迪特尔的表哥说的；迪特尔的表哥是听迪特尔说的。

就这么七绕八绕，阿西娜听到了艾克的话：“力的影响取决于你在哪儿。”这句古怪的话可把阿西娜弄糊涂了，不过到最后她才意识到，原话一定被误传了。想了一会儿后，她断定，艾克原话一定是：“保时捷的表现要看它是什么车型。”

我们会看到，阿西娜最初听到的言论是对的，粒子间的距离不同，发生的物理过程也不同。本章的内容是，这些在距离不同的粒子间发生的物理过程是怎样联系起来的，为什么像粒子质量、作用强度等物理量都要依赖于粒子的能量？这种对能量和距离的依赖关系，超出了经典力学中力对距离的依赖关系。例如，在经典物理学里，像引力强度一样，电磁场强度与作用物体间距离的平方成反比（平方反比定律）。但量子力学通过影响相互作用的强度，使得不同距离（以

及能量)的粒子似乎在以不同的电荷相互作用,从而改变了距离依赖关系。

力随着距离的增大变弱或变强,从而产生虚粒子。所谓虚粒子是一种短暂存在的粒子,是量子力学和不确定性原理的结果。它会与规范玻色子相互作用并改变力,使得力的作用要依赖于距离。这就好像是阿西娜从朋友那里听到的艾克的话一样,随着传播路途的延长而被改变了。

量子场论告诉我们怎样计算虚粒子对力与距离和能量的依赖关系的影响。这些计算的一大成就是,解释了强力为何如此强大;还有另一有趣的发现:一定潜藏着一个大统一理论。在这一理论中,除引力之外的其他三种力尽管在低能量上差别巨大,但在高能量上却可能融合在一起,成为一个统一的力。我们即将探索这两个成果以及作为其基础的量子场论的思想和计算。

在看随后的几章时,请记住我们要讨论的不同的能标:大统一能标大约是1 000万亿GeV;而普朗克能标,比它还要大1 000倍,在这一能量上,引力变得强大起来;弱力能标则相对要小得多,当今的实验就在此能量上进行,它只有100~1 000 GeV。弱力能标相比大统一能标,就如一个玻璃球的直径与地球和太阳之间的距离相比。因此,有时我会称弱力能标为低能量^[47],尽管从实验角度来讲,这已经很高了,但相比大统一能标和普朗克能标,它实在是太小了。

从大到小

有效场理论将我们第1章讨论的有效理论的思想应用到量子场论,它们只集中研究有望探测到的能量和距离尺度。在特定能量或距离尺度适用的有效场理论,“有效地”描述了我们需要考虑的能量和尺度。它只关心那些当粒子具有某一特定(或更低)的能量时可能出现

的力和相互作用，而忽略了不可达到的高能量。如果物理过程或粒子只出现在高不可攀的能量中，那么它便不会追根究底。

有效场理论的一大优点在于，即便你不知道在小距离上发生了什么相互作用，你仍可以研究在你感兴趣的尺度上有意义的物理量。实际上，你只需要考虑（在理论上）能探测的物理量即可。调颜料时，你并不需要知道颜料的分子结构，但也许你想了解它那些比较容易得到的基本属性，如颜色和质地。有了这些信息，即便你不知道颜料的微观结构，仍可以根据它们的相关属性分类，并预知当你将这些颜料涂在画布上时，会呈现出什么颜色。

但是，如果你知道了颜料的化学构成，物理规律就会让你推测出它们的一些属性。作画（使用有效理论）的时候，你不需要这些信息；但是，如果是要制作颜料（从一些更为基本的理论得出有效理论的参量），你就会发现它们非常有用。

同样的道理，如果不知道小距离（高能量）的理论，你就不能得出可测量的量。可是，如果你了解了微观领域的详情，量子场理论就会告诉你如何联系适用于不同能量的不同的有效理论。它让你从一种有效理论的量导出另一有效理论的量，如质量或作用强度。

计算物理量是如何依赖于能量或距离的方法是由肯尼斯·威尔逊（Kenneth Wilson）在1974年首次提出的，它有一个非同寻常的名字：重整化群。物理学除了对称以外，还有两个最为强大的工具，那就是有效理论的概念和重整化群，它们都涉及不同长度和能量尺度的物理过程。“群”是一个突出的数学名词，不过其数学根源与此并不相干。

“重整化”指的是，每到一个你感兴趣的距离尺度，都暂停下来，总结自己的所得。你决定哪些粒子和哪些相互作用是与你感兴趣

的特定能量相关的，然后对理论的所有参量使用一个新的标准——即重新标定计量单位。

重整化群的观点类似于第2章的观点，当时，我们讨论了以低维语言解释高维理论的可行性，并把含有一个卷曲维度的两维理论看作只有一维。维度卷曲时，我们忽略发生在额外维度内部的所有细节，并假设所有事物都可以用低维语言来描述。我们的新“重整化”就是可以在关注大尺度时使用的四维描述。

通过类似的过程，我们可以由任意适用于微小尺度的理论导出一个适用于较大尺度的理论：**确定你所关心长度的最小值，然后“洗掉”与更小尺度相关的物理量。**这样做的一种方法是，找到你准备忽略的量的平均值，这些量的细节只有在你决定忽略的更小尺度上才会产生影响。如果有一张方格纸布满了刻度模糊的小点，你可以很精确地知道一些小点的平均阴影密度，然后找到大一些的点来产生同样的阴影效果。在日常生活中粗略观察时，我们的眼睛会自动地完成这一步骤。

如果你的观察精确度有限，为了作出可观测量的有用计算，你也不需要知道在更小距离上发生了什么。最为有效的办法是在理论中选择与你的精度水准相一致的“像素大小”。这样，你就可以忽略那些不能生成的重粒子和不可能发生的小距离相互作用，而集中精力来计算在能达到的能量上会产生的粒子和相互作用。

但如果你确实知道适用于更小距离的更精确理论，那么你就可以用这些理论算出你感兴趣的有效理论的量——即精确度要求较低的有效理论。就如用刻度模糊的小点一样，当你从一个精确度很高的微观的有效理论转向另一个精确度稍低的理论时，从根本上来讲，你是变换了理论分析的“像素大小”。重整化群告诉我们如何计算这些小距

离的相互作用可能对长距离理论的粒子所产生的影响。这样，你就将物理过程从一个长度或能量尺度外推到了另一个。

虚粒子，真实粒子鬼魅般的孪生兄弟

重整化群作出这种推想，靠的是把量子力学过程和虚粒子的作用考虑在内。作为量子力学的结果，虚粒子是真实粒子鬼魅般的孪生兄弟，它们神出鬼没，倏忽即逝。虚粒子与真实粒子有着相同的电荷、相同的相互作用，但能量看上去却不同。例如，高速运动的粒子会携带巨大的能量，但虚粒子的速度可以高得惊人却没有能量。事实上，虚粒子可以具有任意能量，只是与相对应的真实的物理粒子所携带的能量不同。如果能量相同，它就成了真实粒子，而非虚粒子了。虚粒子是量子场论的一个奇异特征，要作出正确预言，它是不可或缺的。

虚粒子

作为量子力学的结果，虚粒子是真实粒子鬼魅般的孪生兄弟，它们神出鬼没，倏忽即逝。虚粒子与真实粒子有着相同的电荷、相同的相互作用，但能量看上去却不同。事实上，虚粒子可以具有任意能量，只是与相对应的真实的物理粒子所携带的能量不同。如果能量相同，它就成了真实粒子，而非虚粒子了。

那么，这一显然不可能存在的粒子为何会出现？倘若不是不确定性原理，外借能量的虚粒子根本不可能存在。不确定性原理允许粒子的能量产生错误，只要它的存在时间短到根本无法测量。

不确定性原理告诉我们，如果要无限精确地测量能量（或质量），那么我们必须花费无限长的时间——粒子的寿命越长，对其能量的测量就越准确。但是，倘若粒子寿命太短，而其能量不可能被无限精确地测定，那么它的能量与一个真实的、短暂存在的粒子就可以产生暂时的偏离。事实上，由于不确定性原理，粒子做什么都不会受

到管辖，只要它们能做到。只要没人观察，虚粒子就会无所顾忌，为所欲为（阿姆斯特丹的一个物理学家甚至建议称它们为荷兰人）。

你可以把真空看成能量库——虚粒子就是暂时借来一些能量而在真空出现的。它们倏忽一闪，很快又消失在真空中，带走借来的能量。这些能量可能会重回它们起源的地方，也可能会转移到其他位置的粒子。

量子力学的真空可是一个繁忙之地，即使从定义来讲真空是空的，但量子作用却使它产生了大量的、时隐时现的虚粒子和反粒子——虽然稳定、持久的粒子并不存在。所有粒子-反粒子对在原则上都可能生成，尽管其存在时间短暂，短到无法直接看见。无论多么短暂，我们仍很在意虚粒子的存在，因为它们毕竟在长寿粒子的相互作用中留下了印迹。虚粒子有其可测量的结果，因为它们影响了进入和离开作用区域的真实粒子的相互作用。在其短暂的生命里，虚粒子会穿行于真实粒子间，然后消失，将能量归还真空。因此，虚粒子就像中间媒介，影响着稳定、持久粒子的相互作用。

例如，前文图7-1中的光子实际上就是一个虚光子，它的交换产生了经典的电磁力。它不具备一个真实光子的能量，当然也不必具备，它只需存在足够的时间来传递电磁力，并使真实的带电粒子发生相互作用。图11-1给出了另一例虚粒子。在这里，光子进入作用区域，产生了一个虚拟电子-正电子对，然后，这一正负电子对在另一位置被吸收；就在它们被吸收的地方，真空出现另一光子，带走由中介的正负电子对暂时借来的能量。现在，我们来研究这种由相互作用产生的引人注目的结果。



图11-1 粒子-虚粒子转化图。一个真实的物理粒子可以转化成一个虚电子和一个虚正电子，然后，它们再转回成一个光子。左边简图和右边费米图示都诠释了这一过程。

相互作用强度为什么会依赖于距离

我们知道，力的强度依赖于粒子相互作用中涉及的能量和距离，虚粒子在这一依赖中发挥了作用。例如，两个电子相隔很远时，电磁力的强度就小（记住，量子力学中，力的下降要超出经典力学中电磁力对距离的依赖）。虚粒子的作用和力对距离的依赖都是真实的，理论预言与实验恰好相符。

有效理论的量，如力或作用强度，取决于参与粒子的能量和距离，这是由量子场论的一个特点引起的，物理学家乔纳森·弗林（Jonathan Flynn）将其戏称为“无政府主义原理”（anarchic principle）^[48]。这一原理是量子力学的结果，它告诉我们所有可能发生的粒子相互作用都会发生。在量子场论里，不被禁止的所有事情都会出现。

每一个特定的粒子群相互作用的单独物理过程，称为一个路径。一个路径可能会涉及、也可能不涉及虚粒子。涉及时，我称这一路径为量子贡献。量子力学告诉我们，所有可能的路径都会对相互作用的最终强度作出贡献。例如，真实粒子可以转化成虚粒子，它们会相互作用，然后再转换成其他的真实粒子。在这一过程里，原来的真实粒子可能会重新出现，也可能变成另外的真实粒子。虽然虚粒子存在的

时间非常短暂，我们无法直接观察，但它们仍会影响真实、可观察粒子间相互作用的方式。



想阻止虚粒子促进相互作用，就如同把秘密告诉朋友而又不想让另一个朋友知道一样。某个“共同的”朋友迟早会泄露你的秘密，把这消息告诉另一个朋友。即使你已经把这个秘密告诉了那个朋友，但当你们共同的朋友与他讨论这一秘密时，仍旧会影响他对这件事的看法。事实上，他的看法就是每个人与他谈论后形成的最终结果。

在传递力时，不仅真实粒子间的直接作用，有虚粒子参与的间接作用同样扮演了重要的角色。就如你朋友的想法会受到与他谈论的每个人的影响，粒子间的最终作用是所有可能贡献的总和，包括来自虚粒子的贡献。由于虚粒子的重要性要取决于涉及的距离，因此，力的强度也就要取决于距离。重整化群会告诉我们怎样计算虚粒子在相互作用中的影响，所有中间虚粒子的影响会叠加起来，要么增强、要么减弱规范玻色子的相互作用强度。

在相距较远的粒子间作用时，间接作用扮演了更重要的角色。距离的增大就好比你把秘密告诉了更多的“虚”朋

友，尽管你并不确信哪个朋友会泄露你的秘密，但你告诉的朋友越多，秘密泄露的可能性就越大。虚粒子会影响相互作用的最终强度，量子力学告诉我们，只要存在一个包含虚粒子的路径，它们就一定会产生影响，且虚粒子对强度的影响取决于力被传递的距离远近。

但实际的重整化群计算更为聪明，因为它还加上了朋友之间互相谈论的贡献。与虚粒子贡献更为相似的比喻是，在一个庞大的官僚机构里传播一条信息。最高层领导下达一条信息，很快就会被传播开来。但低层的人想要传达一个信息，可能就要经过上司的审查：若是一个较低职位的人发出一条书面信息，那么它可能还会辗转多个部门，层层盖章，最终才能到达目的地。在这种情况下，官僚机构的每一层次都要把这一信息传阅一遍，然后才能送达更高的一层。只有最后传到高层主管那里，信息才会被发布。在这种情况下，最终出现的信息早已不再是原来的样子——它已经过了多个官僚的层层过滤。

如果你把虚粒子当作官僚连，那么更高级别的官僚对应的就是一个具有更高能量的虚粒子——高层的信息会被直接传播，而低层信息却要经过层层关卡

的审查。量子力学的真空就是一个光子要遭遇的“官僚机构”，每一次相互作用都要经过虚粒子的层层审查，结果是能量变得越来越小。

就如在官僚机构中一样，在各个层次（或距离）上都有一些岔路，有些路径会绕过由虚粒子设置的“官僚”关卡，而有些会遭遇许多虚粒子，使传播距离越来越长。较短距离（较高能量）相比那些较长距离的传递会遭遇更少的虚粒子过程。但是，在虚粒子过程与官僚层次间存在着一个明显区别：无论路径有多复杂，官僚机构中，任何一条信息都只有一条特定的路径；而量子力学却认为会有多种路径，而且作用的最终强度应是所有可能发生的路径共同贡献的结果。

设想一个光子从一个带电粒子行驶到另一带电粒子，在途中它会转换成一个虚电子-正电子对（见图11-2），量子力学告诉我们它往往会发生。而有虚电子-正电子对参与的路径，就影响了光子传递电磁力的效能。

这还不是量子力学会出现的唯一过程。每对虚电子-正电子本身就可以发出光子，光子也会转化成其他虚粒子，依此类推。交换光子的两个带电粒子间的距离决定了“信使”光子在真空中会与粒子发生多少次这样的相互作用，以及这些作用的影响会有多大。电磁力的强度就是将所有的“官僚障碍”（在或远或近的距离上，虚粒子可能参与

的量子力学过程) 考虑在内后, 光子可能采取的多条路径的最终结果。由于光子可能遇到的虚粒子数量要取决于它所行驶的距离, 因此, 光子的作用强度就要取决于与它相互作用的带电物体间的距离。

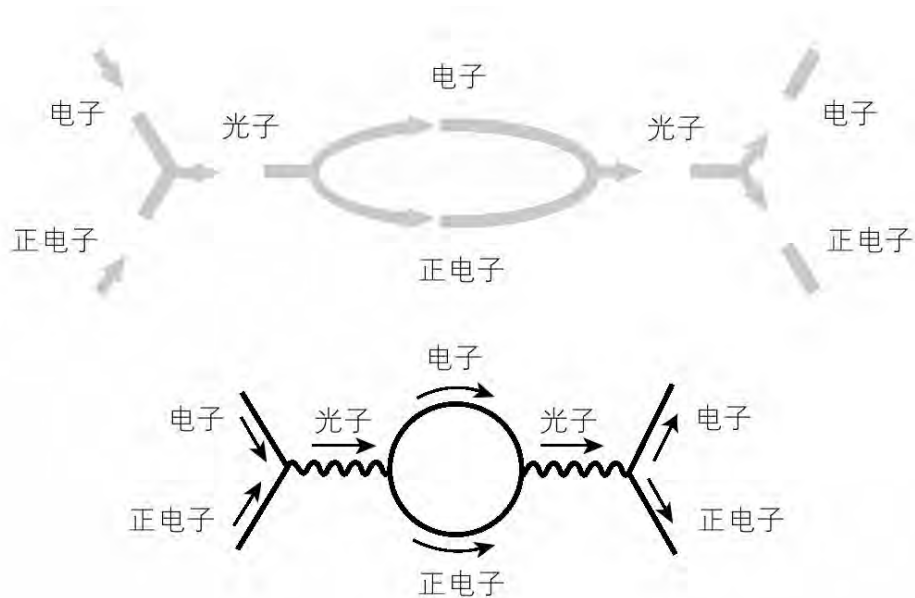


图11-2 电子-正电子散射的虚拟修正。由左至右为: 一个电子和一个正电子湮灭形成一个光子, 光子接着分裂成一个虚电子-正电子对, 然后再湮灭成一个光子, 光子再转化成电子和正电子。中间的虚电子和正电子就影响了电磁力的强度。

计算表明, 当所有可能路径的贡献都被加起来之后, 真空会削弱光子由电子带来的信息。对电磁作用削弱的直观解释是: 异性相吸, 同性相斥。因此, 一般来说, 虚正电子比虚电子离电子更近, 虚粒子的电荷由此削弱了初始电子电磁力的完整作用。量子力学作用屏蔽了电荷, 这意味着光子和电子间的相互作用强度会随距离增加而降低。

在远距离上的真正电磁力, 比经典的近距离电磁力要小, 因为在近距离传递力的光子会更常采用不包含虚粒子的路径。远距离传递力的光子必须穿越虚粒子的浓雾, 减弱其强度, 而近距离行驶的光子则不必这样。

不仅仅是光子，所有传递力的规范玻色子，在抵达目的地的途中都要与虚粒子发生相互作用。粒子与其反粒子的虚粒子对，在真空中会自发生成并被吸收，从而影响相互作用的最终强度。这些虚粒子会暂时拦截传递力的规范玻色子，改变其总的作用强度。计算表明，弱力的强度与电磁力一样，会随距离的增大而降低。

但虚粒子并不总是阻碍相互作用的，令人惊讶的是，它们有时还会反过来促进其作用。20世纪70年代初，哈佛大学的戴维·波利策（David Politzer）——当时还是西德尼·科尔曼（即提出这一问题的人）的研究生，普林斯顿大学的戴维·格罗斯（David Gross）和他的学生弗兰克·维尔切克（Frank Wilczek）以及荷兰的杰拉德·特·胡夫特（Gerard 't Hooft）都计算出了同样的结果，证明强力的表现与电磁力恰好相反。在远距离上，虚粒子不仅没有屏蔽强力使其变弱，反而加强了胶子（传递强力的粒子）间的相互作用，正因如此，强力在远距离上才名副其实。格罗斯、波利策及维尔切克因对强力的重要见解而获得2004年诺贝尔物理学奖。

这一现象的关键在于胶子本身的性质。胶子与光子的一个重要差别在于，胶子之间也会相互作用。胶子可以进入作用区域，转化成一对虚胶子，从而影响力的强度。与所有虚粒子一样，这些虚胶子也只是存在片刻，但是随着距离的增大，它们的影响会累积起来，直到强力变得异常强大。计算结果表明，虚胶子在相距较远的粒子间会大大地增加强力的强度，粒子间的相隔距离越远，强力越强。

强力会随着距离的增大而加强，相比电荷的屏蔽，这确实有悖于我们的直觉。粒子怎么可能相距越远，反而相互作用越强？大多数的相互作用都要屈服于距离。我们需要通过实际的计算来让大家看到这些，但现实生活中也有这样的例子。

如果有人通过一个官僚机构传达某条信息，而某个中层管理人员对其重要程度并不确信，他就有可能大张旗鼓，把本来只需一张便条传达的事，变成了一项重要指令。信息经中层领导这样篡改，其影响远远地超出了由发起者直接传达的效果。

距离越远，力反而越强，特洛伊战争也是其中一例。根据《伊利亚特》（*Iliad*）记载，特洛伊战争的起因是特洛伊王子帕里斯决定带走斯巴达国王墨涅拉奥斯的王后海伦。如果帕里斯当初能直接与墨涅拉奥斯决战来争夺海伦，而不是和海伦私奔到特洛伊的话，那么，希腊与特洛伊的战争也就早早结束，不会载入史册了。一旦墨涅拉奥斯与帕里斯相距遥远，中间再经别人添油加醋，大肆渲染，敌对状态就变得异常强大起来，最终演变成了旷日持久的希腊-特洛伊战争。

尽管非常令人惊讶，但强相互作用会随距离的增大而加强，这足以解释强力许多独特的特征。它解释了为什么强力这么强，能将夸克束缚在一起，形成质子和中子；为什么夸克会被困在喷射流中。强力在远距离上变得如此强大，以至于经受强力的粒子永远不可能彼此远离，像夸克这种通过强力相互作用的基本粒子从来就没有孤立存在过。

一对相距甚远的夸克和反夸克会贮存大量的能量，这些能量足以强大到在它们之间生成更多的夸克和反夸克，而不会保持孤立。如果你试图把两者之间的距离拉得更远，真空中就会生成新的夸克和反夸克，就如波士顿的交通，如果你紧跟在一辆车后面，一旦没有跟上，落下了一点距离，刚够一车长，那么必然会有车从别的车道上并进来。在最初的夸克周围，总会有一些新的夸克和反夸克在逡巡，一旦有夸克和反夸克比刚出现时随从要少些，它们必然会替补上去——别的夸克和反夸克似乎总在左右等候。

远距离上的强力非常强大，它们根本不允许通过强力作用的粒子彼此分开。强力作用下的粒子总是由许多带强力色荷的其他粒子护卫，形成强力色荷中性组合。结果就是，从来不曾出现独立的夸克，我们看到的只有紧紧束缚在一起的强子和喷射流。

大统一理论

大统一理论

1974年，乔治和格拉肖提出了一个大胆设想：弱力、强力和电磁力这三种力会随着距离和能量而变化，当能量达到极高时，它们会融合成一个统一的力。乔治和格拉肖称这一理论为大统一理论。

上节的内容给我们讲述了有关弱力、强力和电磁力对距离的依赖关系。1974年，乔治和格拉肖提出了一个大胆设想：这三种力随着距离和能量而变化，当能量达到极高时，它们会融合成一个统一的力。乔治和格拉肖称这一理论为大统一理论（GUT）。如第7章的讨论，强力对称可以对调三种色夸克，而大统一力的对称会作用于所有类型的标准模型粒子，使它们可以对调。

根据乔治和格拉肖的大统一理论，在宇宙演变早期，温度极高，能量极大——温度要超过100亿亿亿K，能量超过1 000万亿GeV——除引力之外的三种力各自的强度都同样大，因此，这三种力便融为一体，统称为“力”。

随着宇宙的演变，温度逐渐降低，统一的力会分裂成三种各不相同的力。每种力对能量的依赖也不同，通过它们不同的能量依赖，渐渐演变为我们现在所知的三种力。尽管起初是一个统一的力，但在低能量上，由于虚粒子对它们产生的不同影响，它们就有了不同的作用强度。

这三种力就像由同一个受精卵发育成的三胞胎，最终成长为三个性格各异的青年：其中一个可能留着染了色的朋克头；一个留着水手样的小平头；而另一个则像艺术家一样扎着小辫子。但不管现在外形差异有多大，他们仍有着相同的DNA，小时候是很难让人分清的。

宇宙发展早期，这三种力也是难以区分的。但它们最终由于对称的自发破缺而分裂。正如希格斯机制使弱电对称破缺，只留下电磁力没有改变一样，它也打破了大统一力的对称，留下了我们现在看到的三种各不相同的力。

在高能量上统一的作用强度是大统一理论的一个先决条件，这就意味着，能量函数所代表的不同作用强度的三条线，肯定会在某一个能量值上相交。但我们已经知道，除引力外的三种力，其强度是怎样随着能量发生变化的，而且，量子力学告诉我们，远距离就等于低能量，而短距离则意味着高能量^[49]，那么前几节的结果就可以以能量来理解。在低能量上，电磁力和弱力不如强力强大，但在高能量上，它们会增强，而强力则会减弱。

换种方式来说，除引力外的三种力的强度，到了高能量上更加和谐统一，它们甚至可能合并为一个统一的力。这就意味着：能量函数所代表的作用强度的三条线在高能量上会相交于一点。两条线相交于一点，没什么可兴奋的——彼此靠近时，它们必然会发生；但三条线相交于一点，要么是极端巧合，要么这种相交必然蕴含着某种深意。如果三种力真的相交，这一作用强度就可能表明，在高能量上只有一种力——这样，我们就有了一个统一的力。

尽管统一至今仍只是一种设想，但如果能够实现，对于我们更为简洁地描述世界必然会迈出巨大的一步。由于统一原理饶有趣味，物理学家研究了三种力在高能量上的强度，看它们究竟是否能够融合。1974年之前，没有人能够精确测得引力之外三种力的作用强度，霍华德

• 乔治·斯蒂芬·温伯格和海伦·奎因（Helen Quinn，当时还是哈佛大学的一位博士后，现在是斯坦福线形加速器中心的物理学家，并任美国物理学会主席）利用当时所能达到的并不完善的测量，使用重整化群计算推算了高能量上的力的强度，他们发现，代表引力之外的力的三条线似乎真的会在某一点相交。

1974年，在乔治-格拉肖有关大统一理论的著名论文里，开篇是这样的：“我们提出的一系列假设和推想，最终不可避免地会得出一个结论……所有作用于基本粒子的力（强力、弱力、电磁力）都是有着同一强度的、同一基本作用的不同表现。也许我们的假设未必正确，推想也不具备意义，但我们的理论设计独特、结构简洁，这足以引起人们的重视。”这些语言如此谦虚低调，是因为乔治和格拉肖并不真的认为独特和简洁能成为他们的理论对世界作出正确描述的充足证据，他们还希望得到实验证实。

尽管由标准模型推想超出10万亿倍的能量，需要有巨大的信心，这一能量从没有人直接探索过，但他们意识到，他们的推测会产生可验证的结果。在论文中，乔治和格拉肖解释了他们的大统一理论“预言质子会产生衰变”，而这一预言可以尝试由实验来验证。

乔治和格拉肖的统一理论预言：质子不会永远存在，经过很长一段时间后，它们会产生衰变。在标准模型里，这是不会发生的：夸克和轻子根据它们经受的力，总是能被区分开来；但在大统一理论里，力从根本上都是一样的，因此，就如上夸克通过弱力会变成下夸克一样，通过一个统一的力，夸克也能够变成一个轻子。这就意味着，如果大统一观点正确，宇宙中夸克的净数量不会保持不变，夸克可以转化成轻子，使得质子（由三个夸克合成）产生衰变。

因为在一个联结了夸克和轻子的大统一理论里，质子会产生衰变，所以，我们熟知的所有物质最终都是不稳定的。但是，质子的衰

变速度相当缓慢——其寿命会远远超过宇宙的年龄。这意味着质子衰变的迹象无论有多明显，我们几乎还是没有机会能够探测到它：它的发生频率实在是太低了。

为了找到质子衰变的证据，物理学家不得不建立一些持续时间极长的大型实验来研究大量的质子，通过这种方式，即便一个质子不大可能产生衰变，但用大量的质子，便能大大地增加探测到其中某个质子产生衰变的机会。就如即使我们赢得彩票的可能性极小，但如果我们买大量彩票，中奖机会就会大大增加一样。

物理学家真的建立起了这样一些多质子的大型实验室，其中包括欧文-密歇根-布鲁克黑文（IMB）实验室，位于南达科他州的一个地下盐矿里；神冈实验室，位于日本神冈，水槽和探测仪都深藏于地下1 000米的废矿井里。虽然质子衰变过程极少发生，但是，如果乔治-格拉肖大统一理论正确，这些实验早就能发现证据了。不幸的是，这一宏大的抱负终未实现，还没有人发现这样的衰变。

这并不一定排除统一的可能。事实上，幸亏有更为精确的力的测量，我们现在知道，由乔治和格拉肖提出的最初模型几乎肯定是错误的，而只有标准模型的一个延伸版本才可能将力统一起来。结果，在这些模型里质子的寿命甚至更长，质子衰变不可能被探测到。

现在，我们并不确信力的统一是否是自然界的一个真实特征。如果是，它又有什么意义呢？计算表明，在后面将讨论的模型里，有好几个都会发生力的统一，其中包括超对称模型，霍扎瓦-威滕额外维度模型，以及我和拉曼·桑卓姆创建的弯曲的额外维度的模型。额外维度模型尤其引人入胜，因为它有可能将引力也兼并进来，真正实现四种已知力的统一。这些模型之所以重要，还因为最初的统一模型认为，在弱力级别之上不会再发现新的粒子，而只有具有大统一质量

的粒子，^[50]而另外这些模型则显示，即便在弱力能量级别之上产生许多新粒子，统一仍旧可能发生。

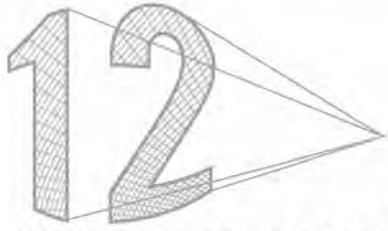
但是，无论力的统一有多么神奇，物理学家还是因其理论贡献分为两大阵营：有的喜欢自上而下的研究方式，有的选择自下而上的研究。大统一理论的观点体现了自上而下的方法，乔治和格拉肖作出了一个大胆的假设：在1 000~1 000万亿GeV之间不存在有质量的粒子，并在此假设之上推想出一个理论。在粒子物理学界，大统一是争议的第一步，而弦理论的争议至今一直甚嚣尘上，两种理论都是由测得能量的物理定律去推想至少超出1 000万亿倍的能量。乔治和格拉肖后来开始怀疑弦理论和大统一理论所体现的自上而下的方法，此后，他们转变观念，扭转方向，开始专注于低能物理的研究。

尽管大统一理论有一些吸引人的特征，但我并不确信它们的研究最终形成对自然的正确认识。在我们已知的能量和推想的能量之间存在着一条巨大的鸿沟，这中间究竟会发生什么，有太多可能供我们想象。无论哪种情形，除非我们发现质子衰变（或其他一些预言）——如果它存在，否则我们都不可能确定力在高能量上是否真的会统一起来。在此之前，这一理论只能是一个伟大的理论设想。

-
- 虚粒子与真实的物理粒子有着相同的电荷，但其能量不同。
 - 虚粒子只能存在短暂的时间，它们从真空中暂时借得能量，真空是没有任何粒子的宇宙状态。
 - 物理过程中的量子贡献，源自虚粒子和真实粒子的相互作用。这些虚粒子时而出现，时而消失，作为真实粒子间的媒介，其贡献会影响真实粒子的相互作用。



- 由无政府主义原理我们得知，在探讨粒子属性时，必须将量子贡献考虑在内。
 - 在大统一理论中，一个单一的高能量的力转化成低能量上除引力外的三种已知力。三种力要统一，它们在高能量上必须有着相同的强度。
-



等级问题：唯一有效的涓滴理论

WARPED PASSAGES

高速公路是搏命者的天下，你可要当心。

如果碰巧捡到了点什么，那你一定要收好了。

鲍勃·迪伦 (Bob Dylan)

与上帝赌博

艾克开着他华丽的新保时捷跑车撞到了电线杆上，不幸意外身亡。可是在天堂里，艾克是幸福的：他可以一直玩游戏了，其实艾克骨子里就是一个赌棍。

一天，上帝他老人家亲自邀请艾克玩一个非常奇特的游戏：上帝让他写下一个16位数字。上帝要掷骰子，那种天堂里特有的20面的骰子。与我们凡间通常所见的6面骰子不一样，这个骰子有20个面，从0~9的10个数字每个出现两遍。上帝解释道，他要掷16次骰子，把每次掷的结果依次写下来形成一个16位的数字。如果上帝和艾克写的都是同一个数字——也就是说，所有数字顺序相同，那么，上帝就赢了；如果数字不完全相同——假设有任何一位数不对，那么艾克就赢了。

上帝开始摇。他掷出的第一个数是“4”，这与艾克所写数字的第一位恰好相符，艾克的数字是“4 715 031 495 526 312”。上帝能恰好摇出同样的数，让艾克感到很惊讶，因为这概率只有1/10，但他相信，第二个、第三个肯定就不会一样了，因为上帝连续掷出两个数字都相符的概率只有1/100。

上帝再次摇骰子，然后第三次，他分别掷出一个“7”和一个“1”，又对上！他继续掷，这下艾克可是目瞪口呆了：上帝掷出的所有数字都对了。这一概率是一亿亿分之一，上帝怎么可能会赢呢？

艾克有点儿生气（在天堂里，人不可能火气太旺），他问，像这样根本不可能的事是怎么发生的？上帝圣明地答道：“只有我才可能赢，因为我是无所不知、无所不能的，但是，你一定也曾听说过，我不喜欢掷骰子。”

说完，空中出现四个大字——“禁止赌博”。艾克愤怒了（当然，只是一点点儿），他不仅输了这场游戏，还失去了赌博的权利。

到现在为止，你已经学到了有关粒子物理学的大量知识，并了解了物理学家建立标准模型的许多理论元素。在解释许多不同的实验结果时，标准模型发挥了异常强大的作用，但是，它的基础是极不稳定的，这基础本身就是一个意味深长的谜题，而它的谜底几乎必然会引发对物质基本结构的新的认识。本章我们就将探索这一谜题，也就是物理学家所说的等级问题（hierarchy problem）。

这一问题并非指标准模型预言与实验结果不符，关于电磁力、弱力、强力的质量和电荷都已得到了精确验证；在CERN、SLAC和费米实验室进行的实验，也都证实了标准模型对于已知基本粒子相互作用和衰变速度的预言；标准模型中强力的强度也已无神秘之处；而且，它们之间的关系还启发人们想到了一个潜在的大统一理论的观点；再者，希格斯机制也完美地解释了真空怎样打破弱电对称，使W、Z规范玻色子以及夸克和轻子获得质量。

等级问题

为什么希格斯粒子的质量，以及由此带来的弱力规范玻色子的质量如此微小？等级问题告诉我们，在推想极端高能物理过程之前，我们至少还有一个低能的问题迫切需要解决。

但无论表面多么其乐融融的家庭，如果深入观察，你总会发现紧张的暗流在涌动。尽管表面呈现的是一幅和谐、幸福的景象，但暗地里却隐藏着毁灭性的家族秘密。标准模型就有这样一个隐痛。如果你不加批判地认为，电磁力强度、弱力强度和规范玻色子都采用实验测得的值，那么所有东西都会符合预言。但我们很快会看到，质量参数

（决定基本粒子质量的弱力级质量）虽然已被精确测量，但它比物理学家根据一般理论思考得出的质量要小16个数量级。任何物理学家要依据高能理论推算弱力级能量（从而推算所有的粒子质量），都会发现它彻底错了，质量似乎是无中生有的。这一谜题——等级问题，是我们对粒子物理学认识的一个巨大漏洞。

我在引言里把等级问题解释成引力为何如此微弱的问题。但现在我们看到，这一问题还可以阐释为：为什么希格斯粒子的质量，以及由此带来的弱力规范玻色子的质量如此微小？这些质量要与它们的测量值相符，标准模型必须加进一个附加参数，而这个参数就如有人要随机地掷出一个16位数字，结果必须与艾克的数字恰好相符一般不可能。尽管标准模型有许多成功之处，但它却要依靠这一荒谬、不合情理的参数，才能得出已知基本粒子的质量。

本章解释了这一问题，也解释了为什么我和多数粒子理论学家都认为它那么重要。等级问题告诉我们，无论是什么使得弱电对称产生破缺，它都会比第10章呈现的两个希格斯场的例子更为有趣。所有可能的解决方法都会涉及新的物理原理，而答案将可能引导物理学家发现更基本的粒子和定律。找到是什么发挥了希格斯场的作用并打破弱电对称，会给我们展现一些新的、最为丰富的物理现象，或许在我有生之年，它们就会得到确定。几乎可以肯定的是，在大约1 TeV的能量上，一定会出现新的物理现象。检验各种假说的实验也已准备就绪，不出10年，不论有什么实验发现，我们对基本物理定律的认识，都将发生巨大的改变。

等级问题告诉我们，在推想极端高能的物理过程之前，我们至少还有一个低能的问题迫切需要解决。在过去大约30年里，粒子物理学家一直在探寻一种理论结构，能够预言和维护弱力级能量，这是弱电对称破缺相对较低的能量。我们许多人都认为，对于等级问题必会

有一个解决方法，它将提供最好的线索，告诉我们超越标准模型会有什么。很快我将讲述一些理论，要理解这些理论的动机，我们首先要简单地了解一下这个虽然专业却很重要的问题。对解决方法的探求已引导科学家们开始研究一些新的物理概念，如后面几章探讨的那些概念，而问题的解决几乎必然会修正我们现在的观点。

在探讨等级问题的一般形式之前，我们首先来看大统一背景之下的等级问题，这比较好理解，而且，我们正是在大统一理论里首先认识到了这一问题。然后，我们要在更大（也更为普遍）的背景之下来看这个问题，看它为什么会最终归结为引力比其他已知力更微弱的问题。

等级问题，大统一理论里的依赖性问题的

假设你去看望一位个子很高的朋友，他的身高大约是1.95米，而你发现他的同胞兄弟却只有1.5米，你是不是会感到非常惊讶呢？你会以为，朋友和他的兄弟有着相似的遗传基因，个子也应该差不多高。现在再设想，事情可能更离奇：你到了朋友家里，发现他哥哥的身高比他要高10倍或是矮10倍。这才真的会让人感到无比奇怪呢！

除非有充分的理由，我们并不认为所有的粒子都会有着相同的属性，所以我们通常认为，经受同样力的粒子性质也应是相似的，例如，我们认为它们的质量也应该差不多。就如你有足够的理由认为同胞兄弟身高应该不会差很多一样，粒子物理学家也有充分的科学依据，期待粒子质量在同一理论里也应是相似的，如在大统一理论里。但是，大统一理论里的质量却绝非相同：即便是经受相似力的粒子，其质量也有着巨大的差异，这可不是10倍的差别：它们之间质量的差超过了10万亿倍。

大统一理论的问题在于：尽管打破弱力对称的希格斯粒子必须很“轻”，大致应是弱力级质量，但大统一理论却提出一个对应的通过强力与其他粒子发生相互作用的希格斯粒子。大统一理论的这个新粒子却必须极重，大约是大统一级质量。换句话说，这两个本应由一种对称（大统一力的对称）相联的粒子，其质量却有着天壤之别。

在大统一理论里，这两个不同却互相关联的粒子必须同时出现，因为在高能量上弱力和强力应是可以互换的，这是大统一理论背后的整个观点——所有的力都是一样的。因此，当强力和弱力统一起来时，经受弱力的所有粒子，包括希格斯粒子，都必须伴有另一个经受强力的粒子，并且与原来的希格斯粒子有着类似的相互作用。但是，这一与希格斯粒子相关联的、经受强力的新粒子，却存在着一个大问题。

这个与希格斯粒子为伴的强荷粒子能同时与夸克和轻子产生相互作用，因此能使质子衰变——甚至比大统一理论预言的速度还要快。要使质子发生衰变，强相互作用的粒子必须在两个夸克和两个轻子之间进行交换；为避免过快衰变，它们则必须极重。而目前质子的寿命限制告诉我们，强荷的希格斯粒子伙伴，如果自然界中存在的话，其质量必须接近于大统一标度的质量，大约为1 000万亿GeV。假使这一粒子存在却没有这么重，那么不等你看完这句话，你和这本书就都将灰飞烟灭。

但是，我们已经知道，弱荷的希格斯粒子必须很轻（大约250 GeV）才能使弱规范玻色子质量与实验测得的质量相符。这么一来，实验设定的限制告诉我们，希格斯粒子质量一定与其经受强力的伙伴粒子质量有着天壤之别。在一个统一的理论里，假定与弱荷希格斯粒子有着相似相互作用的强荷希格斯粒子必须具有不同的质量，否则这世界就不会是我们看到的样子。这两个质量间的巨大差异——一个是另一个的10万亿倍，真的难以解释，尤其是在一个统一的理论里，不论强荷粒子还是弱荷粒子，都假定有着相似的相互作用。

在大多数统一理论里，要让一个粒子重而另一个粒子轻，唯一的办法就是引进一个巨大的附加参数，任何物理原理都不会预言这么大的质量差异。精心选择一个参数是使它有效的唯一办法，而这一数字的精确值必须达到13位，否则，要么质子会快速衰变，要么弱力规范玻色子的质量就会变得很大。

粒子物理学将这一必要的附加参数称作微调，微调就是为了得到期望的精确值而对参数进行的调整。选择“调”这一词，就像是为了让钢琴敲出正确的音符而对琴弦进行的调试。如果你想让一个只有几百赫兹的音频达到13位数的精确度，那么你监听的时间则必须达到100亿秒，即1 000年。13位数的精确度实在是千年难遇的。



有关微调的比喻，我还可以举出其他一些例子，但我肯定它们听起来都缺乏真实性。例如，在一家大公司里，一人负责开支，一人负责收入，假设他们从来不曾交流和沟通，而到年底时，公司的支出应当恰恰等于收入——不可超过1美元，要不然这公司就会倒闭。是的，这个例子实在是太不切实际了。而它不可能切合实际，任何一个符合常理的情形都不会依赖于微调，没人会把自己的命运（或是企业的命运）悬在那几乎不可能发生的偶然巧合上。但是，几乎所有包含一个轻希格斯粒子的大统一理论，都存在这样的依赖性问题。如果一个理论的物理预言这么敏感地依赖于一个参数，那它根本不可能成为一个完善的理论。

但是，最简单的大统一理论要得到一个小质量的希格斯粒子，唯一的办法就是在理论里加进一个附加参数，大统一理论模型没有提供更好的方法。对于在四维中统一的大多数模型来讲，这是一个严重的问题，正因如此，许多物理学家，包括我自己在内，对力的统一都持怀疑态度。

而等级问题愈发严重：即便你不需要根本的解释，而乐于简单地假定一个粒子极轻而另一粒子极重，你还是会遇到所谓量子力学贡献（或简称量子贡献）带来的问题。这些量子贡献必须加在经典的质量之上，才能确定希格斯粒子在真实世界中的真正的物理质量，而这些贡献通常会远远大于希格斯粒子所要求的几百GeV的质量。

我先提醒一下，下一节里有关量子贡献的讨论，虽说是以虚粒子和量子力学为基础，却可能并不符合人的直观理解，所以不要尝试以经典的类比去想象，我们要思考的就是一个纯粹的量子力学效应。

对希格斯粒子质量的量子贡献

上一章我们讲到，通常一个粒子不会毫无妨碍地穿越空间，虚粒子时而出现，时而消失，因此影响了原本粒子的路径。量子力学告诉我们，对任何一个物理量，我们必须把所有这些可能路径的贡献都加进来。

我们已经看到，这些虚粒子的存在使得实验测得的力的强度对于距离的依赖完全符合预言。使得力对能量产生依赖的这一同类型的量子贡献，还会影响到质量的大小。但就希格斯粒子的质量来说，与力的强度不同，所以虚粒子的作用与实验对理论的要求似乎并不相符，它们似乎太大了。

因为希格斯粒子会与一些重粒子发生相互作用，它们的质量高达大统一级别，所以希格斯粒子所采取的路径包含了由真空中分裂的虚重粒子及其反粒子，希格斯粒子在其行驶途中会暂时地转化成这样一些粒子（见图12-1）。在真空中忽生忽灭的重粒子会影响希格斯粒子的运动，是它们导致了大量的量子贡献。



图12-1 在大统一理论中，重粒子对希格斯粒子质量的虚贡献。希格斯粒子会转变成虚重粒子（大统一质量），然后它们再转回成希格斯粒子。左边简图和右边费曼图都说明了这一过程。

量子力学告诉我们，如果要确定希格斯粒子实际拥有的质量，我们必须在那条没有虚粒子的路径上加进这些有着虚重粒子的路径。问题在于，包含了虚重粒子的路径导致的贡献，会使希格斯粒子达到大统一理论里重粒子的质量——比我们所需质量高13个数量级。所有这些来自虚重粒子的巨大贡献，都必须加到希格斯粒子质量的经典值上，产生的物理值还必须符合测量。如果要使弱规范玻色子的质量正确，这一值大约应是250 GeV。这就意味着，即便任一GUT质量贡献都要大至13个数量级，但当我们把所有这些巨大贡献加进质量时，最终的答案应该接近于250 GeV，如果有些贡献为正，则有些必然为负。哪怕只有一个虚重粒子与希格斯粒子发生作用，都会不可避免地出现问题。

如果像上一章那样，我们把虚粒子看作官僚机构的成员，那么，它们的职责就像美国移民局（INS）的员工一样，是为了拦截来自可疑

人员的信件，而现在，他们却仔细地审查所有信件。本来是一个两层的系统，可以让一些信件快速通过，只拦截另外的一些，现在呢，所有信件都被同样对待。同样的道理，希格斯机制要求虚粒子“官僚”能够让有些粒子重，而让包括希格斯粒子的另外一些粒子轻，但是，就像工作热情过高的官员一样，包含虚粒子的量子路径对所有粒子的质量都有同等的贡献。因此我们期待所有粒子，包括希格斯粒子在内，都具有大统一标度的质量。

如果没有新的物理理论，那么解决希格斯粒子质量过大问题的唯一（且并不令人满意的）方法是：假设它的经典质量值会恰好抵消巨大的量子贡献（这样，经典值就有可能为负）。尽管单个的贡献都很大，但理论中决定质量的参数必须使所有的贡献相加起来的数值很小，这就是上节里我提到的微调。

这虽然可以想象，但要在现实中发生几乎是不可能的。这一问题不单单是加进一个参数稍作调整就能得到正确质量的问题，而是因为这一参数太大了，结果又要求无比精确：精确值少于13位则会导致结果完全错误。要明确的是，这一离奇的参数不同于精确测量的某个量，如光速。通常情况下，定性预言并不依赖于任何一个参数取什么特定的值。只有一个值能契合已测得的精确量，但即便一个参数稍有不同，世界也不会产生很大变化。也就是说，牛顿引力常数（它确定了引力强度）的值即使1%的误差，结果根本不会有什么太大的变化。

而在大统一理论里，参数的稍微一点变化足以彻底毁掉理论预言，不论定量的预言还是定性的预言。打破弱电对称的希格斯粒子的质量异常敏感地依赖于一个参数。对这个参数的几乎所有数值，大统一质量和弱力质量之间的等级分化都不会存在，因而也就没有依赖这个等级的结构和生命。即使这一参数仅出现了1%的偏差，希格斯粒子

的质量也会变得很大，弱规范玻色子以及其他粒子的质量也都会大得多，而标准模型的结果也就根本不是现在这个样子了。

粒子物理的等级问题

上一节里，我们看到了一个很大的谜团：大统一理论的等级问题。而真正的等级问题更糟：尽管是大统一理论首先惊醒物理学家意识到了等级问题，但即使在没有GUT质量的粒子理论中，虚粒子仍会对希格斯粒子质量引发过大的贡献，甚至连标准模型都值得怀疑。

问题在于，将标准模型与引力结合起来的理论会包含两个差别极大的能量尺度。

一个是弱力能标，即250 GeV，在此能量上，弱电对称破缺。当粒子能量在此标度之下时，弱电对称破缺的效果是很明显的，弱规范玻色子和基本粒子都有质量。另一能量是普朗克能量，高达 10^{19} GeV，比弱力能标大 10^{16} 倍，即1亿亿倍。普朗克能量决定了引力作用的强度：牛顿定律说，强度与其能量的平方成反比。因为引力强度很小，普朗克质量（由 $E=mc^2$ 与普朗克能量相联）就会很大，巨大的普朗克质量就等同于极其微弱的引力。

到现在为止，普朗克质量还没有进入我们粒子物理学的讨论范畴。因为引力太过微弱，对大多数的粒子物理计算，它都可以忽略不计。但这正是粒子物理学一直想要解答的问题：为什么引力如此微弱，以至于在粒子物理计算里都可以忽略？等级问题换一种方法来问就是：为什么普朗克质量如此巨大？与粒子物理尺度相关的质量都不超过几百GeV，为什么普朗克质量却要大出1亿亿倍？

为方便比较，你可以想象两个低质量的粒子，如两个电子间的引力作用。这种引力吸引相比两个电子之间的斥力要弱1万亿亿亿亿倍，只有当电子的质量比实际再重100万亿亿倍时，这两种力才有可比

性。而这是一个巨大的数字——就好比要把整个可见宇宙都填满曼哈顿岛，你所需要的岛的数量。

普朗克质量要远远大于电子以及我们所知的其他粒子的质量，这表明引力比其他已知力要微弱许多。但为什么在力的强度之间会有这么巨大的差异——或者，换句话说，为什么普朗克质量相比其他已知粒子的质量会如此巨大？

对粒子物理学家来说，普朗克质量与弱力质量之间的巨大差异达到1亿亿倍，这真的让人难以接受。这一数字，比从宇宙大爆炸起到现在经过的分钟数都要大；若在地球和太阳之间摆满玻璃球，它是所需玻璃球数量的1 000倍；如果用硬币来核算美国财政赤字，那么它就是这些硬币数量的100倍！为什么描述同一物理系统的两个质量会有如此巨大的差异？

如果你不是物理学家，无论这些数字大得有多离奇，你可能也感觉不到这一问题本身有多重要，毕竟，并不是所有事情都是我们能够解释的，也许这两个质量的差异我们也找不出什么原因。但是，实际情况更糟：不能解释的不仅仅是巨大的质量比例问题，我们在下一节会看到，在量子场论里，所有与希格斯粒子发生相互作用的粒子都会参与一个虚过程，使希格斯粒子质量的值增大到普朗克质量，高达 10^{19} GeV。

事实上，任何一个诚实的物理学家，假设他知道引力的强度，而对弱力规范玻色子已测得的质量一无所知，如果你请他用量子场论来估算希格斯粒子的质量，他预言的希格斯粒子质量（由此也得出弱规范玻色子质量）也不会大出1亿亿倍。这就是说，他会由计算得出结论，普朗克质量与希格斯质量（或者说由希格斯粒子质量来确定的弱力级质量）之比更接近于1，而不是1亿亿！他所计算的弱力质量会接近于普朗克质量，这样，所有粒子都将变成黑洞，而我们了解的粒子

物理知识也将不复存在。尽管无论对弱力质量还是普朗克质量的数值，他都没有先验的预期，但他可以用量子场论估算二者的比值——而他必会大错特错。显然，这里存在着一个巨大的差别，下一章我们就解释其原因。

高能量的虚粒子

普朗克质量进入量子场论计算的原因非常微妙。正如我们所见，普朗克质量决定了引力强度，根据牛顿定律，引力与普朗克质量的值成反比。而引力如此微弱则告诉我们，普朗克质量非常巨大。

通常来讲，在做粒子物理学方面的预言时，我们可以忽略引力，因为对于一个质量大约为250 GeV的粒子，引力作用完全是可以忽略不计的。如果你真的需要解释引力效应，你可以系统地兼容它们，但通常不值得去费事。在后面的几章里，我会给出一个全新的截然不同的图景，其中高维引力非常强大，不能被忽略，但在传统的四维标准模型里，忽略引力是惯常而合理的做法。

但是，普朗克质量还有另外一个作用：在一个可靠的量子场论的计算里，它是虚粒子所具备的最大质量值，如果粒子携带的质量大于普朗克质量，计算将不再可信，广义相对论也不再可信，只能由一个更为综合的理论（如弦理论）来代替。

可如果粒子（包括虚粒子）质量小于普朗克质量，惯常的量子场论仍然适用，依据量子场论的计算应当也是可信的。这就意味着，即使计算涉及了质量几乎高达普朗克量级的虚顶夸克（或任何其他虚粒子），结果也是可靠的。

这里的等级问题是，来自极高质量的虚粒子对希格斯粒子质量的贡献会大至普朗克量级，比我们需要的希格斯粒子质量，即能给出正

确的弱力质量和其他基本粒子质量的质量，要大出1亿亿倍。

设想一个路径，如图12-2所示，在这一路径中，希格斯粒子将转化成一对虚顶夸克-反顶夸克，我们会看到，对于希格斯粒子的质量，这一贡献太大了。事实上，与希格斯粒子发生作用的任何类型的粒子都可能会是一个虚粒子，而其质量^[51]达到普朗克量级，所有这些可能路径的结果就形成了对希格斯粒子质量的巨大贡献。而希格斯粒子必须要远小于这一质量。

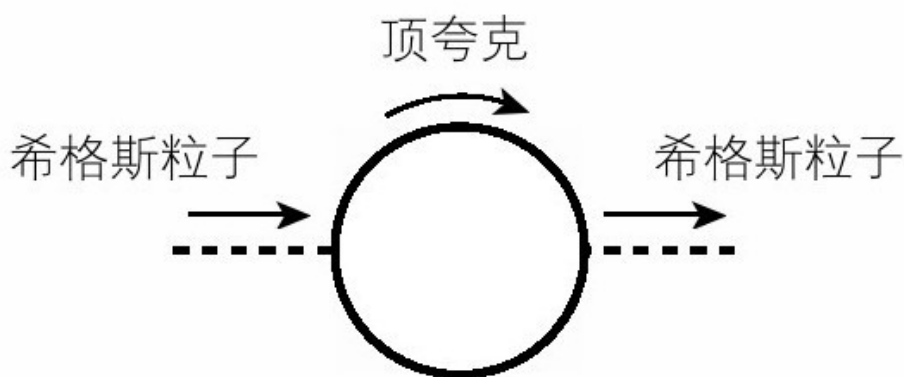


图12-2 虚顶夸克和虚反顶夸克对于希格斯粒子质量的贡献。希格斯粒子可以转化成一对虚顶夸克和一对虚反顶夸克，它们会对希格斯粒子的质量产生巨大的贡献。

在目前这种状态下，粒子物理学就像是一个太过高效的“涓滴式”理论。在经济学里，财富的等级分化并不难达到。利益均沾的涓滴经济学从来就没能增加穷人的财富，更不用说提高上流社会的经济水平了。但在物理学里，财富的转化却太有效了。如果一个粒子质量很大，量子贡献则会告诉我们所有基本粒子的质量都应该很大，结果所有的粒子都具有质量。但我们由测量得知，在我们的世界里，高质量（普朗克质量）和低质量（粒子质量）是同时存在的。

如果不对标准模型进行修正或推广，粒子物理只有通过神奇的经典质量值才能使得希格斯粒子获得小质量，而这一神奇值必须极大（且有可能是负值）才能完全抵消巨大的量子贡献。所有的质量贡献相加必须等于250 GeV。

正如我们前面探讨的大统一理论一样，要做到这些，质量必须是一个经过微调的参数，而这一微调参数必须极大且精确得惊人，才能使希格斯粒子恰好得出一个很小的最终质量。来自虚粒子的量子贡献和经典贡献两者之间必有一个是负值，而且两者量级几乎相同。正、负两个数值都要达到16个数量级，相加起来必须得到一个很小的数。这一微调要求的精确度必须达到16位，比你要让一支铅笔的尖朝下立起来所需做的微调要求还要高，这就像让一个人随机选择，在那场游戏里打赢艾克一样。

要保证希格斯粒子质量很轻，那么标准模型就需要微调。粒子物理更希望一个模型不需要涉及这种高要求的微调，尽管作为无奈之举，我们可能会做一些微调，但我们讨厌这样。微调几乎是一种耻辱，反映了我们的无知。虽然不太可能的事常会发生，但你越想它发生，它却越少发生。

等级问题是标准模型所面临的最为迫切的问题，但让我们乐观一点来看：等级问题提供了一个线索，让我们去发现究竟是什么扮演着希格斯粒子的角色并打破了弱电对称。取代两个希格斯场理论的任何理论，都应该自然地包括或预言一个很低的弱电质量——否则这一理论根本不值得考虑。许多潜在理论都能与我们看到的物理现象和谐相容，但很少有理论触及等级问题，并以令人信服地避免了微调的方式把轻质量的希格斯粒子包括在内。

将力统一起来的任务虽然艰巨，但这一来自高能物理的理论引人入胜、充满了诱惑；而解决等级问题，是我们必须完成的一项实实在

在的任务，它将激励着低能量研究领域的进步。使得这一挑战愈发振奋人心的是，任何解决等级问题的理论都会产生可由LHC测量的实验结果。实验者们期待，能在LHC的实验里找到质量在250~1 000 GeV之间的粒子。如果没有其他粒子，也就没有办法解决这一问题。我们很快会发现，解决等级问题的实验结果有可能是超对称伙伴，也有可能是我们将在以后讨论的在额外维度航行的粒子。

-
- 尽管我们都知道希格斯机制赋予粒子质量，但体现希格斯机制的最简单的已知例子只有通过一个巨大的附加参数才会奏效。在这一简单的理论里，规范玻色子和夸克的质量要比其实际质量大出1亿亿倍。等级问题问的就是，为何事情并非如此？
 - 等级问题起因于弱力质量与巨大的普朗克质量之间的差异（见图12-3），后一质量对引力非常重要——其巨大的质量值告诉我们引力非常微弱。因此，换一种方式来问等级问题就是：“引力为什么这么微弱，要远远地小于其他三种已知力。”



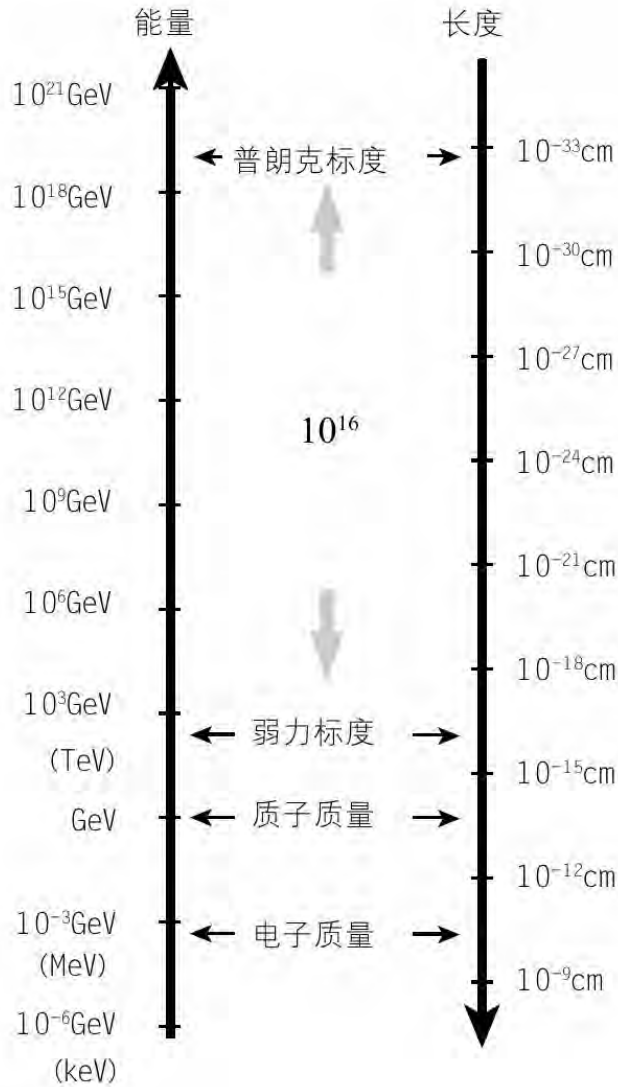


图12-3 弱力质量与普朗克质量的差异。等级问题即为什么普朗克级能量要远远地大于弱力级能量。

- 任何解决等级问题的理论都可以用实验来验证，因为它们对以高于弱力能标的能量运行的对撞机都有实验意义。大型强子对撞机很快就将探索这样的能量。



超对称：超越标准模型的大飞跃

WARPED PASSAGES

你为我而生，我生来为你。

吉恩·凯利 (Gene Kelly) 《雨中曲》

天堂里同样危机四伏

艾克第一次到天堂时，先被领进了一个培训课堂。在那里，他了解了天堂的规则。令他惊讶的是，右翼宗教团体的理念居然是正确的，实际上，家庭准则是他所处的这一新环境的基础。权力机构很久以前就确立了一个传统的家庭结构，来保证世代的划分及婚姻的稳定：上层的男子必然要娶下层的女子；一个风姿俊秀的人必然匹配一个形貌古怪的人；城里的姑娘必然要嫁乡下的小伙子。所有人，包括艾克在内，对此安排都很满意。

但是，艾克后来才得知，天堂里的社会结构也并不总是坚如磐石。起初，曾有过一些危险的、能量强大的潜入者威胁到社会的等级基础，但是，在天堂里，大多数问题都能得到解决。上帝为每人派了一个天使，天使与他们护卫的人一同英勇协作，消除了等级隐患，维护了社会的正常秩序。正是这样，艾克现在才能够乐在其中。

即便这样，天堂也并非安然无恙。结果呢，天使反倒成了自由派，没有条文规定他们必须属于哪一代。一些反复无常的天使，曾经英勇地捍卫了等级制度，现在却威胁要摧毁天堂的家庭准则。艾克感到惊恐，尽管传说中天堂是那么的美好，但他却发现，这里也是一个危机四伏的地方。

“超”在物理学术语里比比皆是：我们有超导体、超冷却、超饱和、超流体，还有超导体对撞机（SSC）——如果不是在1993年被国会否决，它就会成当今能量最高的对撞机了。当然，带“超”的物理学术语还有很多。因此，当物理学家发现时空对称本身也有一个更大的“超”形式时，你可以想象他们有多兴奋。

“超对称”的发现真是令人惊讶。最初创建超对称理论的时候，物理学家还以为他们早就知道了空间和时间的所有对称。时空对称是我们比较熟悉的，在第9章中我们已见证过这种对称，它告诉我们，单从物理定律无法辨别自己处在什么位置、面对什么方向或现在是什么时刻。例如，假设你在加州或纽约打篮球，篮球的轨迹不会因为你在球场的哪一边而改变。

1905年，随着相对论的诞生，空间对称变换又囊括了运动速度和方向的变换。物理学家以为，这就是最后的对称了，没人相信还会有未发现的涉及时间和空间的对称。1967年，杰弗里·曼杜拉（Jeffrey Mandula）和西德尼·科尔曼这两位物理学家证实，不再存在其他对称，人们的这一直觉由此形成了定式。但是，他们（以及所有人）都忽略了基于非常规假设的一种可能。

本章我们将介绍超对称——一种能够将玻色子和费米子交换的新型的奇异对称变换。现在，物理学家能够构建包含超对称的理论，虽然超对称作为自然的对称，还只是一种假设，因为至今没有人在我们周围的世界里发现它。不过，物理学家还是有两个主要的理由认为它可能存在于我们的世界里。

超对称

一种能够将玻色子和费米子交换的新型的奇异对称变换。超对称不一定能够解释弱力质量与普朗克质量巨大差异的根源，但确实消除了对希格斯粒子质量的巨大量子贡献。

第一个理由是超弦理论。超弦理论包含了超对称，是唯一已知可能重现标准模型粒子的弦理论形式。如果没有超对称，弦理论就不大可能描述我们的宇宙。下一章我们将对此进行全面的探讨。

第二个理由是超对称理论有可能解决等级问题。超对称不一定能够解释弱力质量与普朗克质量巨大差异的根源，但确实消除了对希格斯粒子质量的巨大量子贡献。等级问题是一个十分令人伤脑筋的问题，面对这一问题，很少有提议能够经得起实验验证和理论推敲。直到额外维度理论被提出成为可能的解决方法之前，超对称是唯一有希望的答案。

至今还没人知道，现实世界中究竟是否存在超对称，而我们现在所能做的就是估量这些候选理论及其结果。只有这样，当实验达到高能量时，我们就会有充分的准备来辨认，支撑标准模型的究竟是什么物理理论。所以，我们现在就来看看可供选择的理论都有哪些。

费米子和玻色子：一种不太可能的对称

在一个超对称的世界里，每一个已知粒子都会与另一个粒子配成一对——即它的超对称伙伴，通过超对称变换，两者可以对调。超对称变换将费米子换成它的玻色子伙伴，或把玻色子换成它的费米子伙伴。我们在第6章中看到，玻色子和费米子是两种截然不同的粒子类型，量子力学理论根据自旋将它们区分开来。费米子是半整数自旋，而玻色子是整数自旋。整数自旋是常见物体在空间旋转都可采取的数值，而半整数自旋是量子力学才有的奇异特征。

在超对称理论里，所有的费米子都可以转换成它的玻色子伙伴，而所有玻色子也都可以转换成它们的费米子伙伴。超对称是描述这些粒子理论的一个特征。如果你弄混了描述粒子行为的方程，通过超对

称变换对调了玻色子和费米子，结果方程仍会保持不变。在对称变换前后作出的预言不会存在差别。

乍看上去，这种对称根本不符合逻辑。对称变换应该是使系统保持不变，但超对称变换对调的是明显不同的两种粒子：费米子和玻色子。

尽管我们不会想到对称能够混淆差异这么大的东西，但有几队物理学家还是证明这是有可能的。20世纪70年代，欧洲和俄罗斯的一些物理学家证明：对称可以对调差异很大的粒子，而且在玻色子和费米子对调的前后，物理定律可以保持不变。

这一对称与我们前面所讨论过的对称有点不同，因为它所对调的物质明显有着不同的属性。但是，如果玻色子和费米子以相同数目出现，对称就有可能存在。打个比方，假设有相同数量、不同大小的红色玻璃球和绿色玻璃球，同一种颜色的玻璃球大小都不同。假如你和朋友玩一个游戏，你拿红色玻璃球，朋友拿绿色。如果红、绿玻璃球恰好配对，无论选择哪种颜色，你都不会占优势。但是，如果任一给定大小的红、绿球数量不等，这就不是一场公平游戏了。选择红色还是绿色，对你很重要，而如果你和朋友交换颜色，游戏进展就会大大不同。要想对称成立，每一个大小的玻璃球必须既有红色，也有绿色，而同一大小的两种颜色的玻璃球数量必须是相等的。

同样的道理，只有当费米子和玻色子能够恰好配对时，超对称才有可能。正如要进行对调的玻璃球必须大小相同一样，配对的玻色子和费米子也必须有相同的质量和电荷，而且它们的相互作用也必须由同样的参数控制。换句话说，每一个粒子与它的超对称伙伴必须有相似的属性。如果一个玻色子经受强力，那么它的超对称伙伴也必须经受强力；如果有相互作用涵盖了一定数量的粒子，那么相关作用也应涵盖它们的超对称伙伴。

物理学家之所以认为超对称这么令人兴奋，其中一个原因就是：如果它果真能在我们的世界被发现，那么它将是近一个世纪以来首个新型的时空对称发现。因此，它才被称作“超级的”。在此我将不做数学解释，知道超对称能够交换不同自旋的粒子，就足以推想出一种联系。因为自旋不同，玻色子和费米子在空间旋转时会有不同变换，为了弥补这一差别，超对称变换必须包括空间和时间。

不要以为这就意味着你能够描绘单个超对称变换在物理空间中的样子，即使物理学家也只能从数学描述及其实验结果来理解超对称。我们很快将看到，这着实异常精彩。

超对称的发展历史

如果你想，完全可以跳过这段，因为本节将追溯历史，不会介绍重要的新概念。但是，超对称的发展是一个很有趣的故事，部分是由于：它充分展现了新观点的可变通性以及弦理论和模型构建有时所表现出的相辅相成的关系。弦理论激励着超对称的研究，而超弦理论——最有希望描述现实世界的理论，得到了许多人的认可，只是因为它来自对超引力（包括了引力的超弦理论）的洞察。

1971年，法国物理学家皮埃尔·拉蒙（Pierre Ramond）首次提出超对称理论。他将自己的理论置于一个两维背景之下：一维空间，一维时间，而不是我们（曾经）认为我们所生活的四维世界。拉蒙的目标是找一种使弦理论里包含费米子的方法。由于技术原因，弦理论的最初形式只含有玻色子，但费米子对希望于描述世界的所有候选理论都是必不可少的。

拉蒙的理论包含了两维的超对称，并最终发展为费米子弦理论。这是他与安德烈·内沃（Andre Neveu）和约翰·施瓦茨（John Schwarz）共同创建的。拉蒙的理论是在西方世界出现的第一个超对称理

论，当时苏联科学家也发现了超对称，但由于冷战原因，他们的论文并不为西方世界所知晓。

因为四维量子场论远比弦理论的基础坚实，因此，一个明显的问题就是：超对称在四维世界中是否可能实现？但是，由于超对称被精密地织入时空结构中，要由两维推广到四维，并不是一项轻而易举能完成的任务。1973年，德国物理学家朱利叶斯·韦斯（Julius Wess）和意大利物理学家布鲁诺·朱米诺（Bruno Zumino）创建了一个四维超对称理论；在当时的苏联，两位物理学家沃尔科夫（Volkov）和阿库洛夫（Akulov）也独立得出了另一四维超对称理论。但是，冷战再一次阻断了观念的交流。

这些先驱们创建了四维超对称理论后，立即引起了更多物理学家的关注。但是，1973年的韦斯-朱米诺模型并不能完全兼容标准模型的所有粒子，还没有人知道怎样将承载力的规范玻色子加进四维超对称理论。1974年，意大利物理学家瑟吉奥·费拉拉（Sergio Ferrara）和布鲁诺·朱米诺解决了这一问题。

2002年，在我们参加完弦理论会议，乘火车由剑桥去往伦敦时，瑟吉奥告诉我如果不是超空间（时空的抽象扩展，另外包含了一个费米维）的形式，要发现正确的理论几乎是不可能的。超空间是一个极其复杂的概念，我不打算描述它。重要的是，这一全然不同类型的维度（并非空间维度）在超对称的发展中发挥了极为重要的作用。这一纯理论的工具今天仍在继续简化超对称的计算。

费拉拉-朱米诺理论告诉物理学家怎样在一个超对称理论里包含电磁力、弱力和强力，但超对称理论仍没有包含引力。所以，超对称理论仍存在一个问题：它是否能够兼容剩下的这一作用力。1976年，瑟吉奥·费拉拉、丹·弗里德曼（Dan Freedman）和皮特·范·纽文豪

生 (Peter Van Nieuwenhuizen) 三位物理学家创建了超引力理论，解决了这一问题。这是一个包含了引力和相对论的复杂的超对称理论。

有趣的是，在形成超引力论的同时，弦理论也在独立地发展着。在弦理论的一项重要理论成果里，斐迪南多·格里奥兹 (Ferdinando Gliozzi)、乔尔·谢尔克 (Joel Scherk) 和戴维·奥利弗 (David Olive) 等人发现了一个稳定的弦理论，这是雷蒙与内弗、施瓦茨等人创建的费米子弦理论的自然发展。结果发现，费米子弦理论包含了一类人们以前只在超引力中才会遇到的粒子。新粒子与引力子的超对称伙伴——引力微子有着相同的属性，最终证明它果然就是引力微子。

由于超引力的同步发展，物理学家抓住了这两个理论的共同元素，并很快认识到，超对称就存在于费米子弦理论中。这样，超弦理论诞生了。

在下一章里，我们将回到弦理论和超弦理论。现在，我们将重点来看超对称的其他重要应用：它对粒子物理和等级问题的作用。

标准模型的超对称延伸

如果能将已知粒子配对，超对称将是最经济、最吸引人的。但是，想要变成现实，标准模型必须包含相同数量的费米子和玻色子——而它却不符合这一标准。这就告诉我们，如果宇宙是超对称的，那么它一定还包含着更多的新粒子。事实上，它所包含的粒子数量，至少是我们迄今实验测知的两倍。标准模型里的所有费米子——三代夸克和轻子，一定都配有它们新的还未被发现的玻色子超对称伙伴；而规范玻色子——传递力的粒子，也一定有它们的超对称伙伴。

在一个超对称宇宙里，夸克和轻子的伙伴应是一些新的玻色子，因为物理学家喜欢一些古怪（却系统）的命名方式，便称它们为超夸

克和超轻子。总的来说，费米子的玻色子超对称伙伴都与它有着相同的名字，只是在其前面加了一个“超”（S）。例如，电子与“超电子”配对；顶夸克的伙伴是“超顶夸克”。每一费米子都有一个玻色子超对称伙伴，即与它配对的超“费米子”。

这些粒子的属性与它们的超对称伙伴严格一致：玻色子超对称伙伴与它们对应的费米子有着相同的质量和电荷，也有相关联的相互作用。例如，如果电子的电荷是-1，那么超电子的电荷也是-1；如果中微子通过弱力相互作用，那么超中微子也是如此。

如果宇宙是超对称的，那么玻色子一定也有它的超对称伙伴。标准模型里的已知玻色子都是力的承载者：光子、弱玻色子W和Z、胶子，它们的自旋都是1。超对称的命名方式规定，新的费米子超对称伙伴，应与它们对应的玻色子有相同的名称，只是加一后缀“微子”。因此，规范玻色子的费米子伙伴就是“规范微子”：胶子的费米子伙伴是胶微子，希格斯子对应的是希格斯微子。就如玻色子超对称伙伴一样，费米子超对称伙伴也与它对应的玻色子有着相同的电荷和相互作用。而且，如果是绝对超对称，还会有相同的质量（见图13-1）。

你可能会觉得奇怪，既然从没人发现过超对称伙伴，物理学家为什么对超对称的存在深信不疑？某些同事的信心也常令我感到惊讶。但是，除了超对称从未在自然界中被发现过外，还有许多原因让人对其存在持怀疑态度。瑟吉奥·费拉拉是第一批从事超对称研究的物理学家之一，在去往伦敦的火车上，他对我说：“像这么一个令人称奇又引人入胜的理论构建，在研究宇宙的物理学理论中没有发挥任何作用，实在是令人难以置信的。”这代表了许多物理学家的观点。

其他物理学家并不完全相信对称的完美，但他们相信超对称主要是因为标准模型的超对称延伸所带来的益处：与非超对称的理论不

同，它们维护了轻希格斯粒子和质量的等级。

	粒子	超对称伙伴
	轻子	超轻子
例子	电子	超电子
	夸克	超夸克
例子	顶夸克	超顶夸克
	规范玻色子	规范微子
例子	光子	光微子
	W 玻色子	W 微子
	Z 玻色子	Z 微子
	胶子	胶微子
	引力子	引力微子

图13-1 粒子和其超对称伙伴。

超对称与等级问题

标准模型里的等级问题也可以这样问：希格斯粒子为什么这么轻？虚粒子对其质量的量子贡献如此巨大，怎么可能会有这么轻的希格斯粒子？这庞大的量子贡献告诉我们，只有加进一个庞大的、不合情理的参数，标准模型才会发挥作用。

标准模型的超对称延伸的一大优点是，当既有来自粒子也有来自超对称伙伴的虚粒子贡献时，超对称保证不会出现那种使轻希格斯粒子几乎不可能的巨大量子贡献。超对称理论的相互作用只有那些关联的费米子和玻色子的相互作用。正是由于这些限制，超对称理论不含有对粒子质量的大量子贡献问题。

在超对称理论里，对希格斯粒子质量贡献的虚粒子不仅是标准模型的虚粒子，而且其虚的超对称伙伴也有贡献。由于超对称的这一显著特征，两种贡献相加总是零。给希格斯粒子质量量子贡献的虚费米子和玻色子恰好相等，这样，它们各自的大贡献便保证能互相抵消。费米子的贡献值为负，恰好能抵消玻色子的贡献值。

图13-2阐释了这样的抵消。其中有两种虚粒子，一个是虚顶夸克，另一个是虚超顶夸克，它们都会给希格斯粒子质量带来很大贡献。但在超对称理论里，由于粒子和相互作用之间的特殊关系，来自顶夸克和超顶夸克的对质量的大量子贡献相加为零，这样它们就被抵消了。

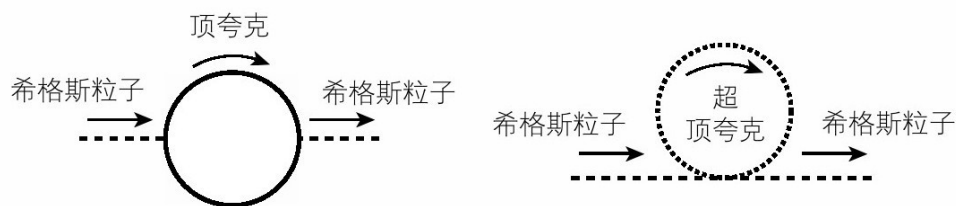


图13-2 在超对称理论中，希格斯粒子质量得到的量子贡献既来自粒子，也来自超对称粒子。在这一图示里，左图是一个虚顶夸克和一个虚反顶夸克；右图是一个虚超顶夸克和一个虚反超顶夸克。两个图例看上去不同，是因为费米子和玻色子的相互作用是不同的。但是，当两个图例里给希格斯粒子质量作出的贡献相加时，它们便相互抵消了。

在非超对称理论里，巨大的量子贡献会摧毁低能量上的弱电对称破缺，除非引入一个巨大的不太可能的参数使大量子贡献相加得到一个很小的数值。但标准模型的超对称延伸则保证，图13-2所示的任何潜在不稳定影响的总和都为零。希格斯粒子的小经典质量值保证了其中包含量子贡献的真实质量仍然会很小。

超对称就像是标准模型的一个灵活而又稳定的基础。如果把标准模型的微调看作将铅笔尖朝下立起来所需做的平衡，那么超对称就像是一根将铅笔固定的细线。或者，如果你把等级问题看作移民局官员的越职审查，由此延误了许多信件，那么超对称伙伴就像是保护公民自由的检察官，他们给移民局官员以一定的限制，让大多数信件得以迅速通过。

由于常见的虚粒子贡献与超对称粒子贡献相加为零，超对称就保证了虚粒子的量子贡献不会将低质量粒子从理论中排除出去。在超对称理论里，即使将虚贡献考虑在内，被假定应该轻的粒子（如希格斯粒子）仍会保持很轻。

破缺的超对称

超对称破缺

超对称仍存在一个严重问题：很显然，世界不是超对称的。这并不意味着我们必须放弃超对称的观点，但它意味着，假设自然中存在，超对称不可能是一种绝对对称，就如伴随弱电统一作用力的局域对称一样，超对称也必须破缺。

尽管超对称有可能解决希格斯粒子质量的巨大虚贡献问题，就目前我所提出的形式，超对称仍存在一个严重问题：很显然，世界不是超对称的。怎么可能是呢？如果存在与已知粒子相同质量、相同电荷的超对称伙伴，它们早就被发现了，但没人见过超电子或是光微子。

这并不意味着我们必须放弃超对称的观点，但它意味着，假设自然中存在，超对称不可能是一种绝对对称，就如伴随弱电统一作用力的局域对称一样，超对称也必须破缺。

理论推理显示，超对称可以被没有完全相同质量的粒子和它们的超对称伙伴所打破；微小的超对称破缺效应就能将它们区分开来。粒子与其对应的超对称伙伴之间的质量差异，由超对称的破缺程度来控制。如果超对称只是微弱破缺，那么质量差异会很小；如果破缺严重，则差异很大。事实上，粒子与超对称伙伴的质量差异大小，是描述超对称破缺程度的一种方式。

几乎在所有超对称破缺模型里，超对称伙伴的质量都要比已知粒子大。这很幸运，只有这样，超对称才能与观察结果相符，才能解释我们为什么还未发现它们。重粒子只有在高能量下才能产生，那么可以想见，如果超对称存在，对撞机还没有达到足够的能量生成它们。因为实验探索的能量已达到几百GeV，至今仍未能发现超对称伙伴的事实告诉我们，它们如果存在，质量至少有那么大。

超对称伙伴超过什么质量才能不被探测到，这要取决于那一特定粒子的电荷和相互作用。相互作用越强，粒子越容易生成，因此，为了避免被探测到，相互作用较强的粒子一定会比那些较弱的粒子更重。如今的实验对大多数超对称破缺模型的限制告诉我们，如果超对称存在的话，所有超对称伙伴的质量一定要超过几百GeV才不会被发现。那些要经受强力的超对称伙伴，如超夸克，甚至会更重——其质量至少有1 000 GeV。

希格斯粒子质量之谜

正如我们看到的，给希格斯粒子质量的量子贡献，在超对称理论里并不是一个问题，因为超对称保证了它们相加为零。但是，我们也看到，若要在现实世界存在，超对称必然会产生破缺。因为，在超对称破缺模型里，超对称伙伴与其对应的标准模型粒子质量不同，所

以，给希格斯粒子质量的贡献就不像精确超对称时那般严格平衡。因此，当超对称破缺时，虚粒子贡献不再恰好抵消。

但是，只要给希格斯粒子质量的量子贡献不那么大，标准模型就不必依赖微调或附加参数。即使超对称破缺——只要作用不大，标准模型仍能包含一个轻希格斯粒子。即便超对称有轻微破缺，也足以消除由高能虚粒子带来的普朗克质量的巨大贡献，有少量的超对称破缺，还无须超乎寻常的抵消作用。

我们希望超对称破缺足够小，这样就可以保证超对称伙伴与标准模型粒子之间的质量差异很小，以避开附加参数。结果是，虽然来自虚粒子与其超对称伙伴的量子贡献并不等于零，但也不会远远超出超对称破缺质量差。这告诉我们，粒子与其超对称伙伴之间的质量差异大约应是弱力级质量。在这种情况下，给希格斯粒子质量的量子贡献也大约应是弱力级质量，这正是希格斯粒子所应具备的恰当质量。

因为标准模型的已知粒子都很轻，超对称伙伴与标准模型粒子之间的质量差异应与超对称伙伴的质量差不多，因此如果超对称解决了等级问题，那么超对称伙伴的质量也不会超出250 GeV很多。

如果超对称伙伴质量与弱力级质量大约相同，给希格斯粒子质量的量子贡献也不会太大。在非超对称情况下，量子贡献超过了16个数量级，所以需要—个庞大的、令人难以接受的附加参数来维持希格斯粒子的小质量；而超对称世界则不同，几百GeV的超对称破缺质量，不会给希格斯粒子质量产生过大的量子贡献。

希格斯粒子以及其所有超对称伙伴的质量都不能超过几百GeV（为不致引起大量的量子贡献）。这一要求，再加上实验已在开始寻找质量大约为几百GeV的超对称伙伴，都告诉我们，如果超对称在自然界中存在并解决了等级问题，那么超对称伙伴的质量一定是只有几百GeV。

这非常令人振奋，因为超对称的实验证据似乎触手可及，很快就会在粒子对撞机里出现。现在的对撞机能量只要再增大一点，就足以达到能够生成超对称伙伴的能量。

LHC所要探索的就是这一能量范围，它要找的是质量达到几百GeV的粒子。如果LHC没有发现超对称，那么就意味着超对称伙伴还是太重了，无法解决等级问题。由此，超对称方案就将会被排除。

但是，如果超对称能够解决等级问题，这将是一项意外的实验收获。探索能量达到1 TeV (1 000 GeV) 的粒子加速器会发现，除了希格斯粒子外，还有许多标准模型的超对称伙伴，我们会看到胶微子、超夸克、超轻子、W微子、Z微子和光微子，所有这些新粒子都会与标准模型粒子有相同的电荷，只是要更重一些。有了足够的能量和对撞，这些粒子都应该出现。如果超对称正确，我们会看到它“梦想成真”。

超对称：称量证据

这就留给我们一个突出的问题：自然界中存在超对称吗？陪审团还未入场，没有更多的证据，一切都只能是猜想。但此刻，辩方和控方都持有对各自有利的论据。

我们已提到了相信超对称的两个强有力原因：等级问题和超弦理论。支持超对称的第三个有力证据是，力有可能在标准模型的超对称延伸里统一起来。正如第11章讨论过的，电磁力、弱力和强力的相互作用强度要依赖于能量。尽管乔治和格拉肖原来发现标准模型的力统一了，但更为准确的测量显示，标准模型的统一并不是很有效。图13-3上图显示的是作为能量函数的三种作用强度的图。

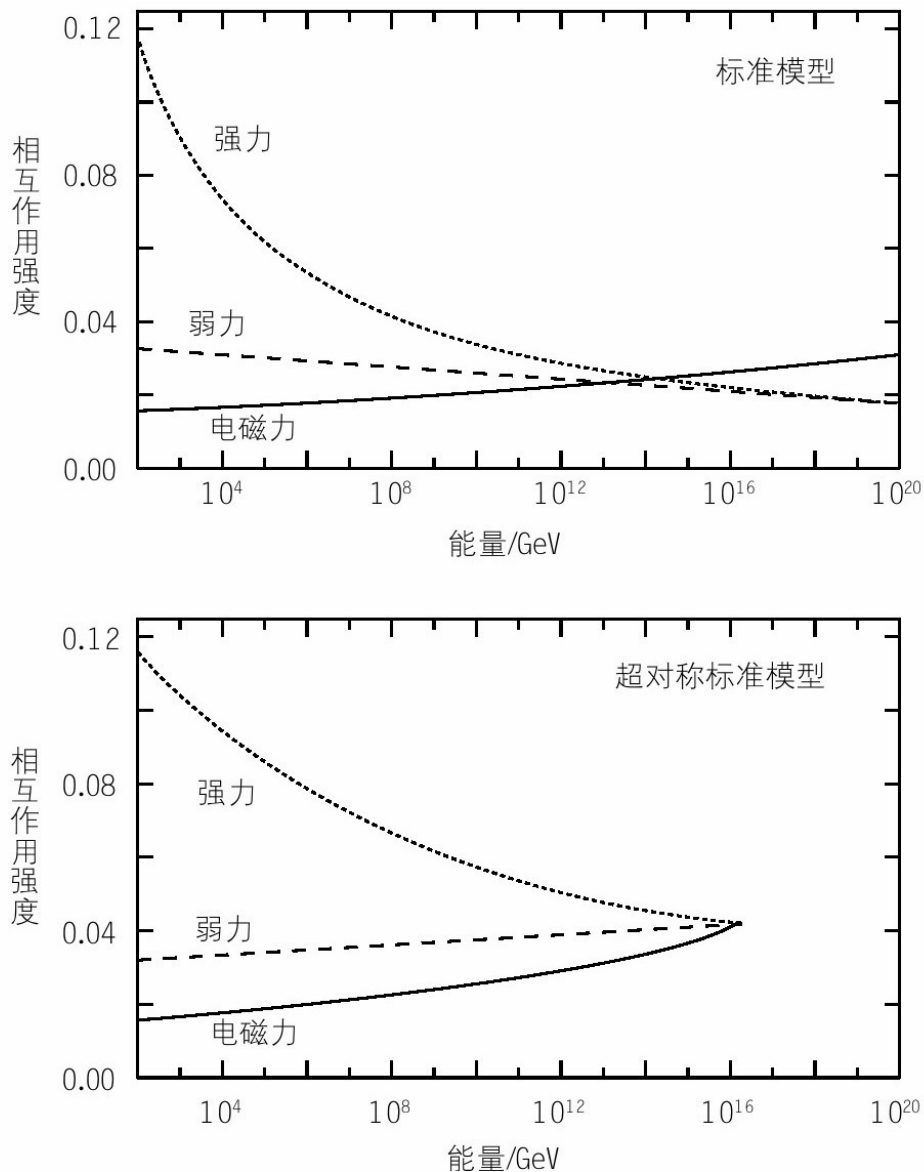


图13-3 标准模型与超对称标准模型。上图表现的是，在标准模型里，作为能量函数的电磁力、弱力和强力的作用强度：三条曲线逐渐靠近，却并未相交于一点。下图表现的是，在标准模型的超对称延伸里，同样的三种力作为能量函数的作用强度：在高能量上，三种力的强度是一样的，这表明三种力有可能统一为一种力。

但是，超对称引出了通过这三种力相互作用的许多新粒子，这改变了力对距离（或能量）的依赖关系，因为超对称伙伴还会作为虚粒

子出现。这些额外的量子贡献进入重整化群计算，并影响了电磁力、弱力、强力等相互作用强度对能量的依赖。

图13-3下图显示的是，当将虚的超对称伙伴作用包括在内时，相互作用强度随能量会产生的变化。值得注意的是，有了超对称，三种力似乎比以前更为精确地统一了起来。这比早先的统一尝试更为重要，因为我们现在对相互作用强度的测量更为准确。三条线的相交可能属于巧合，但它也可以被看作支持超对称的证据。

超对称理论的另一个优点是，它们包含了暗物质的一个有力的候选者。暗物质是遍布于宇宙的不发光物质，是通过其引力作用发现的。即便宇宙中有大约1/4的能量储存于暗物质中，但我们仍不了解它究竟是什么。^[52]一个不会衰变且有恰当质量和作用强度的超对称粒子，可能会是合适的暗物质候选者。事实上，最轻的超对称粒子不会衰变，它们可以具有恰当的质量和相互作用，从而成为构成暗物质的粒子。它可能是光微子，即光子的伙伴；或者，在我们后面将探讨的额外维度图景里，它也可能是规范玻色子W的伙伴——W微子。

但是，超对称并非无懈可击。对超对称最强有力的反驳是，无论希格斯粒子还是其超对称伙伴，都尚未被发现。虽然超对称伙伴也许很快就会被发现，但如果说超对称解决了等级问题，而我们竟尚未观察到它，那么就完全难以理解了。实验已达到了几百GeV的能量，尽管超对称伙伴肯定更重一点，但并没有理由如此。事实上，从解决等级问题的角度来讲，更轻的粒子更为适宜。如果说超对称解决了等级问题，那么为什么我们至今仍未能找到超对称伙伴？

从理论方面探究，超对称也不十分令人信服。因为它的破缺方式，仍是一个未解之谜。我们知道它肯定是自发破缺，但就如标准模型和弱力对称的破缺一样，我们尚不清楚究竟是哪些粒子引起了破

缺。很多人提出了一些新奇的观点，但我们仍在期待一个更加令人满意的四维理论。

第一次接触超对称时，从模型构建角度来看，我发觉它似乎是太过简单了。因为没有量子贡献，超对称理论似乎可以包含完全不相干的质量，即使我们不明白质量差异为什么会出现，它们也不会带来任何问题。从模型构建角度来看，这很令人失望，因为对于还未确定的基本理论来说，它没有提供任何线索，而且，我们在模型构建过程中没有遇到任何挑战，这也显得很无趣。

但是，在这之后我遇到了超对称的“味”问题，这说明它肯定不对。事实上，你很难让一个对称破缺理论的具体细节产生效力。这一问题虽然微小，但非常重要。对于超对称破缺的简单理论，味对称问题是一个重要障碍。所有超对称破缺的新理论都集中在这一点上，我们会在第17章看到，为什么说额外维度里的超对称破缺是一种可能的解决方法。

现在我们回顾一下，标准模型里费米子的味是有着相同电荷却质量不同的三代不同的费米子。例如，上夸克、奇夸克和顶夸克；再如，电子、 μ 子和 t 子。在标准模型里，这些粒子的身份是不会改变的。例如， μ 子永远不会与电子直接相互作用，它们只能通过弱规范玻色子的交换间接地作用。尽管 μ 子可以衰变成电子，这只是因为衰变产生了一个 μ 子中微子和一个电子中微子（如图7-7所示）， μ 子永远也不会不经释放相关的中微子而直接转化成电子。

特定类型的轻子的身份不会发生变化，对此，物理学家的表达方式是，电子或 μ 子数将保持不变。我们指定电子和电子中微子的电子数为正，而正电子和电子反中微子的电子数为负；指定 μ 子和 μ 子中微子的 μ 子数为正，而反 μ 子和 μ 子反中微子的 μ 子数为负。如果 μ 子和电子数保持守恒，那么 μ 子便永远也不可能衰变成电子和光子。

如果真是那样，我们会以正 μ 子数和零电子数开始，而以正电子数和零 μ 子数结束。事实上，从没有人见过这种衰变。就我们所知，所有的粒子相互作用，电子和 μ 子数都将保持不变。

在超对称理论里，电子和 μ 子数守恒告诉我们，尽管如 μ 子和超 μ 子一样，电子和超电子也可以通过弱力相互作用，但电子永远也不会直接与一个超 μ 子相互作用。不论出于什么原因，如果电子和超 μ 子配对，或是 μ 子与超电子配对，就会引发在自然界根本不可能出现的相互作用，比如， μ 子会衰变成一个电子和一个光子。

这里的问题是，尽管在真正的超对称理论里不会发生味改变的相互作用，但一旦超对称破缺， μ 子和电子数守恒即得不到保证。在一个超对称破缺的理论里，超对称相互作用会改变电子和 μ 子的数量——这有悖于实验结果。这是因为，超对称伙伴的大质量玻色子并不像它们对应的费米子一样，身份感很强。在超对称理论里，它们所拥有的质量允许超对称玻色子可以相互混淆身份。例如，对应 μ 子的，不仅可以是超 μ 子，还可以是超电子。但超电子与 μ 子的配对，就可能会产生我们从未见过的各种衰变。任何有关自然的正确理论里，改变 μ 子或电子数的相互作用一定很微弱（或者根本不存在），因为我们从未发现过这种相互作用。

夸克也会遇到类似的问题：当超对称破缺时，夸克的味不会守恒，而且会导致危险的世代混杂，也就是篇首故事里艾克担心的那样。自然界里确实会发生某些夸克的混杂，但其程度却远远要小于超对称破缺理论的预言。

这种味改变的相互作用在自然界里极少出现，解释这一问题，是超对称破缺理论面临的一个严峻考验。不幸的是，大多数超对称理论都不能解释为什么这种味改变的效应并没有出现。这是不允许的：理论要与自然一致，必须禁止这种身份混淆。

如果你仍不太明白，也许听到以下这一事实，你就宽心多了，许多物理学家最初也有同样的感觉，而且也没有把超对称的味问题看得有多重要。简单地说，看法的不同可以根据地域来划分：欧洲人不像美国人看得那么严重。

多年来，我们一直在从其他背景下思考味问题，了解这一问题多么难以解决；而许多人从一开始就忽略了“无政府主义原理”的含义，因而不理解我们为什么小题大做。现在在西雅图核理论研究所工作的一位杰出的物理学家戴维·卡普兰（David B. Kaplan，我读研究生时的第一个合作者），从1994年密歇根国际超对称会议回来后，就向我讲述了他的沮丧：会上，他对听众解释了自己提出的味问题的解决方案，过后才发现，根本没几个人认为那有什么问题！

这很快就发生了变化，大多数人现在都意识到了这一问题的严重性。我们很难找到一个超对称破缺理论，能够既不改变粒子身份，又能给所有的超对称伙伴必要的质量。要成功解决等级问题，超对称理论面临着一个巨大的困难：怎样使对称破缺，而又能防止产生味改变。 μ 子、电子（和夸克）数量的不守恒听上去有点过于专业，但它的确是超对称破缺的一大障碍，真的很难防止超对称伙伴的互相转变，对称通常是无力阻止它的。

因此，我们再次返回主题：有对称的理论是完美的，但描述我们可见世界的对称破缺应该同样完美。超对称为什么破缺？又是如何破缺的？只有一个完美的超对称破缺模型被构建起来后，我们才能够迎接理论挑战，理解超对称。

这并不是说超对称一定是错的，或者说它根本不能解决等级问题。但这确实意味着，超对称理论要成功地描述世界，还需要其他元素。很快我们将看到，这其他元素可能就是额外维度。

-
- 超对称从根本上使粒子增加了一倍：在此理论里，对应每一个玻色子，超对称都会配给它一个费米子伙伴；而每个费米子，都会对应一个玻色子伙伴。
 - 如果没有超对称，量子力学很难维持一个小质量的希格斯粒子，从而使标准模型发挥作用。在额外维度理论出现之前，超对称是解决这一问题的唯一已知办法。
 - 超对称不一定告诉我们希格斯粒子为什么轻，但它确实使轻希格斯粒子的假设听起来有一定道理，从而解决了等级问题。
 - 标准模型粒子与它们的超对称伙伴对希格斯粒子的巨大虚贡献相加为零，因此，在超对称理论里，轻希格斯粒子就不再成问题。
 - 即便超对称有可能解决等级问题，但它也不可能是精确对称的。倘若是这样，超对称伙伴就会与标准模型粒子有相同的质量，我们早就该在实验中发现超对称的证据了。
 - 如果存在超对称伙伴，它们的质量一定会比其对应的标准模型粒子质量大。高能对撞机只能生成一定质量的粒子，这些对撞机可能还没达到足够的能量生成它们，这就解释了为什么我们还没能找到它们。
 - 一旦超对称破缺，就会产生味改变的相互作用。这些过程将夸克或轻子变成了有着相同电荷的另一代的夸克或轻子（只是质量发生了改变）。这是一个很奇怪的过程——它会改变已知粒子的身份，而这在自然界里极少发生。但大多数超对称



-imension

探索大揭秘

破缺理论都预言这会经常发生——远远超过了我们的实验所见。



WARPED

第四部分

弦理论，神秘膜宇宙的创世者

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



我的世界在琴弦上。

弗兰克·西纳特拉 (Frank Sinatra)

艾克四十二世的维度历险

快速前进1 000年。

艾克四十二世正在摆弄他的新设备：Alicxur 6.3型，这可是他刚从太空网买回来的（伊卡洛斯·拉什莫尔三世对速度和机械的挚爱显然经过多代传到了他的身上）。Alicxur可以让用户观察任何大小的东西，可以看到极小，也可以看到极大。艾克四十二世肯定，他所有买了Alicxur的朋友，必然会不约而同地首先去尝试大的尺度，调至百万秒差距的恢宏浩大的场景，这样就能超越已知宇宙，看进外太空了。艾克四十二世却想：“我还不知道极度微小的尺度上都发生了些什么呢。”因此，他决定先到微观世界游历一番。

但艾克四十二世是个急性子，他可没耐心去看那厚厚的一本说明书。管他呢，先闯进去看看再说。他把号码盘调到了 10^{-33} 厘米的设置上，直接就按下了“出发”键，根本没注意到最小尺度范围上覆盖的红色标识。

可令他惊恐的是，他忽然觉得头晕目眩，因为他闯进了一个急剧颤动着的陡峻的世界，那里到处都是弦，而熟悉的世界完全不见了。相反，眼前出现的是一条条盘旋的蛟龙，它们在尽情地翻腾，从一个表面上探出来，再一头扎回去；或者绕着自己翻过来，再转回去，打成一个个环。艾克四十二世跌跌撞撞、拼尽全力地摸索

着“停止”键。还好，赶在自己失去知觉之前他及时地按下了“停止”键，一切才恢复了正常。

安下神之后，艾克四十二世才意识到，也许他应该先看一下说明书的。他翻到了“警示”部分，上面写着：“你的Alicxur 6.3型新设备只适用于大于 10^{-33} 厘米的尺度范围。我们还没有纳入弦理论研究的最新成果，因为直到去年，物理学家和数学家才将其预言与现实世界联系起来。”

艾克四十二世感到很失望，原来只有新出的7.0型才纳入了最新成果。但艾克很快就赶上了弦理论的最新进展，他把自己的Alicxur进行了改造，加大马力，然后再也没有晕过机。

爱因斯坦的广义相对论是里程碑式的，有了它，物理学家对引力场便有了更深入的理解，并且对引力的计算精确到了令人难以置信的地步。相对论成了物理学家预言所有引力系统演变的工具——甚至包括整个宇宙的演化。但是，尽管它所有的预言都很成功，相对论不可能是引力的最终结论，到了极端微小的距离，广义相对论不再适用。在极端微小的尺度上，只有新的引力范式才可能成功。许多物理学家相信，这一范式必然是弦理论。

如果弦理论是正确的，那么它将包容广义相对论、量子力学和粒子物理学的成功预言，但也会将物理学拓展到其他理论所无能为力的距离和能量范围。弦理论现在还不够完善，我们还无法衡量它的高能预言，也无法证实它在这些难以捉摸的尺度和能量范围内的效力。但弦理论的确有几个明显的特征，让人对这一幅充满希望的景象信心倍增。

现在，让我们来看看弦理论和它戏剧性的发展历程。它的高潮是1984年的“超弦革命”，那时物理学家证明，弦理论的碎片能天衣无缝地拼接起来。超弦革命只是开始，现在众多的物理学家积极地投入到了这一轰轰烈烈的研究项目中。本章和随后的几章，我们将回顾弦理论的历史及其新近的令人振奋的发展成果。但我们也会看到弦理论仍

面临着众多艰巨的挑战，物理学家只有解决了这些难题，才可能用它对我们的世界作出预言。

早期的躁动

量子力学和广义相对论在很大的距离范围内和谐相容，包括实验所能观察到的任何尺度。尽管两个理论都应该适用于所有的尺度范围，但对于哪一测量范围由哪一理论主导，物理学家已形成了一种共识。两者都尊重对方在各自指定领域的权威，所以和谐地分享着这些领域：广义相对论适用于大质量的延伸物体，如恒星和星系；但引力对原子的影响是可以忽略的，所以，在研究原子时，你可以放心大胆地忽略广义相对论；而在原子大小的距离上，量子力学是非常关键的，因为它对原子的预言至关重要，而且与经典力学的预言大不相同。

但是，量子力学和相对论也关系也并非完全和谐。在普朗克长度（ 10^{-33} 厘米）这一极小尺度上，这两种截然不同的理论从未有过充分的“交流”。

由牛顿的万有引力定律，我们知道了引力强度与质量成正比，与距离的平方成反比。即使在原子尺度上引力非常微弱，但引力定律告诉我们，在更小的尺度上，引力作用是非常强大的。引力不仅对大质量延伸的物体非常重要；对极端靠近、间距只有普朗克长度的物体，也同样重要。如果我们想要对这些不可测量的微小尺度作出预言，量子力学和广义相对论都会有它们重要的贡献，但两种理论的贡献却是互不相容的。在这一角逐激烈的领域里，量子力学和广义相对论的计算不能相互协调，无论量子力学还是引力都不能被忽略，其预言必然要失败。

只有逐渐地让时空产生弯曲的平滑引力场存在时，广义相对论才会发挥作用。但量子力学告诉我们，能探测或影响普朗克长度的任何东西都具有巨大的动量不确定性。探索普朗克长度的能量会引致混乱的力学过程，例如，虚粒子的高能爆发，这将摧毁广义相对论描述的一切希望。根据量子力学，在普朗克长度上，空间几何不再是逐渐弯曲，而是未来艾克遇到的那种地形——时空狂乱地起伏跌宕，一会儿绕成一个个圆环，一会儿又伸出一条条枝蔓。在这样一个桀骜不驯的世界里，广义相对论毫无用武之地。

当然，广义相对论也不能退场让量子力学自由发挥，因为引力在普朗克长度上会产生极强的作用力。尽管在我们熟悉的粒子物理能量上，引力是微弱的，但在探索普朗克长度的高能量上，它是强大的。^[53]普朗克能量——探索普朗克长度的能量，正好是这样的能量尺度，所以不能再将引力当作弱力而忽略了。在普朗克长度的水平上，引力不能忽略。

事实上，在普朗克级的能量上，引力是一大障碍，因为我们根本不可能使用传统的量子力学进行计算。任何足以探索 10^{-33} 厘米的能量都会被禁闭一切的黑洞所吞噬，只有量子引力理论能告诉我们里面究竟发生了什么。

在极微小的尺度上，量子力学和引力迫切需要一个更为基本的理论。由于两者之间的冲突，除了引进另外一个“仲裁者”来取代它们，我们别无他法。而新的统治体系必须给量子力学和广义相对论以足够的自由，主宰它们各自互不干涉的领域，但同时又要有足够的权威，掌管这片两个旧理论都不能控制的争议区域。弦理论可能就是答案。

量子力学和引力的互不相容，还反映在传统引力对引力子高能相互作用的合情合理的预言上。引力子是量子引力理论中传递引力作用

的粒子。

根据经典引力理论，引力通过引力场在两个庞大物体之间传递，就如电磁力的传递一样：根据麦克斯韦经典电磁场理论，电磁力由一个带电粒子通过经典电磁场传递给另一带电粒子。但电磁力的量子场理论——量子电动力学，以光子的交换重新阐释了经典电磁力。^[54] QED这一关于光子的理论，是经典电磁理论兼容量子力学效应的延伸。

与此类似，量子力学规定，引力的传递也必须有一个粒子，这一粒子就是引力子。在量子引力理论里，两物体之间引力子的交换会重现牛顿定律的引力作用。尽管引力子还没有被直接观察到，但因为量子力学承认它们的存在，所以许多物理学家对此深信不疑。

引力子独特的自旋特征将对我们非常重要。因为引力子传递与空间和时间内在相联的引力，因此它与所有其他已知的力的承载者（如光子）有着不同的自旋。这里，我们不去细究它的原因，但引力子是已知唯一自旋为2的无质量粒子，不像其他规范玻色子那样自旋为1，也不像夸克和轻子那样为1/2。寻找额外维度理论的可信证据时，自旋为2这一事实非常重要。正如我们很快将看到的，引力子的自旋也是我们认识弦理论潜在意义的关键。

但是，量子场论对引力的描述是不完整的。没有一个引力子的量子场论能预言它在所有能量上的相互作用。当引力子的能量高至普朗克级能量时，量子场论就彻底崩溃了。

理论推理显示，在低能量上无关紧要的额外引力子的相互作用在高能量上却很重要；可量子场论的逻辑不足以揭示它们是什么，或如何解释它们。如果我们忽略在低能量上无关紧要的相互作用，不恰当地使用一个引力的量子场论对极高能的引力子作出预言，那么我们得到的结论是：引力子相互作用的发生概率会大于1——这显然是不可能

的。在普朗克能量，或（根据量子力学和狭义相对论）在 10^{-33} 厘米的普朗克长度上，引力子的量子力学描述显然失败了。

普朗克长度比质子小19个数量级，小到如此地步，物理学家本可以置之不理，但它关乎一个根本问题，只有一个更为全面、普适的理论才有可能作出解答。例如，当今的宇宙学理论猜想，宇宙的起源是一个普朗克长度大小的小球，但我们尚不明了大爆炸的“爆炸”。对宇宙的后期演变，我们了解了很多，但它是怎么开始的，我们还不知道。推导出在小于普朗克长度范围适用的物理定律，会帮助我们了解宇宙的最初演变。

再者，还有许多关于黑洞的未解之谜。这些重要的未解问题包括：在黑洞的视界发生了什么？在奇点又发生了什么？视界是一个不归之地，没有任何东西能够脱逃它；而奇点在黑洞的中心，广义相对论不再适用。另一未解之谜是，掉进黑洞的物体信息是如何储存的？与我们感受到的引力作用不同，黑洞里面的引力效应很强，在寻常的平坦空间看来，那就像是能量高达普朗克能标的物体产生的效应。如果不能找到一个可以和谐包容量子力学和广义相对论的理论——一个在 10^{-33} 厘米普朗克长度上的量子引力理论，我们将永远无法解答黑洞问题。黑洞展示了某些只有量子引力理论才能解决的强引力效应。弦理论是已知最有希望的候选者。

弦是一切之源

有关物质的本质，弦理论的观点与传统粒子物理学大不相同。根据弦理论，物质的已知最根本不可分结构是弦——振动的一维能量环或线段。通常的物质，如小提琴的弦，都由原子构成，而原子由电子和原子核构成，原子核又由夸克构成。这些弦则不一样，事实上恰恰相反：它们是最根本的弦。这意味着所有东西，包括电子和夸克，都

是由它们的振动构成的。根据弦理论，小猫玩的纱线球是由原子构成的，而其本质上是由弦的振动构成的。

弦理论大胆地假设粒子是由弦的和谐共振产生的，对应每一个粒子都会有一个基本弦的振动，而振动的特征就决定了粒子的属性。因为弦可以有多种振动方式，由此就产生了多种类型的粒子。理论学家最初以为，基本弦只有一种类型，它构成了所有的已知粒子。但就在几年前，这一图景发生了变化。现在我们相信，弦理论里包含了多种不同的、独立的弦，每种弦都可能有多种不同的振动方式。

弦只会沿着一个方向延伸，在任一特定时刻，你只需一个数字就能确定弦上的一点，因此，根据我们对维数的定义，弦就是一维（空间维度）物体。但是，就如我们现实世界里的琴弦一样，它们也可以卷起或绕成一个圈。事实上，弦有两种类型：一种开弦，有两个端点；一种闭弦，即没有端点的闭合圈（见图14-1）。

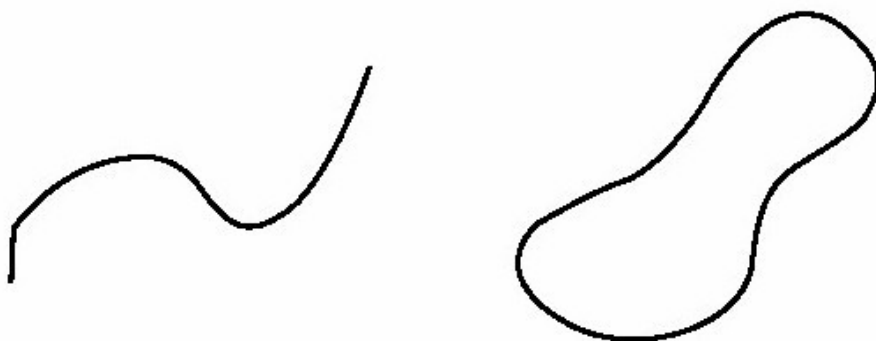


图14-1 开弦和闭弦。

弦将形成的粒子类型取决于弦的能量以及由此激起的具体的振动模式。弦的模式就如小提琴琴弦的共振模式，你可以把这些振动当作能结合成所有已知粒子的基本单位。以这种逻辑来描述，粒子就是合唱，而相互作用就是和声。如果不用琴弓，小提琴不可能发出各种各

样的声音，同样地，弦理论的弦并不总能形成所有的粒子。就如琴弓会使小提琴弦产生各种不同的振动方式一样，能量会激起弦的模式，当弦具备足够的能量时，就会产生不同的粒子类型。

对于开弦和闭弦，共振方式就是沿着弦的长度所进行的整数倍的振动次数。图14-2显示了这样几种方式。在这些模式里，波会以整数倍上下振动，而所有的振动都是在弦上发生的。在开弦里，波的振动到达弦的终点后，接着返回，如此循环往复；而闭弦上的波，则缠绕着闭合弦圈上下振动。任何其他形式的波——那些不能完成整数倍振动的波，都不能存在。

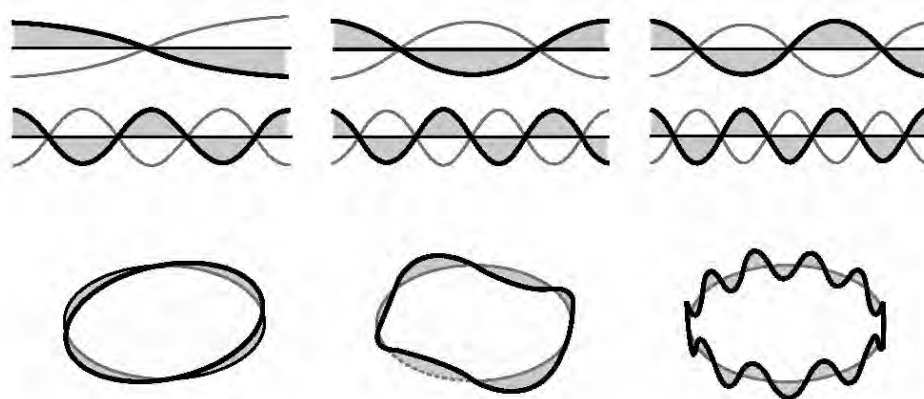


图14-2 开弦（上）和闭弦（下）的一些振动方式。

最终，弦的振动方式决定了粒子的所有属性，如质量、自旋和电荷。一般说来，具有相同自旋和电荷的粒子有很多不同的复本和质量。因为有无限种模式，所以一个弦就有可能生成无数的重粒子。已知的粒子相对较轻，都是由振动最少的弦形成的。我们熟悉的轻粒子，如通常的夸克或轻子，可能就属于没有振动的模式。但高能量的弦会以多种方式振动，因此，弦理论的突出特征在于它的重粒子，它们是由更高的振动模式形成的。

可是，更多的振动就需要更高的能量。弦理论里由更多的振动产生的额外粒子很有可能极重，生成它们需要大量的能量。因此，即使弦理论正确，它的新奇结果也很可能难以探测。我们不期望能以可得到的能量生成新的重粒子，但我们期望弦理论和粒子物理能在可探测的能量上产生同样的可观察结果。如果有关额外维度的某些最新成果是正确的，那么这一景象将发生变化。但现在，让我们先来回顾一些传统的弦理论，然后再探讨额外维度模型。

起源：强力而非引力

超弦理论

弦理论最初出现是在1968年，它本意是要描述强子的强相互作用。超弦理论超越最初弦理论的一个关键优势在于，它包含了半整数自旋的粒子，这使它有可能描述标准模型的费米子，如电子和各种类型的夸克。

到未来艾克的时代，弦理论是可以炫耀其悠久的历史了，但出于科学的目的，我们将把故事局限在20世纪和21世纪早期。现在我们把弦理论当作有可能协调量子力学和引力的理论，但最初它却有着完全不同的应用。这一理论最初出现是在1968年，它本意是要描述强子的强相互作用。那一理论并不成功；就如在第7章里看到的，现在我们知道强子是由夸克通过强力紧紧束缚在一起形成的。但弦理论却得以存续——不是作为强子的理论，而是作为引力理论。

尽管它对强子的描述是失败的，但通过检验强子弦理论所面临的几个问题，我们可以看到关于引力的弦理论的几个良好特征。尤为引人注意的是，强子弦理论的失败之处恰恰是量子引力弦理论的补偿特征（至少不是障碍）。

最初的弦理论遇到的第一个问题是，它含有一个“快子”（tachyon）。人们最初把快子当作行驶速度超过光速的粒子（这一词汇来

源于希腊语的“tachos”，意为“速度”）。但现在我们知道，快子表明了包含它的理论的不稳定性。令科幻迷遗憾的是，快子不是出现在自然界的真实物理粒子：如果你的理论包含一个快子，那么你对它的分析肯定会出错。一个包含了快子的系统最终会转变成一个不存在快子相关低能系统，这一含有快子的系统不能维持足够长的时间来产生任何物理效应，它只是错误的理论描述的一个特征。在确定真正的物理粒子和作用力之前，首先需要对相关的没有快子的稳定结构作出一个理论描述。如果没有这样的稳定结构，理论将是不完整的。

有快子的弦理论似乎是不合情理的，但没人知道以什么方式消除快子从而构造一个弦理论。这就意味着来自弦理论的预言，包括那些有关粒子而非快子的预言，都是不可靠的。这样，可能你就会以为我们有足够的理由放弃强子弦理论，但物理学家把希望寄托在快子不是真实的这一预言上，有人认为这只是在构造理论时所产生的数学近似问题，但事实可能并非如此。

但拉蒙、内弗和施瓦茨发现了弦理论的另一种超对称形式：超弦理论（superstring theory）。

超弦理论超越最初弦理论的一个关键优势在于，它包含了半整数自旋的粒子，这使它有可能描述标准模型的费米子，如电子和各种类型的夸克。但超弦理论还有另外一个收获，即它不含有困扰原版弦理论的快子。超弦理论似乎在各种情形下都更有希望，它没有妨碍其进步的快子不稳定性。

最初的强子弦理论存在的第二个问题是：它包含了一个自旋为2的无质量粒子。计算显示，没有办法能将它消除，而又从未有实验者发现过这种烦人的粒子。如果无质量的粒子相互作用如强子一般强烈，那么实验者早就该发现它们了。这样，强子弦理论似乎陷入了困境。

谢尔克和施瓦茨转变了弦理论的方向，他们证明：

令强子弦理论束手无策的自旋为2的粒子，事实上可能成为引力弦理论的无上荣耀，自旋为2的粒子可能就是引力子。他们继续研究发现，自旋为2的粒子正如引力子应有的表现一样。弦理论包括引力子的一个候选者，这一关键性的发现，使得弦理论成为量子引力理论的潜在理论。没有人能以粒子描述说明怎样构造一个在所有能量都适用的内在一致的引力理论，而弦理论的描述似乎能实现我们的愿望。

另有证据显示，尽管强子弦理论不会奏效，但谢尔克和施瓦茨对引力的弦理论探索可能走对路了。在第7章里我们看到，斯坦福线形粒子加速中心的弗里德曼、肯德尔和泰勒演示了电子由原子核的剧烈散射，表明其中存在的是坚硬、点状的物质，即夸克。这一实验与第6章里描述的卢瑟福的散射实验本质是相同的：在卢瑟福实验里，剧烈的散射让我们发现了一个坚硬的原子核。而这一案例让我们想到，原子核里是点状的夸克，而不是蓬松的、伸展的弦。

可是，弦理论的预言与SLAC的实验结果并不相符：弦永远也不会引起急剧的散射，只有坚硬、紧致的物体才可能做到。因为在既定时间内，只有某些弦发生相互作用，弦的碰撞会更加轻柔，这一相对不那么剧烈、较为温和的散射便终结了强子弦理论。但是，换一个角度，从量子引力理论的视角来看，这一属性似乎颇有前途。

在以粒子描述引力子的理论里，引力子的相互作用在高能量上过于强烈，而一个更完善的理论，不该预言这么强烈的高能引力子的相互作用。弦的引力理论正好做到了这一点。在弦理论里，延伸的弦取代了点状的粒子，这就保证了引力子在高能量上的相互作用不会那么剧烈。与夸克不同，弦不会产生有坚硬内核的散射过程，弦的相互作用发生在一个延伸的区域，有点儿“拖泥带水”，更为松散。这一属性意味着弦理论可能会解决引力子荒谬、不合情理的高频率相互作用

问题，正确预言高能引力子的相互作用。弦的高能软碰撞是引力弦理论可能正确的另一重要标志。

综上所述，超弦理论包含了费米子、承载力的规范玻色子和引力子——我们了解的所有类型的粒子，它不包含快子。而且，超弦理论包含了一个引力子，它的量子描述在高能量上可能是有意义的，看来弦理论有可能描述所有的已知粒子。这是一个颇有希望的关于世界的候选理论。

高潮：超弦革命

即使对于解决如量子引力这么深奥的问题来说，超弦理论仍算得上是极为大胆的一步。引力弦理论预言了无数我们不知道的粒子，而且弦理论极难用计算来分析。解决量子引力问题，我们需要付出多么高昂的代价啊：一个有着无数新粒子的理论，一个几乎无章可循的数学描述。20世纪70年代投入弦理论研究的人，要么极其坚定，要么就是头脑发热。在为数不多的几个人当中，谢尔克和施瓦茨在这条前程未卜的路途上，不畏风险，英勇地探索着。

1980年，谢尔克不幸辞世，施瓦茨继续坚持他的研究。他与当时的另一位（可能是唯一的一位）“皈依者”、英国物理学家迈克尔·格林（Michael Green）协作，共同算出了超弦理论的结果。施瓦茨和格林发现了超弦的一个奇异特征：

只有在十维（其中九维空间，一维时间）中，它才有意义。其他任何数量的维度，不可接受的弦的振动方式都会产生明显不合情理的预言，比如一些过程的负概率，这会涉及一些不该存在的弦的模式。而在十维中，所有这些不该存在的模式都被消除了，任何其他维度的弦理论都没有意义。

要明确的是，弦本身是沿着一个空间维度延伸的，但会穿越时间，这就是拉蒙在最初发现超对称时所研究的两个维度。但正如我们所知，一个点状的物体（它在空间维度上没有延伸，因此空间维度为零）能够在三维空间自由移动；一根弦（它有一个空间维度）也能在一个比自身有着更多维度的空间里四处移动。可以想见，弦可以在三维、四维或更多维度的空间里移动。计算表明，正确的数量是十维（包括时间）。

维度太多并非超弦才有的新奇特征。早期形式的弦理论（不含有费米子或超对称）曾有过二十六维：一维时间，二十五维空间。但早期的弦理论存在其他问题，如快子问题，而超弦理论却充满了希望，值得人们去探索。

即便如此，弦理论还是被人们严重地忽视了。直到1984年，格林和施瓦茨展示了超弦理论的一个令人惊讶的特征，其他许多物理学家才相信他们的研究是很有前途的。这一发现，与我们很快将看到的另外两个重要进展，使得弦理论终于成为物理学的主流。

格林和施瓦茨的研究所探讨的现象被称作“反常问题”。由名称就可想到，第一次发现时，“反常”给人们带来了多大的震惊。最早研究量子场的物理学家想当然地以为，经典理论里的所有对称在量子力学的延伸里也应当保持，这一延伸是一个同时包括了虚粒子作用的更为全面的理论。但事情却并不总能如此。1969年，史蒂文·阿德勒（Steven Adler）、约翰·贝尔（John Bell）和罗曼·贾基夫（Roman Jackiw）证明：

即便经典理论保持对称，包含虚粒子的量子力学过程有时也会打破对称，这种对称的打破被称作反常，包含了反常的理论就是“反常的”。

反常对力的理论极为重要。在第9章里我们看到，一个成功的力的理论要求存在内部对称，这必须是精确对称，不然就没有办法消除规范玻色子的多余极化，力的理论也就没有意义了。因此，与力相关的对称必须没有反常存在——即所有对称破缺的效应总和为零。

这对力的任何量子理论都是一个强大的约束。我们现在知道，对于标准模型里夸克和轻子的存在，这也是一个最有说服力的解释。单个的虚夸克和轻子会导致打破标准模型对称的反常量子贡献，但夸克和轻子的量子贡献的总和为零。正是这一神奇的抵消作用，使标准模型能连贯一致。如果想要标准模型里的力有意义，轻子和夸克都是必不可少的。

反常对弦理论是一个潜在的问题，毕竟它也包含了力。1983年，理论学家路易斯·阿尔瓦雷兹-高梅（Luis Alvarez-Gaume）和爱德华·威滕（Edward Witten）证明：这种反常不仅仅出现在量子场论中，也出现在弦理论中。这一发现似乎将终结弦理论的历史，使其成为虽然有趣却遥不可及的观点。弦理论不像是能够保持必要对称的理论，它有潜在的反常可能，在一片怀疑的氛围中，格林和施瓦茨甩出一记清脆的响鞭，证明弦理论可以满足避开反常所需要的限制条件。他们计算了对所有可能反常的量子贡献，证明对特定的力，反常奇迹般地相加为零。

格林和施瓦茨的结果令人这么吃惊，其中一个原因是，弦理论允许存在许多很复杂的量子力学过程，而每一个看起来都有可能产生对称破缺反常。但格林和施瓦茨证明，对这些十维超弦理论中的所有可能，对称破缺反常的量子力学贡献的总和为零。这就意味着，弦理论计算所要求的许多抵消实际真的发生了，而且，这些抵消发生在十维里，我们已经知道这一维数对超弦理论是非常特殊的。这一发现太

神奇了，许多物理学家确定这种协同绝非巧合。反常消除是支持十维超弦理论的强有力的论据。

再者，格林和施瓦茨完成他们的研究恰逢时机。多年来，物理学家一直在徒劳地寻求一个能推广标准模型以兼容超对称和引力的理论，他们已经在考虑寻找新的东西了。他们不会忽略格林和施瓦茨关于超对称理论的发现，它有可能重现标准模型的所有粒子和力。即使弦理论的冗余结构令人厌烦，但超弦却在其他可能更为经济的理论失败的地方获得了成功。

另外两个重要的进展，使得弦理论很快被纳入了真正的物理学当中。其中一个是由普林斯顿团队，包括戴维·格罗斯、杰夫·哈维（Jeff Harvey）、埃米尔·马丁尼克（Emil Martinec）和瑞安·罗姆（Ryan Rohm）等在1985年提出的理论，他们称之为“杂化弦理论”（heterotic string）。“杂化”一词来源于植物学，意为“异配优势”（hybrid vigor），指的是那些杂交品种比其祖先具有更优越的属性。在弦理论里，振动模式可以顺时针或逆时针地沿着弦运动。使用“杂化”一词是因为，向左运动的波与向右运动的波被区别对待，因此，这一理论比其他已知的弦理论形式包括了更为有趣的力。

杂化弦理论进一步证实了：格林和施瓦茨发现的不存在反常且在十维可以接受的理论是很特别的。

他们发现了几组作用力，其中包括在弦理论里已证明可能的所有力；还有一组力，以前从未有人（在理论上）证明它也是弦理论的组成部分，这组力正是格林和施瓦茨证实的没有反常的新型的力。有了杂化弦理论，另外的这一组力——涵盖标准模型里的力——被证明不仅在弦理论里是可能的，而且可以实现。杂化弦是联系弦理论与标准模型的一个真正突破。

还有另一成果最终巩固了弦理论的显要地位，这一发现解释了额外维度对弦理论的必要性。它很好地显示了超弦理论是内在一致的，并体现了标准模型里的力。但如果你被困在错误的空间维数上，这就不那么有趣了。超弦假设了十维，我们周围的世界看起来只有四维（包括一维时间），那么，其他多余的六维就需要解释。

物理学家现在认为答案可能就是维度的卷曲——即第2章里讲的卷曲的、无法被观察到的极小维度。但是，起先额外维度的这种卷曲似乎并非处理弦理论里额外维度的恰当方式，问题在于：有了卷曲维度的理论不能重现第7章讨论的弱力的重要（而惊人的）特征：弱力以不同方式对待左旋和右旋粒子。这不仅仅是一个技术细节，标准模型的整个结构都要依赖于这样的事实：左旋粒子是唯一经受弱力的粒子，否则标准模型的几乎所有预言都将失效。

尽管十维弦理论能将左旋和右旋粒子区别对待，但一旦那6个额外的维度卷曲起来，就不对了。结果是，四维有效理论总是包含恰好配对的左旋和右旋粒子，作用于左旋费米子的所有力也会作用于右旋粒子，反之亦然。如果弦理论不能找到摆脱困境的好办法，那么它就只能被排除。

1985年，菲利普·坎德拉斯（Philip Candelas）、加里·霍洛维茨（Gary Horowitz）、安迪·斯特罗明格（Andy Strominger）和爱德华·威滕发现了一种更为微妙、复杂的额外维度卷曲方式，这一卷曲的维度被称作卡拉比-丘流形（Calabi-Yau manifolds）。其中细节非常复杂，但基本来说，卡拉比-丘流形留下了一个四维理论，能够区分左右，且能重建标准模型的粒子和力，包括宇称不守恒的弱作用力，而且，把额外维度卷曲成卡拉比-丘流形还保持了超对称^[5]。有了这一突破，超弦理论就开始施展才能了。

在许多大学物理系里，超弦理论超越了粒子物理学，超弦革命更像是一场政变。因为超弦理论纳入了量子引力并包括了所有的已知粒子和力，许多物理学家甚至把它当成万物基础的终极理论。事实上，20世纪80年代，弦理论被冠以“终极理论”（TOE）或“万能理论”的头衔。弦理论甚至比大统一理论的抱负还要远大：在比大统一能量还要高的能量上，物理学家希望以弦理论将所有的力（包括引力）统一起来。即便没有任何观察结果支持弦理论，许多物理学家仍然相信，弦理论将兼容量子力学和引力的潜在可能就足以支持它的卓越。

挫折：旧制度的忍耐

如果弦理论正确，世界最终由基本的振动弦构成，那么粒子物理学所有理论都要被抛弃了？答案是一声响亮的“不”。弦理论的目标是在小于普朗克长度的尺度协调量子力学和引力，我们相信那是新理论接管的尺度。因此，在传统的弦理论（不同于额外维度模型的变体）里，弦的大小大约应是普朗克长度。这就告诉我们，在传统的弦理论里，粒子物理学和弦理论的差异只能出现在微小的普朗克长度，或等价地说，在超高的普朗克能量上，引力应是极其强大的。长度这么小而能量这么高，弦不可能在实验能达到的能量上排除粒子描述。

在低于普朗克能标的能量上，粒子物理学实际上就已足够了。如果弦如此微小，无法探测其长度，那么弦也就等同于一个粒子，实验看不出其差别。对我们来说，弦的一维长度就如前面讨论的卷曲额外维度一样是不可见的，除非我们有仪器探知 10^{-33} 厘米的尺度，这样的弦小得根本看不到。

可以理解，弦理论和粒子物理学在可达到的能量上应该是一样的。不确定性原理告诉我们，研究小距离的唯一方法是使用高动量

粒子，它们的能量非常高。因此，如果没有足够的能量，就无法辨别其是细长的弦还是点状的粒子。

从原则上讲，通过寻找弦理论预言的许多新粒子——对应于弦的多种可能振动而形成的粒子，我们能够找到支持弦理论的证据。这一策略的问题在于，由弦产生的大多数粒子极重，其质量达到了普朗克标度 10^{19} GeV。相比实验已探测到的粒子，这一质量是极为庞大的，而实验测得的粒子最重也就是大约200 GeV。

由弦的振动产生的额外粒子之所以这么重，是因为弦的张力非常大。张力即弦对抗拉伸的能力，它决定了弦是否易于振动并产生重粒子。普朗克级能量决定了弦的张力，这个张力是弦理论为重现正确的引力子相互作用强度（由此也就重现了引力本身）所要求的。弦的张力越大，产生振动所需要的能量就越大（这就好比一根较紧的琴弦比一根较松的琴弦更加难以弹拨或更换），这一高能量也等于是由弦产生的额外粒子的大质量。这些普朗克级质量的粒子实在太重，现在（或者，最有可能是未来）运行的粒子实验还不能生成它们。

因此，即便弦理论正确，我们也不大可能找到它预言的许多额外的重粒子。当今实验的能量要低16个数量级，因为额外的粒子这么重，从实验中找到弦存在的证据的希望非常渺茫。也许我后面讨论的额外维度模型就是一个例外。

但在大多数的弦理论图景里，因为弦的长度这么小，而张力又这么大，所以即便弦的描述正确，在加速器能达到的能量上，我们还是不能找到支持弦理论的证据。只注重对实验结果作出预言的粒子物理学家，可以安全地使用传统的四维量子场论，而忽略弦理论，仍能得到正确的结果。只要你所关注的尺度大于 10^{-33} 厘米（或者，能量低于 10^{19} GeV），我们以前探讨过的关于粒子物理学的低能量结果都不会改变。既然质子的大小是 10^{-13} 厘米，而现今加速器能达到的最大能量

大约是1 000 GeV，那么我们完全可以放心大胆地认为粒子理论的预言就已经足够了。

即使这样，专注于研究低能现象的粒子物理学家也还是有足够的理由关注弦理论。弦理论引入了许多新的数学和物理观点，那是没人会从其他方面想到的，比如膜理论和其他额外维度的观念。即使在四维里，弦理论也开辟了一条道路，让我们对超对称、量子场论以及量子场论模型可能包括的力都产生了更为深入的了解。当然，如果弦理论确实能够给出一个十分一致的关于引力的量子力学描述，这项成就将令人肃然起敬。即使对那些只关注实验可测现象的人来说，弦理论的这些益处也使它非常值得探索。尽管我们难以探测到弦（甚至是不可能的），但由弦引发的理论观点也许关系着我们的世界。

很快我们将看到，它们会是些什么。

尾声：挑战与机会并存

1984年，正值“超弦革命”的高峰期，当时我正在哈佛大学读研究生。我很快发现，初出茅庐的物理学家面临着两大选择：要么跟随威滕和格罗斯从事弦理论，他们那时都在普林斯顿大学；要么在乔治和格拉肖所在的学院做粒子物理学家，与实验结果保持密切联系，不过他们都在哈佛大学。对相同问题感兴趣的物理学家如此格格不入，真让人难以置信，但两个阵营对怎样进行研究的确有很大的分歧。

哈佛大学的焦点仍旧是粒子物理学研究，那里的许多物理学家十分排斥弦理论。粒子物理学和宇宙学的许多问题仍旧未找到答案，在陷入弦理论可能的数学雷区之前，我们为什么不先解决这些问题呢？物理学家能接受把物理学延伸到不可探测的领域吗？有这么多杰出的物理学家，还有许多振奋人心的观点，都告诉我该怎样使用更为传统的方法来超越粒子物理学的标准模型，我没有理由跳转阵营。

但是，在别的地方，物理学家却深信有关超弦理论的所有问题很快都将被解决，而未来（以及当今）物理学界必然是弦理论的天下。超弦理论初登舞台，许多人认为只要投入足够的人力和时间，弦理论最终会得出与已知物理学一致的结果。在1985年有关杂化弦理论的论文里，格罗斯和他的同事写道：“虽然还有许多工作要做，但要从杂化弦理论得出所有已知物理的结果似乎并无不可逾越的障碍。”弦理论志在成为终极理论，普林斯顿大学冲在最前沿。那里的物理学家深信弦理论就是通往未来的道路，整个系的粒子物理学家无一不投入到了弦理论研究中——如今这仍是普林斯顿大学有待改正的一个错误。

而今，我们不好说弦理论面临的障碍是否是“不可逾越的”，但它肯定是极富挑战的。许多主要的未解问题仍未找到答案，物理学家和数学家迄今所创立的方法还不足以应对弦理论的这些未解问题，我们似乎需要一个新的数学机制或一个新的基本途径来解决它们。

乔·波尔钦斯基在他被广泛使用的弦理论教科书里写道：“就其宏大结构来讲，弦理论就如真实世界一样。”在某些方面它确实如此：弦理论包括了标准模型的粒子和力，当其他维度卷曲时，还可以缩减至四维。但是，尽管有诱人的证据表明弦理论能兼容标准模型，但寻找标准模型的理想候选者的计划在历经20年的探索之后仍然遥遥无期。

物理学家原本希望弦理论能对我们世界的本质作出一个独特的预言——一个为我们的可见世界量身打造的预言，但现在弦理论可以产生许多可能的模型，每一模型都含有不同的作用力、不同的维度及不同的粒子组合，我们想找到与可见世界恰好相符的那一个以及为什么是那一个，而现在没人知道该选择哪种。总之，没有一个看上去完全正确。

例如，卡拉比-丘流形能够决定基本粒子有几代，其中一种可能就是标准模型的三代，但没有一个唯一的、无可争议的卡拉比-丘候选者。尽管弦理论学者最初希望卡拉比-丘能够选出一个共同认可的形状和一个唯一的物理定律，但他们很快就失望了。斯特罗明格对我讲过他如何在一个星期内找到一种卡拉比-丘流形，并认为它是唯一的。但他的同事加里·霍洛维茨又发现了其他几个候选者，后来斯特罗明格从丘成桐处得知，卡拉比-丘流形有成千上万个类型。现在我们知道建立在卡拉比-丘空间基础之上的弦理论可以含有几百代。那么，究竟哪一个卡拉比-丘流形才是正确的呢？如果正确，又为什么正确？虽然我们不知道弦理论的一些维度一定是卷曲的或看不见的，但弦理论学家还是有必要确定一个原则，告诉我们卷曲维度的大小和形状。

弦理论不仅仅包含波沿着弦的多次振动而产生的新重弦粒子，还包含一些低质量的粒子。如果它们存在且正如弦理论预言的一样那么轻，那么我们应该可以认为实验是能够探测到它们的。许多以弦理论为基础建立起来的模型都包含了比我们在低能量实验观察到的更多的轻粒子和作用力，但我们尚不明瞭哪些才是正确的。

要使弦理论与我们的现实世界对应起来，的确是一个非常复杂的问题，我们必须弄清楚：由弦理论得出的引力、粒子和作用力为什么应该与我们世界里已知正确的东西相一致？但这些关于粒子、作用力和维度的问题，与对宇宙能量密度的过高估计问题相比，就是小巫见大巫了。

宇宙即使没有粒子也可以拥有能量，这种能量叫作真空能量。根据广义相对论，这种能量有一个物理结果：它会使空间产生伸缩。正的真空能量会加速宇宙的膨胀，而负能量则使它坍缩。爱因斯坦在1917年首先提出这一能量，目的是给广义相对论方程找到一个恰当的解。在这一方程里，真空能量的引力作用会与物质的引力作用相抵消。尽管由于多种原因，其中包括爱德温·哈勃在1929年观测到的宇宙膨胀，爱因斯坦不得不放弃了这一观点，但却无法从理论上证明这种真空能量在我们宇宙中不该存在。

事实上，天文学家最近在我们的宇宙中测量了真空能量，并发现了一个很小的正值（它也称作暗能量或宇宙学常数）。他们发现，遥远的超新星比我们预期的更为黯淡，除非它们在加速离开。超新星测量以及对大爆炸期间产生的遗留光子的深入观察告诉我们：宇宙在加速膨胀，这是真空能量有一个很小正值的证据。

这一测量是令人振奋的，但它也引出了一个更为重要的谜题：加速非常缓慢，这告诉我们虽然真空能量的值不等于零，却极其微小。测得真空能量的理论问题是，它远远小于人们的估计。根据弦理论估测，能量应该要大得多，如果真是那样，这一能量导致的就不仅仅是难以估量的超新星加速。如果真空能量很大，宇宙可能早就坍塌了（若是负值），或者会很快膨胀至虚无（倘若是正值）。

宇宙的真空能量为什么会这么小？而我们知道它必须这么小。弦理论必须给出解释，而粒子物理学也没有答案。但是，与弦理论不同的是，粒子物理学没有意图要成为一个量子引力理论——它远没有那么雄心勃勃。不能解释能量的粒子物理学模型最多是不尽如人意，而弦理论如果将能量搞错却会被排除。

能量密度何以如此微小，这是一个全然无解的问题。有的物理学家认为，对此不会有一个真正的解释。尽管弦理论是一个独立的理论，只有一个参量——振动弦的张力，但弦理论学家还是不能用它预言宇宙的大多数特征。大多数物理理论会包含一些原理，它们允许你确定在众多可能的物理结构中，这一理论会实际预测哪一种。例如，大多数系统最终都安顿下来，归于有着最低能量的结构。但这一标准似乎并不适用于弦理论，它看上去会产生无限多个具有不同真空能量的不同构形——而我们不知道自然偏爱哪一个（如果有的话）。

人择原理

那些物理学家认为，我们生存的这个世界其真空能量值令人这么不可思议，是因为任何更大的值都会妨碍星系和宇宙结构的形成——也就不可能有我们。人择原理大大地偏离了最初弦理论要预言宇宙所有特征的目标。

有些弦理论学家不再试图找到一个唯一理论。他们考察了卷曲维度的可能大小和形状以及宇宙可能含有的不同能量值，然后得出了这样一个结论：弦理论只能够画出一个大致景象，它描述了我们生活的众多可能的宇宙。这些弦理论学家认为弦不会预言一个唯一的真空能量，他们认为宇宙里有着许多互不相联的区域，有着不同的真空能量值，而我们就生活在含有恰当值的这一部分。在众多可能的宇宙里，只有能产生结构的那个才能够（且实际）容纳我们。那些物理学家认为，我们生存的这个世界其真空能量值令人这么不可思议，是因为任何更大的值都会妨碍星系和宇宙结构的形成——也就不可能有我们。这一推理有一名称：人择原理（anthropic principle）。

人择原理大大地偏离了最初弦理论要预言宇宙所有特征的目标，它告诉我们不必对小能量值作出解释——宇宙里存在着许多互不相联的区域，有着多种真空能量值，但我们只生存在那为数不多的能形成结构的其中一个里。这一宇宙的能量值异乎寻常地小，只有极个别形式的弦理论才能预言这么小的值，但我们只能存在于这样有着极小能量的宇宙中。这一原理很可能被未来的成果推翻，或可能被更深入的研究所证实，但不幸的是，它很难（甚至不可能）被验证。如果人择原理是一个世界的答案，这一图景显然是令人失望的，我们不会满足于此。

无论是何种情形，弦理论就其现在的发展状态来说，肯定不能预言我们世界的特征，即便从其根本形式来讲这是一个独立的理论。我们需再次面对这一问题：如何把一个完美对称的理论与我们世界的客观现实联系起来？最简单的理论阐述太过对称：许多维度、许多粒子及许多作用力，我们明知它们一定是不同的，但看起来却有着同等的

地位。为了与标准模型、与我们的可见世界相联系，这一秩序必须被打乱。对称破缺之后，根据哪一对称产生了破缺、哪些粒子变重、哪些维度使它们彼此区分，这一单独的理论又呈现出多种不同的形式。

弦理论似乎是一套设计精良却并不合身的衣服，就其目前的状态而言，你只能把它挂在衣架上，赞叹它精致的做工和细致的纹理——它实在是太美了，但是却无法穿上它，除非做必要的剪裁。我们希望弦理论能够囊括我们对世界的所有了解，但是“老少皆宜，胖瘦兼顾”的尺寸往往对谁都不适合。现在，我们甚至不知道是否有恰当的工具来对弦理论进行适当的“剪裁”。

因为我们并不真正知道弦理论的含义，而将来能否知道的问题我们也没把握。一些物理学家干脆把弦理论定义为能在小距离解决量子力学和广义相对论矛盾的任何理论。当然，大多数弦理论家相信弦理论和正确的理论就是一回事，或者至少是密切相关的。但显然还有很多东西需要了解。现在就断言弦理论是描述这一世界的终极理论未免为时过早，或许更为精密的数学工具能使科学家真正理解弦理论，也或许将弦理论的思想用于周围世界积累的物理认识会提供重要的线索。物理学家和数学家迄今所创立的方法还不足以应对弦理论的未解之谜，我们似乎需要一个新的、更为根本的工具。

不管怎么说，弦理论是一个引人瞩目的理论。关于引力、维度和量子场论，它引出了许多重要的见解，是我们已知的最有希望成为一个自洽的量子引力理论的候选者。而且，弦理论还带来了许多令人难以置信的数学成果，但是20世纪80年代提出的要将其与世界联系起来的想法还有待实现。我们并不了解弦理论蕴含的大部分深意。

公平地说，粒子物理学的问题也并未得到解答。许多在20世纪80年代的未解问题至今仍没有答案，这些问题包括：对基本粒子质量巨大差异的根源作出解释；找到解决等级问题的正确方法。而且，模型

构建者仍在等待实验线索来告诉我们在超越标准模型的众多可能中，哪一个才正确地描述了真实世界，只有等我们探索了TeV以上的能量，才可能确定地回答我们最关心的那些问题。

如今，相比20世纪80年代，弦理论学家和粒子物理学家对于自己的理解水平都有了更为清醒的认识。我们试图解决的是一些困难的问题，它们必然需要耗费时日，而这是些充满了兴奋的时日，尽管有许多未解问题（或许正是由于这许多未解问题），但我们有足够的理由保持乐观。现在物理学家对粒子物理学和弦理论的结果有了更深的领悟，而那些思想开明的物理学家会从两派所取得的成就中汲取对自己有益的东西，这正是我和我的一些同事所选择的中间立场——而且，它已带来了许多振奋人心的成果，我们很快就将看到。

-
- 就如光子传递电磁力一样，引力子是传递引力作用的粒子。
 - 根据弦理论，世界的根本物质是弦，而不是点状的粒子。
 - 后面的额外维度模型不会明确地用到弦理论；在超过极小普朗克长度的距离（ 10^{-33} 厘米），粒子物理学就已足够了。
 - 然而，由于弦理论引进了许多新的观念和分析工具，因此，即使是在低能量上，弦理论对粒子物理学也是非常重要的。



-imension
探索大揭秘



辅助通道：膜的狂飙突进

WARPED PASSAGES

薄膜里的狂乱，大脑里的狂乱。

柏树山乐队 (Cypress Hill)

十一维宇宙

未来艾克决定再次进入微小的普朗克领域。幸运的是，经他改良的Alicxur运转正常，他顺利到达了布满了弦的十维宇宙。因为急于探索新环境，艾克抬动曲轴，启动了他刚从Gbay买来的超空间驱动器。那些弦着了魔似地纠缠、碰撞，艾克看得入了迷。

尽管艾克担心自己的Alicxur会出故障，但由于对这个新奇的世界仍充满好奇，因此，他加大了手上的力量，继续拉动驱动杆。弦碰撞得更为频繁、更为剧烈了，但随着他手上力量的不断加大，驱动轴不断抬高，他进入了一个从未见过的全新的世界。艾克甚至不能判断时空是否完好无损，但他继续拉动曲轴。可奇怪的是，他竟安然无恙。

可是，现在他周围的环境完全变了。艾克所待的地方不再是他开始进入时的十维宇宙，而是一个充满了粒子和膜的十一维宇宙。而且，虽然这听起来很奇怪，但在这个新的宇宙里，所有东西彼此间作用并不那么强烈。艾克回过头来看看他的控制盘，发现超空间驱动杆被神秘地重置，又回到了低处。艾克感到迷惑不解，而且

有点儿绝望，他再次抬起操作杆，竟发现自己又回到了开始的十维宇宙。艾克检查他的控制系统，发现驱动杆再次回到了低处。

艾克还以为他的Alicxur失灵了，所以急忙查看最新版的说明书，才知道他的设备运行良好——十维弦理论里的高位超空间驱动与十一维宇宙里的低位超空间驱动是同样的。

说明书上没有讲明如果超空间驱动既不高也不低时会发生什么，艾克只好登录空间站，把自己加进等待名单，期望得到一个改进版来解决这一问题。但是，Alicxur设计者只给出了这样的许诺：发行日期将在1 000年内的某个时候。

在当今物理界，你可以说“弦理论”的名字是个错误。事实上，理论学家迈克尔·达夫（Michael Duff）诙谐地说，“弦理论”指的是“以前曾被当成弦的理论”。如今，弦理论已不仅仅是关于只在一个空间方向上延伸的弦的理论，还包含了在二维、三维或更多维度上延伸的膜。现在我们知道，膜可以延伸至超弦理论所包含的任何数量的维度。它与弦一样，也是超弦理论的一个重要组成部分。理论家早先忽略它是因为，他们研究弦时，弦的相互作用强度“操作杆”还很低，而且膜的相互作用也不那么重要，最终却发现膜竟是失踪的几片拼图，找到它们才能神奇地完成整个拼图。

在本章里，我将讲述膜是如何从一件只是好玩、不被人看重的稀奇事，最终发展成弦理论故事的主角的。我们会看到，自20世纪90年代中期以来，膜以多种方法帮助解决了弦理论里一些令人迷惑的问题，膜帮助物理学家理解了弦理论里不可能起源于弦的神秘粒子的来源。当物理学家将膜考虑在内时，他们发现了“对偶理论”（dual theories）——一些看似非常不同，实际却产生了同样物理结果的理论。篇首故事里讲到的就是本章将探索的对偶性的一个值得关注的例子：十维超弦理论与十一维超引力理论之间的等效性。十一维超引力理论是一个只包含膜却没有弦的理论。

本章还将介绍M理论，一个既包括超弦理论又包括超引力论的十一维理论，其存在是通过对膜的认识推测出来的。没人真正知

道“M”代表什么——这一术语的创建者，爱德华·威滕，有意不对它作出解释，但有人说它是“膜”，有人说是“魔幻的”，也有人说是“神秘的”，而此时我要说，M理论还是一个“迷失的理论”，它还只是推测，还未被充分理解。然而，即使M理论仍留有许多未解问题，由膜所取得的进展仍解释了一些理论联系，它们要求M理论具有更为复杂、涵盖更广、更加神奇的结构。这就是当今弦理论学家都在研究它的原因。

本章更新了始于20世纪80年代的弦理论图像，呈现了20世纪90年代以来物理学家提出的一些更为先进的观点。这些资料大多并不是膜在粒子物理学应用中的核心，而以后的膜宇宙猜想也不会明确地依赖于下面我将描述到的现象，因此你尽可以跳过去。但是，如果你喜欢，可以借此机会了解弦理论的某些令人瞩目的进展，它们在很大程度上将膜恰到好处地铺展在了弦理论的版图上。

膜的兴起

在第3章里，我们看到膜会在某些空间维度里延伸，可未必是所有的维度。例如，即便体空间会包含多个维度，但一个膜可能只会在三个空间维度里延伸。额外维度到膜这里可能就会终止，换句话说，膜是额外维度空间的边界。我们还知道膜可以“收容”只沿着它的维度运动的粒子。即使存在许多额外维度，局限在膜上的粒子也只能在那个膜所占据的有限区域内运动，而不能自由地探索整个额外维度空间。

现在我们看到，膜不仅是一个位置，它本身就是一个物体。膜，就像薄膜一样，是真实的东西。

膜可以是松弛的，这时它们会弯曲和运动；膜也可以是紧绷的，这时它们可能处于静止状态；膜可以携带电荷并通过力相互作用；而

且，膜也会影响到弦和其他物体的行为。所有这些属性告诉我们，膜在弦理论里是必不可少的，弦理论的阐述要前后一致就必须包含膜。

1989年，当时在得克萨斯大学的吉恩·戴（Jin Dai）、罗布·利（Rob Leigh）和乔·波尔钦斯基，以及捷克物理学家彼特·霍拉瓦（Petr Horava），分别独立地从数学上在弦理论的方程里发现了一种特殊类型的膜，叫D-膜。闭弦会绕成一个圈，而开弦则有两个自由的端点，这些端点必定有终止的位置。在弦理论里，开弦端点所能允许的位置就是D-膜（“D”指的是彼得·狄利克雷 [Peter Dirichlet]，19世纪德国数学家）。体可以包含不止一个膜，罗布·利和霍拉瓦发现所有的开弦都必须在膜上终止，弦理论则会告诉我们这些膜应拥有什么样的维度和属性。

有些膜会向3个维度延伸，但其他膜会伸向4个或5个甚至更多的维度，事实上，弦理论包含的膜可以延伸的维度直至9个。弦理论的惯例是使用它们的空间维度而非时空维度的数量来标识膜。例如，一个3-膜是在三维空间（四维时空）里延伸的膜。当我们要看膜在可见世界里的结果时，3-膜是非常重要的。但是，对于本章讨论的膜的应用，其他维数的膜同样有着非常重要的作用。

弦理论里产生了不同类型的膜，它们的区别不仅仅表现在它们延伸的维数，还表现在它们的电荷、形状以及一个重要的特点：张力（我们将很快讨论到）。我们不知道这些膜是否存在于真实世界，但我们知道弦理论所说的这些膜确实是可能的。

刚发现它们的时候，膜只不过是一种让人感觉新奇的东西，那时没有人会考虑相互作用或运动的膜。如果弦只是像弦理论学家最初猜想的那样微弱地相互作用，那么D-膜将是绷紧的，处于静止状态，对弦的运动或相互作用没有任何影响。

而如果膜对体里的弦不作回应，那么它们的存在则纯属多余。它们将只是一个位置，相对于弦的运动和相互作用，它们不过就像长城于你的日常生活，丝毫不会有任何影响。而且，物理学家不想在一个实际的弦理论里考虑膜，因为膜违背了他们的直觉经验：所有维度天生都是一样的。而膜使得某些维度异于其他——那些沿着膜延伸的维度与远离膜延伸的维度是不同的，而已知的物理定律将所有方向都同等对待。弦理论为什么会有不同呢？

我们还期望在空间任一点的物理定律与其他所有点都是同样的，而膜也不遵循这种对称。尽管沿着某些维度膜会无限延伸，但在其他方向上，它们又会静止于一个固定的位置，因此它们不会向整个空间延伸。在那些膜的位置固定的方向上，离膜3厘米与离膜将近1米或800多米就出现了不同：想象一张浸满了香水的膜，离它是远是近，你立刻就能辨别出来。

出于这些原因，弦理论学家最初忽略了膜，但在膜被发现5年之后，它们在理论界的地位开始迅速提升。1995年，乔·波尔钦斯基不可逆转地改变了弦理论的进程，他证明膜是一种动态的物体，是构成弦理论所必不可少的，且很可能在其最终构建中发挥着关键作用。波尔钦斯基解释了在超弦理论里存在着哪些类型的D-膜，并证明这些膜会携带电荷，因此能相互作用。



弦理论里的膜有一定的张力。膜的张力就像鼓面：在被敲击之后，它会弹回其本来静止的位置。如果膜的张力为零，那么膜便不会有任何抵抗，轻轻地碰触就会让它产生强大的作用；而相反，如果膜的张力无限大，那么它就是一个静止而非动态的物体，敲击不会产

生任何作用。因为张力是有限的，就如其他所有带荷物体一样，膜便可以产生起伏和运动，并对力作出反应。

膜的一定张力和非零荷说明它们不仅仅是一些位置，还是物体：带荷意味着它们会相互作用；而张力说明它们会运动。就像蹦床一样，当其表面受压和产生反弹时，会与周围环境发生相互作用，膜也会运动和相互作用。例如，无论是蹦床还是膜都会产生弯曲，都会影响它们的环境：蹦床的影响是通过推动人和空气；而膜的影响则是通过推动带荷物体和引力场。

如果膜在宇宙里存在，那么它们对时空对称的破坏便无异于太阳和地球对空间对称的破坏。太阳和地球占据着特定的位置，当相对于太阳和地球进行测量时，三维空间的所有位置并非都是一样的。但是，即便宇宙状态并不对称，物理定律却仍保持了三维空间的时空对称。就这一方面来看，膜引起的扰乱并不比太阳和地球更严重。如在空间中有着固定位置的其他所有物体一样，膜会打破时空的某些对称。

简单一想就会发现这并不是坏事，毕竟，如果弦理论正确地描述了自然，那么并非所有维度都是天生一样的。我们熟悉的三个空间维度看上去一样，但额外维度必然不同，如果它们相同，那也就不是“额外”的了。从物理世界的角度出发，时空对称的打破能够帮助解释为什么额外维度是不同的：膜可能正确区分了我们体验和了解的三个空间维度与弦理论的额外维度。

在后面的章节里，我将探讨有三个空间维度的膜，并描述它们对现实世界一些潜在的激进意义。但本章的后半部分，我要集中讲述为什么膜对弦理论这么重要——事实上，它们激发了1995年的“第二次超弦革命”。下节讲述了膜为什么在过去10年里一直处于弦理论的前沿，以及为什么我们认为它仍将继续处于理论前沿。

收紧的膜与失踪的粒子

正当乔·波尔钦斯基致力于D-膜的研究时，他在圣塔芭芭拉的同事安迪·斯特罗明格也在同时思索着p-膜——爱因斯坦方程的一个神奇的解是：它们在某些空间方向无限延伸，而在另外的维度又表现得像黑洞，困住所有靠近的物体。而D-膜则是开弦终止的曲面。

斯特罗明格告诉我，每天午餐时间他和乔·波尔钦斯基都会讨论各自研究的进展：斯特罗明格会谈论p-膜，而乔·波尔钦斯基则谈论D-膜。虽然他们两人都在研究膜，但像其他所有物理学家一样，他们最初以为这两种膜是两种不同的东西。最终，乔·波尔钦斯基意识到，原来它们是同一回事。

斯特罗明格的工作显示了他正在研究的p-膜对于弦理论是非常重要的，因为在某些时空几何里，它产生了新粒子。弦理论有一个令人瞩目却有悖于直觉的假设：粒子是由弦的振动模式产生的。即便这一假设正确，弦的振动模式却未必能够解释所有的粒子。斯特罗明格证明，可能会有其他一些粒子并不依赖于弦。

膜有不同的形状、形式和大小。尽管我们只集中注意到膜是弦终止的地方，但膜本身也是独立的物体，能够与它们周遭的环境互相作用。斯特罗明格研究的p-膜包裹着一个极微小的卷曲空间区域，他发现这些紧紧包裹的膜可以表现得像粒子一样。这样的膜可以比作一个系紧的套索：当你把一个绳圈套在柱子或牛角上拉紧时，绳索就会变

紧。膜也一样，它可以包裹一个卷曲的空间区域，如果那一空间区域很小，那么包裹它的膜也会很小。

这些微小的膜与我们熟悉的现实物体一样，也是有质量的，而且质量会随着体积的增大而增大：东西越多越重（如铅管、灰尘或樱桃一样），越少越轻。一个卷曲的膜包裹着一个极微小的空间区域，那它也会非常轻。斯特罗明格的计算显示，在极端情况下，当膜微小到你可以想象的极限时，这一微小的膜看上去就如一个新的无质量粒子。斯特罗明格的结果非常重要，因为它表明即便是弦理论最基本的假设——所有东西都是由弦产生的，也并不总是成立。在粒子谱中，膜同样作出了贡献。

而乔·波尔钦斯基在1995年令人瞩目的发现是，这些由微小的p-膜产生的新粒子，用D-膜也可以解释。事实上，在他确立D-膜重要意义的论文里，乔·波尔钦斯基证明D-膜和p-膜实际是同一回事。在弦理论与广义相对论作出同样预言的能量上，D-膜融入了p-膜。尽管乔·波尔钦斯基和斯特罗明格起初并未意识到，但最终发现他们实际在研究同一件东西。这一结果意味着D-膜的重要性将不再有疑问：它们与先前的p-膜同样重要，而那些p-膜对弦理论的粒子谱是必不可少的。而且，有一种绝佳的方式来理解p-膜和D-膜为什么是等效的，其基础是一个微妙且重要的观念——对偶性。

十维超弦理论与十一维超引力论的对偶性

在过去的10年里，对偶性是粒子物理学和弦理论里一个最为激动人心的概念。在量子场论和弦理论所取得的最新成果里，它都发挥了重要作用。而且，正如我们很快将看到的，它对膜理论也有着特别重要的含义。

当两个理论是有着不同描述的另一理论时，就是对偶的。1992年印度物理学家阿修克·森（Ashoke Sen）首先认识到了弦理论里的对偶性。在研究中，他得出：如果一个理论的粒子和弦被交换，理论仍保持不变。他的依据是物理学家克劳斯·曼通宁（Claus Montonen）和戴维·奥利弗（David Olive）在1977年提出的对偶性观点。20世纪90年代，当时在罗格斯大学的以色列物理学家内森·塞伯格（Nati Seiberg）也证实，在两种不同的超对称场理论之间，虽然有着看似大不相同的作用力，却存在着显著的对偶性。

为了理解对偶性的重要意义，我们最好先了解一下弦理论学家通常是怎样计算的。

弦理论的预言要依赖于弦的张力，但它们也会依赖于一个被称作弦耦合的数值，它决定了弦相互作用的强度：它们或者轻轻接触，对应于弱耦合；或者密切联系，对应于强耦合。如果我们知道了弦耦合的值，我们就可以只为这个特定的值研究弦理论，但是，到目前我们还不知道弦耦合的值，就只能寄希望于当我们对弦相互作用强度作出预言时能够理解这一理论，然后就能发现哪一个才是有效的。

问题在于，自弦理论出现伊始，强耦合的理论似乎就是不可捉摸的。20世纪80年代，人们只了解了关于弦的微弱相互作用的弦理论（我们用“微弱”来描述弦相互作用的强度，但不要被这一词语误导——它与弱力无关）。弦的相互作用强烈时，所有计算都将变得异常困难。一个系得很松的结总会比一个系紧的结更容易解开，同样地，只有弱相互作用的理论要比有着强相互作用的理论更容易驾驭。当弦之间的作用非常强烈时，它们将成为一堆乱麻，让人很难理出头绪。物理学家尝试了各种各样巧妙的方法来计算强相互作用的弦，但始终未找到能应用于现实世界的非常有效的方法。

事实上，不仅仅是弦理论，所有的物理理论都是在相互作用较弱时更容易理解。这是因为，如果对于一个可解理论——通常没有相互作用来说微小的扰动或改变，那么你就可以使用叫作微扰理论（perturbation theory）的技法去解释。

微扰理论允许你从没有相互作用的理论开始，逐步改进计算，从而回答在微弱相互作用理论下的问题。微扰理论是一个系统的程序，它告诉你怎样一步一步地细化计算，直至达到所需的精确水准（或者直到你感到疲惫）。

在一个难解的理论中，使用微扰理论来逼近一个量，就好比调色的过程：假设你渴望得到一种难以形容的、蓝色中略带绿意的颜色——类似于地中海在其最美时的色彩，那么你可能就会以蓝色开始，然后渐渐地、一点一点地加入少量的绿色，偶尔再稍稍调进一点蓝色，直至最终（几乎）达到你想要的颜色。这种调色过程就是为获得尽可能近似的理想色彩所进行的一种渐进的过程。同理，微扰理论是由已知该怎样解决的问题入手，通过一种渐进程序，逐步接近你要研究的正确答案的一种方法。

然而，试图找到一个有关强耦合理论问题的答案，更像是要重现杰克逊·波拉克（Jackson Pollack）的油画效果而随意地倾倒颜料一般：每次你倒出一点颜料，整幅画就会彻底改变一次！在经过12次重复之后，你的画一点也不会比在8次重复之后更接近理想的效果。事实上，在你每次倒出颜料时，你并不希望它会给画造成太大的改变，完全盖过你上一次的努力，以致你每次都好像要完全从头开始一般。

当一个可解理论被强相互作用微扰时，微扰理论同样是无效的。就如你想重现一幅现代派代表作会徒劳无功一样，当试图接近一个强相互作用理论的数值时，即使是全方位的努力也不会成功。只有当相互作用很微弱时，微扰理论才是有效的，计算也会在掌控之内。

有时，在某些例外的情形中，即使微扰理论无效，你仍能理解一个强相互作用理论的定性特征。例如，系统的物理描述在大致轮廓上可能像一个弱相互作用理论，即便细节大不相同，但在更多情况下，对有强相互作用的理论也根本无能为力。即使在定性特征上，强相互作用系统也常常截然不同于表面上看似微弱的相互作用系统。

因此，对于强相互作用的十维弦理论的解决可以有两种可能：没人能解决它，对它无能为力；或者，至少在大致轮廓上，强相互作用的十维弦理论能够与弱耦合的弦理论看起来一样。可矛盾的是，在某些情况下，这两种可能都不对。在一种被称作IIA的特种类型的十维弦理论里，强相互作用的弦理论与弱相互作用的弦理论看上去毫无相似之处。但是，因为它可以计算，是一个可驾驭的系统，所以我们仍可以研究它的结果。

1995年3月在南加州大学召开的弦理论年会上，爱德华·威滕让所有与会观众大吃一惊，因为他证明了如下观点。

在低能量上有着强耦合的十维超弦理论完全等同于人们原以为完全不同的十一维超引力论，即包含引力的十一维超对称理论。而在这一个等效的超引力理论里，物质的相互作用是微弱的，因此微扰理论完全可以适用。

这似乎自相矛盾，它意味着可以用微扰理论来研究原来的强相互作用的十维超弦理论。在此，不是把微扰理论用于强相互作用弦理论本身，而是用于一个表面上全然不同的理论：弱相互作用的十一维超引力论。剑桥大学的保罗·汤森德（Paul Townsend）以前也注意到了这一令人瞩目的结果，它意味着：尽管从外面看两者截然不同，但在低能量上，十维超弦理论与十一维超引力论实际上是同一理论。或者，如物理学家所说，它们是对偶的。

我们仍可以用调色的比方来说明对偶性。假设我们以蓝色开始，然后慢慢加进绿色对其“微扰”，那么对这种混合颜料的正确描述是：蓝色微带一点儿绿意。但反过来，假设我们加进的绿色颜料并非小量的“微扰”，而是大量的绿色颜料——如果这个量远远超过了原来的蓝色颜料，那么对这种混合颜料更好的“对偶”描述就是：绿色微带一点儿蓝。可见，选择哪种描述完全要取决于被加进的每种颜色的量。

同样的道理，当相互作用的耦合很小时，一个理论可能会有一种描述，但当耦合足够大时，微扰理论在原本的描述里便不再有效。但是，在某些非同一般的场合里，原来的理论可以换上完全不同的“外形”，以使微扰理论能够适用，这就是对偶性。

这就好像有人把一餐五道菜的所有原料一股脑地塞给了你：即便所有原料都很齐全，但你可能并不知道该从哪里开始。为了准备这一餐，你必须弄清楚：每种原料是为哪道菜准备的；一种调料配这种食物是什么效果，配另一种食物又是什么效果；该用什么烹调方法；什么时候上什么菜。但是，如果承办宴席的人把同样的原料事先安排妥当，然后把它们按沙拉、汤、开胃菜、主菜、甜食等配置好给你，我想任何人都能用它们准备一桌盛宴。同样的原料如果事先做好安排，那么准备宴席就从一项复杂的任务变成了小菜一碟。

弦理论里的对偶性就是以这种方式运作的：尽管强相互作用的十维超弦理论看起来完全不可驾驭，但对偶性描述会自动地把所有东西组织成一个可以应用微扰理论的理论。在一种理论里难以驾驭的计算，在另一理论里则会变得容易起来。即便在一个理论里耦合太大不能使用微扰理论，但在另一理论里的耦合就可以足够小，使你可以进行微扰计算。但是，我们还不能完全理解对偶性，例如当弦耦合既不很大也不很小时，没人知道该怎样进行计算。但是，当其中的一个耦

合要么很小要么很大时（相对应的另一个就是要么很大要么很小），那我们就能进行计算了。

强耦合的超弦理论与弱耦合的十一维超引力论的对偶性告诉我们，通过使用一个貌似完全不同的理论进行计算，即便在强相互作用的十维超弦理论里，你也可以算出你想知道的所有事情。由强相互作用的十维超弦理论预言的所有东西都可以由弱相互作用的十一维超引力论萃取出来，反过来也是一样。

使得对偶性如此不可思议的一个特征是：两种描述都只包含局域相互作用——即与邻近物体的相互作用。即使对应的物体在两种描述里都存在，两种描述都有局域相互作用，对偶性也只是一个真正意外和有趣的现象。然而，维度不仅仅是点的集合，还是根据事物远近对事物进行组织的方式。一堆电脑文件可能包含了我想知道的一切，而且也算得上是一套组织得当的文件，但是，只有当信息被前后连贯地组织起来且包含了相关的周边信息后，它才能成为一个凝练的描述。正是在十维超弦理论和十一维超引力论里都存在的局域相互作用，才使得两种理论里的维度以及理论本身变得有意义起来。

十维超弦理论与十一维超引力论的等效被剑桥大学的保罗·汤森德和当时在得克萨斯A&M大学的迈克尔·达夫（Michael Duff）所证实。长期以来，很多弦理论学家一直拒绝并诋毁他们对十一维超引力论的研究——他们不明白，当弦理论明显是未来最有希望的物理理论时，达夫和汤森德为什么要浪费时间去研究这一理论？在威滕的发言之后，弦理论学家只能承认，十一维超引力论不仅有趣，且与弦理论具有同等的价值！

在由伦敦乘飞机回国的途中，我知道了对偶性令人惊讶的结果引起了多少人的关注。一个同行的旅客看到我在读一些物理学的论文，便走过来问我，宇宙究竟是十维的还是十一维的？我感到有点儿惊

讶，因为他是一个摇滚音乐家。不过，我还是回答了他并给他解释，在某种意义上，十维和十一维都对。因为两种理论等效，都可以被认为是正确的。惯例是给出一个理论的维数，只要它有弱相互作用的弦，也因此有较低的弦耦合物理值。

与标准模型里的力相关的耦合，其强度是我们能够测量的，可与此不同的是，我们还不知道弦耦合的大小。它可能很微弱，在这种情况下，我们可以直接使用微扰理论；但它也可能很强烈，这样的话，我们使用对偶理论里的微扰也足以应付。如果不知道弦耦合的值，我们就无法知道在应用于现实世界时，两种描述中的哪一种更简单地描述了弦理论。

在1995年的弦理论会议上，关于对偶性还有更多的惊喜。在此之前，大多数弦理论学家以为超弦理论有五种形式，每种形式都含有不同的力和相互作用。但在会上，威滕（在他之前，还有汤森德及另外一位英国物理学家克里斯·赫尔[Chris Hull]）证明了超弦理论各对形式之间的对偶性。在1995—1996年间，弦理论学家证实，所有这些十维弦理论彼此之间都是对偶的，而且与十一维超引力论也对偶。威滕的发言激发了一场真正的对偶性革命。有了这一来自膜特征的其他条件，5种明显不同的超弦理论被证明为同一种理论的不同表现形式。

因为弦理论的各种形式实际上都是一样的，威滕肯定必然有一个单一的理论能够兼容十一维超引力论和形式不同的弦理论，无论它们是否只包含弱相互作用。他将这种新的十一维超引力论命名为M理论——即我在本章中开始提到的理论。

由M理论你可以得到超弦理论的所有已知形式，但M理论还将已知形式延伸至我们尚不明了的领域。它有可能给出了一个更为统一与连

贯的超弦图像，并最终实现了弦理论的宿愿，使其成为一个量子引力理论。

但是，要实现这一目标，我们还需要更多的信息和模型来充分领会M理论。如果超弦理论的各种已知形式是考古遗址中挖掘出的陶瓷碎片，那么M理论就是我们一直探寻的、将碎片拼接起来的神秘工艺品。还没有人知道构建M理论的最佳途径，但现在弦理论学家已经把它当作了首要的目标。

对偶性到底是什么

本节我将详述上面提到的在十维超弦理论与十一维超引力论之间的对偶性。以后我不会再用到这一解释，因此你尽可以跳到下一章。但是，因为本书是关于维度的，谈一点儿有着不同维度的两个理论之间的对偶性并未完全脱离主题。

一个特点使得对偶性的存在更为合理：两个理论里总有一个包含强相互作用的物体。如果相互作用强烈，你很难直接推导出理论的物理含义。尽管一个看上去是十维的理论要由一个完全不同的十一维理论给出最佳描述非常奇怪，但是，想想你的十维理论里包含的物体有那么强烈的相互作用，以至于你根本无法预言会发生些什么，这似乎就没那么奇怪了。毕竟，我们已然输掉了所有的赌注。

但是，关于不同维数理论之间的对偶性仍有很多令人费解的特征。在十维超弦理论和十一维超引力论之间对偶性的这一特殊情形，乍看起来似乎存在着一个极为根本的问题：十维超弦理论包括了弦，而十一维超引力论里则没有。

物理学家用膜来解决这一问题，即使十一维超引力论不包括弦，但它包括二维的膜。不同的是，弦只有一个空间维度，而2-膜有两个

维度（正如你猜想的一样）。现在，假设十一维中的一个维度卷曲成了一个极小的圆，这样，包含一个卷曲的圆形维度的二维膜看上去就像一根弦，如图15-1所示，卷曲的膜看起来只剩下一个空间维度。这意味着，有了卷曲的维度，即使原来的十一维超引力论并不包含弦，但它看上去确实包括了弦。

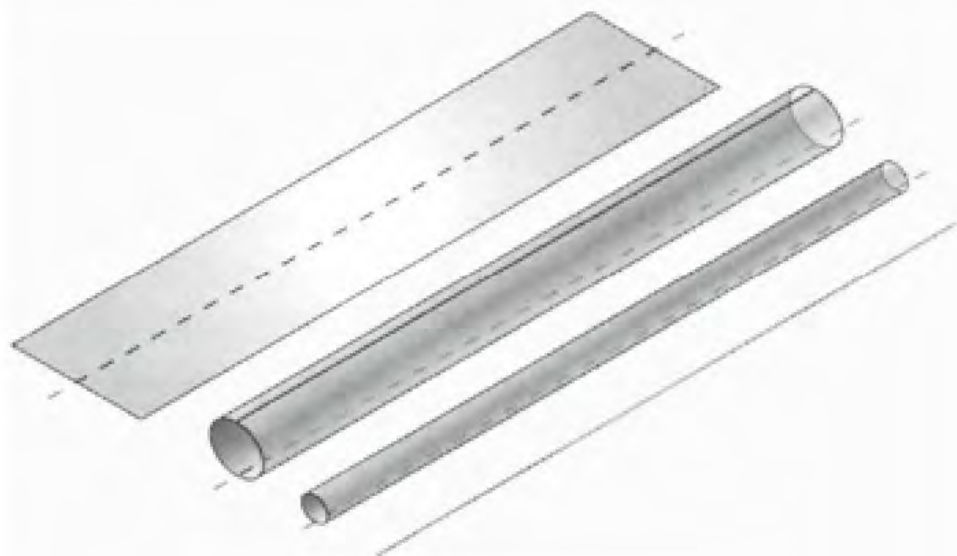


图15-1 2-膜图示。有着两个空间维度的膜，其中一维卷曲成了一个很小的圆，它看上去就如一根弦。

这听上去有点自欺欺人，因为我们已经说过，在远距离和低能量上，一个有着卷曲维度的理论看起来总是会包括更少的维度，因此，当你发现一个有着卷曲维度的十一维理论表现得就如十维理论一样时，就不会感到惊讶。如果你想证明这些十维与十一维的理论都是等价的，为什么仅仅研究十一维理论中一个卷曲的维度就够了呢？

答案的关键在于，我们在第2章里只说明了卷曲维度在远距离和低能量上是看不见的，而威滕在1995年的会议上做了更深入的阐释。他证明：即便在近距离上，一个有着一个卷曲维度的十一维超引力论与十维超弦理论也是完全等价的。当一个维度卷曲时，如果你靠近了

看，仍然可以区分出沿着这个维度的不同位置的点。威滕证明，对偶性理论里的所有事物都是等效的，甚至包括那些有足够能量去探索小于卷曲维度距离的粒子。

在有着一个卷曲维度的十一维超引力论里的所有东西——甚至是微小尺度和高能的过程与物体，在十维超弦理论里都能找到其对应物。而且，不管维度卷曲成任意大小的圆圈对偶性都成立。以前我们看卷曲维度时，只是说小卷曲维度不会被注意。

但是，不同维数的理论怎么可能是一样的呢？毕竟，空间的维数是我们确定一个点所需要的坐标的数量。只有当超弦理论总是用额外数字来描述点状物体时，对偶性才可能成立。

对偶性的关键就在于，在超弦理论中有一种特别的新粒子，只有明确了它在九维空间里的动量及电荷值，你才能确定它。而在十一维超引力论里，你需要知道在10个空间维度里的动量。注意，即便在一种情形下你有9个维度，而在另一种情形里是10个，在这两种情形里，你都需要明确10个数字：一种情况下是9个动量值和1个电荷值；另一种情况下是10个动量值。

常见的不带电荷的弦与十一维超引力论里的物体不匹配。因为在十一维超引力论的时空里，定位一个物体需要11个数字，因此只有带电粒子才有其十一维的“配偶”。而十一维超引力论里粒子的伙伴变成了膜——即带电的点状膜，叫作D0-膜。弦理论和十一维超引力论是对偶的，因为对应在十维超弦理论里的每一个既定电荷的D0-膜^[56]，都有一个相应的、特定的十一维动量的粒子，反之亦然。十维和十一维理论里的物质（以及它们的相互作用）恰好对等。

尽管在某一特定方向上，电荷与动量似乎大不相同，但如果在十一维超引力论里，每个特定动量的物体，都能够与十维超弦理论里的

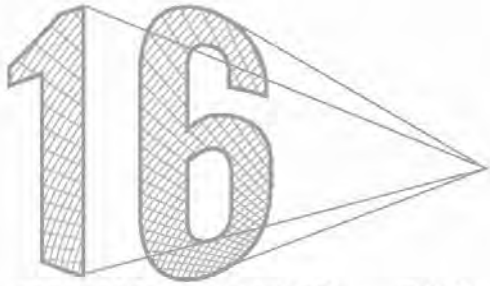
某个特定电荷的物体相匹配（反之亦然），那么这一数字究竟是被称作动量还是电荷，就由你来做主了。维数是独立的动量方向的数量——即一个物体可以在其中穿行的不同方向的数量。但是，如果沿着一个维度的动量可以由一个电荷所取代，那么维度的数量就没有被明确界定。最佳的选择要由弦耦合的值来确定。

这一令人震惊的对偶性是证明膜有建设性意义的最早分析。不同的弦理论要互相匹配，膜是必须的附加成分。但弦理论里的膜之所以在物理学理论的应用中非常重要，是因为一个关键特征：它们能束缚粒子和力。下一章我们就将解释其中的原因。

-
- 弦理论是一个错误的名字：弦理论里还包含了高维的膜。在弦理论里，开弦（不能将自己绕成圈的弦）必须在膜上终止，这种膜便是D-膜。
 - 最近十几年，在弦理论的许多发展成果中，膜发挥了重要作用。
 - 对偶性表明，表面不同的弦理论形式实际上都是等效的。在对偶性的证明中，膜发挥了关键作用。
 - 在低能量上，十维超弦理论与十一维超引力论——包含超对称和引力的十一维理论——是对偶的，一种理论里的粒子与另一种理论里的膜相对应。
 - 本章有关膜的结果与后面的讨论无关，但这些结果确实说明了在弦理论学界由膜所引发的一些兴奋。
-



-imension
探索大揭秘



熙熙攘攘的通道：膜宇宙

WARPED PASSAGES

欢迎你来到这个世界，在这里，时间是静止的，没有人离开，也没人想离开。

金属乐队 (Metallica)

天堂膜与监狱膜

艾克对天堂越来越失望了，他原指望这里的环境是自由、宽松的，可事实并非如此：这里禁止赌博，禁止使用金属和银制器皿，甚至连吸烟都不允许！而限制令中最为苛刻的是，天堂本身就被限制在天堂膜上，但天堂里的居民却不可以自由穿行到第五维度。

天堂膜上的所有人都知道有个第五维度，而且还存在着其他膜。事实上，居住在天堂膜的人私下里常议论一些声名狼藉的人物，他们就被囚禁在不远处的监狱膜里。但是，监狱膜里的囚犯听不到天堂膜里散播的有关他们的流言，因此在膜上和整个体里一切都还算相安无事。

从“对偶性革命”的角度来看，你可能会以为，对于想把弦理论与可见世界联系起来的人来说，膜可谓一个福音。如果所有不同形式的弦理论实际上是同一个东西，那么物理学家也就不必再面对令人却

步的难题，去寻求自然据以选择的理论。既然外表不同的所有弦理论本质都一样，那么也就无所谓孰优孰劣了。

尽管看上去我们很快就要找到弦理论与标准模型的联系了，但事情并没那么容易。虽然膜在对偶性里发挥了至关重要的作用，减少了弦理论的形式，但实际上却增加了标准模型的可能表现方式。这是因为，膜能包容理论学家最初创建弦理论时没有考虑到的粒子和力。

因为膜存在的类型及其在弦理论高维空间里的位置有许多可能，可想而知，在弦理论里实现标准模型就会有更多新的方式，这都是物理学家以前没有想过的。标准模型的力并不一定是由一种基本弦形成的，它们可能会是弦在不同的膜上延伸而产生的新力。尽管对偶性告诉我们弦理论最初的五种形式是相同的，而弦理论中可以想见的膜宇宙的数目却极为庞大。

要找到唯一一个标准模型的候选者似乎一点儿也不比从前更容易。意识到这一点以后，弦理论学家关于对偶性的狂热渐渐冷却了下来，可是像我们这些试图从可见物理现象中寻找新见解的人却似进了天堂。现在，力和粒子都被困在膜上，有了这一新的可能，该是我们重新审视粒子物理学出发点的时候了。

膜能够束缚粒子和力，这一特点是其可能被应用于可观测物理的关键，本章的目的就是让你简单了解它是如何作用的。我将首先解释弦理论的膜为什么会束缚粒子和力，然后我们会了解膜宇宙的观点，以及由对偶性和弦理论发展而来的、最早为人们所知的膜宇宙。在下一章里，我们将继续探索我认为最为令人振奋的膜宇宙的一些特点，以及它们在物理理论里的可能应用。

被困在膜上的粒子

正如来自杜伦大学的广义相对论学者露丝·格雷戈里（Ruth Gregory）所说：弦理论里的膜“满载着”粒子和力。这就是说，总是有一些粒子和力被困在膜上。就如一只圈养在家里的猫从来不会冒险走出家门一样，那些被困在膜上的粒子也从不会离开膜。它们不能离开，它们的存在必须依附于膜。即使它们移动，也只是在膜上移动；即使彼此作用，也只是在膜延伸开的空间维度里作用。从被困于膜的粒子角度来看，如果不是引力或可能与它们相互作用的空间里的粒子，世界也就只有膜的维度。

现在我们来查看弦理论是如何将粒子和力困在膜上的。设想只有一个D-膜漂浮在高维宇宙的某个地方，根据定义，开弦的两个端点一定在一个D-膜上，这个D-膜就是开弦开始和终止的地方。开弦的两个端点未必固定在一个地方，但必定是在膜上，就如铁轨一样，虽然它们限制了轮子，却可以让轮子在上面行进。同理，膜就是一些固定的表面，弦的端点虽然被困于其上，却仍可能有限度地移动。

因为开弦的振动形成了粒子，两个端点被困于一个膜上的开弦的模式就形成了被困于这个膜上的粒子，这些粒子只能在膜延伸的维度里穿行和相互作用。结果发现，膜束缚的弦所产生的粒子就是可以传递力的规范玻色子。之所以这样说，是因为它有着规范玻色子的自旋（是1），而且因为它的作用方式就如规范玻色子一样。这样一个被困于膜上的规范玻色子会传递力，这种力则作用于被困在膜上的其他粒子，计算显示：在接受端的粒子总是受到这种力的影响。事实上，终止于膜上的任何弦的端点都会表现得像一个带荷粒子。正是这些被困于膜上的力和带荷粒子的存在告诉我们，弦理论里的D-膜会“满载着”带荷粒子和作用于它们的力出现。

如果结构里不止存在一个膜，那就会有更多的力和更多的带荷粒子。例如，假设有两个膜：在这种情况下，除了被困在每个膜上的粒

子外，还会有一种新粒子，它是由两个端点各位于一个膜的弦产生的（见图16-1）。

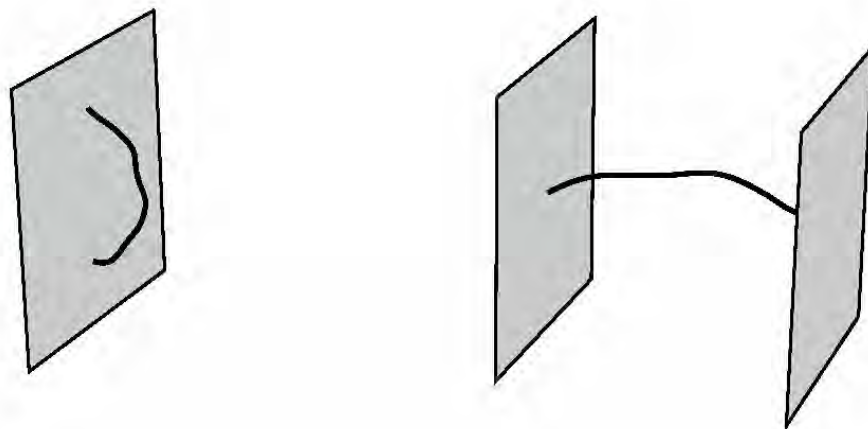


图16-1 膜上的弦生成的粒子。开始和终止于一个膜上的弦生成的是一个规范玻色子，而两个端点各位于一个膜上的弦会生成一种新型的规范玻色子。如果两个膜分隔开来，规范玻色子就有了非零质量。

如果两个膜在空间中隔开，那么在它们之间延伸的弦所生成的粒子就会很重。由这种弦的振动模式所产生的粒子，其质量会随着两膜之间距离的增大而增大，这个质量就像是弹簧被拉伸所积聚的能量——拉伸越厉害，积聚的能量就越大。同样的道理，在两个膜之间延伸的弦所能生成的粒子质量与两膜之间的距离成正比。但是，当弹簧处于静止状态时，它是松弛的，没有积聚任何能量。同样地，如果两个膜没有被隔开——即它们在同一位置上，那么在每个膜上都有一个端点的弦所产生的最轻的弦粒子就是无质量的。

现在，我们假设两个膜出现了重叠，由此就生成了一些无质量粒子。其中一个无质量粒子是规范玻色子，它不同于两个端点在同一个膜上的弦所形成的规范玻色子，而是一种全新的玻色子。这种新的无质量粒子，只有当两个膜碰巧重叠时才会产生，它所传递的力既可能作用于一个膜上的粒子，也可能作用于同时在两个膜上的粒子。与其

他所有力一样，膜上的力也有一种相关的对称。在这种情况下，对称变换就是两个膜的交换。

当然，如果两个膜真的在同一位置上，你可能就会感到奇怪，为什么要把它们看作两个物体呢？你的想法很对：如果两个膜重叠，那么你完全可以把它们想象成一个膜，这个新型的膜在弦理论里是存在的。它是两个秘密交会的膜，有着所有这两个膜的属性。它束缚了以上讨论的所有不同类型的粒子：既有两个端点各在两个膜上的开弦形成的粒子，也有端点在同一膜上的闭弦形成的粒子。

现在，假设许多膜重叠在一起，由于弦的两端可能被困于任何一个膜上，那么这样就可能生成多种新型的开弦（见图16-2）：在不同的膜之间延伸的开弦，或是两个端点在同一张膜上的闭弦，这些弦的振动模式意味着新的粒子。这些新粒子包括了新型的规范玻色子和新型的带荷粒子；与这些新的力相关的对称交换是重叠在一起的各个膜。

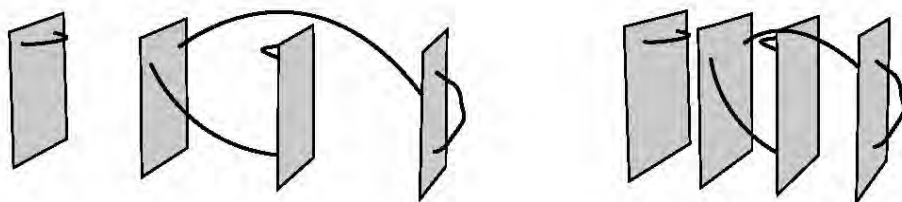


图16-2 新型粒子的形成。在同一个膜上开始和终止的弦或是在两个膜之间延伸的弦都会形成规范玻色子。当膜重合在一起时，对应于一个弦在重合膜上开始和终止的各种方式，就会有新的无质量规范玻色子生成。

因此，膜真的是“满载着”力和粒子出现的，膜越多意味着可能性越多。如果涉及的是几组不同的膜，甚至可能产生更为复杂的情形：位于不同地点的膜承载的会是完全不同的粒子和力，被束缚于这组膜上的粒子和力可能会完全不同于被束缚于另一组膜上的粒子

和力。例如，形成我们的粒子连同电磁力，都被限制在了一个膜上。因此，我们会经受电磁力，而被局限于远处另一个膜上的粒子就经受不到——那些遥远的粒子对电磁力无动于衷。反过来讲，被限制于遥远膜上的粒子所经受的新奇的力，也是我们所完全感受不到的。

这种情形的一个重要性质（后面我们还会用到）是：在不同膜上的粒子之间不会直接发生作用。相互作用是局域性的，即它们只发生在同一位置的粒子之间，分隔在不同膜上的粒子相距遥远，不能直接相互作用。



你可以把体，即整个高维空间，比作一个庞大的网球场，里面正同时进行着几场独立的比赛。每一个场地上的球都会越过球网来来回回，而且可能会跑到场地里的任何地方，但是，每场比赛都相对独立地进行着，因此球只能待在各自的场地里。就如一个只能在规定场地里存在的球一样，只有参与比赛的两位网球选手能够接触到它，被膜束缚的规范玻色子和其他困在膜上的粒子也只能与它们自己膜上的物质相互作用。

但是，如果有粒子和力能够自由地穿梭于体空间中，那么分置于不同膜上的粒子就能够通过它们相互交流了。这样的体粒子可以自由地进出一个膜，偶尔可能会与一个膜上的粒子相互作用，也可能在整个高维空间中自由穿行。

存在着不同的膜和在它们之间穿梭的体粒子，这一情形就如在一个庞大的体育场里同时进行着几场比赛，不同的选手都是同一个教练。因为要同时注意几场比赛的进程，这个教练会不停地从一个场地转到另一个场地。如果一个选手想与其他场地的选手交流，他可以告诉教练，让教练代为传达。比赛时，选手之间不能直接交流，但他们可以通过能在不同场地间自由行动的一个人帮助传话。同样地，体粒子能与一个膜上的粒子相互作用，随后又与另一个遥远膜上的粒子相互作用，由此，它就使得被束缚于不同膜上的粒子能间接地“交流”起来。

下一节我们会看到，引力子——传递引力作用的粒子，就是这样一种体粒子。在一个高维构成中，它能在空间里自由穿行，并与所有地方的粒子相互作用，无论它们是在膜上还是在膜外。

逃出重围的引力子

与其他所有力不同，引力不会被束缚于一个膜上。被膜束缚的规范玻色子和费米子都是开弦的产物，但在弦理论里，传递引力作用的粒子——引力子，是一种闭弦。闭弦没有端点，因而也就没有端点能把它钉在一个膜上。闭弦的振动模式所形成的粒子能在整个高维空间里不受约束地行动。我们知道引力是由闭弦粒子传递的，因而再次与其他的力区分开来。与规范玻色子或费米子不同，引力子一定是能在整个高维空间里自由穿行的，没有办法能将引力局限于低维空间里。

在后面的章节里我们会看到，令人称奇的是，引力会局限于一个膜的附近，但我们不可能真的将它束缚在一个膜上。

这就意味着，尽管膜宇宙可以将大多数粒子和力束缚于膜上，但它们却不能困住引力。这是一个很好的属性，它告诉我们，即使整个标准模型被困在一个四维的膜上，膜宇宙总会包括高维物理。如果有一个膜宇宙，那么里面的所有东西都会与引力相互作用，在整个高维空间里，引力处处都能被感受到。我们很快就将看到，引力这一与众不同的特征为什么可以解释“相比其他力，它为何如此微弱”的问题。

膜宇宙真的存在吗

一旦物理学家认识到膜在弦理论里的重要意义，膜很快便成为集中研究的焦点。尤其是，物理学家急于要了解它们对粒子物理学以及宇宙学概念的潜在意义。而现在，弦理论还不能告诉我们膜在宇宙里是否存在，倘若存在，又会有多少。我们只知道膜是弦理论里一个基本的理论片段，没有它，弦理论就不能成为一个完整体。既然我们知道了膜是弦理论的组成部分，我们肯定还想知道，在真实世界里它们究竟是否存在，倘若存在，它们的结果又是什么？

膜可能存在这一事实，又为宇宙的构成增添了更多的可能性，其中有些甚至会与我们看到的物质属性密切相关。在听到露丝·格雷戈里所说“满载着粒子”的膜时，弦理论学家阿曼达·皮特（Amanda Peet）调侃道：“膜开辟了以弦为基础构建模型的新天地。”1995年之后，膜成了模型构建的一种新工具。

到了20世纪90年代末，许多物理学家，包括我在内，都拓展了自己的视野——把膜包括在内。我们自问：“如果有一个高维宇宙，我

们已知的粒子和力不能在其中所有的维度自由运动，而是局限于一个低维度膜上的少数几个维度里，结果会是怎样？”

膜宇宙图景为我们时空的总体性质提供了更多的可能性。如果标准模型粒子被束缚在膜上，那么我们也是一样，因为我们及周围的宇宙都是由这些粒子构成的。而且，并非所有的粒子都处在同一个膜上，因此还可能会有不为我们所知的、全然不同的新粒子经受着与我们已知的不同的力和相互作用。我们观察到的粒子和力可能只是一个更为浩瀚宇宙的一小部分。

康奈尔大学的两位物理学家亨利·泰伊（Henry Tye）和祖拉·卡布沙兹（Zurab Kakushadze）创造了“膜宇宙”一词来标识这样的图景。亨利告诉我，用这个词他就能一下子描述出宇宙包含膜的所有可能方式，而不必坚持其他可能。

弦理论学家试图得到一个有关世界的唯一理论，膜宇宙数量的激增就有点令人心烦，但同时也令人兴奋。它们是我们生活的世界的所有真实可能，而且其中某个很可能真的描述了我们的世界。因为粒子物理学的法则在更高维宇宙里可能不同于物理学家原来的想象，对标准模型的那些令人费解的特征，额外维度提出了许多新的解答方法。虽然这些观点只是推想，但能解决粒子物理学问题的膜宇宙很快将在对撞机实验里得到检验。这就意味着，是实验而不是我们的偏见将最终决定这些观点是否适用于人类世界。

我们将探讨这些新的膜宇宙，它们会是什么样子？它们的结果又是什么？我们不会把自己局限在明显由弦理论得出的膜宇宙里，而是要探讨已把新观点引进了粒子物理学的膜宇宙模型。物理学家还远远不能理解弦理论的意义，虽然还没有人找到一个有着特定的粒子和力或一定能量分布的弦理论例证，但仅仅为此就排除一些模型未免为时尚早。这些膜宇宙将被当作弦理论探索的目标。在第20章讨论弯

曲等级模型时，我和拉曼·桑卓姆把它当作一种可能的膜宇宙引进，并由弦理论推导出来。

后面的几章，我将介绍几个不同的膜宇宙，每个都呈现了一个全新的物理情境。第一个说明膜宇宙怎样避开“无政府主义原理”；第二个说明维度比我们以前设想的要大得多；第三个说明时空弯曲可以满足物质有不同的大小和质量；最后两个说明，即使无限大的额外维度在时空弯曲时也可能看不见，甚至时空在不同的地方也可能看上去会具有不同的维度。

我将提出几个模型，因为它们都是真实的可能，其中每个都包含了一些新特征，这是物理学家不久之前还认为不可能的。在每章的末尾，我将总结每个模型的意义及其突破传统观念的地方。当然为了首先得到一个整体了解，你可以先读这些“探索大揭秘”，它们概括了该章介绍的特定模型的意义。

在进入这些膜宇宙之前，我将首先简要介绍第一个为人们所知的膜宇宙，它是由弦理论直接推导出来的。皮特·霍扎瓦和爱德华·威滕在他们探索弦理论对偶性的过程中，偶然发现了这一膜宇宙——我们就以他们的名字首字母为它命名为“HW”模型。我之所以要提到这一模型，一是因为它本身就很有趣，二是因为它有几个性质预示了我们即将看到的其他膜宇宙的特征。

霍扎瓦-威滕模型

图16-3是HW膜宇宙的图示。这是一个由两个平行膜作边界的十一维世界，其中每个膜都有9个维度，它们包围着一个有10个空间维度（十一维时空）的体空间。HW宇宙是最早的膜宇宙理论，在HW中，两个膜各含一组不同的粒子和力。

两个膜上的力与第14章介绍的杂化弦理论的力是相同的。那个理论是由格罗斯、哈维、马丁尼克和罗姆发现的，其中沿着弦向左或向右移动的振动会有不同的相互作用。这些力一半被束缚在其中的一个边界膜上，另一半被束缚在另一个边界膜上，两个膜上都束缚了足够的力和粒子。可以想象，每个膜都包含了标准模型的所有粒子（因此也包含了我们）。霍扎瓦和威滕假定，标准模型的粒子和力都停留在其中一个膜上，而引力和同为理论组成部分的其他粒子，虽然在我们的世界里还没见过，但在整个十一维体空间里，它们能自由地穿行于两个膜之间。

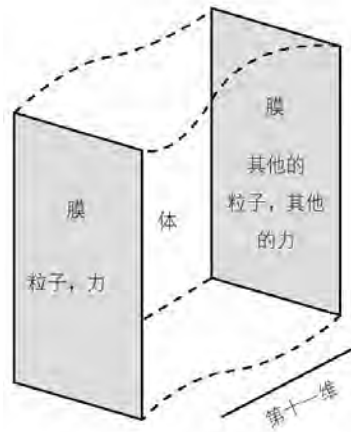


图16-3 霍扎瓦-威滕膜宇宙简图。两个有着9个空间维度的膜（我们用二维来表现）在第十一时空维度（第十空间维度）上被分隔开来，体里包含了所有的空间维度：沿着两个膜的空间方向延伸的九维，及在两膜之间延伸的另一维。

事实上，HW膜宇宙不仅与杂化弦有相同的力——它就是杂化弦，只不过是强耦合的。这是对偶性的另一个范例，在这种情形里，以两个膜作为第十一维（空间第十维）边界的十一维理论与十维杂化弦对偶。也就是说，当杂化弦的相互作用非常强烈时，最好把它描述成有2个边界膜和9个空间维度的十一维理论。这与前一章我们探讨的十维超弦理论和十一维超引力论之间的对偶性如出一辙，但是，在这个例子里，第十一维没有卷曲，而是被束缚在两膜之间。十一维理论再次与

十维理论对等起来，只不过一个理论是强相互作用，而另一理论是弱相互作用。

当然，即使标准模型粒子限制在膜里，理论的维度还是比我们在周围世界看到的更多。HW膜宇宙如果要与现实世界相对应，其中的6个维度一定是看不见的，霍扎瓦和威滕认为那6个维度卷曲成了极微小的卡拉比-丘流形。

一旦那6个维度卷曲起来，你就可以把HW宇宙当作一个有着四维边界膜的五维有效理论。这幅有两个边界膜的五维宇宙图像非常有趣，许多物理学家都曾进行了深入的研究。我和拉曼就将伯特·奥弗鲁特（Burt Ovrut）和丹·沃尔德莱姆（Dan Waldram）两位物理学家在研究HW有效理论时所使用的一些技巧，应用于我将在第20章和第22章讨论的不同的五维理论。



HW膜宇宙里一个引人入胜的因素是，它容纳的不仅是标准模型的粒子和力，而且还有一个完整的大统一理论。因为引力源自更高的维度，因此在这一模型里，引力就有可能与其他力在高能量上有相同的强度。

HW模型给出了三个原因，说明膜宇宙于真实世界物理是非常重要的。第一，它包含的不仅仅是一个膜，这意味着，它可以包含因为束缚膜之间的间隔而只能发生弱相互作用的粒子和力。被限制在不同膜上的粒子相互交流的唯一方式，就是通过它们与空间粒子的共同

作用。这个特点对我们下章将探讨的隔离模型是非常重要的。

膜宇宙的第二个重要特点是，所有的膜宇宙都给物理学引进了新的距离尺度。这些新的尺度，如额外维度的大小，对力的统一问题或等级问题的解决可能有重要意义。在这两个理论里，问题的核心都是：在同一个理论里，为什么会有差异这么大的能量和质量级别，量子效应为什么不能将两者统一起来？

最后一点是，膜和空间能够承载能量。这个能量可以由膜和高维的空间所贮存；它不依赖于在场的粒子，与所有形式的能量一样，它会使体空间产生弯曲。我们很快将看到，这种由遍布于空间的能量引起的时空弯曲结构对膜宇宙是非常重要的。

HW膜宇宙当然还有许多诱人的特点。但弦理论在重现已知现象时所存在的许多问题，它也同样存在：因为维度非常微小，霍扎瓦-威滕理论很难由实验检验；为了逃避观测，许多看不见的粒子必须很重；其中的6个维度必须卷曲起来，但卷曲维度的大小和形状都没有确定。

沿着这些线索继续下去，也许我们会不期然地发现能够正确描述自然的弦理论形式，我们并不完全排除这种可能，不过我们真的需要非常幸运才可能看到它发生。但粒子物理学的问题也在向我们招手，我们需要考察在有着额外维度和沿着其中某些维度延伸的膜的世界

里，如何回答这些可能被解决的问题。这就是本书其余部分将探讨的内容。

-
- 在弦理论的框架里，膜宇宙是可能的。弦理论的粒子和力可能被束缚在膜上。
 - 引力与其他力不同：它永远不会被限制在膜上，而会穿越所有的维度散播。
 - 弦理论若要描述世界，它会包含许多膜。在这个意义上，膜宇宙是自然而然的。



-imension
探索大揭秘



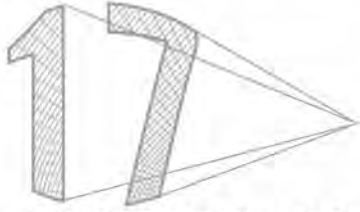
WARPED

第五部分

额外维度宇宙假说

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



人迹罕至的通道：多重宇宙与隔离

WARPED PASSAGES

现在请你走开，（因为）你已不再受欢迎。

格洛丽亚·盖罗（Gloria Gaynor）

天堂里的危机

尽管天堂膜里明令禁止赌博，但艾克还是积习难改。由于三番五次地违反禁令，他被判监禁，到监狱膜里服刑。监狱膜沿着第五维度与天堂膜遥遥相望。即使被关在监狱膜里，艾克还常常试图偷偷地与他的伙计们取得联系。但两膜相距遥远，通讯非常困难，他只能退而求其次，寄希望于可以在空间里走动的过路信使，但他们大多根本无视他的请求。偶尔有那么几个终于肯停下来为他传话到天堂膜里的，也总是迈着悠闲的步子，速度慢得让人心焦。

而与此同时，天堂膜里正酝酿着一场危机。那些护卫天使，曾英勇地捍卫了等级，而现在却无视其他居民的家庭观念，即将破坏稳定，引发世代混乱。堕落的天使认为，所有匹配都可以接受，他们还煽动大家都从另一代里抢一个战利品作伴侣。

艾克得知这一危机时惊恐无比，下定决心要挽救局势。不过艾克忽然意识到，他与天堂膜的联络被迫使用的缓慢、刻意的方式，正可以用来满足居住在天堂膜里那些不守规矩天使的极度膨胀的自负心理。这还真是一个明智的点子。幸好有艾克的干预，天使们不再威胁扰乱天堂秩序了。尽管艾克依然要服刑，但天堂膜里获救的居民将他誉为了永远的神话。

本章探讨的主题是隔离。额外维度对粒子物理学来说非常重要，而隔离就是其中的一个理由。粒子在物理上被分隔在不同的膜上，通过把不同的粒子限制在不同的环境中，隔离有可能解释区分粒子的不同特征。隔离也可能是“无政府主义原理”（认为所有东西都会相互作用）并不总是正确的原因。如果粒子隔离在额外维度，那么它们就不太可能相互作用。

原则上讲，粒子可以被隔离在三个空间维度里的。但就我们现在所知，三维空间里的所有方向和所有地点都是相同的：已知物理学定律告诉我们，粒子可能位于我们所见的三个维度的任何一个地方，因此不存在三维的隔离。但是，在多维空间里，光子和带电物体并不一定存在于任何地方，额外维度引进了一个分隔粒子的方法：不同的粒子类型可能被分隔在由不同的膜所占据的不同空间里。由于在额外维度里并非所有地点都一样，通过把不同的粒子类型限制于不同的膜上，额外维度便提供了一种隔离粒子的方法。

隔离粒子的理论可能会解决许多问题，艾克的故事指的就是隔离在超对称破缺中的应用——这是我对额外维度的首次尝试。由于超对称破缺模型通常会引起一些我们不希望的相互作用，因此四维理论正面临着严重的问题，而被隔离的超对称破缺看上去显然更有希望；隔离还可能解释粒子的质量为什么各不相同，在额外维度模型里为什么没有质子衰变发生。本章我们就将探索隔离及其在粒子物理学中的应用，我们会发现即便我们原以为适用于四维时空的观点，如超对称，在额外维度的背景下，也有可能成功。

我的额外维度历程

作为物理学家，我们是幸运的：我们有许多交流机会，能够与同行见面并分享一些激发灵感的科研思想。但是，粒子物理学每年召开

的会议和研讨会数量如此之多，以至于究竟该参加哪一个实在是太令人难以选择了：有一些是重要的聚会，你可能有机会听到别人的新近研究，也可以宣讲自己的最新成果；有一些会议则相对较短，只持续两三天，与会的物理学家汇报的是高度专业领域里的重要新成果；还有一些会议是时间更长的研讨，物理学家在这些研讨会上开始或者完成与同事的合作。这些会议召开的地点常常那么吸引人，让你根本不想错过。

虽然牛津是一个好地方，但1998年7月我在那里参加的超对称会议，恰当地说还应属于第一类。多年来，超对称一直被认为是解决等级问题唯一可能的方法，因此它慢慢地发展成了一个重要的研究领域，物理学家每年都要聚集在一起探讨本领域里的最新进展。

而牛津会议却令人们颇感意外：会上最有趣的话题不是超对称，而是新出现的观点——额外维度。其中最令人振奋的一个发言是关于大额外维度的，即第19章的主题。其他的发言都谈到了弦理论里额外维度的前景，还有一些探讨的是额外维度的实验意义。在芝加哥理论学家杰夫·哈维的发言题目里，你就能清晰地看出这些观点充满了新奇和假设性：他和后来的几个发言人将他们的发言戏称为《梦幻岛》（*Fantasy Island*）；费米实验室的理论学家乔·莱肯（Joe Lykken）做的幻灯片里，甚至在其中一张上有个小人直指“Da膜，Da膜”（不用说，那些没有看过20世纪70年代美剧的人，是感觉不到其中的幽默的。它是一个关于纹身的笑话，因欢迎“Da飞机”到梦幻岛而闻名）。

且不管这些笑话，从牛津超对称会议返回时，我就开始思索有关额外维度的问题：为什么粒子物理学的问题会在一个额外维度的世界里得到解决呢？尽管我对正是热点的大额外维度持怀疑态度，也不打算去研究它，但我非常相信膜和额外维度会成为构建模型的重要工

具，甚至有可能解释某些令简单的四维理论无可奈何的神秘的粒子物理学现象。

那一年，我计划在波士顿度过后半季。我当时的惯例可不是这样：波士顿的大多数理论物理学家，包括我在内，每年夏天大部分的时间都在旅行，参加各种会议和研讨。但这次我决定留在家里休息，想想新的观点。

拉曼·桑卓姆，当时在波士顿大学做博士后，那年夏天也决定待在波士顿。以前开会的时候或我们到彼此的学院访问时，我经常遇到他，我们甚至同时在哈佛大学做过博士后。得知拉曼也在思考额外维度时，我想，与他谈论一下我的观点和问题也许会有不一样的火花。

拉曼是一个有趣的人，大多数物理学家在其事业早期都会研究一些相对安全的问题——大家都共同关注的问题，这样其实更容易取得进步，而拉曼却坚持要研究他认为最重要的问题，即便它极度困难或完全不被他人所关注，他也在所不惜。虽然他的天赋是有目共睹的，可由于他的偏执，始终不能获得一份终身教职，而只能第三次继续他的博士后职位。这时，拉曼开始思考额外维度和膜，他的兴趣与物理学界的其他人终于不谋而合了。

我们的合作开始于麻省理工学院的托斯卡尼尼分店，这是在麻省理工学院学生服务中心的一个冰激凌店，有上好的咖啡和冰激凌供应（很遗憾它现在已关闭了）。托斯卡尼尼是一个交流思想、激荡观点的理想场所，没有限制，不受打扰，同时还能让人尝到沁人心脾、激发科研灵感的美味。

那些日子，我们品味着咖啡闲聊，随着秋天的脚步一天天走近，研究也逐渐成形。到8月份的时候，为了记录讨论的细节，我们需要的黑板越来越大。那时我在麻省理工学院做教授，办公室里的黑板太小，我们就会逛到“无边的走廊”（贯穿麻省理工学院主建筑的长长过道）去寻找空教室。

我们研究的具体问题就是隔离在超对称破缺中的应用。这一观点是把引起超对称破缺的粒子从标准模型里隔离出来，由此阻止它们之间不该有的相互作用（见图17-1）。我们选择“隔离”，是为了区分粒子被不同的膜所分隔的模型与当时十分流行的所谓“隐藏区域”的

超对称破缺模型。在隐藏区域模型里，超对称破缺粒子与标准模型粒子的相互作用很微弱，但并未实际隐藏（这有点儿名不副实），因此它们才会以现实世界不能接受的某些方式相互作用。



图17-1 超对称破缺模型。在这一超对称破缺模型里有两个膜：标准模型粒子在一个膜上，打破超对称的粒子被隔离在另一个膜上。两个膜都有三个空间维度上，它们被第五时空维度，也即第四空间维度分隔开来。

开始时，我对我们的观点热情很高，而拉曼却持怀疑态度；可随着时间的推移，我们的角色不断转换，但总是一个热情，一个冷静，就这样，我们很快穿越了许多研究的迷雾，最终到达了我們思考的物理学问题的核心。有时我们甚至会太过仓促地放弃一些观点，但通常总会有一个人能坚持足够长的时间，使一个观点取得进展。

弗朗西斯·培根与伽利略一起被认为是现代科学方法的奠基人，前者讲道：“为了保证结果的准确，你必须保留一定的怀疑，而同时又要取得进展，这是多么困难。”一边怀疑它的正确性，一边又怎么可能认真地去看待一个观点并深入研究它的结果？如果有足够长的时间，一个人可能在这两种态度之间辗转徘徊，最终得出正确答案。但是，当我们两个人都持相反态度时——这常常是几个小时，甚至是几分钟的事，那么我们很快就会放弃一个虽然有趣却是错误的观点。

可是，我们开始的观点，即通过隔离来防止超对称理论里不该有的相互作用，在我看来似乎就应该是正确的。四维里没有办法给出一个让人信服的解释，而额外维度似乎能为构建一个成功的模型提供必要的工具。可是，直到夏季即将结束的时候，我和拉曼才充分领悟了隔离及其对超对称破缺的作用，并最终对其意义达成了共识。

自然与隔离，信息封锁王国

隔离之所以重要，是因为它能够阻止由无政府主义原理引起的问题，那个未经证实的原理指出，在四维量子场论里，所有能发生的作用都会发生。但这一原理的问题在于，理论预言的相互作用和质量间的关系在自然界里是不存在的。一旦虚粒子被包括进来，即使是经典理论（没有考虑量子力学的理论）不会出现的作用也会出现，可见虚粒子引发了所有可能的相互作用。



有一个类比可以解释其中的原因：假设你告诉阿西娜明天有雪，阿西娜又告诉了艾克，那么即使你没有与艾克直接交流，你的信息仍会影响艾克明天的穿着——由于你的“虚”建议，他可能会穿带帽子的外套。

同样地，如果一个粒子与一个虚粒子相互作用，而这个虚粒子又与第三个粒子相互作用，那么最终的结果就是：第一个粒子和第三个粒子也产生了作用。无政府主义原理告诉我们，即使在经典理论中不存在，涉及虚粒子的过程是必然要发生的，而这些过程常常会引发不该有的相互作用。

粒子物理学理论里的许多问题都源于无政府主义原理。例如，由虚粒子引起的对希格斯粒子质量的量子贡献就是等级问题的根源，希格斯粒子采取任何路径都会受到重粒子的暂时干扰，这些干扰增大了希格斯粒子的质量。

我们在第11章中还看到了关于无政府主义原理的另一个例子：在大多数有超对称破缺发生的理论里，虚粒子都会引发不该有的相互作用——我们由实验得知不会发生的相互作用，这些作用会改变已知夸克和粒子的身份。这种味改变的相互作用在自然界中要么不存在，要么很少发生，若想让一个理论有效，我们必须消除这些作用——即无政府主义原理告诉我们会发生的作用。

虚粒子并不一定会导致这些不该有的预言。有一种情形，即当对一个物理量巨大的经典力学和量子力学贡献相互抵消时，理论就不会预言不该有的相互作用，但这一情形不太可能发生。即使经典贡献和量子贡献各自都非常大，我们仍然可以想象两者相加能得到一个可以接受的预言。但这种应对问题的方法几乎肯定只是替代真正解决方法的权宜之计。没有人会真的相信这种精确而偶然的抵消是不存在某些相互作用的根本解释，我们只是勉强用这种“幸运”的抵消来辅助我们忽略这些问题，继续其他方面的理论研究。

物理学家相信，只有当相互作用的抵消方式符合物理学家认为自然的观念时，相互作用才算真正从理论中消除了。在日常生活中，“自然”指的是那些不经人为干涉、自然而然发生的事；而对粒子物理学家来讲，“自然”指的不仅仅是发生的事情——它还意味着，如果某件事会发生，它不应给人留下任何迷惑。对物理学家来说，只有意料之中的事才是“自然”的。

无政府主义原理和量子力学引发的不该有的相互作用告诉我们，一个支持标准模型的基本理论若要正确，其理论模型里必须纳入新的概念。对称之所以重要的一个原因就是，它们是在四维世界里保证不出现不该有的相互作用的唯一自然的方法。关于哪些相互作用会发生，对称从根本上提出了一个另外的法则。借助下面的类比，你就能很容易地领会这一现象。



假设你要布置6套餐具，而这6套餐具摆设必须都是一样的，也就是说，你的摆设要允许存在一种对称变换，可以将其中的任何两套餐具对调。如果没有这种对称，你可以给这个人两把叉子，另一个人3把，而另一个人又是一副筷

子；而有了对称的限制，你只能给所有6个人都摆设同样数量的刀叉、勺子和筷子——你不可能给这个人两把刀子，而给另一个人3把。

同样的道理，对称告诉我们并非所有的相互作用都会发生。即使有许多粒子能够相互作用，如果经典的相互作用保持对称，量子贡献通常也不会产生打破对称的相互作用。即使你包括了涉及虚粒子的所有可能的相互作用，但只要你开始没有打破对称，那就不会引起任何对称的破缺（只有第14章里提到过的极少见的反常现象例外）。在你的餐具摆放中如果必须保持对称，那么无论你怎么变换，加进水果勺也好，再加进切牛排的刀也好，那么你最后的摆放总归都是一样的。同样地，即使把量子力学效应算在内，也不会引发与对称不符的相互作用。如果在经典理论里对称没有打破，则没有任何路径让粒子产生破坏对称的相互作用。

不久以前，物理学家还一直以为对称是避开无政府主义原理的唯一方法。

在享用了足够的冰激凌之后，我和拉曼发现，隔开的膜又是一种方法。我当初之所以认为额外维度这么有希望，一个关键的原因就是，除了对称之外，它们又给出了另一个理由，这说明受限或异常的

作用也可能是自然的。将不想要的粒子隔离起来可以阻止不该有的相互作用发生，因为被隔离在不同膜上的粒子通常是不会发生相互作用的。

因为相互作用总是发生在当地——只有在同一地点的粒子才会直接相互影响，所以在不同膜上的粒子之间的相互影响不会很激烈。被隔离的粒子可以与其他膜上的粒子产生联系，但只能通过一个可以在两膜之间穿行的、在其间作用的粒子。就像在监狱膜里的艾克一样，不同膜上的粒子只能通过激发一个中间媒介，以有限的途径彼此之间取得联系。即使这种间接的相互作用能够发生，其影响也是极其微小的，因为体空间里的中介粒子，尤其是那些有质量的介子，根本不能穿越很远的距离。

隔离

隔离之所以重要是因为它能够阻止无政府主义原理引起的问题。将不想要的粒子隔离起来可以阻止不该有的相互作用发生，因为被隔离在不同膜上的粒子通常是不会发生相互作用的。

被隔离在不同地方的粒子之间的相互作用受到了制约，这就如同在一个外来信息受到封闭的国家里，政府小心地掌控着边境和媒体，我权且称之为“信息封锁国”。在国内，人们要获得外面的信息，只有通过想方设法进入该国的外国游客，或者是通过走私进来的报纸和书籍。

被隔离的膜以同样的方式给我们提供了一个避开无政府主义原理的平台，这样便增加了一套自然淘汰的工具，以保证不该有的相互作用不会发生。隔离方法的另一个优点是，它能保护粒子不受对称破缺的影响。只要对称破缺的发生距那些粒子足够远，它就几乎不会对它们产生影响。

当对称破缺被隔离起来时，就好似传染病人都被限制在一个规定的区域内一样，传染病源就被切断了；或者用另一个比方来说，如果没有一个从中干预的信息传播者，无论外部世界发生的事件有多么严重，对信息封锁国都不会产生任何影响。如果没有边境渗透者，信息封锁国就能够独立于外部世界自施其政。

隔离，超对称的破坏者

我和拉曼在1998年夏天研究的一个特别问题是，隔离怎样在自然界运作才能产生恰好具备我们观察到的性质的超对称破缺宇宙。我们发现，超对称能完美地保护等级问题，并保证对希格斯粒子质量的巨大量子力学贡献相加为零。但正如我们在第13章里看到的，即使自然界中存在超对称，可为了解释我们为什么只见到粒子而没有见到超对称伙伴，超对称必须被打破。

不幸的是，对称破缺的大多数模型都预言了不会在自然界发生的相互作用，这样的模型不可能正确。我和拉曼想找到一个物理学原理，大自然也许就是用它来杜绝不该有的相互作用，这样我们就能将其纳入一个更成功的理论中。

我们将精力集中于膜宇宙背景下的超对称破缺，膜宇宙能够保持超对称，但就如在四维世界里一样，当理论的一部分包含了不能维持超对称的粒子时，超对称就会产生自发破缺。我和拉曼意识到，如果引发超对称破缺的所有粒子都与标准模型粒子分隔开来，超对称破缺的模型就不会出现很严重的问题。

我们假设，标准模型的粒子被限制在一个膜上，而能引发超对称破缺的粒子被隔离在了另一个膜上。我们发现，在这样一种构成中，量子力学可能引发的危险相互作用就未必会出现。除了由在空间穿行的中介粒子所传递的超对称破缺效应外，标准模型粒子的相互作用与

在未破缺的超对称理论里是一样的。由此，就如在严格对称的理论中一样，与实验不符的不该有的味改变相互作用就不会发生。空间粒子既与超对称破缺膜上的粒子相互作用，又与标准模型膜上的粒子相互作用，它们精确决定了究竟哪些作用才可能发生——它们不一定包含那些不被允许的作用。

当然，必然有某个超对称破缺会被传递给标准模型粒子。如果超对称破缺没有被传递给它们，就没有东西能够使超对称伙伴的质量增大。虽然我们并不确切地知道超对称伙伴的质量值，但实验限制以及超对称在保持等级分化中所发挥的作用已告诉我们，它们的质量大概应是多少。

实验限制告诉我们超对称伙伴质量之间的定性联系，大致来说，所有的超对称伙伴的质量都应是相同的，这个质量大约应是弱力级质量——250 GeV。我们需要保证超对称伙伴的质量在这个范围之内，同时还要防止不应有的相互作用发生。所有片段都必须恰巧符合隔离的超对称破缺理论，这样，理论才有可能正确。

我们的模型成功的关键在于，寻找能将超对称破缺信息传递给标准模型粒子并给超对称伙伴以恰当质量的中介粒子。但我们还要确保这些中介粒子不会引致不可能的相互作用。

引力子

引力子既能与超对称破缺膜上的粒子发生相互作用，又能与标准模型膜上的粒子发生相互作用，而且引力子的作用是我们所熟知的——它们遵循引力理论。

引力子，无论高能粒子在哪都能与之发生相互作用的空间粒子，似乎就是理想的候选者。引力子既能与超对称破缺膜上的粒子发生相互作用，又能与标准模型膜上的粒子发生相互作用，而且引力子的作用是我们所熟知的——它们遵循引力理论。我们能够证明引力子

作用在产生超对称伙伴必要的质量时，不会引发可能导致夸克和轻子混淆其身份的相互作用——我们已知这种相互作用在自然界中不会发生。由此，引力子看起来是一个颇有希望的选择。

我和拉曼算出了超对称伙伴的质量，这要遵从信使引力子，我们发现，尽管组成要素很简单，但计算却复杂得令人吃惊。经典力学对超对称破缺质量的贡献竟然是零，只有量子力学效应才会传递超对称破缺。意识到这一点，我们将引力子引发的超对称破缺交流称作反常调解，选择这一名称是因为，正像我们在第14章里讨论的反常一样，特定的量子力学效应打破了本应保持的对称。重要的是，因为超对称伙伴的质量依赖的是已知的标准模型的量子力学效应，而不是未知的高维相互作用，所以我们能够预测超对称伙伴质量的相对大小。

我们花了好几天的时间才把它厘清，这意味着在同一天里，我可能一会儿失望一会儿欣喜。记得有天晚上吃饭时，我意识到了一个错误，因此，困扰了我一整天的问题终于迎刃而解了，我完全沉浸在了自己的思考中，把一起吃饭的同伴都吓坏了。最终，我和拉曼发现，如果引力能够传递超对称破缺，隔离的超对称破缺效果好得出奇。

所有的超对称伙伴都有了恰当的质量，规范微子和超夸克之间的质量关系也在我们希望的范围之内。尽管并非所有事情都如我们最初希望的那般简单，但超对称伙伴质量之间的重要联系都在恰当的范围，不会引发不可能的相互作用，这在其他的超对称破缺理论里可是一个大问题。只要再稍做修正，事情就将发挥作用。

而最值得欣慰的是，由于我们预言的超对称伙伴质量与众不同，我们的观点可以得到验证。隔离超对称破缺的一个重要特点是：

即使额外维度异常微小，大约 10^{-31} 厘米，只是普朗克长度的100倍，但它仍会产生可见的效果。这有点超乎我们的常规认识。从常识来看，只有大得多的维度才可能通过修正的引力定律或新的重粒子产生可见的结果。

尽管当额外维度非常小时，我们确实无法看到以上任何一种实验结果，但引力子以一种特殊的方式将超对称破缺传递给规范微子，我们可以由已知的引力作用和发生在超对称理论里的已知相互作用来计算。隔离超对称破缺模型预言了不同规范玻色子的超对称伙伴即规范微子的质量比，而这些质量是可以测量的。

这非常令人振奋，如果物理学家发现了超对称伙伴，他们就能确定它们之间的质量关系是否符合我们的预言。寻找这些规范微子超对称伙伴的实验正在加紧准备，要在位于伊利诺伊州费米实验室的质子-反质子对撞机——Tevatron中进行。如果幸运，我们将在几年的时间里看到结果。

最后，我和拉曼都有理由相信我们发现了有趣的东西。但我们仍存一丝疑虑：我有点儿担心，像这么有趣的观点，如果正确的话，早该有别人发现了。我们还需保证在我们的模型里没有留下任何隐藏的缺陷。拉曼也认为，这么一个好主意怎么可能被忽略呢？但他相信它是对的，只是担心我们可能忽略了物理学文献里任何相似的观点。

拉曼的担心不无道理，超对称破缺的反常调解在大约同一时间也由其他物理学家独立发现了，其中有CERN的吉安·朱迪切（Gian Giudice）、马里兰大学的马库斯·卢蒂（Markus Luty）、伯克利的村山齐（Hitoshi Murayama）、比萨的里卡多·拉塔兹（Riccardo Rattazzi）。那年夏天，他们也在一起工作。我们的论文发表一天后，他们的也发表了。他们的研究在我看来真的很奇妙，我不明白两组物理学家怎么会在同一个夏季经历同样艰辛的思路历程。

但拉曼的猜测是正确的：其他人可能会有同样的兴趣。事实上，我们都是正确的。尽管观点类似，但他们的发展却与额外维度的动机无关，而如果没有额外维度，反常调解的质量只能是空中楼阁。里卡多曾大度地对我们共同的朋友、物理学家马西莫·波拉提（Massimo P

orrati) 说，我和拉曼做得更好，并非因为我们的反常调解形式更为正确，而是因为我们本就有一个所有人都会关注的理由，这个理由就是额外维度。如果没有额外维度，超对称破缺不会被隔离，而反常调解的质量就会被更大的效应所淹没。

自此，物理学家继续研究超对称破缺的隔离模型。他们发现了把这个与其他更早的观点结合起来的方法，构建了可能代表现实世界的、更为成功的模型。有人甚至找到了将隔离观点延伸回四维时空的方法。

模型多得数不胜数，我只提两个我认为特别有趣的观点：第一个观点是由拉曼和马库斯·卢蒂合作提出的，他们用弯曲几何的观点（第20章的描述）重新阐述了在四维时空隔离的效果。用这些观点，他们创建了一类新的四维超对称破缺模型。

第二个有趣的观点是规范微子调解。在这一观点里，对称破缺的传递不是通过引力子，而是通过规范玻色子的超对称伙伴——规范微子。要使这一观点发挥作用，规范玻色子和它的伙伴不可能被困在膜上，它们必须能自由地穿越体空间。拉曼提醒我，规范微子调解实际就是我们以前忽略的众多观点之一。

而优秀的模型构建者戴维·卡普兰 (David E. Kaplan)、格兰姆·克里卜斯 (Graham Kribs)、马丁·施马尔茨 (Martin Schmaltz)，以及扎卡赖亚·查科 (Zacharia Chacko)、马库斯·卢蒂、安·纳尔逊 (Ann Nelson) 和爱德瓦多·庞顿 (Eduardo Ponton) 都证明我们很可能是操之过急，规范微子调解在传递超对称破缺质量时效果会更好，同时还保持了隔离超对称破缺的所有优势。^[57]

闪光质量

隔离对称破缺是模型构建的一个有力工具，现实世界可能包含分隔的膜，而通过此假设来构建模型，物理学家可以探索大量的可能。

上一节解释了味改变相互作用问题是如何在超对称理论里获得解决的。而模型构建者面临的另一个极具挑战的问题是：**为什么会有不同质量、不同味的夸克和轻子？**希格斯机制赋予了粒子质量，但每一味的质量值却是不同的。只有每味粒子与发挥希格斯粒子作用的东西发生不同的相互作用，这才有可能。由于每种粒子类型的三味，如上夸克、奇夸克、顶夸克都有着完全一致的规范相互作用，它们具有不同的质量就很神秘，一定有某种东西使得它们彼此质量不同，但粒子物理学的标准模型却没有告诉我们它是什么。

我们可以构建模型来解释不同的质量，但所有模型几乎总会含有改变味粒子身份不该有的相互作用。我们需要的是能够安全地将味粒子区别开来的东西，而不要产生这些有问题的相互作用。

尼玛·阿卡尼-哈麦德（Nima Arkani-Hamed）与德裔物理学家马丁·施马尔茨假定不同的标准模型粒子被局限在隔开的膜上，并认为它们能解释一些质量。尼玛与萨瓦斯·迪莫普洛斯（Savas Dimopoulos）发现了一个更简单的模型：他们假定存在一个标准模型粒子都被限于其上的膜，膜上粒子之间的相互作用将所有的味粒子同等对待。但是，如果只有这种将所有味一视同仁的味对称相互作用，粒子的质量就应该是完全相同的。很显然，只有能将粒子区别对待的东西才可能解释质量的不同。

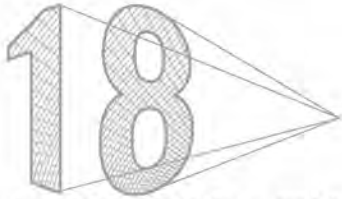
尼玛和萨瓦斯认为，引起味对称破缺的其他粒子被隔离在另外的膜上，与隔离超对称破缺的情形一样，只有通过空间的粒子，味对称破缺才能被传递给标准模型粒子。如果有许多体粒子与标准模型粒子相互作用，其中每个都会从不同距离由一个不同的膜传递味对称破缺，这样他们的模型就能解释标准模型里味粒子的不同质量。由远处

的膜传递的对称破缺，比由附近的膜传递的对称破缺引起的质量要小，尼玛和萨瓦斯将他们的观点取名为“闪光”以强调这一事实。就如当光源离得较远时光会显得暗淡一样，由遥远的膜引起的对称破缺效应会更小。在他们的图景里，夸克和轻子之所以不同味，是因为它们每个都在与不同距离的不同膜相互作用。

额外维度和隔离都是解答粒子物理学问题令人振奋的新奇方法，但它未必会止步于此。最近我们发现，甚至在研究宇宙演变的宇宙学中，隔离也发挥了重要作用。很显然，我们仍必须去发现包含隔离粒子的宇宙（或多重宇宙）的所有特征，而新的观点仍将继续涌现。

-
- 粒子可能被隔离在不同膜上。
 - 即使微小的额外维度也可能对可观察粒子的属性产生影响。
 - 被隔离的粒子并不一定会服从无政府主义原理，因为遥远的粒子不能直接相互作用，所以并非所有的相互作用都会发生。
 - 在一个引起超对称破缺的粒子与标准模型粒子被隔离开的模型里，超对称破缺可以不引起使粒子变味的相互作用。
 - 被隔离的超对称破缺是可以验证的。如果高能对撞机里能够生成规范微子，我们就可以对比规范微子的质量来确定它们是否与预言相符。
 - 被隔离的味对称破缺可能有助于解释粒子质量的差异。
-





泄露秘密的通道：四维世界的高维来客

WARPED PASSAGES

我窥视着，但它一直未曾出现。

我等你给我，最好的纪念。

我如此想念你，而你我却从未谋面。

比约克 (Bjork)

交换生平方K

阿西娜不得不承认，她非常想念艾克。虽然她常常觉得他很烦人，但没了他，阿西娜感到很孤独。这时，有一个交换生平方K要来住些日子，阿西娜已经在翘首以待了。可邻居们闭塞的观念令她吃惊，他们竟然害怕平方K的到来。尽管他与大家讲的是同一种语言，行为也不会有什么两样，可就现在这种气氛，单他来自国外这一事实就足以让人们警惕了。

阿西娜问她的邻居为什么这么担心，他们答道：“如果他把那些更加不安分的外国亲戚请来怎么办？他们会不会行为不端，只遵守自己国家的法律？如果他们都来了，谁知道会发生些什么呢？”

阿西娜赶忙告诉他们，平方K和他的亲戚不会待很长时间的，因为他们都不太喜欢安静，只有热闹聚会上那种热烈喧嚣的气氛才有可能把他的家人吸引来。可不幸的是，阿西娜的这番解释让邻居们的疑虑更重了。意识到自己措辞不当后，阿西娜赶忙改口说，外国客人在其短暂的来访期间，一定会遵守当地法律的。邻居们信了，并和阿西娜一起去欢迎平方K的到来。

本书开始时，我解释了额外维度是如何隐藏起来的：它们可能卷曲成很小的形状或可能被膜束缚起来，让人难以察觉。但额外维度真的有可能将其性质完全隐藏，以至于没有任何物质特征能将它与四维世界区分开来吗？这让人很难相信。即便卷曲的维度非常小，令我们相信世界是四维的，但高维世界一定包含着一些新的元素，能让我们将它与真正的四维世界区别开来。

如果额外维度存在，它的痕迹也一定存在。这些痕迹就是卡鲁扎-克莱因（KK）粒子^[58]，KK粒子是额外维度宇宙的附加成分，它们是高维世界的四维印记。

如果KK粒子存在且足够轻，那么高能对撞机就能生成它们，而它们会在实验数据里留下印记，高维侦探——实验者们，就能将这些线索拼凑起来，把数据转换成高维世界的确凿证据。本章将讲述卡鲁扎-克莱因粒子，以及我们相信它们存在于高维世界中的原因。

神秘的卡鲁扎-克莱因粒子，额外维度的印记

即使体粒子在高维空间穿行，我们仍可以用四维语言来描述它们的属性和相互作用。毕竟，我们不曾直接见过额外维度，因此所有东西在我们看来都好像是四维的。就如平面国的居民一样，他们只看到的到两维，当一个三维球体穿过他们的世界时，他们看到的就只是两维的圆形碟片；而在我们看来，虽然那些粒子来自高维世界，看上去却好像只在三个维度里穿行。这些来自额外维度的新粒子，在我们看来是四维时空^[59]的额外粒子，就是卡鲁扎-克莱因粒子。如果我们能测量和研究它们的所有性质，那么它们就能揭示有关高维空间所有未知的秘密。

卡鲁扎-克莱因粒子

来自额外维度的新粒子，在我们看来是四维时空的额外粒子。卡鲁扎-克莱因粒子是高维粒子在四维的表现形式，正如通过将许多共振模式重叠重现一根琴弦发出的所有声音一样，选择恰当的KK粒子来代替，也可能重现一个高维粒子的行为。

卡鲁扎-克莱因粒子是高维粒子在四维的表现形式，正如通过将许多共振模式重叠重现一根琴弦发出的所有声音一样，选择恰当的KK粒子来代替，也可能重现一个高维粒子的行为。KK粒子完全能代表高维粒子及其穿行的高维几何的特征。

为了模拟高维粒子的行为，KK粒子将必须携带高维动量。每个在高维空间穿行的体粒子都将被KK粒子的四维有效描述所代替，这些KK粒子有着模拟这个高维粒子的正确动量和相互作用。一个高维宇宙既包含我们熟悉的粒子，也包含它们相对应的携带由卷曲空间的具体性质所决定的额外维动量的KK粒子。

但四维描述不包括有关高维位置或动量的信息，因此当我们由四维视角来看时，KK粒子的高维动量必须有一个另外的名称。狭义相对论所规定的质量和动量的关系告诉我们，额外维度的动量在四维世界里就可以被看作质量。因此，KK粒子与我们熟知的粒子一样，只是其质量反映了它们的高维动量。

KK粒子的质量由高维几何所决定，但它们的电荷与已知的四维粒子是相同的。这是因为，如果已知粒子源于高维时空，那么高维粒子就必须与已知粒子携带同样的电荷，这对模拟高维粒子行为的KK粒子也一样。因此，对应我们熟悉的每个粒子，应该会有许多电荷相同的KK粒子，但质量却各不相同。

例如，如果一个电子在高维空间穿行，那么它就会有一个携带同样负电荷的KK伙伴；如果一个夸克在高维空间穿行，那么它也会有一个对应的KK粒子，与夸克一样经受强力。KK伙伴与我们熟知的粒子有相同的电荷，但质量要取决于额外维度。

确定卡鲁扎-克莱因粒子质量

为理解KK粒子的起源和质量，我们不应局限于早先考察过的不可见卷曲维度的直观感觉。为简便起见，我们先来看一个没有膜的宇宙，其中每个粒子从根本上讲都是高维的，而且能自由地在所有方向运动——包括那些额外的方向。具体来说，让我们设想一个只有一个卷曲成圆圈的额外维度的空间以及在这个空间里穿行的基本粒子。

倘使我们生活在牛顿经典力学主宰的世界，那么卡鲁扎-克莱因粒子就可能具有任何值的额外维度动量，由此，质量也会是任意的。可是，我们生活在一个量子力学的宇宙里，这样情况就不同了。量子力学告诉我们：正如只有小提琴的共振模式才能形成琴弦发出的声音一样，KK粒子在重现高维粒子的运动和相互作用时，只有量子化的额外维度动量才有贡献；也正如琴弦的音符依赖于琴弦的长度，KK粒子的量子化额外维度动量也取决于额外维度的大小和形状。

KK粒子携带的额外维度动量，在我们的四维世界看来，就是KK粒子与众不同的质量模式。如果物理学家发现KK粒子，这些质量就会为我们揭示额外维度的几何形状，例如，如果只有一个卷曲成圆圈的额外维度，这些质量就能告诉我们额外维度的大小。

寻找KK粒子在一个卷曲维度的宇宙里的容许动量（从而质量）的做法，就像你从数学上来确定小提琴的共振模式的方法，也像玻尔确定原子的量子化电子轨道所使用的方法。量子力学将所有粒子都与波相联系，只有那些能以整数倍在一个额外维度圆圈上振动的波才是被允许的。我们确定了允许的波，然后用量子力学把波长和动量联系起来，额外维度的动量就告诉我们KK粒子被允许的质量，这正是我们要知道的。

恒定的波（这个波根本不会振动）总是被允许的，这个“波”就像是一个完全静止的、不起涟漪的池塘表面，或者像一根未经拨动的琴弦。这个概率波在额外维度的任何地方都是同样的值，因为一个平坦概率波的值是恒定的，所以与这个波相关的KK粒子不会偏好额外维度的任何位置。根据量子力学，这个粒子不携带额外维度动量；因而由狭义相对论可知，它也没有额外的质量。

因此，最轻的KK粒子就是这类在额外维度里有恒定概率值的粒子。在低能量上，这是唯一能生成的KK粒子。因为它既没有额外维度的动量也没有额外维度的结构，因此很难与一个有着相同质量、相同电荷的常见四维粒子区分开来。在只有低能量的情况下，一个高维粒子根本不可能在收紧的卷曲维度里运动，换句话说，低能量不能生成可以区分我们的宇宙与高维宇宙的额外KK粒子。因此，低能过程和最轻的KK粒子不能揭示有关额外维度存在的任何迹象，更不用说它的大小和形状了。

但是，如果宇宙包含额外维度，而粒子加速器又达到了足够高的能量，那么它就能生成更重的KK粒子，这些携带了非零额外维度动量的重KK粒子，将是额外维度存在的第一个铁证。在我们的例子里，那些与重KK粒子关联的波有着沿额外维度的圆形结构，这些波缠绕着卷曲的维度，沿着圆周整数次地上下振动。

最轻的KK粒子就是那些概率函数波长最大的粒子。一个波围绕卷曲维度上下振动一次恰好围成一个圆，这就是最大的波长，这个波长是由额外维度的周长决定的（它们大约是同等大小），波长再大则不适合：当它沿着圆圈回到那个点时，波与圆则不再重合。有着这种概率波的粒子就是能“记住”其额外维度起点的最轻的KK粒子。

这个有着非零额外维度动量的最轻粒子，其相关波的波长与额外维度大约同样大小，毕竟，只有小到足以探索极小尺度和相互作用的

东西才会灵敏地探测到卷曲维度的存在。用一个较长的波长来探索额外维度，就好比要用一把尺子去测量一个原子的位置。例如，如果你用光或是其他有特定波长的探测工具来探测一个额外维度，那么光的波长必须短于额外维度的尺寸。因为量子力学将概率波与粒子联系起来，上述有关探测波长的陈述也是粒子性质的陈述。只有波长足够短（根据不确定性原理）而额外维度动量和质量足够高的粒子，才有可能灵敏地探测到额外维度的存在。

有着非零额外维度动量的最轻KK粒子，还有另一个吸引人的特征：当额外维度较大时，它的动量（从而质量）会更小。因为较轻的粒子更容易生成和发现，因此，较大的额外维度应该更容易探测，并且有更多较容易探测的结果。

如果确实存在额外维度，那么它们的证据将不仅仅是最轻的KK粒子，其他一些高动量的粒子，会在粒子对撞机里留下更鲜明的额外维度印记。当穿越卷曲的额外维度时，这些粒子的概率波不止振动一次，因为第 n 个这样的粒子对应的是绕卷曲维度上下振动 n 次的波，这些KK粒子的质量都将是最轻粒子质量的整数倍。动量越高，KK粒子在对撞机里留下的印记则越鲜明。图18-1简单说明了KK粒子的质量值，以及对应这些有质量粒子的两个波。粒子质量与额外维度的大小成反比。

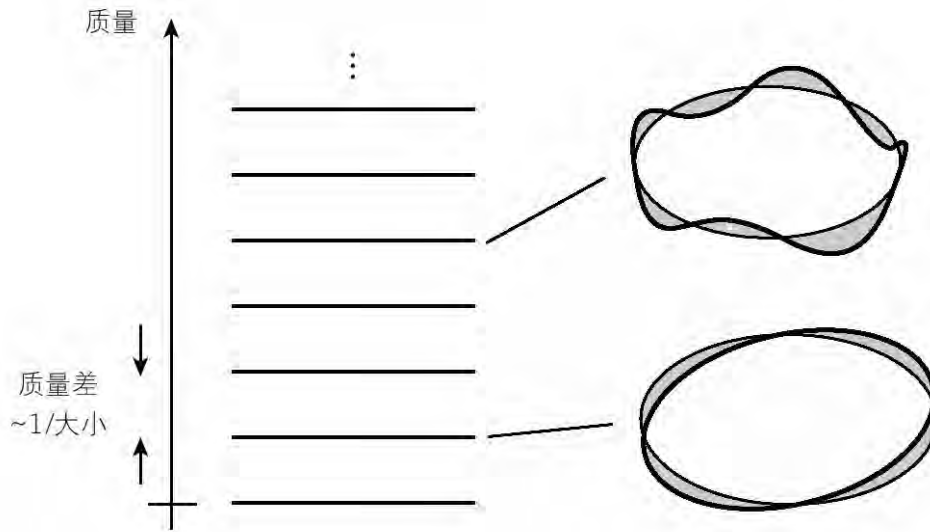


图18-1 KK粒子质量及其对应的波。卡鲁扎-克莱因粒子对应的波围绕卷曲的维度整数次地上下振动，波的振动次数越多对应的粒子则越重。



依次加重的许多粒子就好像一个移民家族的好多代：在美国出生的最年轻一代完全被美国文化所同化，他们讲着完美的英文，一点儿听不出外国口音，也就分辨不出他们的从哪儿来；而上一代人，他们的父母情况就不是这样了：也许他们会带一点儿口音，偶尔还会讲几句他们祖国的谚语；再上一代人的口音则更重，甚至他们的穿戴以及所讲的故事都是从祖国带来的。我们可以说，正是这些早几代的移民拓宽了我们的文化维度，使我们的社会不那么单一，变得更加丰富多彩。

同样地，最轻的KK粒子与一个本质是四维世界里的粒子是没有区别的，只有那些质量更大的“老亲戚”才能揭示额外维度的证据。尽管最轻的KK粒子看上去是四维的，但是一旦达到足够的能量，生成质量更大的“老一代”，它们的证据便会是显而易见的。

如果实验者发现新的与我们熟悉的粒子有相同电荷的重粒子，其质量也彼此相似，那么这些粒子将成为额外维度强有力的证据。如果这些粒子有相同的电荷，而且间隔一个固定的质量差，这很可能意味着我们发现了一个简单的弯曲维度。

但更为复杂的额外维度几何会生成更为复杂的质量类型。如果我们找到足够多的粒子，则KK粒子不仅能揭示额外维度的存在，还将告诉我们额外维度的大小和形状。无论隐藏的维度是什么形状，KK粒子的质量都会依赖于它。无论何种情形，KK粒子和它们的质量都将揭示出关于额外维度性质的许多东西。

实验限制

一直以来，大多数弦理论学家认为：额外维度不会比微小的普朗克长度大太多。这是因为，在普朗克级能量上引力会变得强大，这就需要量子引力理论来接管这一领域，而它可能是弦理论。但普朗克长度远小于我们实验所能探索的尺度，这个微小的普朗克长度对应的（据量子力学和狭义相对论）就是巨大的普朗克质量（或能量）——它是目前粒子加速器所能达到能量的1亿亿倍。普朗克质量的KK粒子如此之重，将远远超出我们能够想象的实验探测范围。

但是，额外维度也许会更大，而KK粒子会更轻。为什么我们不反过来问，有关额外维度的大小，实验都告诉了我们些什么？排除理论偏见，我们又实际了解到些什么？

如果世界是高维的，而且没有膜，那么我们熟悉的所有粒子，如电子，都会有它们的KK伙伴。这些粒子将与我们熟悉的粒子有恰好相同的电荷，携带着额外维度的动量。电子的KK伙伴将与电子一样带负电荷，只是更重。如果一个额外维度卷曲成一个圆，那么这种最轻粒子的质量和电子质量的差额与额外维度的大小便成反比。这就意味着，额外维度越大，粒子质量越小。因为一个更大的维度会形成一个更轻的KK粒子，而实验中还没有发现这样的粒子，所以KK粒子质量的限度就规定了卷曲额外维度被允许的大小。

迄今为止，在运行能量达到约1 000 GeV的对撞机里，还没有出现这种带电粒子的任何迹象。因为KK粒子就是额外维度的印记，我们没有找到它们，这就说明额外维度不可能太大。目前的实验限制告诉我们，额外维度不可能超过 10^{-17} 厘米^[60]，这个尺度是极其微小的——远小于我们能直接看到的任何东西。

额外维度大小的这一限定比弱标度长度大约小10倍，虽然 10^{-17} 厘米很小，但比普朗克长度 10^{-33} 厘米大16个数量级。这意味着额外维度可以比普朗克长度大得多，但仍未被发现。希腊物理学家伊格内修斯·安东尼亚迪斯（Ignatius Antoniadis）最先想到了额外维度可能不是普朗克长度，而是与弱标度长度差不多大小。他一直在思索，把对撞机的能量再提高一点儿会有什么新的发现。毕竟，等级问题告诉我们，在这些能量上肯定会发现某些东西，这是具有弱力级能量和质量的粒子生成的地方。

但即使是以上对额外维度大小的限定，也并不一定总是适用。虽然KK粒子是额外维度的印记，但它们难以捉摸，很难被发现。最近，

我们对KK粒子及其表现的了解又增加了许多，后面的几章就将解释最新的成果：虽然我们认为更大的维度会形成更轻的KK粒子，可为什么一旦将膜考虑在内，额外维度就可能大于 10^{-17} 厘米而仍不被发现？有些模型包括很大的维度（你可能以为这些维度会形成可见的结果），虽然仍是看不见的，但能帮助解释标准模型粒子的神秘属性。在第22章里，我们将看到一个更加出人意料的结果：一个无穷大的维度可以形成无穷多的轻KK粒子，却仍没有留下任何可观察的痕迹。

-
- 卡鲁扎-克莱因（KK）模式是携带额外维度动量的粒子，它们是四维世界的高维来客。
 - KK粒子看上去就与重粒子一样，与已知粒子有着相同的电荷。
 - KK粒子的质量与相互作用要由高维理论来确定，因此它们反映了高维时空的属性。
 - 如果我们能发现和测量所有KK粒子的属性，我们就能知道高维度的大小和形状。
 - 目前的实验限制告诉我们，如果所有粒子都能穿行于高维空间，那么额外维度不可能大于 10^{-17} 厘米。



-imension
探索大揭秘



当它落下的时候，我连一毫米都看不到。

埃米纳姆 (Eminem)

神秘的线索

短暂的来访结束后，平方K离开了，阿西娜只好将大部分时间消磨在附近的网吧里。最近，她发现了一些神秘的网站，这令她非常兴奋。最吸引她的一个是：xxx.socloseandyetsofar.al（天涯若比邻），由名字来看，阿西娜怀疑这个网站是最近的AOB（膜上的美国）与时空预警多媒体合并的结果，但她必须先回家才有时间进行详细研究。

阿西娜一回家就直奔电脑，再次寻找那个奇异的超级链接，在网吧里这很容易找到，可是在家里，她的网络安全卫士却阻止她访问这个层层设防的被禁网站^[61]，于是她换了一个别名“经验丰富的顾问”（Mentor），以掩盖了身份，顺利躲过网管，成功找到了那个神秘的超级链接。

阿西娜总希望平方K发送给她的短讯息就隐藏在网页里，可这些站点很不好理解。她好不容易才找到几个可能有意义的信号，所以决定再深入地研究它们的内容，但愿这合并（与其他有着类似名称的合并不同）能够维持足够长的时间让她弄清楚。

1998年在牛津召开的超对称年会上，斯坦福大学的物理学家萨瓦斯·迪莫普洛斯发表了一个最为有趣的演讲。他讲述了与其他两位物

理学家尼玛·阿卡尼-哈麦德和吉亚·德瓦利（Gia Dvali）共同协作所取得的成果。这个三人组合多样的个性及其观点与他们丰富多彩的姓名可谓相映成趣：萨瓦斯对自己的项目充满激情，他的同事告诉我，他的热情会感染周围所有的人。额外维度令他格外投入，他曾对一个同事说，那些尚未被发现的新的物理学观点让他感觉就像是一个孩子走进了糖果店——不等别人把糖果买走，他就想把它们统统吃掉。吉亚，一个来自格鲁吉亚的物理学家，无论在其物理方法还是大胆的登山运动的壮举中，都敢于冒险。

有一次，他在暴风雪中，被困在高加索的一个山顶上，两天两夜没有吃任何东西；尼玛，一个来自伊朗家庭的物理学家，他充满活力，说起话来绘声绘色，极富煽动性，现在我们是哈佛大学的同事，他常常在走廊里热情地宣讲他所取得的最新进展，鼓动其他人加入他的研究。

萨瓦斯在超对称会议上的发言是关于额外维度的，与超对称毫无关系，却抢了超对称的风头，这颇具讽刺意味。他解释道：

标准模型的基础理论可能是额外维度，而非超对称。如果他的假说成立，不久的将来，实验者们在探索弱力标度时，将有望发现额外维度而不是超对称的证据。

本章呈现的是尼玛、萨瓦斯和吉亚^[62]的观点，这一观点讲述了大的维度将如何解释引力的微弱。从根本上来讲，大的额外维度能分散引力作用，以致引力强度会比我们在没有额外维度时估计的弱很多。他们的模型并不能实际解决等级问题，因为仍必须解释额外维度为什么会这么大。但是ADD认为借助他们的模型，这一新问题会更容易解决。

我们还将探讨ADD所提出的一个相关问题：如果标准模型粒子被限制在一个膜上而不能自由在空间穿行，那么卷曲的额外维度得多大才不会与实验结果相冲突。他们找到的答案是超乎寻常的，在他们那时的论文中，这些额外维度可能是1毫米。

大至1毫米的维度

在ADD的模型里，正如我在第17章里描述的隔离模型一样，标准模型粒子被限制在一张膜上。但是两个模型的目的不同，因此它们的其他特征也就完全不同。隔离模型有一个介于两个膜之间的额外维度，而ADD的所有模型都不止一个维度，且这些维度都是卷曲的。根据不同的细节体现，他们模型里的空间包含有二维、三维或更多的卷曲维度，而且ADD模型只包含了一个膜，标准模型粒子被限制在这个膜上，但膜并不是空间的边界（见图19-1），它只是静止在额外维度里。

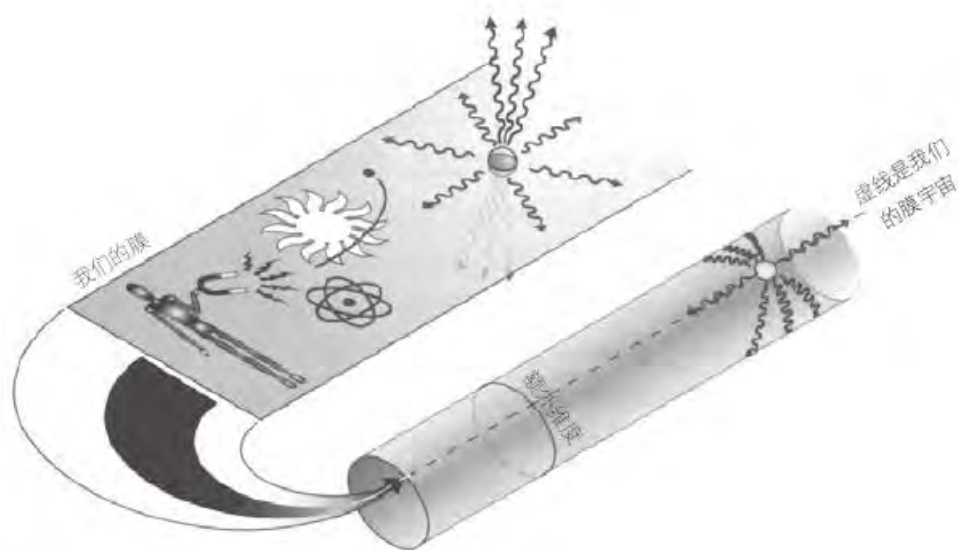


图19-1 ADD膜宇宙简要图示。宇宙的额外维度是卷曲的（且较大）。我们生活在膜上（沿着圆筒的虚线），因此只有引力能感受到额外维度。

如果标准模型的所有粒子都被困在一个膜上，而引力是高维体空间唯一的力，那么额外维度究竟有多大才能隐藏起来？这是ADD想以他们的模型解决的一个问题，而他们发现的答案令许多物理学家都非常吃惊。

上一章我们猜想的额外维度大小是十亿亿分之一厘米，与此形成鲜明对比的是，他们的卷曲的额外维度可能要大至1毫米（现在来说这一确切数字实际有点误导，因为，正如本章后面我们将进一步讨论的，华盛顿大学的物理学家在那以后曾通过实验寻找毫米大小的额外维度，但没有发现。基于他们的结果，现在我们知道，额外维度的大小一定得小于1/10毫米，不然它们就会被排除。可即便1/10毫米大小的额外维度仍是很令人吃惊的）。

你可能会认为，如果有大至1毫米的额外维度（即便再小10倍），我们肯定早就发现了。虽说一个人看不到1毫米大小的物体就该换新眼镜了，可在粒子物理学的尺度上，1毫米可是极为庞大的。

为了弄清1毫米（或者即便是1/10毫米）大小的额外维度有多么超乎寻常，让我们来重温一下我们迄今讨论过的各个长度级别：普朗克长度， 10^{-33} 厘米，远超出所有实验能探测的限度；TeV长度，大约是 10^{-17} 厘米，是我们当前实验所探索的范围，物理学家已在小至 10^{-17} 厘米的距离内探测到电磁力。相比而言，ADD所谈到的尺度几乎是巨大的。如果没有膜的出现，毫米大小的额外维度早就被当作毫无意义的谬论排除掉了。

但是，即便额外维度再大许多，膜使得它们仍旧可以理解。膜能够困住夸克、轻子和规范玻色子，这样，只有引力能够在整个高维空间穿行。ADD的图景假定，除引力之外的所有东西都被限制在膜上；而不涉及引力的东西，即使额外维度极大，看上去也像没有额外维度一样。

例如，你看到的所有东西都像是四维的。你的眼睛能够探察到光子，而在ADD模型里，光子被束缚在膜上，因此你能看到的所有物体看上去都像是只有三个空间维度。如果光子被束缚在膜上，那么无论你的眼镜有多么清楚，你都不能直接看到额外维度的任何证据。

事实上，只有通过极为灵敏的引力探测手段，你才有希望看到ADD图景里毫米级维度的证据。所有常见的粒子物理过程中，如由电磁力传递的相互作用、电子-正电子对的生成、原子核通过强力的束缚，都出现在四维的膜上，因此也就与一个纯粹的四维宇宙没有任何差别。

带荷的KK粒子也不是什么问题。上一章我们解释过，当所有的粒子都在体空间里时，额外维度不会很大，假如它们很大，我们早就该发现标准模型粒子的KK伙伴了。但在ADD图景里，情况却不是这样。因为所有的标准模型粒子，如电子，都被束缚在一个膜上，因此标准模型粒子不会在高维空间穿行，也就不会携带高维动量，从而被局限在膜上的标准模型粒子也就不会有KK伙伴。而如果没有KK伙伴，那么上一章我们讨论过的基于KK粒子的约束也就将不再适用。

在ADD模型里，只有引力子才是唯一必须有KK伙伴的粒子，而引力子是能在高维体空间中穿行的，但引力子KK伙伴的相互作用比起标准模型粒子KK伙伴的要微弱许多。标准模型里的KK伙伴通过电磁力、弱力、强力相互作用，而引力子的KK伙伴则只以引力的强度相互作用——与引力子本身一样微弱。因此，引力子的KK伙伴比标准模型粒子的KK伙伴更难以生成和探测。毕竟，还没有人直接发现过引力子，而如引力子作用一样微弱的KK伙伴，就更难找到了。

ADD意识到，如果额外维度的约束只来自引力，那么在他们的标准模型粒子被困在一个膜上的图景里，额外维度的大小将比上一章建议的大得多。原因在于，引力非常微弱，因此很难用实验探测。对距离

极近的小质量物体而言，如果引力过于微弱，其效应很容易被其他的力所淹没。

例如，两个电子间的引力是电磁力的 $1/10^{43}$ 。地球引力之所以占主导地位，只是由于它的净电荷为零。在小尺度上，重要的不仅仅是净电荷，电荷的分布方式同样重要。为了验证在两个微小物体之间的万有引力定律，引力作用必须从其他力哪怕最微小的力的影响中隔离出来。尽管围绕太阳运转的行星、围绕地球运转的月亮以及宇宙本身的演变显示了引力在大尺度上的形式，但小距离上的引力却难以探测。因此，如果引力是体空间里唯一的力，那么令人惊讶的大额外维度就不会与实验结果有任何冲突。如果在一些维度上粒子被膜束缚，这些维度将很难被发现。

1996年，ADD在撰写他们的论文时，牛顿的平方反比定律已在大约1毫米的距离上得到了验证。这意味着，额外维度可能大至1毫米，而仍没人发现它们存在的证据。正如ADD的论文所说：“我们将MPL（普朗克能量）理解为基本能标（在此，引力作用变得强大起来），当时是基于这样的假设，引力在从它测量的距离……直到普朗克长度 10^{-33} 厘米之间的这33个数量级内，是不需要修正的。”换句话说，1998年，由实验我们无从得知引力在小于1毫米距离上的任何信息。如果相隔距离小于1毫米，引力定律有可能发生变化。比如，随着两个物体渐渐靠近，引力吸引会更加迅速地加强——只是，这至今尚未被人所知。

高维引力，或将成等级问题的解

可能存在大额外维度是一个重要发现。但是，ADD研究大额外维度不仅仅是为了探索抽象的可能性，他们真正的兴趣在于粒子物理学，尤其是等级问题。

正如我在第12章解释过的，等级问题是弱力级质量和普朗克级质量之间的庞大比值，这是与粒子物理和引力相联的两个质量。一直以来，让物理学家感到迷惑的主要问题是，弱力级质量为什么会如此微小？对希格斯粒子质量的庞大虚贡献（达到了普朗克级质量）本应该让它变得很大的。在额外维度进入物理学家的思考之前，所有试图解决等级问题的尝试都围绕着改善标准模型展开，希望找到一个更为全面的粒子物理学的基本理论，以解释为什么弱力级质量相比普朗克级质量要小这么多。

可等级问题涉及的是两个数字之间的庞大差异，问题的症结是普朗克标度与弱力标度为什么相差这么大。因此，等级问题还可以换成另外一种说法：为什么普朗克级质量标度这么大，而弱力级质量标度这么小——这也等于说，基本粒子之间的引力强度为什么这么小？由此，等级问题引出一个思考：是不是引力而非粒子物理学，与物理学家原本设想的有所不同？

顺着这一思路推理下来，ADD提出：试图通过延伸标准模型来解决等级问题实际是误入歧途。依据他们的发现，如果额外维度足够大，同样能够很好地解决问题。他们指出，决定引力强度的基本质量尺度不是普朗克级质量，而是一个小许多的质量标度，接近于1 TeV。

然而，ADD仍留下引力为什么会这么微弱的问题。毕竟，普朗克级质量如此之大的原因是引力非常微弱——引力强度与这一标度成反比。更小的引力基本质量尺度会使引力作用的强度更大。这一问题并非不可逾越，ADD指出，只有高维引力才必须是强烈的。

ADD认为，大额外维度会以某种方式削弱引力强度，使得引力即使在高维很强，但在低维有效理论里会变得很微弱。在他们的研究结果里，引力在我们看来之所以这么微弱，是因为它在一个很大的额外维度空间里分散了。而电磁力、强力、弱力仍旧很强，是因为这些

力被局限在膜上，根本不会被分散。因此，大维度和一个膜就能解释为什么引力要比其他力微弱许多。

尼玛告诉我，当他和合作伙伴弄明白了高维引力和低维引力强度之间的确切关系时，他们的研究出现了转折。这一关系并非新的发现，例如，弦理论学家一直用它把四维引力标度与十维引力标度联系起来。正如我在第16章里简要介绍的，霍扎瓦和威滕在发现引力可与其他力统一时，就用过十维和十一维引力强度之间的关系：一个大的第十一维允许高维引力标度（从而弦的标度）与大统一标度一样低。但在此之前，从没有人意识到只要额外维度大到足以恰当地削弱引力，高维引力就可能足够强大以解决等级问题。尼玛、萨瓦斯和吉亚在研究额外维度一段时间后，知道了该怎样将高维和低维引力联系起来，也明白了其中的非凡意义。

联系高维引力和低维引力

我们在第2章里看到，当我们只探索大于卷曲额外维度大小的距离时，额外维度是不易被察觉的。但是，这并不一定意味着额外维度不会留下任何印记，即使我们看不到它们，它们也仍能影响我们能看到的物理量的值。第17章就给出了这一现象的一个例子，在超对称破缺的隔离模型里，超对称破缺发生在远处的一个膜上，而引力子将破缺传递给标准模型粒子的超对称伙伴，超对称伙伴的质量值反映了超对称破缺的额外维度来源及其通过引力的传递。

现在我们来查看另外一个额外维度影响可测量值的例子。卷曲维度的大小决定了四维引力强度（也就是我们观察到的引力）与它所由来的高维引力强度之间的关系。当卷曲的额外维度包围更大的体积时，其间的引力就会被削弱。



要探明其中的究竟，让我们回到第2章中的例子，把三维花园橡胶管宇宙比作由膜包裹的三维体空间。如果水通过一个小孔进入水管（见图2-9），它首先会从小孔里喷出来，向三个维度喷涌。但是，一旦水充满整个管子，它就只会沿着水管流动——此时，橡胶管就像只有一维。当我们在大于额外维度的范围内测量引力定律时，也是同样的情形。

但是，即使水只会沿着橡胶管的一个维度流动，其压力还是要取决于水管横截面的大小。要理解这一点，让我们设想如果水管再增粗一些会出现什么情况：通过小孔进入水管的水会向更大的空间喷涌，那么水冲击橡胶管壁的压力就会变弱。

如果用水的压力代表引力线，而通过小孔进入水管的水代表由一个有质量物体发出的引力场线，那么，这些引力线就像前面例子中的水一样：它们首先会向三个方向发散，然后，当引力线碰到宇宙的壁（膜）时，它们就会弯曲，然后只沿一个大的维度流动。在橡胶管的例子里，我们发现水管越粗，水的压力就越小。同样的道理，在我们以花园橡胶管模拟的宇宙里，额外维度的面积

决定了引力线在一个低维世界里是如何分散的。额外维度的面积越大，低维宇宙有效理论的引力场强度就越弱。

这一原理也适用于有着任意数量卷曲维度的宇宙中的卷曲维度。额外维度的体积越大，引力线越稀疏，引力强度也越弱。这一点可以用刚才假设的水管高维类比来说明。高维橡胶管里的引力线会首先向所有维度发散，包括卷曲的额外维度，引力线在到达卷曲维度边界之后，就只会沿着低维空间的无限大维度延伸。最初在额外维度里的发散会降低低维空间引力线的密度，因而，在那里感受到的引力强度就会变弱。

回到等级问题

由于引力线在额外维度里变得稀少了，卷曲的额外维度空间体积越大，低维引力则越弱。ADD发现，引力在额外维度的分散可以很广，这就能解释我们所处世界的四维引力为何如此微弱。他们的理由如下：

假设高维理论的引力不依赖于庞大的 10^{-19} GeV的普朗克级质量，而是取决于一个更小的、小16个数量级的1 TeV的能量。他们选择1 TeV来消除等级问题：如果引力在1 TeV能量附近变强，那么粒子物理学中就不会出现质量等级问题，粒子物理和引力的所有东西都将由TeV标度来表征。因此，在他们的模型里，维持一个质量约为1 TeV的合理的轻希格斯粒子就不成问题了。

根据他们的假设，在大约1 TeV的能量上，高维引力与其他已知力相比，将成为一个很强的力。因而，为得到符合我们观察的合理理论，ADD还需要解释四维引力为什么那么弱。他们模型的另一个元素就是假设额外维度极大，我们最终想解释的就是这个大的尺度。但根据他们的提议，卷曲维度包围的是一个庞大的体积，按照上节的逻辑来

看，四维引力就会变得极其微弱。也就是说，我们所处世界的引力这么微弱，是因为额外维度很大，而不是因为有一个大质量引起了微小的引力。我们在四维里测得的普朗克级质量是庞大的（使得引力看起来很弱），只不过是因为引力被庞大的额外维度削弱了。

这些维度必须得有多大？答案要取决于额外维度的数量。因为实验还没能确定存在几个额外维度，ADD模型里额外维度的数量也有多种可能。注意，此刻我们关心的只是大的维度，因此，如果你和你周围的弦理论学家都知道空间维度的数量有9个或10个，你仍可以考虑大维度的数量还有其他可能，并假设所有其他维度都小得可以忽略不计。

在ADD假说里，维度有多大要看它们有多少，因为其体积要取决于它们的数量。如果所有维度都同样大小，那么一个高维区域就会比一个低维区域包围的体积更大，因此对引力的削弱会更多。为了看清这一点，让我们再回到第2章中洒水装置的比喻里，我们有两个洒水装置：其中一个，水只会沿着一条特定长度的线段（一维）喷洒；另一个则以相同的长度为半径洒向一个圆（二维）。你可以看到，一棵植物由一维洒水装置得到的水量肯定比从二维洒水装置中得到的多。水喷洒的维度越多，被分散得便越厉害。

如果只有一个大的额外维度，要满足ADD的假设，它就必须很大。为了能足够地削弱引力，它的尺度必须大至如从地球到太阳的距离，但这是不被允许的：如果额外维度那么大，宇宙在可测量距离的行为将是五维的。我们已经知道，牛顿万有引力定律是适用于这些距离的，如果一个大的额外维度在这样的距离修正引力，显然是可以排除的。

但是，如果只有两个额外维度，那么它们的尺度可以很小，即使小至1毫米，也仍能充分地削弱引力，因此，ADD才会对1毫米的尺度倾注如此大的热情。当然，这不仅因为它就在实验探测的范围之内，还

因为这个尺度的两个额外维度对等级问题很重要：引力会在毫米大小的两个额外维度里发散，并形成我们熟知的微弱引力作用。当然，1毫米仍旧很大，但正如我们早先说过的，引力检验远没有你想象的那么严格。受到ADD图景的启发，人们更加努力地去思考和寻找这个尺度的额外维度。

如果额外维度超过两个，那么引力只会在极小的距离上修正。如果有更多的额外维度，即使它们相对较小，引力也足以被削弱。例如，如果有6个额外维度，它们的大小只需要 10^{-13} 厘米，即十万亿分之一厘米。

如果幸运，即使是这么小的额外维度，我们在不久的将来也能找到某个粒子的证据——不是在我们下节讨论的引力检验里，而是在后面将讨论的高能粒子对撞机实验中。

寻找大的维度

我们该怎样探索引力在小距离的变化？我们该寻找什么？我们知道，如果存在卷曲维度，那么引力强度在小于额外维度大小的距离上会随着距离增大而降低，其降低速度大于牛顿定律的预言，因为引力不止在3个维度里发散。当物体的间距小于额外维度时，适用的则是高维引力。如果一只小虫足够小，能够沿着一个卷曲维度绕圈，那么它就能够体验到额外维度，不仅因为它能在其中游历，还因为引力会在其中向所有维度发散。因此，如果有人能像这只有着异常感知力的小虫一样，探测到如此小距离的引力作用，额外维度就能产生可见影响了。

这告诉我们，如果实验能探测在相当于（或小于）假设的卷曲维度大小距离上的引力，及引力强度在这个距离上与两物体间距的依赖关系，那么它就能研究引力的行为并找到额外维度的证据。但一个对

小距离引力敏感的实验很难实现。引力实在是太微弱了，很容易被其他力，如电磁力所淹没。就如我们早先提到过的，在ADD提出他们观点的时候，实验已在寻找牛顿引力定律的偏差，并证实牛顿定律至少在小至1毫米的距离仍然适用。如果有人能进一步研究更小的距离，就有机会发现ADD提出的大额外维度，它们刚好在实验范围的边缘。

实验者们接受了挑战。受到ADD观点的启发，华盛顿大学的两位教授埃里克·阿德尔伯格（Eric Adelberger）与布莱恩·赫克尔（Blayne Heckel）设计了一个绝妙的实验，其目的是寻找引力在小距离对牛顿定律的偏离。其他人也在研究小距离引力，但这个检验ADD设想的最严密实验。

他们的仪器被安置在华盛顿大学物理系的地下室里，叫作厄缶-沃什（Eöt-Wash）实验，这个名字是为了纪念匈牙利一位研究引力的著名物理学家拜伦·罗兰·凡·厄缶（Baron Roland von Eötvös）。厄缶-沃什小组的实验（见图19-2）由上下放置、互相吸引的两个碟片构成，其上悬挂着一个圆环，圆环及上、下碟片上都穿有小孔，它们之间的摆放位置按牛顿定律设计。如果牛顿定律正确，那么圆环就不会被扭曲，但是，如果有额外维度存在，两个碟片引力吸引的差异就会背离牛顿定律，从而导致圆环扭曲。

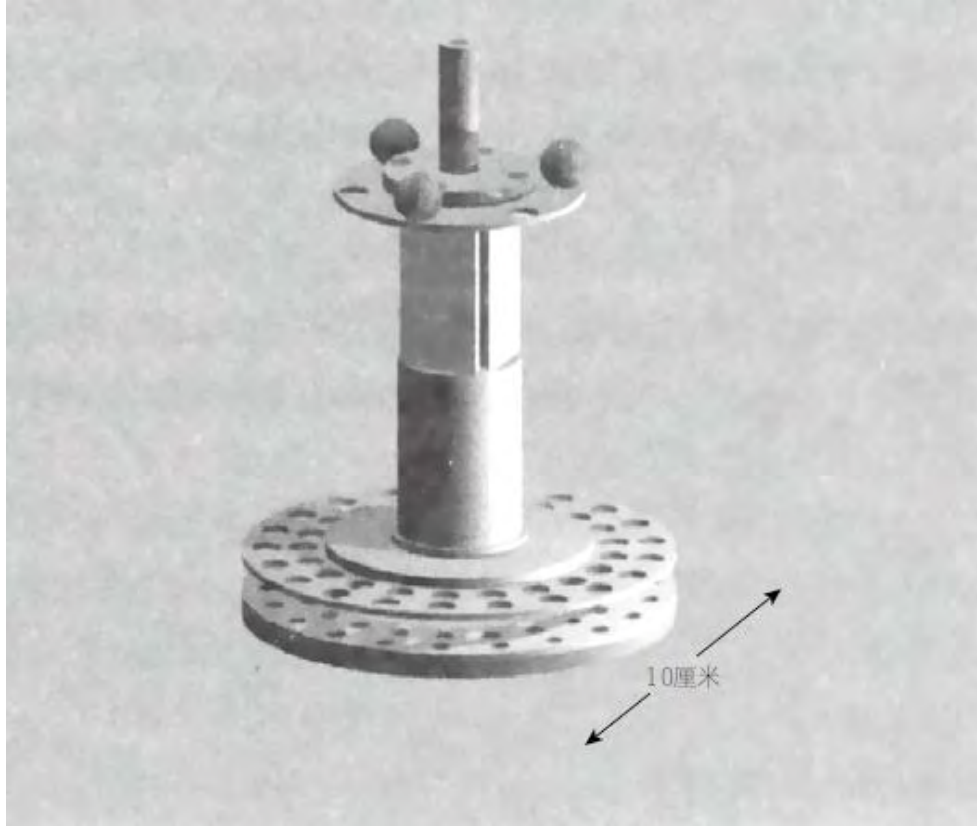


图19-2 厄缶-沃什实验仪器。两个碟片之上悬挂着一个圆环，圆环和碟片上的小孔是为了保证如果引力遵守平方反比定律，圆环不会被扭曲。仪器顶部的3个小球目的是用于校准。

阿德尔伯格和赫克尔没有发现圆环的扭曲，由此他们得出结论，在他们研究的距离上没有额外维度（或其他）效应修正引力作用。他们的实验在前所未有的小尺度上测量了引力，确立了在直到1/10毫米的距离内牛顿引力定律仍然适用。这意味着，额外维度，即便是标准模型粒子被限制在膜上的那些，也达不到ADD提出的1毫米，它们至少要再小10倍。

值得注意的是，毫米大小的维度也被外太空的观察结果所禁止。根据量子力学不确定性原理，与1毫米的尺度相联的是只有 10^{-3} eV的能量，而与1/10毫米相联的是 10^{-2} eV的能量——这两者都是极微小的能量，比生成一个电子所需的能量还要小若干数量级。

如此小质量的粒子可以在我们周围的宇宙和天体（如超新星或太阳）里找到。这些粒子这么轻，如果存在的话，高温的超新星就能生成它们。我们已经知道超新星的冷却速度，也了解它的冷却机制（通过放射中微子），因此，不可能有太多其他低质量的物质被放射。如果能量以其他方式泄露，冷却速度将会太快，尤其是引力子不会带走太多能量。根据这一论证，物理学家证明（不依赖于地球实验），额外维度应该小于大约1/100毫米。

但应该记住，尽管在毫米距离上排除了引力偏差是很难得的，但却并不能验证最近提出的多数额外维度模型。记住，只有具有两个大额外维度的模型才会在1毫米的尺度上产生可观察的效应，如果一个理论有不止两个大的额外维度，且能解决等级问题（或者，如果我们下一章探讨的模型适用于真实世界），牛顿定律的偏离则只有在更小的距离才会出现。

在相距不到1/10毫米的两个物体之间，引力作用会是什么样子？对此，我们并不确信，这从没有人检验过。因此，我们无从知道额外维度在1/10毫米是否会有可能，而这一尺度（如果你曾想过）根本没那么小。相对较大的额外维度，尽管并不一定大至1毫米，但仍是一种可行的选择。要测试这些模型，我们必须等待对撞机实验，这也是下节的主题。

高能粒子对撞机的维度探寻之旅

即便大额外维度超过两个，高能粒子对撞机也非常适合于发现来自大额外维度的KK粒子。在ADD模型里，引力子的KK伙伴总是轻得令人难以置信。如果大维度假说适用于真实世界，那么无论有多少额外维度，引力子的KK伙伴都足以在加速器里生成。这告诉我们，即使维度小于1毫米，当今和将来的加速器探索实验都应该能发现它们。现在的

对撞机所产生的能量已足以生成这些低质量粒子。事实上，如果能量是唯一相关的量，KK粒子早就被大量生成了。

但是，有一个潜在的困难是：引力子KK伙伴之间的相互作用超乎想象地微弱——事实上，它如引力子本身一样微弱。因为引力子相互作用很容易被忽略，以至于对撞机从未在可测速率的水平上生成或发现引力子，因此单个的引力子KK伙伴也没有被发现。

虽然这种推想让人很受挫，甚至我们还以为没有希望了，可实际上，探测到高维KK粒子的前景远比我们想象的更为光明。这是因为，如果ADD假说正确，就会有許多很轻的引力子的KK伙伴，这样就有可能留下它们存在的可探测证据。

如果大额外维度图景正确，那么，即使生成单个KK粒子的概率很小，但生成大量轻KK粒子中的一个，可能性还是很大的。例如，假设有两个额外维度，大约有1 000万亿亿个KK模式轻得足以在运行能量大约为1 TeV的对撞机里生成，那么即使生成其中某个特定粒子的机会极少，但至少生成其中一个粒子的机会总是有的。

这就好像是，有人以非常含蓄的方式向你暗示什么，第一次你听到后没往心里去，可后来有50个人向你重复了同一件事，即便第一次听到这信息时没有在意，可到第50次的时候，这个信息在你脑海里便根深蒂固了。同样，尽管轻KK粒子轻得足以在现今的加速器里生成，但它们的相互作用非常微弱，我们无法探测到单个粒子。可是，当加速器达到足够的能量，能大量生成它们时，KK粒子就会留下可观察的信号。

大型强子对撞机研究的将是TeV能标，如果ADD观点正确，它将极有可能生成KK粒子。这听起来像是一个太过幸运的巧合——无论是KK质量，还是决定KK粒子相互作用强度的质量（MPL）都不是1 TeV，

那么大约1 TeV的能量为什么会与生成KK粒子的机会相关？答案是，大约1 TeV的能量决定了高维引力的强度，而高维引力最终决定了对撞机会生成什么。因为许多引力子KK伙伴的相互作用就等于一个高维引力子的相互作用，而高维引力子在大约1 TeV能量上的相互作用非常强烈，因此，所有KK粒子作用的总和在这一标度上也一定是非常显著的。

实验者已准备用费米实验室的Tevatron寻找KK粒子了。尽管Tevatron达不到LHC那么高的能量，但也达到了实验所需要的能量。而LHC会做得更好，假设ADD的KK粒子存在，LHC找到它们的机会要大得多。

这些KK粒子会是什么样子呢？生成引力子KK伙伴的对撞很像寻常的对撞过程，只不过它看上去会有能量流失。在LHC里，当两个质子对撞时，会生成一个标准模型粒子和一个引力子的KK伙伴。举例来说，标准模型粒子可以是胶子——质子对撞产生一个虚胶子，这个虚胶子可以转化成真实的物理胶子和一个引力子KK伙伴。

但是，单个的KK粒子作用非常微弱，不能被探测到。但如果KK粒子有很多，就可能被发现。探测仪会捕捉胶子，或者更准确地说是围绕胶子的喷射流（见第7章），所以即便记录不到引力子的KK伙伴，产生引力子KK伙伴的过程仍将被记录下来。判别事件额外维度来源的关键就在于这些看不见的KK伙伴，它们带走能量进入额外维度，因而能量看起来就像流失了。实验者们将研究喷射流事件，如果其中被放射胶子的能量少于进入碰撞的能量，他们就能推知一个引力子的KK伙伴生成了（见图19-3）。这与泡利推测中微子的存在有点类似（如我们在第7章里所见）。

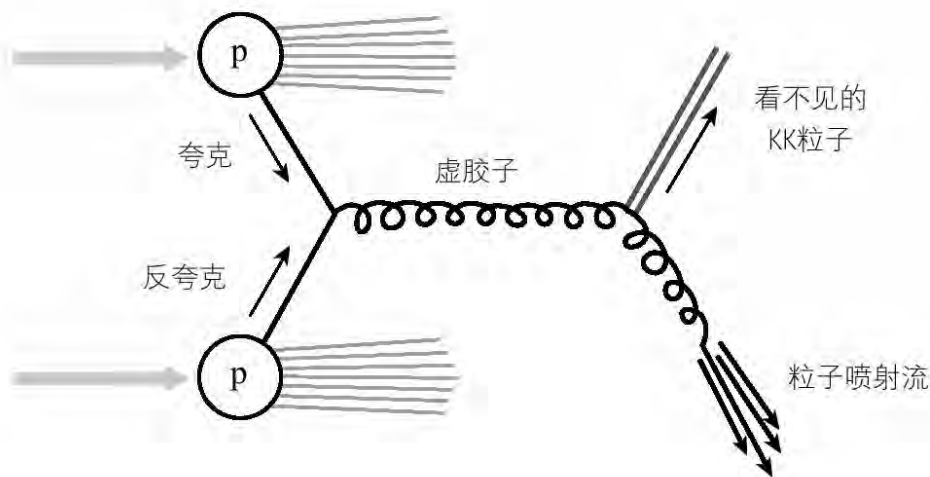


图19-3 ADD模型KK粒子的生成。质子产生对撞，一个夸克和一个反夸克互相消融生成一个虚胶子，虚胶子转化成一个看不见的KK粒子和一个可见的喷射流。灰线是质子对撞时总会大量喷射的一些额外粒子。

因为我们对新粒子唯一所知的就是它带走了能量，所以实际上我们并不能确定加速器生成的就是一个KK粒子，而不是相互作用微弱得探测不到的其他什么粒子。但是，通过对能量流失事件的仔细研究——例如，生成速率与能量的依赖关系，实验者们就有望确定KK粒子的阐释是否正确。

KK粒子将是在我们四维世界里最容易得到的来自额外维度的访客，因为它们可能是代表额外维度的最轻物体。如果我们幸运，ADD模型的其他印记将与它们一同出现，甚至包括更为奇特的物体。如果ADD正确，高维引力将在大约1 TeV上变得强大，这就是说，这一能量远低于我们原以为在惯常四维世界里的能量。如果这样，在大约1 TeV的能量附近就有可能生成黑洞，这种高维黑洞将是我们深入理解经典引力、量子引力以及宇宙形状的敲门砖。如果与ADD观点相关的能量足够低，黑洞的生成将近在咫尺，它们在LHC里就可能形成。

对撞机生成的高维黑洞要远小于宇宙的黑洞，它们的大小与非常微小的额外维度差不多。你大可放心，这种微小的、转瞬即逝的黑洞对我们或我们的星球根本构不成任何危险：在造成任何可能的危害之前，它们就已经销声匿迹。黑洞不会永远存在，它们会通过释放霍金辐射，瞬间蒸发。但是，正如一滴咖啡远比一整杯咖啡蒸发得快一样，小黑洞也要比大黑洞蒸发得快。因此，对撞机能生成的微小黑洞，在瞬间就会蒸发。但是，如果这些高维黑洞生成了，它们就能持续足够长的时间，在探测器里留下可见的存在迹象：它们会有与众不同的表现，因为它们会产生许多粒子，这是你在通常的粒子衰变里看不到的，而且这些粒子会向所有方向喷射。

再者，如果ADD模型正确，奇异的新发现将不止会有黑洞和引力子的KK伙伴。如果ADD和弦理论都正确，对撞机就能在几乎低至1 TeV的能量上生成弦。这仍然是因为基本引力标度在ADD模型里非常低：高维引力将在大约1 TeV的能量上变强，而量子引力会产生可测量的效应。

ADD理论的弦质量不会大到令人根本无法探测的普朗克级质量。如果你把弦当作音符，那么ADD观点里的弦远没有那么高的音调。ADD模型的低音弦所拥有的质量不会比1 TeV大很多，如果我们幸运，它们足以轻得在LHC里生成。那么，能量足够高的对撞就能大量生成这种模型里的轻弦以及一些新的物质：由许多长弦构成的弦球。

虽然这些潜在发现非常吸引人，但不要忘记，LHC的能量很可能接近于（但不会达到）产生弦和黑洞所需的能量。ADD的弦和黑洞是否可见将取决于高维引力的正确能量（当然，也取决于他们的设想是否正确）。

后续影响

ADD的建议是迷人的。谁会想到额外维度会有这么大？谁又会想到它们关系着诸如等级问题那样有着直接意义（至少对粒子物理学家是这样）的问题？但是，这个方案并不能实际解决等级问题，它只是把等级问题换成了另一个问题：额外维度会有这么大吗？对ADD图景来讲，这仍是一个突出的问题。如果没有新的、有待确认的物理原理，维度不应该大得这么超乎寻常。最起码，根据已知理论，要维持ADD假说所需要的大的平坦空间，我们仍需要超对称。在根本上，超对称能够稳定和加固大维度，不然它们将坍缩。而ADD的一个良好特征似乎就是它不再需要超对称，这就有点令人失望。

该理论的另一缺陷在于它的宇宙学意义，若要理论与宇宙演变的已知事实相符，其中的某些数字必须经过谨慎选择。体空间必须含有很少的能量，否则，宇宙学演变就会与观察现象不一致。要解决等级问题，其关键就是要消除对庞大参数的依赖。

然而，许多物理学家乐于严肃地对待额外维度理论，并努力设计方法来寻找它们。实验者们尤其激动，乔·莱肯，费米实验室的一位粒子物理学家，在描述实验者们对于大额外维度的反应时说：“对他们来说，所有‘超越标准模型’的研究都是癫狂古怪的，是超对称？还是大额外维度？谁在乎呢，额外维度不见得会更古怪。”实验者们迫切想找到一些新的东西，而除了超对称之外，额外维度又提供了另一个有趣的可能。

理论家的反应也各不相同：一方面，大额外维度似乎是古怪的，以前从未有人想到过，因为人们想不出额外维度为什么要这么大的理由；另一方面，又没人能找到理由排除它们。事实上，在写第一篇有关大额外维度的论文之前，作者之一的吉亚·德瓦利曾在斯坦福大学谈到过它们。作者们知道他们的提议有多么激进，因此等着他们的发言遭到抨击。

反对并没有那么激烈，这让他们松了一口气，可同时，也让他们感到沮丧——人们怎么可能如此平静地接受一个如此激进的观点？尼玛告诉我，当他们第一次在网上贴出论文时，他们经历了相似的感受：他们原以为会有大量的回复，不料只收到了两个。显然，只有我和里卡多·拉塔兹，一位意大利物理学家，对一些潜在的问题作出了评论。而即便是这两条信息也不是独立发出的：因为同在CERN参观，我们两人刚刚讨论过这篇论文。

后来，当物理学家领会了ADD模型的含义时，他们更加深入地研究了它于真实世界的作用，考虑了引力检验、加速器探索、天体物理结果及宇宙学含义。研究兴趣和风格不同的人，对此反应也不同。

探索标准模型细节的那些物理学家乐于接受一个可能的新观点，无论怎么说，它都是有趣的。令人惊讶的是，更多的抵触来自模型构建者，他们不愿意放弃超对称的观点。多年来，它已深入人心。我们必须承认，这么剧烈地改变标准模型将面临艰巨的挑战。

所有的新模型都必须重现标准模型已经验证的那些特点，而对标准模型作出巨大改变的理论都需经历一段艰难的时日来面对这些挑战。而且，超对称的闪光点——耦合的统一，即在高能量上所有的力都将有同样的强度，将不得被放弃。但是，并未完全投入超对称的年轻物理学家更兴奋：额外维度还是一个新兴的、未被冷落的题目，而且提出了新的挑战 and 开放的问题。

来自弦理论学家的反应同样是复杂的。萨瓦斯·迪莫普洛斯开始他的项目时，他料到额外维度的研究会将弦理论和粒子物理学的距离拉近。弦理论学家确实给予了关注，但大多数人只是把大额外维度当作一个有趣的观点，认为它永远都不会对弦理论产生影响。对弦理论家学来说，问题主要是理论性的：很难理解维度怎么会像ADD设想的那么大。

就我个人的观点，即便额外维度存在，我也不相信它们有这么大。^[63]这既有理论原因（我们很难得出这么大的额外维度），也有现实依据（这很难让宇宙学有所结果），这一观点似乎有点瞎猜。即使是作为主角之一的尼玛，在这一点上也持怀疑态度，但这是一个非常重要的理论观点。这个新的、以前从未被探索过的提议让我们进一步意识到我们对引力和宇宙的形状是多么无知。ADD的论文激发了大量的新思想，无论这一观点最终能否被证明是正确的，它都对物理学家的思想产生了重要的影响。大维度图景引出了有关额外维度的许多新见解，以及有关实验测试的许多新观点。

-
- 如果标准模型粒子被限制在一个膜上，那么额外维度将比物理学家原来以为的大得多：它们可能大至1/10毫米。
 - 额外维度这么大，它们甚至能够解释为什么引力远弱于电磁力、弱力和强力。
 - 如果大额外维度能够解决等级问题，那么高维引力将在大约1 TeV的能量上变得强大起来。
 - 如果高维引力在大约1 TeV的能量上变得强大起来，那么LHC将有机会生成KK粒子。KK粒子将从对撞中带走能量，因此，它们的印记将是有能量流失的事件。





对我很大的东西，对你很小。

如果这是我做的最后一件事，我就要让你知道。

苏珊·薇格 (Suzanne Vega)

一个有关五维世界的梦

阿西娜突然惊醒了。最近她总是重复做着同一个梦，这次的梦又是从她进入兔子洞开始的。在她的梦境里，当兔子宣布“下一站，二维世界”时，阿西娜根本不想理会，就等她念到自己中意的站点。

在3个空间维度的那一站，兔子宣布：“如果你住在这里，现在就到家了。”可无论阿西娜怎么恳求，说自己真的住在这里，而且非常想回家，他就是不肯打开门。

在下一站，穿统一制服的六维居民想要进来，可一看到他们一个个庞大的腰身，兔子立即关上了电梯门，说电梯里装不下他们，要上来只能给他们强行瘦身。^[64]听到兔子的恫吓，他们赶忙转身离开了。

电梯继续着它的奇异旅程。当它再次停下的时候，兔子宣布：“弯曲几何——一个五维世界。”^[65]他温和地将阿西娜推向门边，告诉她：“进入哈哈镜里，它会把你带回家的。”因为兔子曾提到过第五维度，阿西娜觉得这极有可能。而她也别无选择，只能硬着头皮进入，但愿兔子没有捉弄她，说的都是真话。

当你学习一种语言时，会记住哪些词汇要看你最感兴趣和最需要的是什么。例如，我在意大利骑自行车旅行时，学到了各种各样水的表述方法——自来水、瓶装水、汽水、纯净水，等等。同样地，当一位物理学家看到一些新的物理景象时，每个人都有他自己的视点和想法，他可能注意到一个系统的某些特定方面，或发现某些已知现象的不同含义，即便面对的是同样的语言或情景，每个人听到的都是不同的东西。仔细倾听，异常重要。

我和拉曼都对等级问题思考了很多年，但开始合作时，我们并没有想到寻找解决等级问题的更新、更好的办法。如我在第17章里所说，我们在致力于一个隔离的超对称破缺模型的研究。在这一研究过程中，我们不期然地发现：在两个膜之间束缚着一个神奇的弯曲时空几何。

因为我和拉曼的焦点都聚集在粒子物理和引力的微弱上，所以我们很快认识到弯曲几何的潜在意义：如果把粒子物理学的标准模型置于这样的时空中，等级问题便迎刃而解。我不知道我们是否是第一个研究爱因斯坦的这组特殊方程的人，但我们肯定是最先认识到这一惊人含义的。

后面的几章我就将解释弯曲时空以及其他几种引人注目的可能，并说明其结果为什么有时会违背我们的预期。本章集中探讨一个弯曲的五维世界，它将帮助解释粒子物理学里极其重要的质量差异。尽管在四维量子场论里，人们预期粒子的质量大致相同，但在弯曲的高维几何里，情况并不如此。弯曲几何给出了一个框架，其中悬殊巨大的质量会自然而然地出现，量子效应也在掌控之内。在本章描述的这一特定几何里，我们会看到以下现象。

两个平坦的边界膜之间时空严重弯曲，粒子物理学的等级问题被自然而然地消除——根本不需要大的维度，也不需要任何一个庞大的

参数。在这一图景里，一个膜经受强大的引力，而另一个膜不经受。时空在第五维度的方向上急剧变化，从而成功地将与两膜间隔相关联的小数转化为与引力的相对强度相关联的大数（约1亿亿）。

首先，我们将以引力子的概率函数来解释在第二个膜上引力的微弱，这一函数决定了引力子在第五维度上任一位置的相互作用。但我们也会以另一种不同的方式来解释引力的微弱，其基础是弯曲几何本身，而不是引力子相互作用的强度。我们会看到，弯曲几何的一个神奇结果就是，在第五维度上大小、质量甚至是时间都要取决于位置。在这个有两个膜的结构里，时间和空间的弯曲就像是在黑洞视界附近的时间弯曲，但在这一情境里，空间是膨胀的，时间被拉长，在其中一个膜上粒子有很小的质量——这样等级问题就自动消除了。

在讨论过弯曲几何及其对等级问题的意义后，本章最后，我们将探讨这一理论对未来实验的特殊意义，它最令人兴奋的一个方面是：与上一章的大额外维度一样，如果该理论正确，不久它就将在粒子加速器里留下可观察的结果。事实上，我们会看到，它们的印记甚至比我们曾探讨过的能量流失的印记还要显著。引力子的KK伙伴虽然来自高维空间，却是很容易辨认的可见粒子，它将衰变成我们四维空间膜的熟悉粒子。

令人惊讶的弯曲几何

在本章中，我们将探讨的空间几何包含两个膜（见图20-1），它们是第五维度的边界，这一构成与我们在第17章中探讨过的构成类似：都有两个膜，而两个膜之间延伸着第五维度。然而，这是一个全新的理论，粒子和能量的分布都不相同，而且理论中也没有超对称。可是，与那个理论一样，我们都认为标准模型的所有粒子以及引起弱电对称破缺的希格斯粒子，都被限制于其中一个膜上。

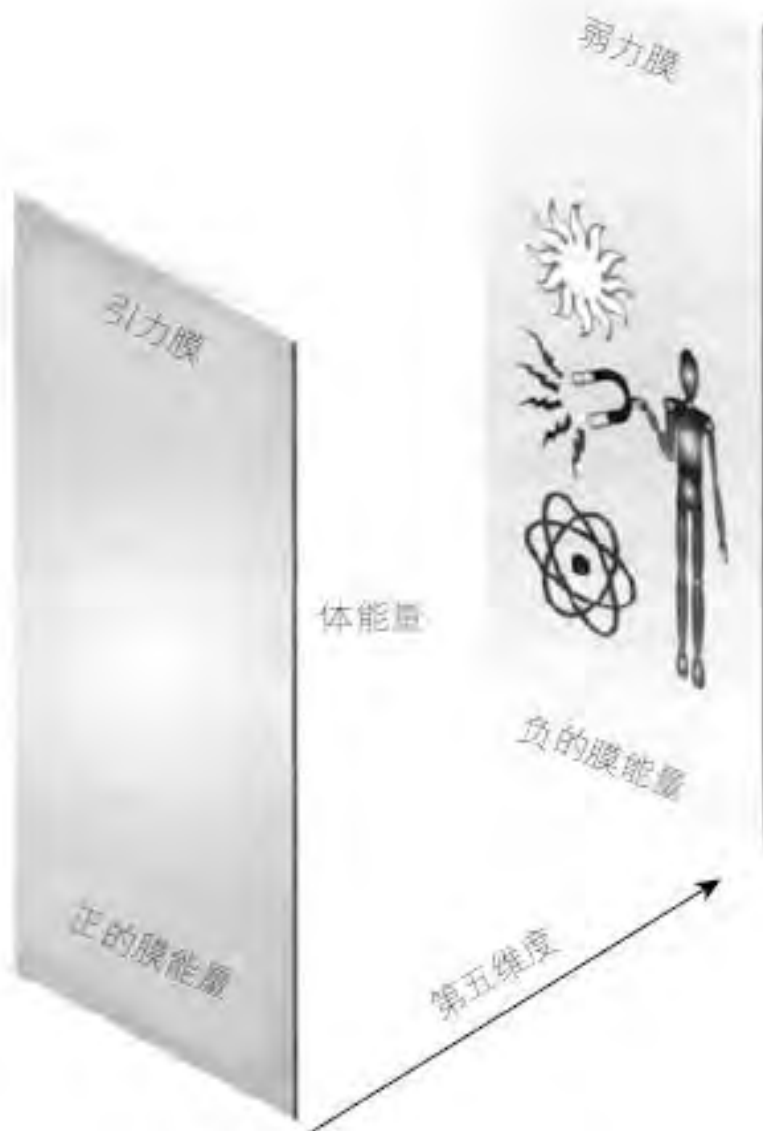


图20-1 有两个膜的五维弯曲几何。宇宙有5个时空维度，而标准模型只居住于四维的膜上（弱力膜）。在这一构成中，时空维度的总数也是5个，而空间维度是4个，其中三维在膜上延伸，而另一维延伸于两膜之间。

如前所述，在这一情景里，我们假设引力是存在于第五维度中唯一的力，这就意味着，如果不是引力，每个膜都将与传统的四维宇宙一样。

被限制于膜上的规范玻色子和粒子会传递力并互相作用，就像第五维度根本不存在一样；标准模型粒子只在有三个平坦维度的膜上穿行，力也只沿着膜的三维平坦表面发散。

然而，引力却不同，因为它没有被限制在膜上，而是存在于整个体空间里。所以在第五维度上，我们处处都能感受到引力，但这并不意味着我们感受到的引力处处都相同。在膜上和在整个五维体空间里的能量会使时空产生弯曲，这就使得引力场产生了巨大的差异。

上一章的大额外维度利用了膜能束缚粒子和力的事实，却忽略了膜本身所承载的能量。我和拉曼并不确信这个假设一定就好，因为爱因斯坦广义相对论的一个核心要素是：能量导致了引力场。这就意味着当膜承载能量时，它们能弯曲空间和时间。我们意图探讨的是只有一个额外维度的宇宙，在这样一个宇宙里，没有任何证据显示膜和体能量可以忽略不计：膜的引力效应不会那么快就消散，因此即使在远离膜的地方，我们仍可以想象时空会有弯曲。

我们想知道，当两个高能量的膜束缚着额外维度空间时，时空将产生怎样的弯曲。我和拉曼解开了这一两膜结构的爱因斯坦引力方程，并假设在空间和膜上都存在能量，我们发现这种能量真的非常重要——它们导致时空产生了极大的弯曲。

在有些情况下，弯曲空间很容易描画，例如一个球体的表面是二维的——你只需知道经度和纬度就可以确定你的位置，但它显然是弯曲的。但是，许多弯曲空间因为在三维空间不容易表现，所以很难勾画。现在我们要探讨的这一弯曲时空就是这样，它是叫作反德西特空间（anti de Sitter space的一个时空组成部分，反德西特空间有负曲率，不像球而更像品客薯片。其名称源于荷兰数学家、宇宙学家威廉·德·西特尔（Willem de Sitter），他研究的是一种正曲率空间，叫作德西特空间（de Sitter space）。尽管现在我们需要这一

名称，但后面当我们将这一理论与弦理论学家一直研究的反德西特空间理论联系起来时，将会涉及它。

在探讨五维时空弯曲的有趣方式之前，我们先看看第五维度两端的两个膜，这两个边界膜是完全平坦的。如果你在其中一个膜上，那么你会困在一个三维加一维的世界里（三维空间，一维时间），^[66]这一世界在三个空间维度里无限延伸，俨然是平坦的时空，没有任何特别的引力效应。

而且，弯曲时空有一个奇异的特点：如果你把自己限制在沿着第五维度的任一薄片上，而不仅仅是两端的膜上，你都会发现这个薄片是完全平坦的。也就是说，尽管在第五维度上，除了两个边界外并非处处都有膜，但如果把自己限于五维空间的任一点上，你由此得到的三维加一维的几何表面看上去都是平的，它与边界的两个大的平坦膜有着相同的形状。如果你把边界膜看作一个长面包两头的外壳，那么在时空第五维度任一点上平坦、平行的四维区域就像从面包里取出的任一平坦的切片。

但不管怎么说，我们探讨的五维时空是弯曲的，四维平坦时空的切片沿着第五维度粘连在一起的方式就可以反映出这一弯曲。我第一次谈起这种几何是在圣塔芭芭拉的卡夫里理论物理研究所。在那里，弦理论学家汤姆·班克斯（Tom Banks）告诉我，从专业角度来看，我和拉曼发现的五维几何是弯曲的。尽管有许多弯曲时空在日常口语中都被说成是弯曲的，但专业术语所指的几何是其中的每一切片都是平坦的，^[67]它们会被一个整体弯曲系数连在一起。弯曲系数是改变整体标度的函数，它随着在第五维度上不同点的位置、时间、质量和能量而不同。弯曲几何的这一神奇特征非常微妙，下一节里我将进一步解释。弯曲系数还反映在引力子的概率函数和相互作用上，我们也将很快进行探讨。

图20-2画的是一个有着平坦切片的弯曲空间。这是一个实心漏斗，我们可以用刀将漏斗切成平整的薄片，但漏斗表面明显是弯曲的。在某些方面，这就像是我们正探讨的五维时空，但这种比喻并不十分恰当，因为在漏斗里，唯一弯曲的地方是漏斗的边界，即它的表面，而在弯曲时空里，处处都存在着弯曲。这种弯曲可以通过在整体上重新标度空间标尺和时钟速度来刻画，它们在第五维度的每一点都是不同的。

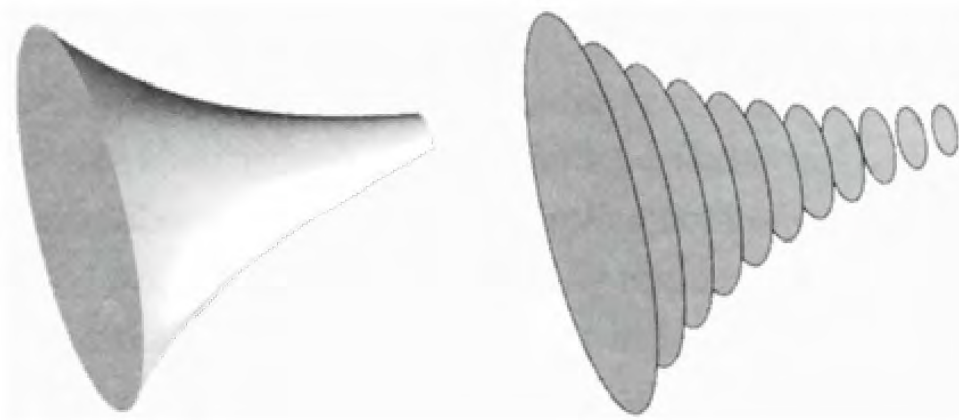


图20-2 一个实心漏斗是由许多平坦切片粘连在一起形成的。

说明弯曲时空曲率的一个更为简单的方法是通过引力子概率函数的形状。引力子是传递引力的粒子，它的概率函数说明在空间某个特定位置找到引力子的概率。引力强度也反映在函数中：它的值越大，引力子在那一点的相互作用越强，引力强度就越大。

对平坦时空而言，引力子出现在各处的概率都一样，因此，平坦时空的引力子概率函数是恒定的；但对弯曲时空，正如在我们所考虑的弯曲几何中一样，情况不再如此，弯曲给我们揭示了引力的形状，当时空产生弯曲时，在时空的不同位置上，引力子概率函数的值是不一样的。

因为在我们的弯曲几何中，时空的每个切片都是完全平坦的，在沿着三个常见的空间维度上，引力子的概率函数没有变化——它只在第五维度上发生改变。^[68]换句话说，即使引力子在第五维度上每个点的概率函数值不同，但只要两点到第五维度的距离相同，它们的值就是相同的。这告诉我们，引力子的概率函数只取决于在第五维度上的位置，但它能完全代表时空弯曲的特征。因为这个函数只随一个坐标的改变而改变，即第五维度坐标，因此很容易描绘。

沿着第五维度，引力子的概率函数如图20-3所示，我们称第一个膜为引力膜，第二个膜为弱力膜。离开第一个膜向第二个膜靠近时，概率函数陡然下降。引力膜与弱力膜不同，前者携带正能量，而后者携带负能量，这种能量分布使得引力子在引力膜附近的概率函数要大许多。

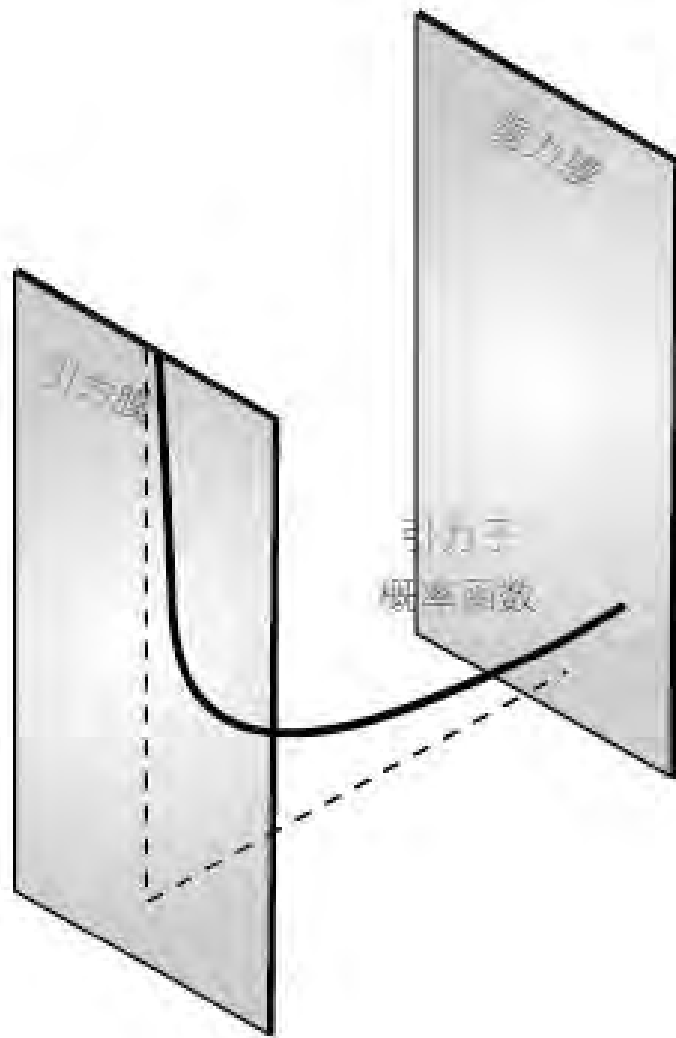


图20-3 引力子概率函数。在远离引力膜靠近弱力膜的过程中，引力子的概率函数呈指数级陡然下降。

概率函数陡降的结果是，在弱力膜附近很难发现引力子，而引力作用正是通过这种粒子的交换才产生的。因此，引力子的相互作用在弱力膜上受到了高度抑制。

引力的强度如此强烈地依赖于第五维度上的位置，以至于在界定弯曲的五维世界两个端点的膜上所经受的引力强度有着异常大的差距。引力在它定域的第一个膜上强，而在标准模型所在的第二个膜上

弱。因为引力子的概率函数在第二个膜上小得可以忽略，因此引力子与被困于这个膜上的标准模型粒子的相互作用也极其微弱。

这告诉我们，有希望在这一弯曲时空中发现可观测质量与普朗克级质量之间的等级差异。虽然引力子无处不在，但它与引力膜上粒子的相互作用远强于与弱力膜上粒子的相互作用。引力子并不总是在那里逗留。引力子在弱力膜上的概率函数极其微小，如果这一图景与真实世界相符，那么这个微小的值就决定了我们所处世界中引力的微弱。

在这个模型里，弱力膜上微弱的引力并不要求两膜之间有巨大的距离。一旦离开引力子概率函数高度集中的引力膜，引力就会呈指数级衰减，从而使弱力膜上的引力变得微弱。因为引力子概率函数陡然下降，引力在弱力膜上（我们居住的地方）被高度抑制。即便两个膜相隔非常近，它也会比没有弯曲时的强度弱1亿亿倍。

理论的这一方面，即两个膜并不需要相隔很远，使得这一模型远比大额外维度要更为现实。尽管大额外维度是对等级问题的诱人诠释，但直到最后，它还留有一个没有解释的庞大数字——额外维度的大小。在我们现在讲述的理论里，即便弱力膜与第一个膜（引力膜）只相隔很短的距离，弱力膜上的引力仍比其他的力要弱多个数量级。

在这一弯曲几何里，两膜之间的距离只需比普朗克长度稍大一点。大维度图景需要引进一个庞大的数字，即维度的大小，而在弯曲几何中，无须一个特别设计的庞大数字来解释等级。这是因为指数会自动地把一个适中的数转化成一个极为庞大的数（指数）或一个极小的数（指数倒数）。

在弱力膜上，引力强度非常微小，它的强度随两膜之间的距离而呈指数级衰减。^[69]不同质量间的比值约为 10^{16} （1万亿），如果弱力膜在引力膜外16个单位距离处，^[70]就可以预见普朗克级质量（这一庞大的质量告诉我们引力很微弱）与希格斯粒子（弱力规范玻色子）质量的庞大比值。这意味着两个膜之间的距离只要比最直观的猜想再大16倍，就足以解释等级问题。16个数量级听起来很大，但比起我们试图解释的1万亿，已经小多了。

多年来，粒子物理学家一直希望为等级问题找到一个指数级的解释，也就是说，我们希望把以前不能解释的庞大数字解释为一个自然发生的指数级函数的结果。现在，通过引入额外维度，我和拉曼发现了一种使粒子物理学能自动兼容一个指数级质量等级的方法。在我们的膜上，即弱力膜上，引力的相互作用比其在引力子概率函数的高峰值上要低许多。因为在我们的膜上引力被弯曲几何所削弱，如果标准模型被局限在弱力膜上，等级问题便解决了。这就是等级问题的解决方案，这一结果正符合我们的期望。

还有一种方法可以用来理解弯曲几何这一令人瞩目的新特征，我们来看看引力是怎样被削弱的。在第19章中，我们以引力线的方式解释了ADD图景里引力的微弱问题：从一个有质量物体放射出的引力线，由于在整个大的维度里发散而被削弱。根据我和拉曼的观点，这种削弱还可以解释为引力子概率函数的结果。

这一概率函数告诉我们引力在空间中是如何发散的。在大额外维度的图景中，引力在额外维度里的每个地方都同样强大，在这种情况下，引力子概率函数就是平坦的。这一平坦的概率函数告诉我们，引力子——传递引力的粒子，遍布于由额外维度包围的庞大区域中，而这个在整个额外维度空间均衡分布的平坦概率函数表明，引力的作用在四维里被大大削弱了。

在我们现在考虑的弯曲的五维时空里，有一种有趣的弯曲：引力子在引力膜和弱力膜两个边界之间的五维空间出现的概率不再处处相同。作为膜所承载的能量和空间能量的自然结果，引力子的分布事实上根本不均等。引力子的概率函数会有所变化：它在一个区域很大，而在其他所有区域都很小。正是这种变化提供了一个分散系数，使得引力在我们的世界里这么微弱。引力在弱力膜上非常微弱是因为引力子在那里的概率函数极小。



现在，我们再回到早先用来解释引力强度随距离减弱的洒水装置的例子。洒水装置喷洒的区域越大（见图20-4中上图所示），水分散得越广。当额外维度很大时，引力在一个大区域里也会被削弱。因此，在一个低质量的四维有效理论里，引力看起来就很微弱。

而弯曲几何与第三个洒水装置类似，它并不是将水均演示地洒向所有方向，而是偏向某个特定区域，即靠近引力膜的区域（见图20-4）。有了这样一个不公正的洒水装置，很显然，除了受偏袒的区域外，其他各处得到的水会更少。如果由受偏袒地方到其他地方所得到的水量呈指数级地下降，那么其他地方的水的比例就很小了，即使它们离水源并不远。显然，由“弯曲的”洒水装置提供的水量，比起向所有区域均衡分布的水量要减少得更快。

结果就是，如果所有标准模型粒子都局限在弱力膜上，那么引力就比其他3种力微弱很多，这也就解决了粒子物理学中的等级问题——即引力为什么比其他力微弱。微弱的引力是引力子概率函数在弱力膜上振幅很小的自然结果，即使弱力膜离引力膜只有相对较小的距离（大约比弦理论偏爱的普朗克长度大10倍）。

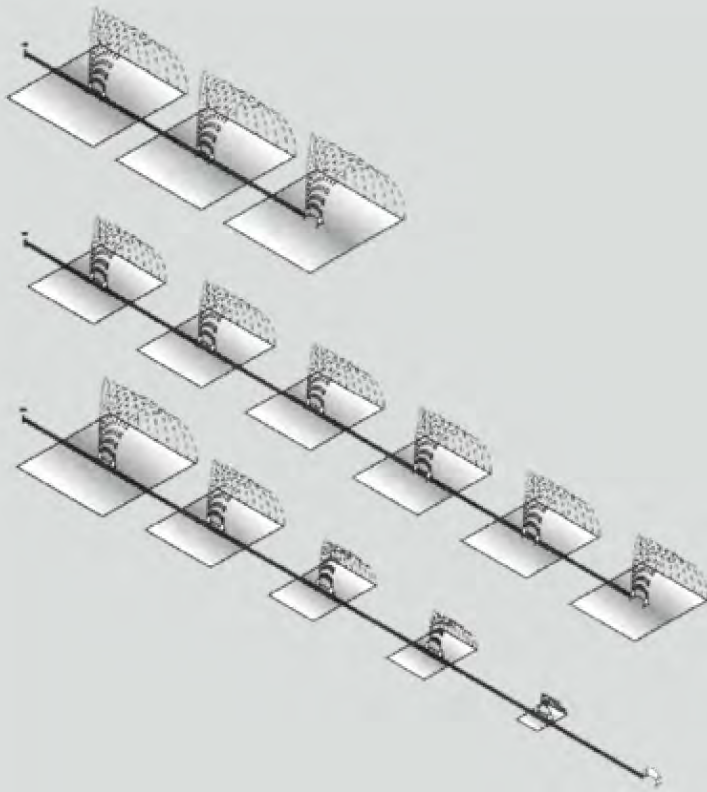


图20-4 三个不同的洒水装置。第一个与第二个相比，我们看到的水管越长，浇在一个特定区域内的水就越少。第三根水管显示，水的分布可以是不均衡的，这样，第一个花园总能得到一半的水量，而第二个花园得到1/4的水量，依此类推。

在这种情况下，浇给第一个花园的水量与水管的长度无关，它总能得到一半的水量。

卷曲宇宙的伸缩

用概率函数的指数级陡降来解释等级问题，已足以让我们理解弯曲的时空。对微弱引力的直观解释是：在弱力膜上不太可能找到引力子。是否接受这个解释要由你自己决定，你也可以跳过这一节，但或许你愿意读下面这一更严格的解释，它进一步探讨了弯曲时空的神奇属性。

我们会在这一节看到：弱力膜上引力的微弱还可以解释为物体随我们从引力膜靠近弱力膜时逐渐变大、变轻的结果。如果阿西娜从引力膜向弱力膜移动（下一章的故事里，她会这么做），她会看到自己在引力膜上的影子越变越大。这个影子会长到非常大——它将增大16个数量级！

我们还会看到，在这个几何空间里，轻、重粒子可以和谐共存：如果在其中一个膜上有普朗克级质量的粒子，那么另一个膜上就只能有弱力级质量的粒子。因此，等级问题不复存在。

我们可以这样来理解它的运作机制：假设你与大多数人一样（至少是那些没读过此书的人），对第五维度一无所知——毕竟，它是看不见的，你泰然地相信自己居住于一个四维世界里，只知道四维引力，并且认为它是由一个寻常的四维引力子传递的。在描述我们所见现象的四维有效理论里只有一个引力，因此也就只有一种类型的四维引力子。所有粒子都会与这同一种引力子相互作用，但那个引力子并不包含粒子在原来的高维理论中位置的任何信息。

这种推理使它看起来就好像是，所有引力子的相互作用都应该是一样的，而不管一个物体来自第五维度的什么地方。毕竟，你并不知道这个物体来自第五维度，甚至根本不知道还有个第五维度。决定引力子作用强度的牛顿引力常数，也将是决定所有四维引力作用强度的唯一的量。但我们在上一节里看到，当你从引力膜向弱力膜移动时，引力作用会越来越弱，这就引出了一个问题：引力强度是怎么容纳物体在第五维度的位置信息的？

这一矛盾的解决方案取决于一个事实：引力与质量成正比。而在沿着第五维度的不同点上，质量可以且一定是不同的。要重现引力子作用在第五维度的切片上渐次减弱，唯一的办法是以不同的方式测量每个四维切片上的质量。

弯曲时空有许多引人瞩目的属性，其中一个**是**：当你从引力膜向弱力膜移动时，**能量和动量会减小**。减小的能量和动量（与量子力学和狭义相对论一致）告诉我们，距离和时间一定会膨胀（见图20-5）。在我所描述的这一几何里，大小、时间、质量和能量都依赖于位置。四维世界的大小和质量所承继的值要依赖于它们在五维世界的起始位置。物理现实看起来是四维的，但测量长度的标尺或测量质量的标度，都取决于五维的起始位置。引力膜或弱力膜上的居民看到的都是四维世界，但他们会测得不同的大小，并预期不同的质量。

在原来的五维理论里，远离引力膜的粒子的引力作用在四维有效理论里较小，因为其质量本身很小——在第五维度的每一位置上，质量和能量都会被重新标度，这取决于与引力子概率函数在那一点的振幅成比例的量。而用以重新标度能量的量，即弯曲系数，随距引力膜变远而减小。事实上，它的图形与引力子概率函数的形状完全一样。因此，在沿着第五维度的每个点上，质量和能量会以不同的系数缩减——弯曲系数决定其将缩减多少。

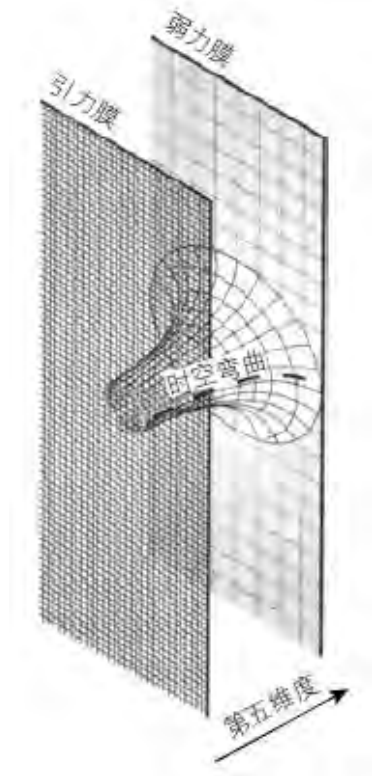


图20-5 随着从引力膜向弱力膜的移动，尺度会增大（而质量和能量则会减小）。

这种重新标度看似随意，实际并非如此，让我们来看一个类似的情景。假设我们要以火车行驶160公里需耗时多久来计量时间，我将这种时间单位称作TT（火车时间）。这是一种很好的时间计量方式，只是时间的确定要看你是在哪里乘火车旅行：火车行驶速度是快还是慢？例如，假设一部电影要放映两个小时，如果在美国一列火车行驶160公里要花费1个小时，那么一个美国观众看完整部电影，火车就要行驶320公里，我们就可说这部电影持续的时间长度是2TT；而在法国乘TGV旅行的观众则认为这部电影长度是6TT，因为法国的高速列车比美国火车要快3倍，法国观众要看完整部电影需列车行驶600公里的路程。

法国的列车20分钟能行驶100公里，而美国的列车要行驶同样的路程则需1个小时，因此，如果你想使美国和法国观众的时间单位相同，对电影的TT长度取得一致，那么你就需要按比例重新标度火车时间。要将法国时间换算成美国时间，你就需要将法国火车时间乘以系数3。

同样地，在弱力膜上（其引力子相互作用远小于引力膜），为了考虑引力的微弱，用以测度能量的标准单位必须重新标度。这种重新标度是通过一个巨大的量： 10^{16} ，即1亿亿。它的含义是，虽然在引力膜上所有的基本质量都应是 M_{Pl} （普朗克级质量），但在弱力膜上，它们大约只能是1 000 GeV，是其 $1/10^{16}$ 。处于弱力膜上的新粒子的质量可能会大一些，兴许是3 000或5 000 GeV，但它们不可能大太多，因为所有的质量都被极大程度上重新标度了。

当所有的质量都被提高至最大质量左右时，就出现了等级问题。如果那个质量是普朗克级质量，那么所有质量都应和这一质量一样大。由于这种重新标度，如果引力膜上所有东西的期望质量是普朗克级质量，那么在弱力膜上的预期质量则是小16个数量级的1 TeV。^[71]这就意味着，希格斯粒子的质量根本不是问题：即便引力非常微弱，我们仍可以期待它的质量约为1 TeV——比普朗克级质量小1亿亿倍。这一解释的根本是重新标度解决了等级问题。

依此类推，在弱力膜上所有的新物质，包括弦在内，其质量大约都应该是1 TeV左右。这告诉我们，这一模型会产生戏剧化的实验结果。在弱力膜上，与弦相关的额外维度粒子要比它在引力膜上轻许多。从发现额外维度的角度来看，弱力膜呈现出了一个绝妙的图景，如果这一观点正确，那么来自额外维度的低质量粒子将触手可及——在弱力膜上会有大量的TeV质量的粒子。

弱力膜上的所有东西都应该比普朗克级质量轻 10^{16} 倍。根据量子力学，质量越小意味着尺寸越大，在从引力膜走向弱力膜时，阿西娜的影子也会变大。这告诉我们，在弱力膜上，弦的大小也不会是 10^{-33} 厘米，而是要大 10^{16} 倍，即 10^{-17} 厘米。

尽管我集中探讨的是一个有着特定弯曲系数的两个膜的情景，但我们考虑的特点很可能比这一特例更具一般性。通过额外维度，我们有很好的理由预期差异很大的质量。粒子物理学中关于质量都应相差无几的直觉意识被打破了，可以预想范围很大的质量。不同位置的粒子自然会有不同的质量，它们的影子会随你的四处游移而变化。在我们的四维世界里，其结果就是跨度很大的尺度和质量，这正是我们观察到的。

进一步发展

当我们以弯曲几何解释等级问题的论文在1999年发表时，大多数同行并没有意识到：这是一个与大维度观点完全不同的全新理论。乔·莱肯对我说：“反应是慢慢建立起来的。起初没有反应，但渐渐地，所有人都明白了这篇论文（还有另外一篇，我在第22章中会讲到）意义重大，包含全新的、原创的观点，且开辟了一个全新的观念领域。”

在我的论文发表后的几个月里，人们总是要我谈谈自己关于“大额外维度”的研究，我不得不一遍遍地纠正，说我们理论的精要恰恰在于维度不是大的！事实上，粒子物理学家马克·怀斯（Mark B. Wise）就曾嘲笑过我在2001年轻子光子会议闭幕式上的发言题目。那是一次重要的粒子物理学会议，会上，实验者们报告了他们重要的实验结果，大会组织者指派给我的题目涵盖了有关额外维度的所有研究，可就是不包括我自己的！

马克和他当时的学生沃尔特·戈德伯格（Walter Goldberger）是最早领会弯曲图景意义的两个人，但他们也发现，我和拉曼的研究结果中还留下了一个潜在的漏洞，需要得到补充。我们原以为膜动力学会自然地隔绝两个膜，但我们并没明确说明两个膜之间的距离是怎样

被确定的。这不仅仅是一个细节问题，我们的理论要解决等级问题，就必须能够很容易地将两膜固定在一个小而有限的间隔上。结果，距离的反指数函数（我们希望它极度微小）而非距离自身很可能会成为一个适中的数。如果这样，弱力级质量与普朗克级质量之间预言的等级差别将会是一个适中的数，而不是（小得多的）这个数的指数倒数——这样我们的解决方案就失效了。

马克和戈德伯格做了重要的研究，弥补了我和拉曼理论的危险漏洞。他们证明：两膜之间的距离是一个适中的数，而这个距离的指数倒数极其微小。这正是我们的解决方案发挥作用所必需的。

他们的观点非常完美，而且最终证明比人们当时所认识到的更有效。结果是，之后所有的稳定模型都与他们的极为相似。

马克和戈德伯格提出，除了引力子之外，在五维体宇宙里还居住着一个大质量的粒子，他们假设这个粒子的属性就像弹簧一样。通常来说，弹簧有一个自然长度，变长或变短都会让其携带能量，使它运动。对他们引进的这个粒子（以及相应的场）而言，场和膜的平衡结构将包括一个适中的膜间距——这又是我们解决等级问题所要求的。

他们的方法依赖于两个相互竞争的效应，其中一种效应偏爱一个较大的膜间距，而另一种则偏爱较小的膜间距，不过结果是一个稳定的折中位置。两种相互抵制的效果结合起来自然地导致了一个两膜模型，在这个模型里，两个膜之间就有适当的距离。

马克和戈德伯格的论文让人们明确了有着弯曲空间的两膜图像确实是等级问题的解决方案。把两膜之间的距离固定下来之所以重要，还有另外一个原因：如果两膜之间的距离没有确定，随着宇宙演变过程中温度和能量的变化，两个膜可能靠近，也可能远离。如果膜间距会改变，或者说，如果五维宇宙的不同边界会以不同的速度膨胀，那

么，宇宙的发展就不会是我们看到的四维世界应有的样子。因为天体物理学家已检验了宇宙发展后期的膨胀，我们最近已知道，宇宙膨胀就表现为它好像是四维的。

使用戈德伯格-怀斯的稳定机制，弯曲的五维宇宙就与宇宙学观察相吻合了。一旦两膜的相对距离固定下来，那么即便宇宙实际是五维的，其发展也会如四维一样。即使有第五维度，但稳定性对于沿着第五维的不同地方有着严格的限制，使它们必须以相同的方式发展演变，这样，宇宙的表现就像它在四维里一样。因为戈德伯格-怀斯的稳定机制可能出现得相对更早些，因此在大部分演变过程中，弯曲宇宙看起来就像是四维一样。

一旦人们领会了稳定性和宇宙学的意义，等级问题的弯曲几何解决方案便开始发挥功效。随后，相继涌现出了其他多种关于弯曲几何的有趣发展，其中一个就是力的统一。在我们探讨的弯曲几何里，所有的力，包括引力，在高能量上都可能统一起来。

弯曲几何与力的统一

在第13章里，我解释了超对称获得青睐的一个重要原因是，它能成功地容纳力的统一，解决等级问题的额外维度理论似乎要摒弃这一有重大潜力的成果了。既然我们没有找到任何结论性的实验证据，例如质子衰变，那这未必就是一个重大损失，因为我们还不肯定统一是正确的。但不管怎么说，三条线交于一点总是神奇的，而且可能预示了某种有意义的东西。即便统一还未被完全确立，我们也不应急于放弃它。

巴塞罗那大学的西班牙物理学家亚历克斯·波马罗尔（Alex Pomarol）发现，在弯曲几何里也会出现力的统一。但他所探讨的情形稍有不同：电磁力、弱力和强力不是被限制在膜上，而是出现在整个五维

体空间里，标准模型的规范玻色子——胶子、W子、Z子和光子没有固定在3+1维的膜上。

根据弦理论，规范玻色子可能被困在一个高维的膜上，或者与引力一样，也可能在空间里。引力子肯定是由一个闭合弦形成的，而规范玻色子和带荷的费米子则不同，它们对应的既可能是开弦，也可能是闭弦——这取决于模型。根据它们的形成来源，规范玻色子和费米子要么被困在膜上，要么可以自由地穿行于体空间中。

在大额外维度的图景里，倘若非引力的力出现在体空间里，它们会过于微弱，不符合我们的观察。在体空间中，力会在一个庞大的空间里散播，因此，就与引力一样，它们会被大大地削弱，而这是不可接受的，因为我们已测知力的强度要远远大于这一理论所作出的预言。

但是，如果像在弯曲几何里一样，额外维度并不大，那么非引力的力很可能会出现在五维体空间中，能够削弱它们的只有额外维度的大小，而不是弯曲——在弯曲图景里，额外维度的尺寸很小。这就意味着，描述世界的真正理论很可能让整个体空间经受所有4种力。这么一来，不仅仅是膜上的粒子，在体空间里的粒子也能够感受到电磁力、弱力和强力，同时还有引力。

如果弯曲图景里的规范玻色子出现在体里，那么它们的能量要远远大于1 TeV，经常来往于体空间的规范玻色子会经受过由低到高的各种能量强度。它们不再被锁在弱力膜上，可以在体里四处穿行，它们所拥有的能量可以高达普朗克级能标。只有在弱力膜上，能量才必须小于1 TeV，因为力将会在空间里，因此可以在高能量上运行，力的统一则成为可能。

这是令人振奋的，因为它意味着在高能量上力能够统一起来，即使是在一个有额外维度的理论里。波马罗尔发现了这一有趣的结果，统一真的出现了，就好像理论是真的四维的一样。

但是，事情发展还在向更好的方向前进：**统一和弯曲等级机制可以结合起来**。波马罗尔证明了力的统一，但他认为超对称解决了等级问题。但等级问题的解决只需要希格斯粒子被束缚在弱力膜上，以使它的质量约等同于弱力级质量——在100 GeV和1 TeV之间，而规范玻色子则不必被困在上面。

在弯曲几何里，等级问题的解决需要保证希格斯粒子保持在较轻的质量水平上，这是因为希格斯场导致了自发对称破缺，而这是所有基本粒子质量的来源，只有弱力对称被打破，规范玻色子和费米子才会拥有质量。只要希格斯粒子具有弱力级质量，弱力规范玻色子的质量最终总会正确。等级问题的弯曲引力解决方案实际上只要求希格斯粒子处在弱力膜上。

所有这些意味着，如果希格斯粒子在弱力膜上，而夸克、轻子和规范玻色子都在体空间里（见图20-6），那么你就既能拥有蛋糕，又可以享用它。这样，弱力标度就会得到保护，大约是1 TeV，但统一仍可以出现在高能量上——在大统一级别上。我以前的一个学生马修·施瓦茨（Matthew Schwartz）和我一起证实了超对称不是唯一能兼容统一的理论——弯曲的额外维度理论也能做到！

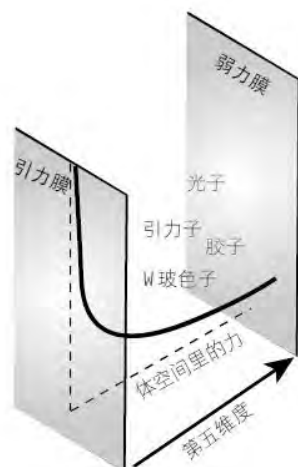


图20-6 体空间里的力。非引力的力也可以出现在体里，在这种情况下，力可以在高能量上统一起来。

实验意义

弱力膜上的自然标度大约是1 TeV，倘若这一弯曲几何图景被证明是我们世界的真实描述，那么瑞士CERN大型强子对撞机的实验结果将是极不平凡的。弯曲的五维时空的印记可能会包括卡鲁扎-克莱因粒子、反德西特空间的五维黑洞以及TeV质量级的弦。

弯曲时空的KK粒子可能是这一空间几何最容易得到的实验先驱，如前所述，KK粒子是有着额外维度动量的粒子，但这一模型的新特点是，因为空间是弯曲而非平坦的，KK粒子的质量将反映弯曲几何的独特秉性。

因为我们唯一确信能够在体里穿行的粒子是四维引力子，现在我们就集中考虑它的KK伙伴。正如在平坦空间里的情况一样，最轻的引力子KK伙伴将是那些在四维世界里根本没有动量的粒子，这个粒子与一个真正源于四维的粒子不可区分：它是能够在看似四维的世界里传递引力的引力子，也是我们在本章仔细研究过其概率函数的引力子。

如果没有额外的KK粒子，引力的表现将与它在真正的四维宇宙里完全一样。在这一图景中，宇宙暗地里是五维的，但表现如四维引力子的粒子不会显露出这一事实。如果没有更重的KK粒子，阿西娜的世界在她看来实际就与四维世界一样。

只有质量更大的KK粒子才能泄露五维理论的秘密，但它们又必须轻得足以生成。在这一理论里，要计算KK粒子的质量却有点棘手。由于几何与众不同，KK粒子将不会像在平坦空间的卷曲额外维度里那样，拥有与维度大小成反比的质量。这一质量将会极为令人惊讶，因为对于我们探讨的小额外维度来说，它将是普朗克级质量。在弱力膜上，任何重于1 TeV的东西都不能存在。因此，我们当然不可能在那里发现有普朗克级质量的东西。既然1 TeV是与弱力膜相关的质量，那么当你将弯曲时空考虑在内，并正确计算出KK粒子拥有大约1 TeV的质量时，就不是什么意外了。像我们假设的那样，当第五维度在弱力膜上终结时，无论是最轻的KK粒子，还是依次加重的KK粒子之间的质量差，最终都会是1 TeV左右。大量KK粒子累积在弱力膜上（因为它们的概率函数在这里达到高峰），它们就拥有了弱力膜粒子的所有属性。

这意味着引力子会有一些重KK伙伴，它们的质量大约为1 TeV、2 TeV、3 TeV等，依据LHC最终能达到的能量，我们极有可能找到其中的一个或多个。与大额外维度图景里的KK伙伴不同，这些KK伙伴相互作用的强度远比引力大得多。

这些KK粒子并不像四维引力子那样只是微弱地相互作用——它们的作用强度要高出16个数量级。在我们的理论里，引力子KK伙伴的作用如此强烈，以至于对撞机产生的任何KK伙伴不仅会立即逃出我们的视线、挟走能量，而且不会留下任何可见信号。相反，它们会在探测仪里衰变成我们能够探测到的粒子，可能是 μ 子或电子，我们可以用它们来重建派生它们的KK粒子（见图20-7）。

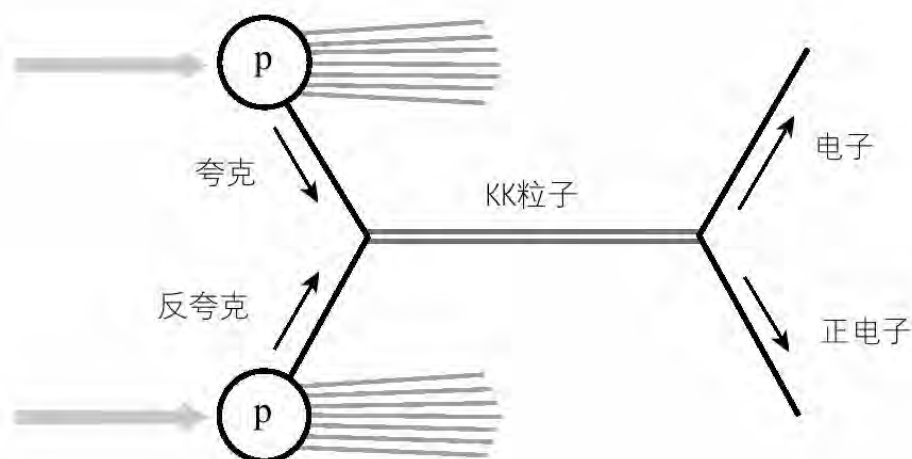


图20-7 KK粒子的生成示意图。两个质子对撞，一个夸克和一个反夸克互相湮灭生成一个引力子的KK伙伴。KK粒子接着衰变成可见粒子，比如一个电子和一个正电子，灰线是质子对撞喷射出的大量粒子。

这是发现新粒子的惯常方法：研究所有的衰变产物，并由此推导出这些产物的来源。如果你发现的东西是你以前不曾知道的，那么它一定是新事物。如果KK粒子在探测器里发生衰变，那么额外维度存在的迹象将非常明显。在我们的模型里，出现的将不仅是能量流失的迹象——它没有任何确定能量流失根源的标志性意义，也不会让我们把这一模型与其他可能区分开来；真正有用的线索是被重建的KK粒子质量和自旋，它们会告诉我们许多关于新粒子身份的信息。

KK粒子的自旋值为2，这将是一个真正的身份证明，它将告诉我们新粒子与引力有关系。一个自旋为2、质量大约为1 TeV的粒子将是额外卷曲维度的有力证据。很少有模型会预言这么重的自旋为2的粒子，即便作出这样的预言，也会有其他明显不同的特点。

如果我们足够幸运，那么除了引力子的KK伙伴之外，实验还可能会生成众多其他的KK粒子，在一个标准模型的大多数粒子处于体空间的理论里，我们还可能看到夸克、轻子和规范玻色子的带荷KK伙伴。

这些粒子将带荷且很重，而且它们最终会告诉我们有关高维世界的更多信息。事实上，模型构建者乔鲍·萨基（Csaba Csaki）、克里斯托夫·格罗琴（Christophe Grojean），路易吉·皮罗（Luigi Pilo）及约翰·唐宁（John Terning）都发现，在一个有着标准模型粒子的额外维度弯曲时空里，即使没有希格斯粒子，弱电对称仍可能被打破，届时实验将发现的带荷粒子就会告诉我们，这一替代模型是否与我们生活的真实世界相符。

一个更加离奇的可能性

我已描述了额外维度的许多离奇属性，但还有更加不同寻常的可能。很快，我们就将看到一个卷曲的额外维度实际上可以无限延伸，但与平坦维度不同，它仍旧是不可见的，而平坦维度若要与我们的观察相符，其大小则必定是有限的。

这一结果真令人无比惊讶，在第22章中，当我们讨论无限大的额外维度时，我们将集中探讨空间的几何形状，而非等级问题。但在这里，我要简要介绍一下，在额外维度无限的图景里同样能解决等级问题。

到目前为止，我们已探讨了一个两膜的模型：引力膜和弱力膜，两者各居第五维度的一个边界。但是，弱力膜却并不一定是世界的尽头（即第五维度的边界）。如果希格斯粒子被限制在位于一个无限大额外维度中央的另一个膜上，这样的模型也能够解决等级问题。

引力子的概率函数在弱力膜上将会很小，引力也很微弱，等级问题将与弱力膜是额外维度边界时一样得以解决。在一个有无限大的卷曲维度模型里，引力子的概率函数会在弱力膜之外继续延伸，但这不会影响等级问题的解决，因为它只依赖于弱力膜上引力子概率函数的极小值。

但是，由于维度是无限的，KK粒子会有不同的质量和相互作用，因此这一模型的实验意义就与我刚刚描述的模型不同。在阿斯本物理中心（一个激发灵感的集会地点，也正是这一原因，许多物理学家都愿意去那里徒步旅行），当我和乔·莱肯第一次谈论它的可能性时，我们并不确信这一观点能否发挥实际的作用。

如果第五维度没有在弱力膜上终止，那么并非所有的KK粒子都会很重（拥有大约1 TeV的质量），有些KK粒子的质量会很小。假设这些粒子可以探测，而实验却没有发现它们，那么这个模型将会被排除。

但我们的模型最终将是安全的，坐在公园的长椅上，望着周围美丽的湖光山色，我算出了KK粒子的相互作用（乔也做了同样的计算，不过，我想他一定是在研究中心的办公室算出来的）。这一结果告诉我们，尽管KK粒子的相互作用会很大，在未来的实验里足以让人产生兴趣，但它们还没大到让人们早该发现的地步。

假设这一模型的KK粒子存在，LHC将来极有可能生成它们，这些粒子看起来不会像大小有限的卷曲额外维度模型里的粒子。无限额外维度模型的KK粒子与那些能在探测器里衰变的KK粒子不同，它们会逃进额外维度（与大维度里KK粒子的表现类似）。因此，如果存在能够解决等级问题的无限大卷曲额外维度和一个弱力膜，实验只能希望发现能量流失的迹象。即便如此，在足够高的能量上，流失的能量也足以向我们暗示有新事物出现。

五维黑洞与五维弦

当LHC投入运行之后，除了KK粒子以外，还有可能发现额外维度的其他一些非凡征象。尽管五维引力作用在常见能量上是微不足道的，但当对撞机生成高维粒子时，五维引力将成为主角。事实上，当能量达到1 TeV时，五维引力的作用将是强大的——它们会淹没四维引力子

的微弱相互作用，在我们生活的弱力膜上（也即进行实验的地方），四维引力子存在的概率将很小。

高强度五维引力意味着，除了五维的弦以外，五维黑洞也有生成的可能。一旦能量达到1 TeV左右，位于弱力膜上或靠近弱力膜的所有东西彼此之间将发生强相互作用。这是因为，在TeV级能量上，引力和额外KK粒子的效果将是巨大的，它们将一起来促使所有事物彼此间相互作用。这种发生在所有已知粒子和引力间的强相互作用不会出现在四维图景里，它们绝对是新事物出现的信号，就如大额外维度情形一样，我们还不知道是否能达到足够高的能量，让我们看到这些新物质。如果相互作用在比1 TeV更大一些（不要高出太多）的能量上强烈发生，肯定能得到实验的验证。

尾声

等级问题的解决与TeV级能量上的实验结果有着很强的联系，但我们将看到的细节还要取决于模型。不同模型会产生不同的实验结果，这一点确凿无疑。

-
- 即便膜本身是完全平坦的，但如果有体空间和膜能量出现，时空将急剧弯曲。
 - 本章我们探讨的模型有两个膜：引力膜和弱力膜，分别处于有限大小的第五维度的两端。空间和膜上的能量使时空弯曲。
 - 单一的额外维度引进了等级问题的一个全新的解决方法，在这一模型里的额外维度并不大，却是极度弯曲的。引力的强度强烈依赖于你在第五维度上的位置，引力在引力膜上强而在弱力膜上弱，我们就位于弱力膜上。



- 由一个认为自己在四维世界的观察者的角度来看，如果物体来自第五维度的不同地方，它们也应有不同的大小和质量，被局限于引力膜上的物体会很重（大约为普朗克级质量），而被局限于弱力膜上的物体要轻得多，质量大约是1 TeV左右。
 - 如果希格斯粒子（而不是规范玻色子）被局限在弱力膜上，那么所有的力都可以统一，而等级问题也能得到解决。
 - 引力子的卡鲁扎-克莱因伙伴会引起非常奇特的粒子碰撞事件，它们会在探测器里衰变成标准模型粒子。
 - 在标准模型粒子处于体空间的模型里，还会生成和发现其他KK粒子。
-



去问爱丽丝吧，那时她有3米高。

杰斐逊飞机乐队

第五维度的旅行A

当阿西娜走出梦幻世界电梯进入弯曲的五维世界时，她吃惊地发现这里居然只有3个空间维度。难道兔子是在捉弄她，假装把她带进了一个有4个空间维度的世界，而实际上却只有三维？到一个看上去与平常没什么两样的世界里旅行，还要靠这种方式，这可真有意思！^[73]

一位当地居民极为殷勤地接待了这位满脸疑惑的新访客：“欢迎来到膜城^[74]，这是让我们觉得无比荣耀的首都，请允许我带领您参观。”阿西娜非常疲惫，又迷惑不解，她脱口说道：“膜城看上去也没什么特别呀，就连市长也与平常的市长没什么两样嘛。”不过，她不得不承认，她实际上并不十分确定，因为她以前从未见过市长。

阿西娜说的市长早就在首席顾问柴郡肥猫的陪同下来了。肥猫的职责是监管这个城市的一些事务，这对他来说易如反掌，因为他的出现总让人出其不意——鉴于肥猫体型庞大，这尤其令人吃惊。肥猫喜欢这样对人们解释：他具备这一技能是因为他可以消失在空间里，可没有一个人明白他的意思。^[75]

肥猫出现在阿西娜身边，问她是否愿意陪他四处转转，并警告阿西娜说，她最好能适应庞大的空间。阿西娜赶忙回答说，她最喜欢的舅舅实际也是非常非常胖

的。肥猫看上去有点儿不相信，但仍同意带上她，他给了阿西娜一块抹了黄油的冰淇淋蛋糕，阿西娜很高兴地吃掉了。然后，他们出发了。

阿西娜感到奇怪：她刚才吃掉的究竟是什么？现在，她感觉自己似乎是在一个五维世界的四维切片上，她一点儿也不比这个薄薄的四维切片厚多少，她惊呼：“我怎么像我的纸片娃娃一样了！只不过多莉是在三维世界里的两维纸片，而我呢，在一个四维空间里，却只有三个空间维度。”

肥猫意味深长地笑了笑，解释道：“现在你已感觉到了我总喜欢称之为‘体’的东西，你仍旧在膜城，但很快就将离开（并长大）。膜城实际是五维宇宙的一部分，但是，第五维度的弯曲如此隐秘，以至于膜城的居民完全没有意识到它的存在。膜城就在五维国的边境上，对此他们一无所知。当你刚刚到达这个城市时，你也得出了一个错误的结论，认为只有三个空间维度。全新的阿西娜，你已脱离了膜的束缚，能自由在第五维度旅行了，现在我可否提议，让我们前往目的地——一个叫作弱力膜的村庄？它就在五维宇宙的另一个边界。”

去往弱力膜村

这是多么奇异的五维之旅啊！离开膜城之后，阿西娜发现自己来到了另一个维度，而且，随着她的移动，她竟在迅速长大（见图21-1）。^[76]一旁观察的肥猫注意到阿西娜脸上疑惑的表情，他宽慰道：“弱力膜就在附近，我们转眼就到。^[77]那是一个可爱的地方，可是如果弱力膜上的居民因为你说有4个空间维度而嘲笑你，不要像你遇到的膜城居民那样大惊小怪。你呢，因为能看穿整个体空间，所以能看到在膜城有一个庞大的影子，比你开始的时候要大1亿亿倍。而其他几乎所有事情在你和他们看来都完全正常。”

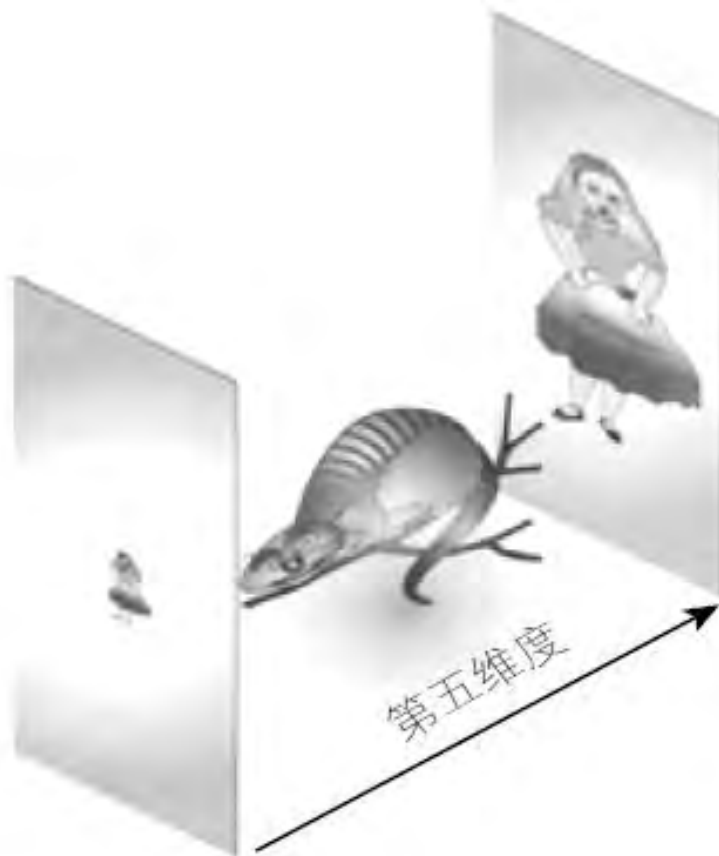


图21-1 弯曲的五维世界。

在穿过体空间由引力膜移向弱力膜的过程中，阿西娜变得越来越大。

但阿西娜一到弱力膜，她就注意到了另一件事：四维引力子悄悄地陪伴了他们一路，现在正轻轻地拍打她的肩膀。他的动作这么轻柔，以至于她差点没有注意到。^[78]

但当引力子唠唠叨叨地发出一连串的抱怨时，阿西娜是不能充耳不闻的：“如果没有顽固的等级强权的影响，弱力膜本该是令人振奋的。全副武装的强力、弱力和电磁力在弱力膜上只允许我拥有最微弱的力量。”引力子不住地抱怨，诉说他在其他所有地方是如何被对待

的，尤其是在膜城，那是一个有着几个不相上下的强力的寡头社会。
[79]而在弱力膜上，引力最受压制，所以这是引力子最不喜欢的地方。
[80]引力子转而问阿西娜，希望她加入他的计划，从统治权威手里把权力夺过来。

阿西娜想，她最好还是赶紧离开，于是她四下寻找兔子洞，却没能找到。终于，她看到了一只小白兔，原指望他能成为一个好向导呢，可弱力膜兔子迈着那令人惊讶的慵懒步伐，不断地重复着生活多美好，后面的日子又有多长。阿西娜意识到这只兔子是哪儿也不会去的，因此，她只好又找到一只心急的兔子，跟随他终于找到了回家的路。一旦她领会了其中的物理学含义，阿西娜就非常喜欢她的梦境了——但应该注意的是，她再也没有吃到冰激凌蛋糕。



从另一个维度，怀着偷窥的意图，让我们再做弯曲时间的游戏。

瓦妮莎（Vanessa）《洛基恐怖秀》（*The Rocky Horror Picture Show*）

重返膜城

阿西娜惊醒过来，她那重复出现的梦境再次把她带进兔子洞。但这次，她请求兔子直接把她送回弯曲的五维世界。

阿西娜重返膜城（或者说，她认为是这样），肥猫很快出现，她急切地求助于他，期望能得到他的梦幻蛋糕，再到弱力膜上进行一次愉快的旅行。可肥猫告诉她，在这一宇宙里，没有弱力膜这样一个地方，这令她感到无比失望。^[81]

阿西娜不相信肥猫的话，她认为在远处肯定有另一个膜。她自恃很了解在弯曲几何里远处的膜上引力较弱，因此自作聪明地想，或许它该叫作“微弱膜”吧。她问肥猫她是否可以去那里。

但她再一次失望了。肥猫告诉她：“没有这么一个地方，你就在膜上，再没有别的膜了。”“越来越奇怪了。”阿西娜想，显然这和以前不是同一个地方了，因为这里只有一个膜。但阿西娜不想这么快就放弃。“我可不可以自己去看看，是否有别的膜？”她请求道，用上了自己最甜美的声音。

肥猫警告她不要这么做，他说：“在膜上的四维引力可并不保证在体里也是同样的四维引力，有一次，我几乎把小命丢在那里，差点就回不来了。”

虽然阿西娜是一个好奇心很强的女孩，有过多次数历险，可她还是听从了肥猫的劝告，将它牢记于心。但她常常想，他是什么意思呢？在膜之外，究竟有些什么？她怎么才能知道呢？

弯曲时空有着非凡的属性，我们在第20章已探讨了一些，包括质量、大小及引力强度是怎样依赖于位置的。在本章中，我将给大家展示弯曲时空一个更为非同寻常的特征：即便世界实际是五维的，它也可以表现得只有四维。在进一步仔细研究了弯曲的时空几何之后，我和拉曼惊讶地发现：一个额外维度即使可以无穷大，也仍可以不被我们发现。

本章中我们将探讨的时空几何与第20章中描述的几乎完全一样，只不过正如故事里讲的那样，这一几何有个明显的特征：它只有一个膜。但这是一个异常重要的特点：因为没有另一个膜作边界，这就意味着第五维度是无限延伸的（见图22-1）。

这是一处明显的差别。自1919年西奥多·卡鲁扎提出空间的额外维度后，近100年以来，物理学家相信额外维度是可以接受的，但它们的大小一定是有限的，它们要么卷曲，要么被限制于两膜之间。无限大的额外维度很容易就被排除了，因为如果引力在这样的维度上无限远地扩散，在所有的距离尺度，甚至我们熟悉的尺度，它看上去都会是错的。人们猜想无限大的第五维度会破坏我们周围所有事物的稳定性，包括由牛顿定律束缚在一起的太阳系。

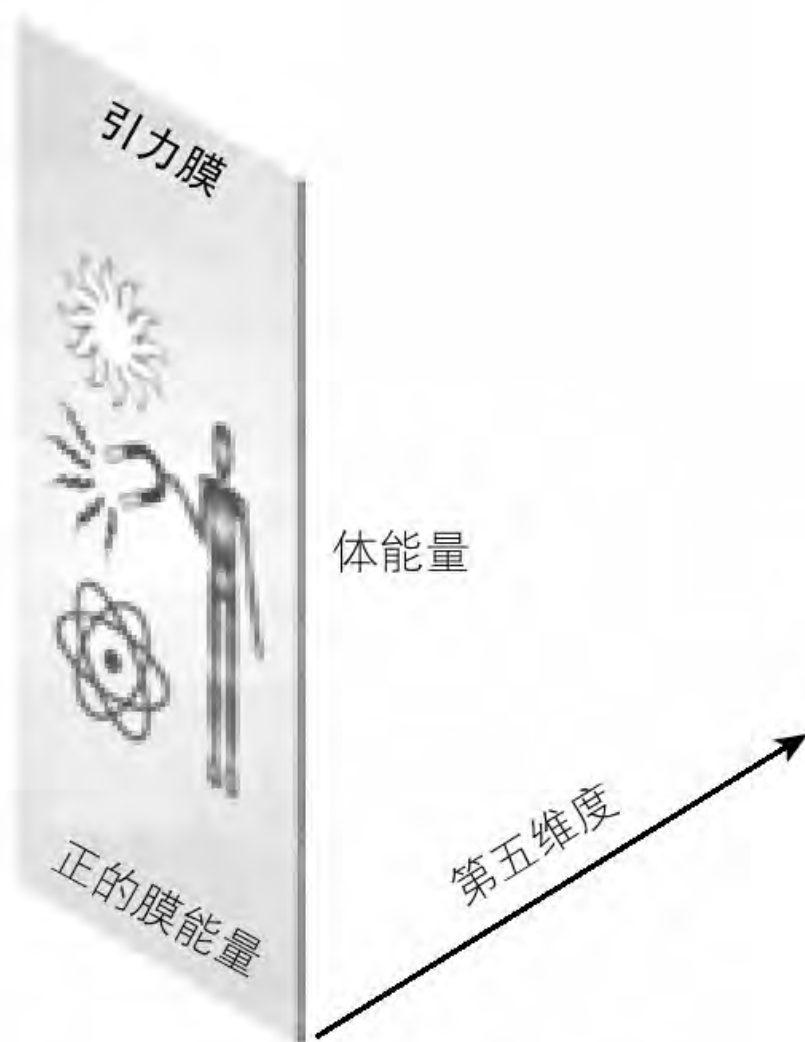


图22-1 只有一个膜的无限弯曲时空。在五维宇宙里，只有一个四维膜，标准模型就居住在这个唯一的膜上。

本章将解释为什么这一推理并不总是成立的。1999年，我和拉曼发现了一个全新的理由来解释额外维度是怎样隐藏的：时空可以极度弯曲，使得引力场高度集中于靠近膜的一个小区域内——因为非常集中，以致一个无限维度的无限延伸也无关紧要。引力并没有迷失在额外维度里，而是聚集于靠近膜的一个微小区域内。

在这一图景里，传递引力的粒子——引力子，就被限制于膜的附近，那就是阿西娜故事里的膜，我称之为引力膜。阿西娜的梦将她带入了这个弯曲的五维空间，其中引力膜彻底改变了时空的性质，从而使时空显得只有四维，虽然它实际上有五维。值得注意的是，一个卷曲的额外维度可以无限延伸，却仍能不被发现，而其他三个无限延伸的平坦维度展示了我们这个世界的物理现实。

改变时空的引力膜

可能你还记得，在第一次讲到膜时，我讲明了不愿意探索遥远的地域与真的被拘禁的区别——真的拘禁是明确禁止你离开被禁锢之地到别的地方旅行。尽管你可能从未去过格陵兰岛，但没有任何法律禁止你到那儿去，只不过我们要去一些地方很麻烦。即使没人禁止我们去这些地方，即使这些地方并不比我们曾去过的其他地方更遥远，但我们仍可能永远也不会去。

再假设一个人的腿受了伤，从理论上讲，只要他想出门，他随时都可以出去，既没有锁也没有栅栏会把他困在家里，可更多的时候，我们还是看到他待在家里。

同样的道理，局域的引力子可以不受限制地到达无穷大第五维度的任何地方，可它仍高度集中在膜的周围，很少出现在远处。根据广义相对论，所有事物——包括引力子，都服从引力。引力子丝毫没有削弱，可它的行为像是受膜的引力吸引，因此总是在膜的附近。因为引力子很少活动到一个有限的区域，所以额外维度可以是无限的，而又不会产生任何危险的效应将理论排除在外。

在我和拉曼的研究中，我们关注的是只有一个额外空间维度的五维时空，因此我们可以集中研究我们即将讨论的局域化机制，它使引力集中在五维时空的一个小区域内。我将假定，如果宇宙有10个或更

多维度，那么局域化和卷曲的某种组合隐藏了其余的维度。那些额外的隐藏维度不会对我将描述的局域化现象产生影响，因此我将忽略它们，而只集中关注对我们的讨论至关重要的5个维度。

在我们的模型里，只有一个膜静止于第五时空维度的一端。它和我在第20章里描述的那两个膜一样，具有反射作用——撞到膜上的东西都会被弹回。因此，任何东西撞到膜上都不会丢失能量。我们现在讨论的模型只包含这一个膜，我们假设标准模型粒子就被限制在这个膜上。注意这一模型与上一章讨论的模型的差别：在上一章的模型里，标准模型粒子是在弱力膜上，而在这里，弱力膜不复存在。标准模型粒子的位置对于时空几何来说无关紧要，但对于粒子物理学来说，它却具有重要意义。

虽然在本章中我们感兴趣的是只有一个膜的理论，但我和拉曼得到的可能存在无限的第五维度的线索，却是两膜弯曲几何的奇特特征。起初我们假设第二个膜有两个作用：一是束缚标准模型粒子；二是让第五维度有一定限度。与平坦额外维度一样，有限的第五维度保证了引力在足够大的距离上会表现为四维时空的样子。

但是，一个奇怪的事实却让我们想到，第二个膜的后一个作用实际上转移了我们的注意力，要模拟四维宇宙的引力，第二个膜不是必需的。

四维引力子的相互作用与第五维度的大小实际上没有任何关系。计算显示，第二个膜无论是在原位置也好，或者离引力膜两倍甚至是10倍距离也好，引力的强度是不变的。事实上，即使我们的模型把第二个膜推至无穷远——也就是说，干脆让它消失，四维引力也将维持不变。如果第二个膜和一个有限的维度是重现四维引力所必需的，那这将根本不可能发生。

这是我们的第一个线索，我们需要第二个膜的直觉是基于平坦维度的，对于弯曲时空未必正确。在一个平坦额外维度，第二个膜对于四维引力是必须存在的。由第20章里洒水装置的比方我们就可以看清这点，一个平坦额外维度对应的是水沿着一个长长的直管均匀地洒向各处（见图20-4），^[82]水管越长，浇到每个花园的水就越少。把这一推理引申出去，假设一个水管无限长，我们会看到，水量被分散得更为稀疏，实际上，在一个面积有限的花园里，根本没有浇到水。同样地，如果引力沿着一个无限延长的不变的维度发散，在这一额外维度上，它将被无限减弱到零。若要引力有四维表现，一个无限额外维度的几何必须要有超越这幅简单的直觉图像的微妙特征。事实上，弯曲时空就给出了这一必要的附加成分。



我们再回到洒水装置的比方上，来看它是怎样运作的，并以此找到上述论点的漏洞。假设你的水管无限长，而你却并非均等地把水浇到所有花园，相反，你可以控制水的分配，保证自己的花园得到充足的水量。要做到这样，你可以把一半的水浇进自己的地里，而剩下的另一半水量分给其他花园，这样，虽然离得较远的花园受到了不公平待遇，但你的花园总能保证得到需求的水量。无论水管延伸到多长，将水洒向无限远的地方，你的花园总能得到一半的水量。有了这种不公平分配，你就能得到你所需要的水量，而水管是无限延伸的，你并不知道它的长度。

同样地，在我们的弯曲几何里，尽管第五维度是无限的，但引力子的概率函数在靠近引力膜的地方总是很大。与上一章里一样，引力子的概率函数在这个膜上达到了高峰（见图22-2），并随着远离引力膜进入第五维度的距离指数式地下降。然而，在这一理论里，引力子的概率函数会无限地延伸下去，但这对靠近膜处引力子概率函数的值不会有任何影响。

像这种陡降的概率函数告诉我们，在远离引力膜的地方找到引力子的可能性是极小的，以至于我们可以忽略第五维度的遥远区域。尽管从理论上讲，引力子可以出现在第五维度的任何地方，但指数级陡降使得引力子概率函数高度集中于引力膜附近，这一情形几乎就像（却又不完全是）有第二个膜把引力子局限在一个有限的区域里。

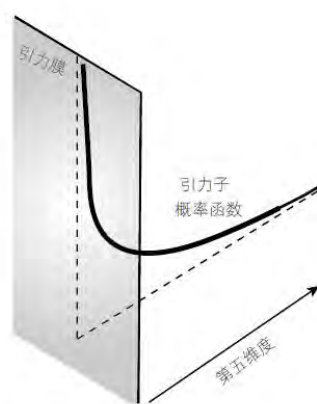


图22-2 引力子在只有一个膜的无限弯曲时空里的概率函数。

在引力膜的附近发现引力子的概率很高，与此对应的是引力场在那里高度集中。我们可以用池塘边密集的贪嘴野鸭来比喻这一情形：通常野鸭并不会均匀地散布于整个池塘里，而常常是聚集在边上，因为那里有爱鸟人投放的面包屑（见图22-3）。因此，对野鸭的分布来说，池塘大小实际上根本无所谓。同样地，在弯曲时空里，引力把引力子吸引在引力膜周围，因此，第五维度的长度实际上也是无关紧要的。

通过想象在引力膜上一个物体周围的引力场，你也能明白为什么第五维度对引力不会产生很大影响。我们已看到，在平坦维度空间里，由一个物体发出的引力场会向所有方向发散。而如果额外维度有限，引力场线先是向所有方向延伸，然后会有一些到达边界，产生弯曲。出于这一原因，如果离一个物体的距离超过了额外维度大小，引力场线就会只沿着低维世界的三个无限维度发散。

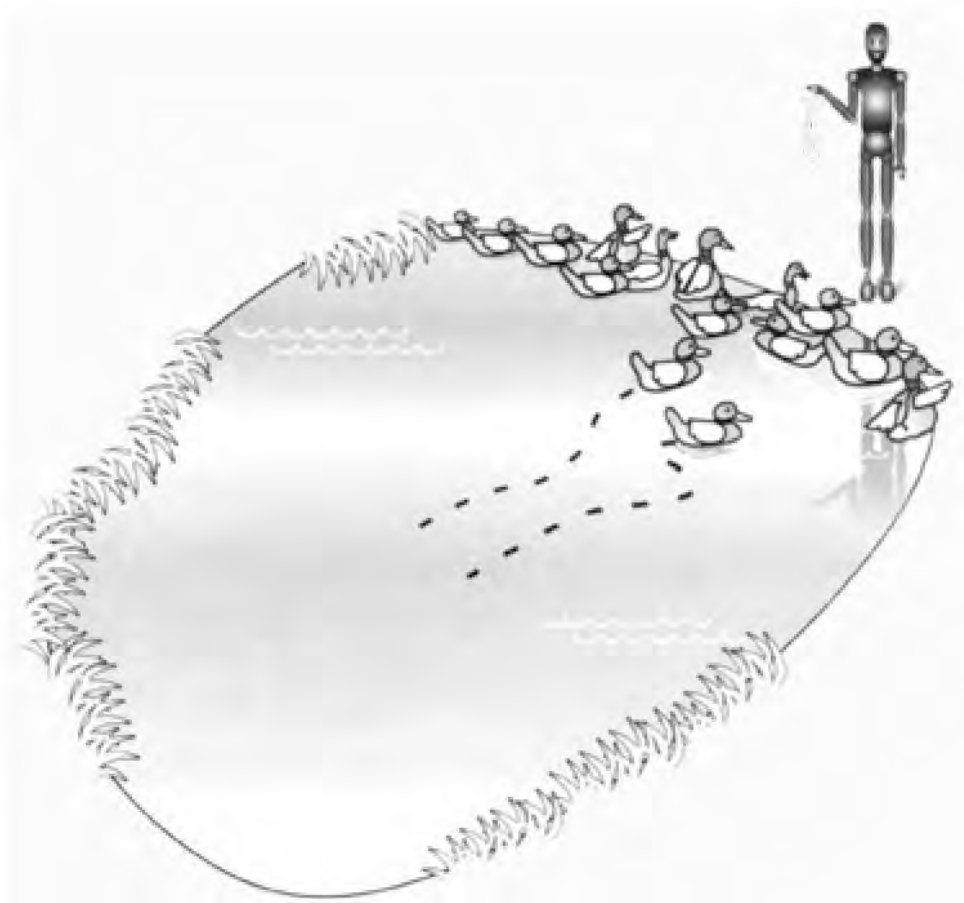


图22-3 聚集在池塘边的野鸭。如果野鸭都集中在岸边，你只要数一数周围的野鸭几乎就等于把所有的野鸭都数全了。

而在弯曲图景里，引力场线不会在所有方向上均等分布，只有在膜上，它们才向所有方向均等延伸。在垂直于膜的方向上，它们几乎没有延伸（见图22-4）。因为引力场线主要沿膜散播，引力场看起来就几乎与一个四维物体发出的场是一样的，它在第五维度的发散非常

小（不会比普朗克级长度 10^{-33} 厘米大很多），我们甚至可以将它忽略。尽管第五维度是无限的，但它对局限在膜上的物体的引力场无关紧要。

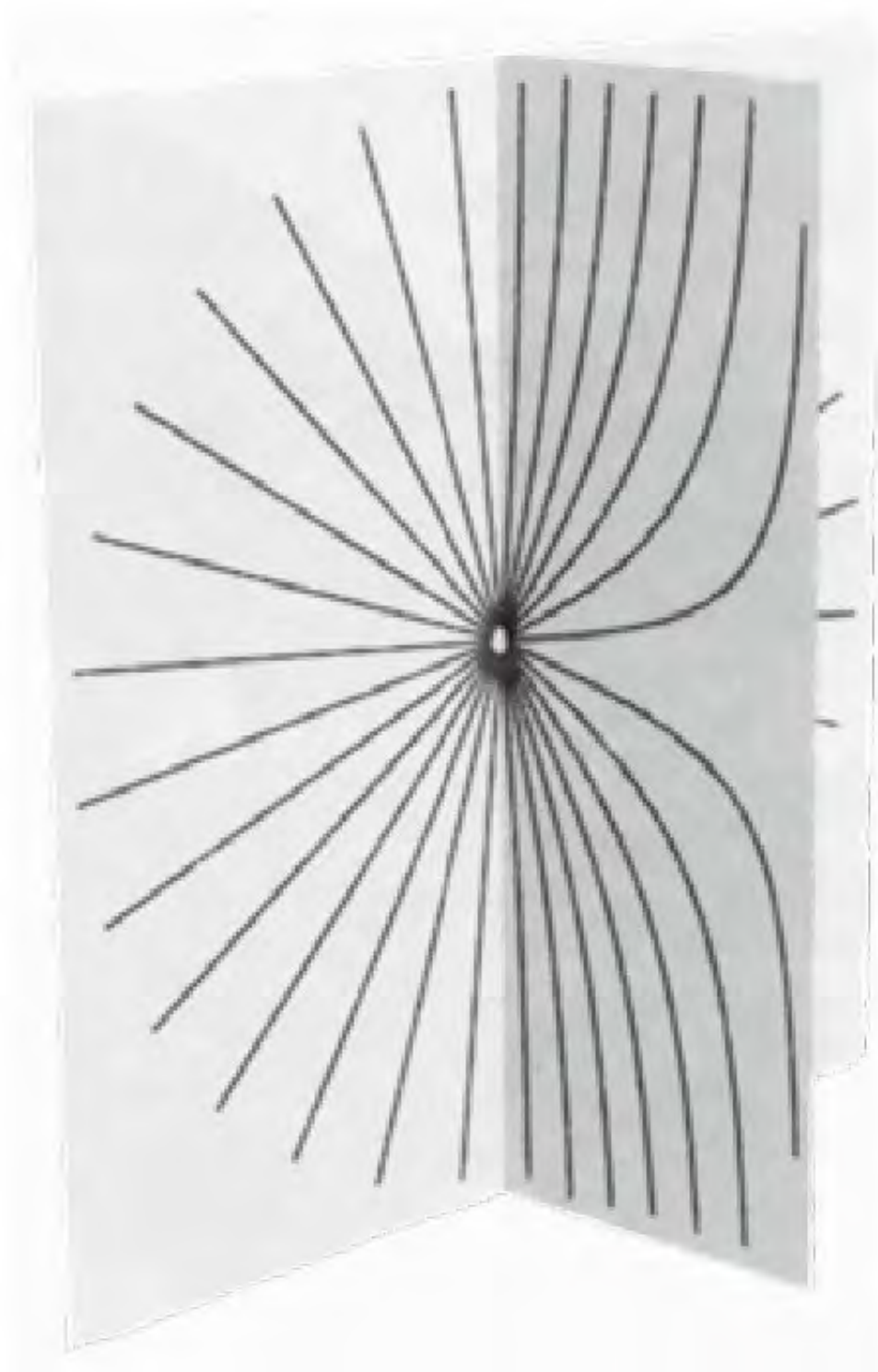


图22-4 弯曲图景里引力场线的分布。在弯曲图景里，引力场线在膜上向所有方向均等发散，但离开膜的引力场线会弯向膜的方向。这样，它们基本上与膜平行，就好像第五维度是有限的。即便是维

度无限，引力场也只是集中于膜的附近，而引力场线的分布就好像只有4个（时空）维度。



你还可以看到拉曼和我是如何解开这个谜题的：为什么第五维度的大小对引力强度无关紧要？回到上面洒水装置的例子，假设我们使水在整个洒水装置里的分配与陡降的引力子概率函数所产生的引力分布类似：在为你的花园浇了一半的水之后，剩下的水再分一半送到邻近的花园，再把剩下的水分一半给下一个花园，依此类推。这样，每个花园得到的都是上一个邻近花园得到的水量的一半。为了模拟第五维度上的另一个膜，我们假设到达某个地点后就停止供水，就如第2个膜会在第五维度的某个点上截断引力子概率函数；而为了模拟无限的第五维度，我们则假设水会无限远地沿着水管的长度继续浇下去。

为了证明在膜附近引力强度与第五维度大小无关，我们就要证明：不管我们在第5个或第10个花园开始停止供水，抑或我们根本不停止供水，那么最近几个花园得到的水量几乎不会因供水的距离受到影响。因此我们设想，如果供水装置在浇完最近的5个花园后停止供水会是什么情形。因为第6个花园之外得到的水量原本就已非常少，所以浇

到最近几个花园的总水量与水管无限延伸浇在这几个花园里的水量只有小量的差别；而如果你在第7个花园之后停止供水，则更是相差无几。用我们这种方式分配水量，几乎所有的水都被浇在了最近的几个花园里，远处的花园，只得到极小比例的水量，这对于近处几个花园得到的水量是无关紧要的。^[83]

在下一章，我将再次用到野鸭的比方，因此，我将以数鸭子的方式来解释同一件事情。有人在池塘边投食面包屑，野鸭会被吸引到岸边，如果你先数周围的鸭子，然后再数稍远一点的，很快，你再继续数几乎就没什么意义了。等到你数离岸边稍远一点的鸭子时，已没几只让你来数了。你根本不需要远离岸边继续数下去，因为只集中数岸边的这些鸭子，根本不用你再费劲了（见图22-3）。

引力子的概率函数在第二个膜之外非常小，以至于第二个膜的位置对四维引力子相互作用强度造成的差异几乎是可以忽略的。换句话说，在这一理论里，由于引力场只集中于引力膜附近，即便没有第二个膜，第五维度的长度对四维引力的强度根本没有影响。而第五维度是无限的，引力看起来仍是四维的。

我和拉曼称我们的图景为局域化引力，这是因为引力子的概率函数只集中在膜的附近。尽管严格来说，因为第五维度是无限的，引力可能泄漏进入第五维度，但由于在远处发现引力子的概率非常低，实际上它并没有这样。空间并未被截断，但所有东西仍集中在膜附近的局部区域。既然引力膜上极少有东西会伸向远处，那么一个遥远的膜对引力膜上的物理过程也就不会有任何影响。引力膜及其附近产生的东西将集中在附近一个局部区域之内。

物理学家通常称这一局部引力模型为 RS_2 ，“RS”代表的是兰道尔和桑卓姆，但“2”就容易让人产生误解——它实际指的是我们有关弯曲空间的第二篇论文，而不是指有两个膜。针对等级问题的有两个膜的图景，被称作 RS_1 （如果我们写作两篇论文的顺序颠倒过来，这名称就不会让人产生迷惑了）。尽管你也可以如第20章后半部分简要介绍的那样，引进第二个膜解决等级问题，但与 RS_1 不同，本章的图景于等级问题未必重要。但是，无论在空间中是否存在着针对等级问题的第二个膜，局域化引力都是一种非常激进的可能，它对长期以来人们认为额外维度必须是卷曲的假设提出了反驳，这有着重要的理论含义。

引力子的卡鲁扎-克莱因粒子伙伴

在上一节，我们讨论了引力子的概率函数，它高度集中于引力膜上。我讲的粒子发挥的是四维引力子的作用，因为它几乎只沿着膜穿行，极少会泄漏进入第五维度。从引力子的角度来看，第五维度对它好像只有 10^{-33} 厘米这么大（由弯曲决定的大小，而弯曲又是由空间和膜上的能量决定的），而不是无限延伸。

虽然我和拉曼对我们的发现感到非常兴奋，但是否彻底解决了问题，我们却并不十分确信。局域化的引力子是否足以产生一个四维有效理论，使引力表现得就如在四维一样？潜在的问题是，引力子的卡鲁扎-克莱因伙伴也有可能形成引力，且由此可能大大地改变引力。

这看起来很危险，通常来讲，额外维度越大，最轻KK粒子的质量越小。对一个有着无限维度的理论来讲，这意味着最轻的KK粒子可能会呈现出任意轻的质量。又因为KK粒子的质量差也会随着额外维度的增大而减小，因此在任一有限的能量上有可能生成无限多种很轻的引力子KK伙伴，所有这些KK粒子都有可能影响引力定律并使它发生改

变。这个问题看起来尤其糟糕，因为即便每个KK粒子相互作用都很微弱，但如果太多KK粒子，那么引力作用仍旧会与四维引力看起来大不相同。

最为重要的是，因为KK粒子极轻，所以它们很容易生成。对撞机运行的能量已足以生成它们。即使常见的物理过程，如化学反应，都可能产生足够的能量形成引力子的KK伙伴，如果KK粒子携带过多的能量进入五维体空间，那么这一理论将被排除。

幸运的是，所有这些担忧最终都没有成为问题。当我们计算KK粒子的概率函数时，我们发现引力子KK伙伴在引力膜上及其附近的相互作用极其微弱。尽管有大量的引力子KK伙伴，但是它们的相互作用都非常微弱，因此不会生成太多的危险，也不会在任何地方改变引力定律的形式。如果说有什么问题，那就是由于这一理论极近似地模拟了四维引力，我们还不知道怎样以实验方法将它与真正的四维世界区分开来。引力子KK伙伴对任何可见现象的影响都可以忽略不计，因此我们没有办法区分4个平坦维度与4个平坦维度再加一个弯曲的第五维度的差别。

由引力子KK伙伴概率函数的形状，你就可以看出它们之间的相互作用有多么微弱。与引力子的概率函数一样，它告诉我们在第五维度的任一位置上发现一个粒子的可能性有多大。在我们的弯曲几何里寻找每个引力子KK伙伴的质量和概率函数时，我和拉曼遵循的基本上是标准的程序，这包括要解决一个量子力学问题。

对于一个平坦的第五维度，量子力学问题（如第6章所述）就是要找到适合卷曲维度的波，使得允许的能量量子化。^[84]对我们的弯曲而无限的第五维度几何而言，量子力学问题显得大不相同，因为我们需要把使时空产生弯曲的膜上和空间里的能量都计算在内。但我们能够修改通常的程序使其适合我们的结构，结果是引人入胜的。

我们发现的第一个KK粒子在第五维度中没有动量，这个粒子的概率函数高度集中在引力膜上，而离开它时则呈指数级下降。这一形状听起来似曾相识：它与我们曾探讨过的四维引力子概率函数是同样的，这个无质量的KK模式就是传达四维牛顿引力定律的四维引力子。

但是，其他KK粒子便不同了。在引力膜附近很难找到这些KK粒子，相反，我们会发现对应在零质量和普朗克级质量之间的任一值，都存在着具备这一特定质量的KK粒子，其中每个粒子的概率函数都会在第五维度的不同地方达到高峰。

事实上，对峰值的位置不同有一个有趣的解释。我们在第20章中看到，在弯曲时空里，为了把所有粒子都置于四维有效理论的同地位，以使它们以同样的方式与引力相互作用，我们需要把第五维度上处于不同位置的所有距离、时间、能量和动量都按比例重新标度。当一个人远离膜旅行时，与每个点相联的都是一个呈指数级变小的能量。这就是在弱力膜上的粒子质量大约应是1 TeV的原因。阿西娜进入第五维度旅行，随着她从引力膜向弱力膜的移动，她的影子越来越大，而她越来越轻。

第五维度上的每个点都能以同样的方式关联一个特定质量，这一质量都通过那一点的重新标度与普朗克级质量相联。概率函数在某个特定点达到峰值的KK粒子，就近似具有那个重新标度的普朗克级质量。当你进入第五维度旅行时，你会陆续遇到概率函数在那里达到峰值的、依次减轻的KK粒子。

事实上，你可能会说卡鲁扎-克莱因图谱展示的是一个高度隔离的社会。重KK粒子在空间的某些区域被排除了，在那里，重新标度的能量太小，不能生成它们；而在那些包含了非常活跃的、高能量粒子的区域，轻KK粒子极少出现。依据它们的质量，KK粒子集中在离引力膜尽可能远的地方，它们的位置就像是十几岁男孩的长裤，只要裤子

掉不下来，口袋越多越好。好在确定KK粒子位置的物理学定律要比当今孩子令人迷惑的时尚更容易理解。

对我们来说，轻KK粒子概率函数的最重要特征是，它们在引力膜上极其微小。这就意味着，你在膜及其附近找到轻KK粒子的可能性极小，因为轻KK粒子会尽可能地避开引力膜，因为在那里将很难生成轻粒子（除了概率函数在引力膜达到峰值的个别引力子之外）。而且，轻KK粒子并不能大大地改变引力定律，因为它们倾向于远离引力膜的地方，因此并不太与束缚在膜上的粒子发生相互作用。

把所有这些串连起来，我和拉曼确定我们发现了一个有效的理论。

集中在引力膜上的引力子导致了四维引力的表现。尽管引力子KK伙伴数量充足，但在引力膜上它们的作用非常微弱，以至于它们的效应根本不会被人注意到。尽管存在着一个无穷大的第五维度，但所有物理定律及过程，包括引力定律，看上去都与四维世界应有的现象一致。在这个高度弯曲的空间里，一个无穷大的额外维度是可能存在的。

如前所述，如果说存在什么问题，那就是从实验角度来看，这一模型让人颇受打击。虽然它看上去令人称奇，但这个五维模型对四维的模拟好得异乎寻常，以至于我们很难将它们区分开来。粒子物理学的实验者们当然将面临巨大的挑战。

然而，物理学家已开始探索能够区分两个世界的天体物理学和宇宙学特征。许多物理学家已探讨了弯曲时空的黑洞，他们仍在继续研究是否存在可以区分的特征，以利用它们来确定我们生活的宇宙究竟是哪一种类型。

迄今为止，我们知道局域化是额外维度在我们宇宙的一种新奇可能。我热切地盼望进一步的发展能够最终确定它是不是我们这个世界

的真实特征。

-
- 如果时空呈现恰当的弯曲，那么一个维度可以无限长但不被发现。
 - 虽然引力并没有被严格地局限在一个特定区域，但它仍会集中于局部。
 - 在局域化的引力理论里，无质量的KK粒子就是局域化的引力子，它集中于靠近引力膜的地方。
 - 所有其他的KK粒子都集中在离引力膜尽可能远的地方，它们的概率函数形状及达到峰值的位置会取决于它们的质量。

书籍免费分享微信jnztxy



-imension
探索大揭秘



我不知道会是哪一天，但总有一天，我们会到达我们想去的地方。

布鲁斯·斯普林斯汀 (Bruce Springsteen)

31世纪里的艾克四十二世

艾克四十二世已做好了变大的准备，他想测试Alicxvr上百万秒差距的超高设置，他能用它来探索银河系以外的地方和整个已知宇宙，到达以前从未有人见过的遥远太空。

Alicxvr把他带到了90、120甚至是130亿光年之外，艾克感到非常刺激。但当他试图飞得再远些时，他的兴奋感消失了，因为信号强度陡然减弱。他把目标设定在150亿光年，但他的探索彻底失败了。因为他根本接收不到任何信息，相反，他听到的是：“信息5B73：您试图接通的视界客户超出了您的寻呼范围，如果需要帮助，请联系当地长途接线员。”

艾克简直不感相信自己的耳朵，这已经是31世纪了，而他的视界服务提供的却仍是有限的覆盖范围。他尝试着与接线员联系，但听到的却是一个录音：“请待在膜上，按顺序等候答复。”艾克怀疑接线员可能永远也不会答复，他才不想傻傻等下去呢。

上一章我解释了为什么弯曲能解放额外维度，允许它无限延伸而不被发现。但是，无穷大的维度并非故事的结局，事实上事情更加离

奇。在本章，我将解释四维（即空间三维加时间一维）引力可能真的是一个局部现象——在远处，引力看起来可能大为不同。我们会看到，不仅空间可能在确有五维的情况下看上去是四维的，而且还有可能我们就生活在五维宇宙中一个孤立的有着四维引力的口袋里。

我们现在将讨论的模型显示，尽管看起来很奇特，可空间的不同区域可以有不同数量的维度。在研究局域化引力的一些令人费解的特征时，我和物理学家安德烈亚斯·卡奇发现了一个时空模型，其中就有这样的情形。我们最终得出的全新、激进的图景让人想到：我们看不到额外维度的原因，可能是因为我们环境的特殊，这依赖性可是人们原来所没想象到的。也就是说，我们可能生活在一个四维溶洞里，而三维空间只不过是一个位置的巧合而已。

回顾

回顾和拉曼合作以来的E-mail记录，我感觉有点儿难以想象，因为中间有那么多分神的杂事，我们究竟是如何完成研究的。研究开始的时候，我正从麻省理工学院去往普林斯顿大学，到那儿做教授。同时，我还在筹划第二年在斯塔芭芭拉举行的为时6个月的研讨会。而拉曼，由于已做了好几期的博士后，他急于获得一个正式的教职，当时正忙于准备谈话和工作申请。真的令人难以相信，他已做了大量工作，我和其他人都试图说服他事情总会好转的，他不应该放弃物理转向其他职业——拉曼显然应该继续他的物理学研究，他绝对该得到一份优秀的教职，可他就是很难找到工作。

从那时起，在杂乱、有趣的物理学问题中间，还夹杂着推荐信请求、会议发言安排、普林斯顿房子的安置还有斯塔芭芭拉会议的组织，另有几封与其他物理学家往来的E-mail谈论我们的工作，却并不很多。尽管 RS_2 论文最终被千百次地引用，且已被广泛认可，但研究最

初接收到的反应却很复杂。过了很长一段时间，物理学家才最终领会其中的含义，并相信了我们。一个同事告诉我，起初人们在等待别人去发现漏洞，所以他们根本不必关注。当然，在普林斯顿大学，人们对于拉曼发言的反应顶多算得上是不温不火。

即便那些真正在听的人也未必立即就相信了我们，与弦理论学家安迪·斯特罗明格的一次谈话对我们有很大的启发，可现在他自嘲说，起初我们说的话他一句都不相信。幸运的是，当时他没有怀疑到根本不听我们讲话，或根本不与我们谈论的程度。

在物理界，有几个人从一开始就理解我们且相信我们是正确的，令我们感到无比幸运的是，史蒂芬·霍金就在其中——而且他毫不迟疑地让物理界听众共享了他的热情。我记得拉曼非常兴奋地告诉我，霍金在哈佛大学所做的著名演讲重点讲述了我们的研究。

另外还有几个人也在致力于相关观点的研究，但直到第二年秋天，在我们的论文发表了几个月后（从我们开始谈论起已是好几个月的事了），整个理论物理界才开始广泛关注起来。事实证明，1999年秋我在圣塔芭芭拉卡夫里理论物理研究所组织的为期6个月的研讨会提供了良好的机会，组织这次研讨会的人还有：戴维·库塔索夫（David Kutasov），芝加哥大学的一位来自以色列的物理学家；米沙·史夫曼（Misha Shifman），明尼苏达大学的一位俄裔粒子物理学家。

研讨会本来的目的是要让弦理论学家和模型构建者聚集在一起，共同探讨两个阵营在诸如超对称和强相互作用的规范玻色子理论等方面的研究。我们事先为研讨会做了详尽的计划，最终却是膜和额外维度的观点产生了巨大的轰动。尽管我们曾希望在弦理论学家和模型构建者之间能形成一种积极的合作关系，但组织之初我们却未料到在会议实际召开时，我们却讨论起了额外维度。

时间的选择实际很凑巧，它提供了一个绝佳的机会，让额外维度的观点变得鲜活起来，也让模型构建者、弦理论学家、广义相对论学家能共享前沿的研究成果。人们进行了激烈的讨论，其中一个主要话题是弯曲几何。最后，弯曲的五维几何让模型构建者和弦理论学家都认真对待起来。事实上，当人们共同合作，致力于相似问题如弯曲几何和其他观点的研究时，两个阵营的界线就变得模糊了。

后来，许多物理学家对弯曲几何的其他方面进行了研究，确立了一些联系并探索了一些细节，这使得局域化引力变得更加有趣了。尽管弦理论家最初对 RS_1 （有两个膜的弯曲几何）不以为然，只把它看作一个模型，但一旦他们开始研究，他们便发现了在弦理论里实现 RS_1 图景的方法。有关黑洞、时间演变的问题、相关几何的问题，以及与弦理论和粒子物理学观点的联系也成了科研的丰产领域。现在局域化引力已进入了不同领域的研究，而新观点仍在不断涌现。

我们的观点被接受之后，不再被认为是错误的了，但有些物理学家又开始走向另一极端：声称我们的理论根本不新鲜，一位弦理论学家甚至总结说，弦理论对卡鲁扎-克莱因模式影响的计算就是确凿的证据，证明我们的理论与弦理论学家早就在研究的一个版本是一样的。这正印证了科学界的一段笑谈：一个新的理论在被接受之前要经历三个阶段：起先它是错的；然后变得显而易见；最后有人会说别人早就发现了。但在这一事例中，当物理学家认识到弦理论的计算比他们想象得还要复杂时，确凿的证据最终化作了一缕轻烟——他们宣称的弦理论答案实际上并不正确。

事实是，融入弦理论的研究令我们所有人都无比兴奋，而且引出了许多重要的新见解。结果，局域化引力与当时大多数重要的弦理论成果都有着强烈的重合：弦理论的研究和我们的研究都包含了一个类似的弯曲几何。事实上，也许因为我们的研究并未直接挑战弦理论模

型，因此，相比模型构建团队，弦理论团队更快地接受和认可了我们研究的意义。虽说起初很凑巧，可这兴许表示我们找对了研究方向。而令人高兴的是，以后拉曼找工作再也没遇到什么困难（现在他是约翰·霍普金斯大学的教授）。

然而，仍有一些人对此持怀疑态度。我和拉曼提出的模型所引出的一些有趣的问题，有人立即回应：局域化会依赖于远处的时空形式吗？当有人试图在超弦理论里找到我和拉曼提出的那种几何例子时，远离膜的引力形式似乎成了障碍。

但这些是不是根本条件？我们想回答的另一个问题是，时空是否处处一定都是四维的？局域化引力使整个五维宇宙表现得就像只有四维引力，事情是否总能这样？或者一些区域看起来是四维的，而另一些区域表现却不同？如果引力膜不是完全平坦的，事情又会怎样？在一个有着不同几何的膜上，局域化会同样有效吗？这些问题是当地局域化引力要解决的，那也是我和拉曼发展的一个理论。

当地局域化引力

空间究竟有几个维度？我们真的知道吗？到现在为止，如果我们断言自己可以肯定地知道额外维度不存在，这未免有些自欺欺人，我想你也会同意我的看法。我们只看到三个空间维度，但实际可能会有更多，而我们未曾发现。

现在你知道额外维度可以隐藏起来，要么因为它们是卷曲的或很小，要么因为时空的弯曲，引力过于集中在一个微小区域，即使维度无限大都不曾被发现。不管是哪种情况，维度是卷曲的也好，是只集中于局部也好，时空表现出的处处都是四维的，无论你位于何处。

在局域化引力图景里，这可能不那么明显。在这一图景里，当你进入第五维度时，引力子概率函数会变得越来越小，如果在靠近膜的地方，引力的表现就好像它真的在四维世界一样。但在别处会是什么样子？答案是，在 RS_2 图景里，无论你位于第五维度的什么位置，四维引力的影响都是无法逃避的。虽然引力子概率函数在引力膜上最大，但所有地方的物体都是通过交换引力子相互作用的。因此，所有物体都会经受四维引力，无论它们在哪里。

所有地方的引力看起来都是四维的，因为引力子概率函数永远不会实际为零——它会永远继续下去。在局域化图景里，远离膜的物体会会有极其微弱的引力相互作用，但即便微弱，引力仍是以四维方式表现。因此，无论你在第五维度的什么位置，牛顿平方反比定律都会成立。

远离引力膜的小但不等于零的引力子概率函数，是我在第20章中提出的等级问题解决方法的根本。在体空间远离引力膜的弱力膜，经受着看似四维的引力，即使那引力非常微弱。这就如在洒水装置的对比里，远离你自家花园的地方，虽然水并不是很多，但总还是有少量的水供应。

假设我们再进一步细想，我们对空间维度确切地知道什么？我们并不知道空间处处都是三维的，在我们能看到的范围内，空间看起来有三维（时空四维），但空间可能延伸出这一范围，直达我们无法看见的领域。

毕竟，光速是有限的，而我们宇宙存在的时间也是有限的。这就意味着我们所能了解的周边空间区域也是有限的，其范围是光自宇宙起源之始所能到达的距离。这并非无穷远，这一区域被称作视界，它是我们能够知道和无法知道的信息的分界线。视界之外，我们一无所知，太空不必都如我们这里一样。随着观察到的距离越来越远，我们

认识到宇宙未必处处都如我们所见，哥白尼式的革命也在不断重复、更新。即使物理学定律处处都是一样的，这也并不意味着它们表演的舞台都一样，有可能膜在我们周围会引致一个与其他地方不同的引力定律。

我们怎能断言了解视界之外的宇宙维度呢？如果这之外的宇宙呈现了更多的维度（或者5维，或者10维，也可能还有更多）都不会出现任何矛盾。我们从最根本处思考，而不是假设所有地方——甚至包括我们无法到达的领域，都与我们的时空构成一样，我们就能推想出真正根本的是是什么，而最终可以想见的、合乎情理的又是什么。

我们只知道自己感知的空间看上去是四维的，如果由此认为宇宙的其他区域也都是四维的，未免有些离谱。一个离我们极其遥远的世界，有可能根本不与我们相互作用——或者有，也只是通过极其微弱的引力信号，那又为什么一定要与我们经历相同的引力和空间？它怎么就不可能有不同类型的引力呢？奇妙的就是，它当然可以。我们的膜宇宙是三维加一维的世界，而外面的区域未必如此。2000年，我和安德烈亚斯·卡奇提出了一个理论，在这一理论里，在膜上和靠近膜的地方，空间看起来是四维的，但远离膜的大多数地方看起来都是高维的（见图23-1），这令我们非常惊奇。

我们称这一图景为当地局域化引力^[85]，因为局域化的引力子只在当地一个区域传递四维引力作用——空间里的其他区域看上去不是四维的。一个四维世界只存在于一个引力“孤岛”上，你所看到的维数取决于你在五维体空间的位置。

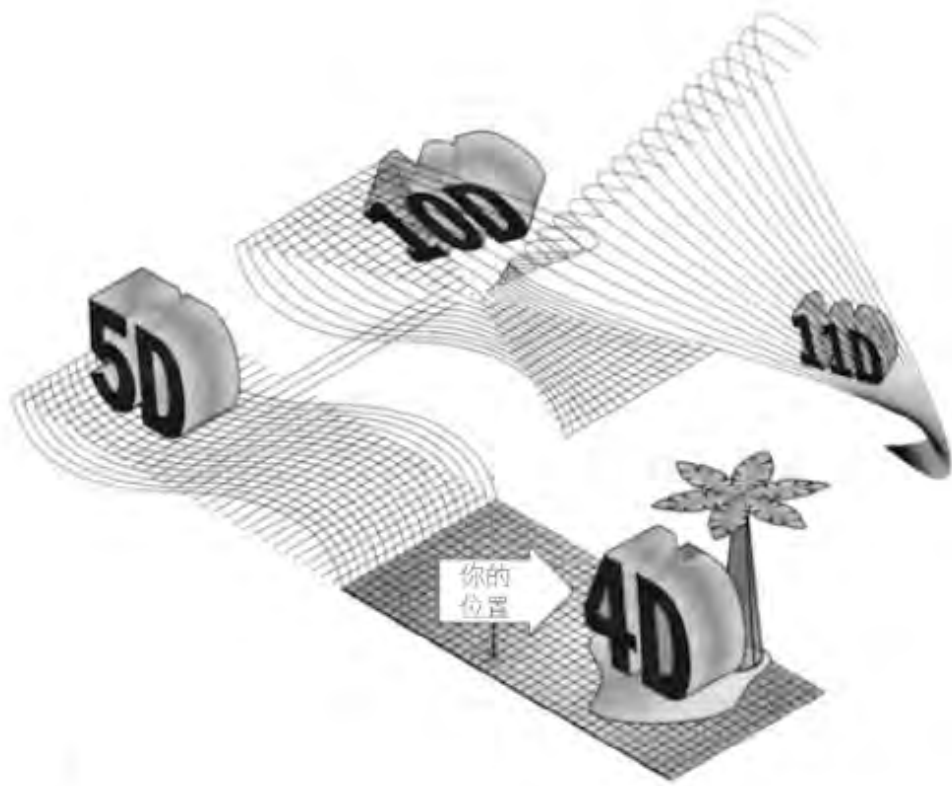


图23-1 我们可能居住在一个高维空间的四维溶洞里。

为了理解什么是当地局域化，我们再回到池塘野鸭的类比。当我说池塘的大小无关紧要时，你可能不以为然。的确，如果池塘真的很大，对岸的鸭子就不会与你这边的鸭子聚集在一起，事实上，如果你能影响到远处的野鸭反倒是很奇怪的。远处的野鸭不会注意到你投食的面包，辽阔的湖面上，它们在另一处遥远的地方闲游，根本不知道你的存在。

当地局域化引力，其根本与此类似：聚集于一个膜上的引力与发生在空间另一遥远区域的事无关。虽然在我和拉曼研究的模型里，有一个引力子概率函数呈指数级迅速衰减却永远不等于零——因此我们处处都能感受到四维引力，但远处的引力表现对膜周围是否存在四维引力的论证并不重要。这是当地局域化引力的根本。

引力子可以集中在一个膜的周围，形成四维引力作用，而不影响远处的引力。四维引力可以是纯粹的局域化现象，只与局部空间相关。

安德烈亚斯是一位优秀的物理学家，一个大好人。有意思的是，他是与我以前在麻省理工学院的一个同事合作进行一个科研项目时，开始考虑这一模型的，他们原意是要挑战我和拉曼的研究（结果令我们非常高兴，他们的合作成功地证明我们的研究成果是正确的）。在项目进展中，安德烈亚斯发现一个模型与我和拉曼创建的那个极为相近，却有着非常奇特的性质。安德烈亚斯在访问普林斯顿大学时，来找我讨论这一模型，最终我们发现它的含义是令人惊讶的。起初，我和安德烈亚斯通过E-mail合作，有时我们到彼此的学校互访，直到我回波士顿后，交流才多了起来，而我们的发现真的令人瞩目。

这一模型与我和拉曼研究的那个非常接近，也是在五维弯曲空间里只有一个膜，但有一处差别，在这一情形下，膜并不是完全平坦的。这是因为，它承载着小量的负真空能量。正如我们看到的，在广义相对论里，有意义的不仅是相对能量，还有总能量，总能量表明时空会怎样弯曲。例如，五维时空里恒定的负能量形成了时空的弯曲，这是我们在前几章讨论过的。但在那一模型里，膜本身是平坦的，而在此，膜上的负能量使得膜产生了略微的弯曲。

膜上的负能量催生了一个更为有趣的理论，但我们实际感兴趣的并非负能量本身——如果我们生活在一个膜上，理论要与观察相符，那么我们的膜应该有小量的正能量。我和安德烈亚斯决定研究这一模型，仅仅是因为它对维数的神奇意义。

为了领会我们的发现，让我们简要回顾一下两膜的构成，以后我们将把第二个膜剔除出去。当第二个膜足够远时，我们发现有两个不同的引力子，各自集中于一个膜的附近，每个引力子概率函数都在其

各自靠近的膜附近达到高峰，而随着你的离开，则呈指数级下降。这两个引力子都不能影响整个空间里的四维引力，它们只在自己聚集的膜的局部区域产生四维引力。在不同的膜上感受到的引力是不同的，它们甚至会有不同的强度，这个膜上的物体不会与另一个膜上的物体通过引力相互作用。

在这一情形里，有两个相距遥远的膜，这就好比在湖泊遥远的对岸也有一个人在喂野鸭一样。那些野鸭甚至可能是不同的品种：你引来的是绿头鸭，而对岸的那个人可能吸引的是木鸭。在这种情况下，对岸就会聚集另一群野鸭，这就好比集中于第二个膜附近的另一个引力子概率函数。出现了两个不同的粒子，且看起来都像是四维引力子，这让我们非常惊讶。通常的物理学原则应该保证只有一个引力理论，而实际上，也确实只有一个五维引力理论。但五维时空却包含了两种不同的粒子，各自在五维空间的不同区域传递着引力，其表现都像四维。两个不同的空间区域看起来都包含了四维引力，但在这些理论里传递四维引力的引力子却是不同的。还有另外一个让我们惊讶的地方：根据广义相对论，引力子是无质量的，就如光子一样，它应以光速行驶，但我和安德烈亚斯发现，其中的一个引力子是非零质量的，不以光速行驶。这真的令人吃惊——但也带来了麻烦：根据物理学文献，还没有一个有质量的引力子能够形成与我们的观察相符的引力。事实上，正如我们在第10章中讨论过的重规范玻色子的情形一样，相比无质量的引力子，有质量的引力子会有更多的极化方向。物理学家通过比较测得的两种不同的引力过程证明，过其他引力子的极化效应并不存在。这困扰了我们好长一段时间。

但模型却超越了传统的智慧。我们刚提出这一模型，纽约大学的物理学家马西莫·波拉提和牛津大学的伊恩·科根（Ian Kogan）、斯塔夫罗斯·莫斯普洛斯（Stavros Mousopoulos）及安东尼奥·帕派佐格罗（Antonios Papazoglou）就发现，在某些特定情形里，引力子

可以有质量，且仍能产生正确的引力预言。他们分析了理论的技术细节，说明了有质量的引力子为什么不符合我们观察到的引力过程的逻辑漏洞。而这一模型还有更为奇特的含义。现在我们考虑如果除去第二个膜会怎么样。结果，物理定律在剩余的那个膜（引力膜）上仍然表现为四维的，尽管有个无限的额外维度。引力膜附近的引力与RS₂模型的完全一样。对引力膜上的事物来说，单一的引力子传递着引力，引力看上去是四维的。但是，这一模型与RS₂有一个重要的区别。

在这一模型里，集中于膜附近的引力子并不能支配在整个空间里的引力，它之所以不同只是由于膜上的负能量。引力子不会与空间所有地方的物体相互作用，它只有在膜上及其附近产生四维引力。远离膜之后，引力看上去不再是四维的。

前面我曾说过，在高维体空间，引力一定无处不在，这似乎就产生了冲突。以前的说法没错，五维引力的确是无处不在的。不过，到现在为止我们所思考的其他额外维度理论里，物理都有四维阐释，而这一理论却不同，只有对膜上或靠近膜的事物，理论才是四维的。牛顿引力定律只在膜上及其附近适用，在别的地方，引力都是五维的。

在这一情形中，四维引力完全是局域化现象，只有在膜的周围才能感受到。由引力表现导出维数，取决于你在第五维度的什么位置。如果这一模型正确，我们必须生活在膜上才能经受四维引力。如果在别的地方，引力都将是五维的。膜是一个四维引力的溶洞——一个四维引力的孤岛。

当然，我们还不知道当地局域化引力是否适用于真实世界，我们甚至不知道额外维度是否存在，倘若存在，它们又变成了什么。但是，如果弦理论正确，就会有额外维度。如若这样，它们就会隐藏起来，要么通过紧紧卷曲，要么集中于局部（或当地局域化），要么就

是两者结合。许多弦理论学家仍相信卷曲才是答案，但由于弦理论产生了许多引力疑难，所以没有人能最后确定。我认为局域化是一个新的选择，当引力集中于局部时，就如卷曲维度的情形一样，物理学定律就当作额外维度不存在。由此，局域化的引力就又给我们增添了一个构建模型的工具，且又扩大了机会，让弦理论找到了一个与观察一致的现实版本。

当地局域化引力模型只集中于我们能明确证实的事物上，我喜欢这种方式。它只表明在我们能够验证的区域，宇宙看起来必须是四维的——而并未说明它必定是四维的。我们的3个空间维度可能只是一个位置的巧合，而这一观点还有待继续探索，但不同的空间区域可能会有不同数量的维度，也并非没有可能。毕竟，每当我们探索到以前未探测到的更小距离时，都会发现新的物理学理论。在大距离上可能也是一样：如果我们生活在膜上，谁知道膜之外会是什么样子？

-
- 局域化引力是一种当地现象，它不依赖于时空的遥远区域。
 - 引力的表现可以使世界看起来在不同区域有不同的维度，因为一个局域化的引力子并不一定会在整个空间中延伸。
 - 我们可能生活在一个孤立的时空口袋中，而它的表现是四维的。





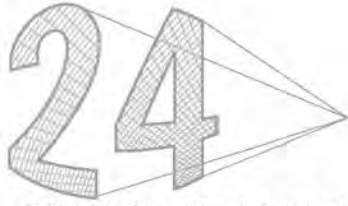
WARPED

第六部分

弯曲的旅行

PASSAGES

UNRAVELING THE MYSTERIES OF THE UNIVERSE'S HIDDEN DIMENSIONS



额外维度：你是在里面，还是在外面

WARPED PASSAGES

可我一直没有找到我要找的东西。

U2乐队

第五维度不见了

阿西娜有关一维世界、膜和五维世界的梦境经过了几代，一直流传到艾克四十二世那里。听到这些故事后，他决定确认一下，这其中是否有一定道理。于是他又搬出了他的Alicxvr，到极小的尺度下探索——不必小到让弦出现，只要小到让他足以看清是否存在第五维度即可。对于他的问题，Alicxvr的回答是，直接将他送进这个五维世界。

但艾克并不十分满意，他还记得以前当他胡乱扳动超空间驱动器时出现过的怪事，于是，他再次拉动操作杆——事情再次发生了急剧的变化。眼前熟悉的事物艾克统统找不到了，他只知道一件事：第五维度不见了。

艾克被搞糊涂了，因此他到太空网上搜索，看看关于“维度”都能找到些什么答案。他浏览了无数个网站，都是从那些更加令人难辨真伪的垃圾邮件里找到的。很快他就意识到，他必须要将搜索范围再明确一下。但搜了半天，始终没有找到一个明确的答案。最后他想，既然一时半会儿也弄不明白维度的根本来源，还不如转而关注时间旅行呢。

物理学进入了一个非凡的世纪，曾只在科幻小说中存在的观点，现如今已经进入理论甚至实验领域。有关额外维度的全新的理论发现，已不可逆转地改变了当今粒子物理学家、天体物理学家以及宇宙

学家对世界的思考。单从发现的数量和速度来看，对一些潜在的神奇可能，我们触及的或许只是其肤浅的表面，而每个观点都正处在自行发展成长中。

但是，许多问题还有待充分解答，我们的路途远未结束。粒子物理学家仍想知道，为什么我们只看到这几种力，是否还有其他的力？我们熟知的粒子的质量和性质的来源是什么？我们还想知道弦理论是否正确，如果正确，它又怎样与我们的世界联系起来？

最新的宇宙观察给我们提出了更多的谜题，这都是我们要面对的。宇宙中大多数的物质和能量是由什么构成的？在宇宙演变早期，是否有一个短暂的快速膨胀时期？如若有，是什么引起了这种膨胀？所有人都想知道，宇宙最初是什么样子。

现在我们知道，引力在不同的距离尺度上可以有非常不同的表现。在很小的距离上，只有量子引力理论，比如弦理论，才可能描述引力；在更大的距离尺度上，广义相对论完美地适用其中；但在宏大的距离尺度上，穿越宇宙的最新发现提出了一些宇宙学谜题，比如，是什么加速了宇宙膨胀？而在更大的距离上，我们则到达了宇宙视界。不过，超出视界之外的地方，我们一无所知。

额外维度理论的一个引人入胜的方面是，它们在不同尺度上会自然形成不同的结果。在这些理论里，引力展示了在小于卷曲维度尺度的小距离的行为，或在小到没有任何效应的曲率处的行为，这些行为不同于在维度不可见或卷曲很重要的更大距离上的行为。这使我们有理由相信，额外维度最终会帮助我们理解宇宙的一些神秘特征。如果我们确实生活在一个多维世界里，那么我们当然不会忽略它的宇宙学意义。这一课题已经有了一些研究成果，但我相信还有更多有趣的结果在等待着我们去发掘。

物理学将去往何方呢？有数不胜数的可能性。但是先让我来描述几个令人迷惑的发现，这将是一些很快就能接近答案的理论。这些谜题都围绕着一个问题，此时提出也许有点令人吃惊，这就是：究竟什么是维度。

究竟什么是维度

你肯定会很吃惊，我怎么会提出这样一个问题？本书的大部分内容都在讨论维度的含义，以及额外维度宇宙假说的潜在意义。但是，既然我曾讲过我们对维度的理解，那就请允许我再简要地回顾一下这个问题。

维数究竟有什么意义？我们知道，维数就是你在空间中确定一个点所需要的量的数目。但在第15章和第16章中，我还给出了几个例子，说明十维理论与十一维理论有时会有相同的物理结果。

这种对偶性表明，有关维度的概念并不像看起来那么严格——它的定义是有弹性的，这使它避开了传统的术语意义。同一个理论存在对偶性描述告诉我们，没有哪一个形式必定是最好的。例如，最好的描述形式甚至其维数，都可能取决于弦耦合的强度。因为没有哪一个理论总是能给出最恰当的描述，所以维数问题也并不总是有一个简单的答案。这种含糊的维度意义以及在强相互作用理论里明显涌现的额外维度，是这10年来最为重要的理论物理学现象。现在，我将列举几个更令人迷惑的理论新发现。它们表明，维度的概念比我们原来预想的更为模糊。

I. 弯曲几何与对偶性

第20章和第22章解释了我和拉曼发现的弯曲时空几何的一些结果。在这一弯曲几何中，物体的质量和大小要取决于其在第五维度的位置，而且引力只局限在膜的附近。但是，这个弯曲时空，在专业上叫作反德西特空间，还有一个更为令人迷惑的特征，这是我必须告诉你的——这一特征引起了有关维数的更深层次的思考。

反德西特空间的另一个显著特征是，它还存在着一个对偶的四维理论。理论线索告诉我们，发生在反德西特五维空间里的所有事情都可以用一个四维对偶理论来描述，而在这个四维理论框架里，有着性质特别的、极强的力。根据这一神秘的对偶性，五维理论里的所有东西都能在四维理论里找到一个类似物，反之亦然。

尽管数学推理告诉我们，在反德西特空间里的一个五维理论就等同于一个四维理论，但我们并不总能知道那一四维对偶理论里确切的粒子内容。但是，现在普林斯顿大学高等研究院的阿根廷裔弦理论学家胡安·马尔达西那（Juan Maldacena），1997年在弦理论中得到了一个类似于对偶性的明确例子，由此掀起了一轮弦理论热潮。他发现了一个有着大量重合D-膜的弦理论版本。

弦在D-膜上强烈地相互作用，既可以由一个四维量子场论来描述，也可以用一个十维引力理论来描述。在这个十维引力论里，其中五维卷曲，剩下的五维位于一个反德西特空间中。

一个四维理论和一个五维（或十维）理论怎么可能有相同的物理含义？比如，一个穿越第五维度的物体，它的等效物是什么？答案是：一个在第五维度上穿行的物体，在四维对偶理论里表现为一个放大或缩小的物体，这就如阿西娜在引力膜上的影子，随着她沿第五维度远离引力膜而不断长大。况且，在第五维度上互相超越的两个物体，在四维空间里对应的是两个物体增大、缩小，然后重叠。

一旦把膜引进来，对偶性的结果就显得更加奇怪了。例如，一个有引力却没有膜的五维反德西特空间等同于一个没有引力的四维理论，但是，一旦你把膜包括在五维理论里，正如我和拉曼所做的那样，这个等效的四维理论立即包含了引力。

那么，这种对偶性是不是表示，我提出的高维理论的弯曲几何是在骗人？当然不是。对偶性确实引人入胜，但它并不能真正改变我所讲过的事情。即便有人找到了一个对偶的四维理论，那这样一个理论也将极难研究：它必须包含大量的粒子，而其相互作用又极为强烈，根本无法使用微扰理论（见第15章）。

一个有着强相互作用的理论，如果没有一个替代的、弱相互作用的描述，几乎不可能得到解释。在这种情况下，这个较易驾驭的描述就是五维理论。只有这个五维理论具有足够简单的形式用以计算，我们由五维角度来思考这一理论才有意义。但即便五维理论更容易驾驭，对偶性仍令我想知道“维度”一词究竟是什么意思。我们知道维数应是你确定一个物体的位置所需量的数量，但我们是否总能确定地知道哪个量是该被计算在内的？

II. T对偶

还有一个原因让我对维度的含义存有疑问，那就是表面不同的两个几何之间的等效性，被称作T对偶。在发现我们讨论过的对偶性之前，弦理论学家早就发现了T对偶。它所交换的两个空间，一个有着极小的卷曲维度，而另一个却有着庞大的卷曲维度。尽管看起来很奇怪，但在弦理论里，卷曲空间的极小体积和极大体积产生的物理结果是一样的。

T对偶适用于有卷曲维度的弦理论，因为在紧缩成一个圆的时空里，存在两种不同类型的闭弦，当一个微小的卷曲维度空间与一个大卷曲维度空间交换时，这两种弦也会被交换。第一种闭弦在绕着封闭维度旋转时会上下振动，这很像我们在第18章里看到过的卡鲁扎-克莱因粒子的表现；而另外一种会缠绕卷曲维度，它可以在卷曲维度上缠一圈、两圈或几圈都有可能。T对偶能够交换大、小两个卷曲维度，因此也会交换这两种类型的弦。

事实上，T对偶是膜一定会存在的第一个线索，没有它们，开弦在对偶理论里就不会有类似物。但是，如果T对偶确实适用，且极小和极大的卷曲维度产生的是同样的物理结果，那么这就再次意味着我们有关“维度”的概念是不完善的。这是因为，如果设想一个卷曲维度的半径无穷大，那么T对偶的卷曲维度就会是一个半径为零的圆——圆根本不存在。这就是说，在一个理论里的无穷大维度与一个少了一维的理论T对偶（因为一个零尺寸的圆不能被算做是一个维度）。因此，T对偶还表明，两个明显不同的空间，其无限延伸的维数不同，却能作出相同的物理学预言。至此，维度的含义变得再次模糊起来。

III. 镜对称

当维度卷曲成一个圆时，T对偶适用，但还有一个对称甚至比T对偶更奇特，即是镜对称（mirror symmetry）。当弦理论的6个维度卷曲成卡拉比-丘流形时，往往就会用到它。根据镜对称，6个维度可以卷曲成两个完全不同的卡拉比-丘流形，而其形成的四维宏观理论却是相同的。

一个特定的卡拉比-丘流形的镜像看上去可以是全然不同的：它可能有不同的形状、大小、卷曲方式，甚至是洞的数量也不同。^[86] 当一个特定的卡拉比-丘流形存在一个镜像时，6个维度会卷曲成其中的任

何一个，物理理论都是同样的。因此，有了镜像流形，两个全然不同的几何产生的是同样的预言。时空再次有了神秘的属性。

IV. 矩阵理论

矩阵理论 (matrix theory) 是研究弦理论的工具，它提供了更神秘的有关维度的线索。从表面上来看，矩阵理论像一个量子力学理论，描述了在10个维度里穿行的Do-膜 (点状膜) 的表现和相互作用。但是，尽管该理论没有明确包含引力，但Do-膜的表现就如引力子一样，因此，即使引力子表面并不存在，该理论最终还是包含了引力作用。

再者，Do-膜理论模拟的是十一维的超引力，而不是十维的，这就是说，矩阵模型包含的超引力似乎就比原来理论描述的要多出一维。这种表现，再加上其他一些数学上的证据，使得弦理论学家相信矩阵理论就等同于M理论，因为M理论也包含了十一维超引力。

矩阵理论的一个尤为奇特的特点是，爱德华·威滕发现：当Do-膜彼此太过靠近时，我们无法明确地知道它们究竟在哪儿。正如矩阵理论的创始人——汤姆·班克斯、威利·菲斯彻勒 (Willy Fischler)、斯蒂夫·申克 (Steve Shenker) 和兰尼·萨斯坎德 (Lenny Susskind) 在他们的论文中所说：“因此，对微小的距离而言，其空间构形无法用寻常的位置来表示。”这就是说，当你试图明确地找到它时，Do-膜的位置根本不是一个有意义的量。

尽管这种奇怪属性使得矩阵理论看似值得研究，但目前将它用于计算仍很困难。问题是，如其他所有包含强相互作用物体的理论一样，还没有人能找到一种方法来解决这些重要问题，而这些问题将帮助我们更好地理解宇宙究竟发生了些什么。即便如此，由于额外维度

的出现，以及当D₀-膜太过靠近时维度的消失，矩阵理论也成了我们对维度含义很疑惑的又一个原因。

之后，我们该思考什么

尽管物理学家已从数学上证明，在不同维数的理论间存在着这些神秘的等效作用，但显然，我们还没有看到事实的全部。我们能确定这些对偶性会适用吗？如果适用，关于时间和空间的性质，它们又能告诉我们些什么？再者，当一个维度既不很大也不很小时（相对于异常微小的普朗克级长度），没人知道最好的描述是什么。或许，一旦我们试图去描述这么微小的东西时，我们的时空观念便彻底崩溃了。

在普朗克级长度上，我们的时空描述还远不充分。我之所以这样认为，其中一个最为强烈的理由就是，即便在理论上，我们都还想不出任何办法来探测这样微小的尺度。由量子力学我们知道，探测微小尺度需要消耗大量的能量，一旦你在微小如普朗克长度 10^{-33} 厘米的区域里投入太大的能量，你就会得到一个黑洞，然后你就无法知道里面究竟发生了些什么了——所有信息都被困在黑洞的视界之内。

最为重要的是，即便你试图投入更大的能量到那一微小的区域里，你仍不能成功。一旦你把大量能量加进普朗克级长度的范围内，这一区域就必然会膨胀——不然你无法加进更多的能量。也就是说，如果你增加能量，黑洞也会变大。因此，实际你不是在制造一个探索研究那一微小尺度的精良工具，而是在让那一区域膨胀，你永远也无法既让它很小同时又能对它进行研究。这就好像你要用激光束来研究博物馆里的一件珍贵的工艺品一样，激光不仅不能帮你探查其中的奥秘，反而会把工艺品灼伤。即使在物理学的思想实验里，你也永远无法看到比普朗克级长度小很多的区域：不等我们到达那里，我们熟悉

的物理学定律便已失效。在普朗克标度附近，传统的时空观念几乎肯定不能适用。

事实如此离奇，我们急需一个更为深入的解释。由这10年令人费解的发现，我们得到的一个最为重要的信息就是，时间和空间一定还有更为基本的描述。对这一问题，爱德华·威滕给出了一句简要的总结：“空间和时间可能是要消亡的。”许多顶尖的弦理论学家也同意这一观点，内森·塞伯格说道：“我几乎可以肯定，时间和空间都是一种幻觉。”而戴维·格罗斯则这样设想：“很有可能空间甚或时间都有各自的组成，它们可能最终成为一个全然不同的理论的自发特征。”不幸的是，有关时空这一根本描述的性质究竟会是什么，至今还没人知道。但是，深入了解空间和时间的根本性质，显然仍是物理学家今后几年将面临的最为艰巨也最具吸引力的挑战。

结语

即将被证实的宇宙真相

正如我们所知，这是世界的终结（我感觉很好）。

美国另类摇滚乐队REM

大团圆结局

艾克四十二世利用他的时光穿梭机回到了过去，他警告艾克三世，如果他继续驾驶保时捷，后面等待他的将是灾难！这一来自未来的客人让艾克三世感到无比震惊，他乖乖地听从了艾克四十二世的警告。他将保时捷卖掉，换了一辆菲亚特，从此过上了一种怡然自得的生活。

能再次与哥哥团聚，阿西娜狂喜无比，看到自己的好朋友迪特尔也非常高兴。可是两个人都感到有点儿纳闷：艾克好像从来就没有离开过呀！阿西娜和迪特尔觉得，艾克对他们讲述的时间旅行纯属虚构。即使在梦里，肥猫也从不曾在时间里轮回，兔子也从未到达过一个有额外时间维度的站点，而量子侦探也会拒绝思考如此古怪的时间行为。但阿西娜和迪特尔喜欢圆满的大结局，因此他们权且放下疑惑，欣然地接受了艾克的幻想故事。

这些年来，虽然物理学的发展令人赞叹，但我们仍不知该如何驾驭引力，也不知该如何用意念隔空取物。或许，在额外维度的研究上大量投入还为时过早。我们不知道该如何将我们生活的宇宙与能在其中进行时间轮回的宇宙联系起来，所以也没人能制造出一台时间机器，而且很有可能无论将来还是过去，都没有人能制造得出来。

虽然像这样的观点还只停留在科幻领域，但我们生活的世界着实是充满了神奇和秘密。我们的目的是要弄清楚：其中的片段是如何连接起来的？它们又是如何发展到目前这种状态的？我们还未弄明白的联系都是些什么？像前几章里我提出的问题，它们的答案又是什么？

即便物质的最根本来源还有待我们进行更深入的理解，但我希望你相信，在我们实验能探测的尺度内，我们已经掌握了自然许多方面的根本性质。即使我们还不知道时空的根本元素，但在远大于普朗克长度的尺度上，我们确实认识了它的属性。在这些领域里，我们可以使用我们已掌握的物理学定律，推导出我已描述的各种结果。有关额外维度和膜，我们遇到了许多出乎意料的特征。这些特征有可能会发挥重要的作用，帮助我们解开宇宙的一些秘密。

额外维度拓宽了我们的视野和想象，让我们看到更多新奇的可能。现在我们知道，额外维度的构成可以以任意多种形状和大小出现：它们可以有弯曲的额外维度，也可以有大额外维度；它们有可能是一个膜，也有可能是两个膜；它们可能让一些粒子在体空间中，而让另外的粒子被束缚于膜上。宇宙比我们能想象的任何东西都更为庞大、更为丰富，也更为多样。

究竟哪一个观点能描述我们的真实世界？我们只能等待真实世界来回答。精彩的是，它可能会给我们提供答案。我描述的这些额外维度模型，最令人激动的性质就是，它们会产生实验结果。这一令人瞩目的事实，其意义无论怎么强调都不过分。额外维度模型（我们原以为它们要么根本不可能，要么根本不可见）能够产生我们看得见的结果，而从这些结果中，我们可能推导出额外维度是否存在。如若这样，我们有关宇宙的观点将不可避免地被改变。

在天体物理学或宇宙学中，额外维度可能会得到验证。物理学家正在发展额外维度宇宙中黑洞的详细理论，研究发现它们尽管与四维黑洞属性类似，却仍存在一些细微的差异。额外维度黑洞可能最终会表现出明显不同的属性，足以让我们分辨出差异。

宇宙学发现还可能告诉我们更多关于时空结构的东西。当今实验探索的是宇宙几十亿年前的样子，许多发现与预言相符，但也存在一些重要的问题。如果我们生活在一个高维宇宙里，它在其早期一定大不相同，其中的一些差异就可能帮助我们解释一些令人费解的现象。物理学家正在研究额外维度对宇宙学的意义，我们有可能了解隐藏在其他膜上的暗物质，或隐藏的高维体空间所贮存的宇宙能量。

LHC是一个巨大的赌注——对物理学家来说，这再好不过了。几乎可以肯定的是，LHC实验会发现新的粒子，它们的性质会给我们一些超越标准模型的新见解。更令人激动的是，没人知道这些新粒子会是什么。

在我从事物理学研究的过程中，我们发现的新粒子一直都是那些理论研究已告诉我们必然会发现的粒子。我并不是不重视这些发现——它们都是令人惊叹的成就，但如果能发现一些全新的、未知的东西将更加令人激动。在LHC投入运行之前，没人真正确信自己集中努力的方向是什么，来自LHC的结果将可能改变我们对世界的看法。

LHC将有足够的能量生成新的粒子类型，它们有望揭示许多秘密。这些粒子可能是超对称伙伴或四维模型预言的其他粒子；也有可能是卡鲁扎-克莱因粒子——穿行于额外维度里的粒子。是否能看到、何时能看到这些KK粒子将完全取决于我们生活的宇宙有多大以及是什么形状。我们是否生活在一个多重宇宙里？这一宇宙的大小和形状会让我们看到KK粒子吗？

针对等级问题的所有模型在弱力级别上都将产生可见的结果。弯曲几何的印记尤其神奇：如果这一理论正确，我们将探测到KK粒子，并由它们留下的线索测知其属性。相反，如果其他额外维度模型正确地描述了世界，那么能量会消失在额外维度里，而我们最终可以通过不平衡的能量探测到那些维度。

当然，我们还无法得知所有的答案，但宇宙的大门很快就将打开：天体物理学对宇宙的探索将比以前更远、更早，也更详尽。在从未探测到的更小尺度上，LHC的发现将告诉我们物质的本质。而在高能量上，我们对宇宙真相的认识也已经要爆满了。

宇宙将慢慢揭开其神秘的面纱，而我，迫不及待地想要一睹芳容。

译者后记

2008年9月10日，欧洲核子研究中心的大型强子对撞机进行了首次实验。无数科学家对此翘首以待，其中就包括本书的作者兰道尔教授。对撞实验也引起了全球许多普通民众的关注，关于高能物理，关于宇宙起源，关于物质本源，人们想知道得更多。本书可以让我们更多地了解近代物理学家的研究成果和面临的未解问题。

大型强子对撞机实验究竟会给我们和科学家提供什么样的答案？实验会发现黑洞，还是额外维度？这是眼下人们正期待的结果，也是作者在本书中探寻的问题。

从这本书中，你会明白什么是相对论，了解粒子物理学、量子力学的发展及其未能解决的问题，也会学到弦理论、超对称、额外维度等物理学家提出的等级问题的新的解决方案。为了让我们最终了解额外维度和高维世界，作者丽莎·兰道尔追溯了理论物理的发展历史，给我们呈现了一份详尽的文献，引领读者从宏观世界游览到微观世界。即便我们并不具备专业知识，也能跟上她的脚步。我们可以看到，早期的物理学家为了揭示物质的最基本构成，进行了怎样孜孜不倦的探索；也明白了正是由于粒子物理学的贡献，我们的现代生活才有如此先进的科技。计算机、电子通信、计算机技术、网络技术，又有哪项离得开这些根本理论的发现？而我们现在使用的万维网恰恰是从欧洲核子研究中心衍生出来的副产品。

兰道尔教授献给我们的是一本科普读物，而非学术论著，所以阅读本书并不要求读者具备高深的专业知识。本书深入浅出、通俗易

懂，作者更多地用比喻来描述一些原本高深的物理现象，使它们与我们日常熟悉的现象联系起来。比如，她用以下例子讲解自发对称破缺：一支铅笔立在圆中心，在那短暂的一刻，所有方向对它都是相同的，但存在着一种旋转对称，当铅笔倒下时，原来的旋转对称则被打破；再如她用层层官僚机构、秘密在朋友间的传播，甚至是特洛伊战争，让我们理解了虚粒子的作用；还有，微扰理论就像是调颜料，为了得到理想的颜色，我们得一点一点循序渐进地加入少量的接近色。

翻译过程中查阅相关资料，有时会碰到论文或教科书之类，由于有太多的计算和专业术语，很多都是我难以领会的，我想，如若那不是给科学家看的，至少也是给物理学专业学生看的。还是回到这本《弯曲的旅行》才能更轻松地解读理论物理学，这才是给我们大众读者看的书。作者把枯燥的理论变得生活化起来，引领我们饶有兴味地读下去。

在本书中，生动有趣的比喻处处可见，池塘边的野鸭、橡胶管里的小虫、身材悬殊的同胞兄弟。作者说大统一的三种力就如同一个受精卵发育成的三胞胎，虽然最终长成了三个性格迥异的青年——一个留着染了色的朋克头，一个留着水手样的小平头，而另一个像艺术家一样扎着小辫子，但他们都有着相同的DNA，小时候让人难以分辨。有了这样的比喻，你还会感觉理论物理学超出我们的理解吗？为了增加阅读的乐趣，抓住读者的注意力，她还在每章的篇首都给出了一个简短的生活故事，由阿西娜和他的哥哥艾克贯穿始终。同时，每章标题下还引用了一段文字，或小诗，或歌词，或短评。从中，我们可以预期本章的内容，由此唤起阅读的兴致。

在书中，我们还能从一个侧面了解到科学家的生活点滴：在欧洲核子研究中心工作的物理学家是幸运的，尽管城里的冬天常常阴云密布，但他们却能沐浴着阳光，皮肤黝黑地度过冬季，因为在附近的山

上就能滑雪、溜冰或是徒步旅行；而有时，他们又会过度专注于自己的思考，甚至过边境时都忘了停车接受检查。拉曼·桑卓姆，作者的一位合作者，尽管两人的研究在史蒂芬·霍金那里备受推崇，但此前他却因难以得到一份正式的教职而不得不做了好几期博士后，甚至差点儿放弃了物理学，好在事情有了转机，他终于找到了工作。由此，我们也看出了一位博士后的无奈。还有，作者在咖啡馆里，品着冰激凌，与同事共同探讨新的论题；或者坐在公园的长椅上，看着湖光山色，得出自己的计算结果。这些点滴，将科学家作为普通人的一面栩栩如生地展现在了我们面前。

翻译该书的过程，也是我学习成长的过程。每翻译完一章，就像是登上了一个山头：极目望去，风光无限。长舒一口气，满怀信心地展望下一山头，又是一片全然不同的风景，使我迫不及待地要去探索。在翻译过程中，尤其日本人名的翻译是一件颇费周折的事，你不可能从英文直接音译成汉字，必须恢复其日语发音，像朝永振一郎是在文献里能查到的，而其他几位如外村彰（Akira Tonomura）等就不那么容易了。

我请教教日语的同事程笙教授，而程教授又不辞辛劳地打国际长途问过几位日本学者，几经周折才最终从早稻田大学的一位教授那里得到确认。由此，我深深敬佩程教授的严谨和不懈，同时深表感激。另一日本老师名为村山齐（Hitoshi Murayama），这一译名是我的爱人张宏伟先生帮助查阅日本网站找到的，在翻译这本书的过程中，他给予了我许多帮助和支持，在此，也对他说声“谢谢”。田南阳老师为本书校稿，实乃字斟句酌，付出良多，是最应表示感谢的人。我要致谢的人还有李绍明老师、季阳老师，正是他们对我的信任和支持，才让我有机会和勇气承担起翻译此书的任务。另外王明岩、杨桂玲、谭慧英也参与了本书的部分工作，在此一并表示感谢。

由于时间紧迫，专业知识有限，书中难免存在疏漏和不足，敬请读者不吝指正。



湛庐，与思想有关……

如何阅读商业图书

商业图书与其他类型的图书，由于阅读目的和方式的不同，因此有其特定的阅读原则和阅读方法，先从一本书开始尝试，再熟练应用。

阅读原则1 二八原则

对商业图书来说，80%的精华价值可能仅占20%的页码。要根据自己的阅读能力，进行阅读时间的分配。

阅读原则2 集中优势精力原则

在一个特定的时间段内，集中突破20%的精华内容。也可以在一个时间段内，集中攻克一个主题的阅读。

阅读原则3 递进原则

高效率的阅读并不一定要按照页码顺序展开，可以挑选自己感兴趣的部分阅读，再从兴趣点扩展到其他部分。阅读商业图书切忌贪多，从一个小主题开始，先培养自己的阅读能力，了解文字风格、观点阐述以及案例描述的方法，目的在于对方法的掌握，这才是最重要的。

阅读原则4 好为人师原则

在朋友圈中主导、控制话题，引导话题向自己设计的方向去发展，可以让读书收获更加扎实、实用、有效。

阅读方法与阅读习惯的养成

- (1) 回想。阅读商业图书常常不会一口气读完，第二次拿起书时，至少用15分钟回想上次阅读的内容，不要翻看，实在想不起来再翻看。严格训练自己，一定要回想，坚持50次，会逐渐养成习惯。
- (2) 做笔记。不要试图让笔记具有很强的逻辑性和系统性，不需要有深刻的见解和思想，只要是文字，就是对大脑的锻炼。在空白处多写多画，随笔、符号、涂色、书签、便签、折页，甚至拆书都可以。
- (3) 读后感和PPT。坚持写读后感可以大幅度提高阅读能力，做PPT可以提高逻辑分析能力。从写读后感开始，写上5篇以后，再尝试做PPT。连续做上5个PPT，再重复写三次读后感。如此坚持，阅读能力将会大幅度提高。
- (4) 思想的超越。要养成上述阅读习惯，通常需要6个月的严格训练，至少完成4本书的阅读。你会慢慢发现，自己的思想开始跳脱出来，开始有了超越作者的感觉。比拟作者、超越作者、试图凌驾于作者之上思考问题，是阅读能力提高的必然结果。



扫码关注湛庐文化，
回复“阅读”
这5种方法，让读过的书变成你的影子

[特别感谢：营销及销售行为专家 孙路弘 智慧支持！]

老 我们出版的所有图书，封底和前勒口都有“湛庐文化”的标志



并归于两个品牌



老 找“小红帽”

为了便于读者在浩如烟海的书架陈列中清楚地找到湛庐，我们在每本图书的封面左上角，以及书脊上部 47mm 处，以红色作为标记——称之为“小红帽”。同时，封面左上角标记“湛庐文化 Slogan”，书脊上标记“湛庐文化 Logo”，且下方标注图书所属品牌。

湛庐文化主力打造两个品牌：**财富汇**，致力于为商界人士提供国内外优秀的经济管理类图书；**心视界**，旨在通过心理学大师、心灵导师的专业指导为读者提供改善生活和心境的通路。



老 阅读的最大成本

读者在选购图书的时候，往往把成本支出的焦点放在书价上，其实不然。

时间才是读者付出的最大阅读成本。

阅读的时间成本=选择花费的时间+阅读花费的时间+误读浪费的时间

湛庐希望成为一个“与思想有关”的组织，成为中国与世界思想交汇的聚集地。通过我们的工作和努力，潜移默化地改变中国人、商业组织的思维方式，与世界先进的理念接轨，帮助国内的企业和经理人，融入世界，这是我们的使命和价值。

我们知道，这项工作就像跑马拉松，是极其漫长和艰苦的。但是我们有决心和毅力去不断推动，在朝着我们目标前进的道路上，所有人都是同行者和推动者。希望更多的专家、学者、读者一起来加入我们的队伍，在当下改变未来。

湛庐文化获奖书目

《大数据时代》

国家图书馆“第九届文津奖”十本获奖图书之一
CCTV“2013中国好书”25本获奖图书之一
《光明日报》2013年度《光明书榜》入选图书
《第一财经日报》2013年第一财经金融价值榜“推荐财经图书奖”
2013年度和讯华文财经图书大奖
2013亚马逊年度图书排行榜经济管理类图书榜首
《中国企业家》年度好书经管类TOP10
《创业家》“5年来最值得创业者读的10本书”
《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
《中国新闻出版报》2013年度好书20本之一
2013百道网·中国好书榜·财经类TOP100榜首
2013蓝狮子·腾讯文学十大最佳商业图书和最受欢迎的数字阅读出版物
2013京东经管图书年度畅销榜上榜图书，综合排名第一，经济类榜首

《牛奶可乐经济学》

国家图书馆“第四届文津奖”十本获奖图书之一
搜狐、《第一财经日报》2008年十本最佳商业图书

《影响力》(经典版)

《商学院》“2013经理人阅读趣味年报·心理学和行为科学类最受关注图书”
2013亚马逊年度图书分类榜心理励志图书第八名
《财富》鼎力推荐的75本商业必读书之一

《人人时代》(原名《未来是湿的》)

CCTV《子午书简》·《中国图书商报》2009年度最值得一读的30本好书之“年度最佳财经图书”
《第一财经周刊》·蓝狮子读书会·新浪网2009年度十佳商业图书TOP5

《认知盈余》

《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
2011年度和讯华文财经图书大奖

《大而不倒》

《金融时报》·高盛2010年度最佳商业图书入选作品
美国《外交政策》杂志评选的全球思想家正在阅读的20本书之一
蓝狮子·新浪2010年度十大最佳商业图书，《智囊悦读》2010年度十大最具价值经管图书

《第一大亨》

普利策传记奖，美国国家图书奖
2013中国好书榜·财经类TOP100

《真实的幸福》

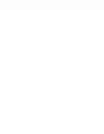
《第一财经周刊》2014年度商业图书TOP10
《职场》2010年度最具阅读价值的10本职场书籍

《星际穿越》

国家图书馆“第十一届文津奖”十本获奖图书之一
2015年全国优秀科普作品三等奖
《环球科学》2015最美科学阅读TOP10

《翻转课堂的可汗学院》

《中国教师报》2014年度“影响教师的100本书”TOP10
《第一财经周刊》2014年度商业图书TOP10



湛庐文化获奖书目

《爱哭鬼小华》

国家图书馆“第九届文津奖”十本获奖图书之一
《新京报》2013年度童书
《中国教育报》2013年度教师推荐的10大童书
新阅读研究所“2013年度最佳童书”

《群体性孤独》

国家图书馆“第十届文津奖”十本获奖图书之一
2014“腾讯网·啖书局”TMT十大最佳图书

《用心教养》

国家新闻出版广电总局2014年度“大众喜爱的50种图书”生活与科普类TOP6

《正能量》

《新智囊》2012年经管类十大图书，京东2012好书榜年度新书

《正义之心》

《第一财经周刊》2014年度商业图书TOP10

《神话的力量》

《心理月刊》2011年度最佳图书奖

《当音乐停止之后》

《中欧商业评论》2014年度经管好书榜·经济金融类

《富足》

《哈佛商业评论》2015年最值得读的八本好书
2014“腾讯网·啖书局”TMT十大最佳图书

《稀缺》

《第一财经周刊》2014年度商业图书TOP10
《中欧商业评论》2014年度经管好书榜·企业管理类

《大爆炸式创新》

《中欧商业评论》2014年度经管好书榜·企业管理类

《技术的本质》

2014“腾讯网·啖书局”TMT十大最佳图书

《社交网络改变世界》

新华网、中国出版传媒2013年度中国影响力图书

《孵化Twitter》

2013年11月亚马逊(美国)月度最佳图书
《第一财经周刊》2014年度商业图书TOP10

《谁是谷歌想要的人才?》

《出版商务周报》2013年度风云图书·励志类上榜书籍

《卡普新生儿安抚法》(最快乐的宝宝1·0-1岁)

2013新浪“养育有道”年度论坛养育类图书推荐奖



延伸阅读

《叩响天堂之门》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂宇宙求索的漫漫历程。
- ◎ 宇宙如何起源？为什么我们要耗资巨额，建造史上最大型的科学仪器——大型强子对撞机？宇宙万物的真相又如何向我们徐徐展开？
- ◎ 科学小白与科学大V都不可错过的年度最佳科普巨作，韩涛、张双楠、陈学雷、朱进、苟利军、吴岩、万维钢、郝景芳等众多顶尖科学家与科学达人挚爱推荐。



扫码直达本书购买链接



《弯曲的旅行》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂神秘的额外维度。
- ◎ 我们了解宇宙吗？宇宙有哪些奥秘？宇宙隐藏着与我们想象中完全不同的维度吗？我们将怎样证实这些维度的存在？

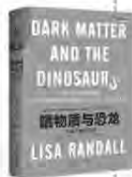


扫码直达本书购买链接



《暗物质与恐龙》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂暗物质以及恐龙灭绝背后的秘密。
- ◎ 暗物质是什么？它是如何让昔日的地球霸主毁灭的？宇宙万物又是如何在看似无关的情况下联系在一起，从而改变了世界的发展的？



扫码直达本书购买链接



《星际穿越》

- ◎ 天体物理学巨擘，引力波领域大师，同名电影科学顾问基普·索恩巨著，媲美霍金《时间简史》。
- ◎ 国家天文台8位天体物理学科学家权威翻译。
- ◎ 国家图书馆“第十一届文津奖”科普奖获奖图书。



扫码直达本书购买链接



[1]对英国读者来说，是一个孩子的攀爬架。

[2]确实有人问过我此类问题。

[3]如果要挑毛病，你会提出反对意见，说山姆也有年龄，因此还应再有一维。在这里，我们假定对于山姆来说数年如一日，因此年龄就成了一个无量量。

[4]应该说，火腿切片是有厚度的，即便很薄，仍是三维的。但它在另外的这一维度上，尺寸非常小，因此我们可以近似地把它们当作二维的。但即使任意换作别的二维薄片，我们仍可以设想，以此方式将它们结合在一起，从而形成三维物体。

[5]同样地，如果书页真是二维的，那么它们就该是绝对没有厚度的无限薄的切片。但现在，这么薄的书页就是最为近似的二维物体了。

[6]在这章及以后的章节里，我们会明确空间维度。引进相对论后，我们会转向时空，把时间当作另一维度考虑在内。

[7]本书中的“有质量”物体是为了与质量为零（且以光速运动）的“无质量”物体相区别。

[8]2004年，红袜队赢得了全美棒球联赛的冠军；而上一次得冠军，是在1918年，一年后卡鲁扎就提出了额外维度——这是多久以前的事了啊。

[9]匀速既给出了速度，也给出了方向。

[10]别误会，我还是很喜欢火车的，但我希望在美国铁路事业能得到更好的扶持。

[11]尽管在美国火车总不能协调好时间，但“美铁”（Amtrak）宣传阿西乐（Acela，穿越东北走廊的快速列车）的广告语“用空间换时间”，似乎确实承认了狭义相对论。但是，“时间”和“空间”并非可以等价交换，尽管“用时间换空间”的口号真正描述了我严重延误的火车旅行，但对于高速列车来说，这可算不上是一个吸引人的广告。

[12]雅诺什·鲍耶是一个天才，尽管他的父亲法尔卡什·鲍耶（Farkas Bolyai）希望他能成为一位数学家，可雅诺什却因为贫穷没能进入学院而参加了军队。他对非欧几何的研究开始也受到了别人的打击，他最终将它发表只不过是因父亲坚持要把它放进自己的一本著作里。法尔卡什是高斯的朋友，他把雅诺什写的附录寄给高斯。但不出所料，这次还是让雅诺什失望了：尽管高斯看到了雅诺什的天资，但在回信中，他只说道：“要我称赞他，结果会成了称赞我自己。因为在过去这30年甚至是35年里，我一直在冥思苦想的正是这些内容，他的全部研究几乎与我的完全一致。”就这样，雅诺什的数学生涯再次受挫。

[13]那是在第一次世界大战期间，史瓦西正在德国军队服役，当时他还在俄国前线上。

[14]这个名字指的是电子，而不是希腊神话中的人物。

[15]“紫外光”的意思是“高频率”。

[16]黑体实际是一个理想化物体，现实物质（如煤）并非完美的黑体。

[17]“……不惜一切代价”的基础是，不能违背热动力学的两个定律。

[18]波长等于普朗克常数 h 除以动量。

[19]尽管在空间中确定一点需要三个坐标，但我们常常将其简化，假定波函数只依赖于一个坐标。这使得我们很容易在一张纸上画出波函数的图像。

[20]托马斯·杨甚至帮助解读了罗塞塔石碑上的古埃及文字。

[21]人们实际能够探测到单个光子，但是要通过精心设计的实验。通常，我们看见的标准光都是由许多光子组成的。

[22]GeV是一种能量单位，我很快就要讲到。

[23]在这一例子里，我们假设水龙头滴水是不规律的。真正的水龙头并不一定这样。

[24]在此，我就不再推算这一数字了。

[25] 以上推理并不能完全真实地解释真正的不确定性原理，因为你永远无法确定，在测量的这一有限的时间段内，测到的就是真正的频率。难道水会永远这么滴下去吗？还是只有在测量的这段时间内在滴？尽管这很微妙，有点难以表现，但即便用一个更为精确的秒表，也永远都不可能做得更好，超过真正的不确定性原理。

[26] 这与前面章节里提到的普朗克长度是同一个量。

[27] 有一定物理基础的人就会知道，它是指轨道倾斜动量。

[28] 尽管被称作“标准模型”，但习惯中存在歧义：有人将假想的希格斯粒子也包括在内，但这里指的只是那些已知粒子，这是我所遵循的惯例。在第10章我们将探讨希格斯粒子。

[29] 猫没有锁骨，脊柱灵活、容易弯曲。因此可以将身体团成一团，以保持其斜向的冲势。事实上，对此，人们至今仍在研究。

[30] 理查德·费曼（Richard Feynman）说过：“从人类历史的长远角度来看，比如，从现在算起到1万年以后，19世纪最为重要的事件毫无疑问是麦克斯韦电动力定律的发现。”

[31] 这是对一个称为电子异常磁矩的量测量而得出的。

[32] 在第11章里，我们会看到，它们还被称作虚粒子。

[33] QED就是将量子场论应用于电磁场。

[34] 在粒子物理学里，这意味着除引力之外的基本的力（即弱力、强力和电磁力）。

[35] 汤姆逊获得这一荣誉，并不仅仅是因为他在科学上的成就，还因为他反对爱尔兰的《国内法令》（*Home Rule*）。

[36] 卢瑟福提出了他的结果，但他知道这样做一定会引起与开尔文爵士的冲突。A. S. 伊夫（A. S. Eve）所著的《卢瑟福传记》里引用了这样一段：“我进入房间，里面有些昏暗，但还是很快就发现开尔文爵士在听众席里。我知道在自己发言的最后，有关地球年龄的一部分必然会遇到麻烦，因为我的观点与他的相悖。令我宽慰的是，开尔文很快就睡着了，但当我讲到最重要的一部分时，我看到这老家伙站了起来，睁开眼睛，恶狠狠地瞪了我一眼！突然我灵感闪现，说道：‘开尔文爵士将地球年龄说短了，是因为没有发现更新的证据。而他那一预言性的论断指的正是我们今晚要探讨的，镭！’看哪，那老家伙的目光几乎要将我烧死。”

[37] 事实上，对弱相互作用的研究还要更早一些，并且人们已知太阳内部会出现核反应机制。但是，直到后来，人们才发现了它与弱力的联系。

[38] 实际上它是一个反中微子，但在此这并不重要。

[39] 这是1934年一次重要的科学会议发给与会者信函里的原话，而泡利为参加一个舞会错过了这次会议。

[40] 1956年，克莱德·考恩（Clyde Cowan）与弗雷德·莱因斯（Fred Reines）在一次核反应中，终于探测到中微子，消除了所有尚存的疑虑。

[41] 有一种方法可以看到，量子力学和狭义相对论对此关系都很重要，即普朗克常数告诉我们，它与量子力学有关；而光速告诉我们，它与狭义相对论有关。如果普朗克常数为零（此时则经典力学适用）或者光速是无限的，则距离为零。

[42] 夸克还是一种德国奶酪，如果它含有凝乳，那这名字就再贴切不过了，因为凝乳漂浮在奶酪里就像是夸克在一个强子里。可惜，我的德国朋友告诉我它不含凝乳。

[43] 现在我们知道有6种。

[44] 这就是“量子色动力学”名称的由来。

[45] 在英国，它们被称为“超级胶子”。

[46] 尽管我是从转换的结果来描述对称，但如往常一样，对称是静态系统的一种属性。也就是说，即使我并不真正地去做任何转换，这一系统仍维持对称。

[47] 这与美国的营销话术正好相反，他们总是把小的东西说成大的。

[48]这是盖尔曼“极权主义原理”（Murray Gell-Mann's term）的修订版本，而我觉得“无政府主义原理”更接近于它所适用的物理情境。

[49]不确定性原理表明，长度的不确定性与动量的不确定性成反比。

[50]这被称作是“荒原假说”。

[51]记住，虚粒子质量与真实粒子的质量是不同的。

[52]宇宙包含着暗能量（这些能量不由任何物质携带），它们构成了宇宙总能量的70%。尽管超对称有可能解释暗物质，但它（或是任何其他理论）却不能解释暗能量。

[53]记住，量子力学关系告诉我们，尽管普朗克长度非常微小，但普朗克级能量却很强大。

[54]实际上，交换的是一个虚光子，而不是一个真实的物理光子。

[55]事实上，卡拉比-丘流形的卷曲恰好保持了适量的超对称，使得弦理论能够重现标准模型的特征。如果有太多的超对称，就不能够有相互作用有别于右旋粒子的左旋粒子。

[56]实际它是Do-膜的一个收紧状态。

[57]约翰·艾利斯（John Ellis）、科斯塔斯·库纳斯（Costas Kounnas）以及德米特里·纳诺波利斯（Dmitri Nanopoulos）早期也曾在弦理论里想到过相关的观点。

[58]即故事里的平方K。KK粒子也被称作卡鲁扎-克莱因模式，这里的“模式”指的是它们的量子化动能。

[59]这是我们通常使用的时空维度的计数方式。在第1章中我们讨论平面国时，由于还没有讲到相对论，所以只把空间维度算在内。

[60]记住，我们的假设是没有膜。在后面的章节里，这一限制会发生变化。

[61]物理学家将论文贴在一个以“xxx”打头的网站里，查阅xxx.lanl.gov时，网络过滤器有时也会禁止人们访问这个网站。

[62]简便起见，我将把他们合称为“ADD”。

[63]如果它们是平坦的（见第22章）。

[64]正如我们在第18章中看到的，额外维度是统一、庞大且平坦的，对这一观点，兔子持怀疑态度。

[65]这种算法包括了时间维度。

[66]当我想强调空间和时间的不同时，有时我使用“3+1”而不是“4”。

[67]确实如此，所有切片都有相同的几何形状，在这种情况下，切片就是平坦的。

[68]记住，第五维度是指第五时空维度，也是推想的第四空间维度。

[69]测量距离的单位由膜上的能量决定，而能量则由普朗克级质量标度所决定。

[70]这一数字以弯曲单位来计量，而弯曲由有膜上和体里的能量所决定。

[71]在物理文献中，普朗克膜和TeV膜或弱力膜的名称是大家共同认可的术语。在下章的故事里，引力膜将变成膜城，而被限制在弱力膜上的粒子质量大约都应该是弱力级质量。

[72]这个题目借用了马丁·加德纳（Martin Gardner）妙趣横生的《注释版爱丽丝》（*Annotated Alice*），在这本书中，他解释了刘易斯·卡罗尔（Lewis Carroll）的小说《爱丽丝梦游仙境》与《爱丽丝梦游仙境2：镜中奇遇记》（*Through the Looking Glass*）里的文字游戏、数学谜题及其他暗示。

[73]膜本身是庞大且平坦的，只有3个空间维度，只有引力才与额外维度发生联系。记住，五维空间里有4个空间维度（还有一个时间维度），而膜有3个空间维度，我仍把时间称作第四维度，而称额外维度为第五维度。

[74]膜城即引力膜。

[75]肥猫与膜城的其他居民不一样，他没有被限制在膜上。

[76]靠近弱力膜，所有东西都变大变轻了。在阿西娜靠近弱力膜而远离引力膜的过程中，她在膜城的影子也在变大。

[77]要解决等级问题，第五维度不需要很大。

[78]引力在弱力膜上是微弱的，在那里，引力子概率函数非常小。

[79]在引力膜上，引力的强度一点也不亚于其他力的作用。

[80]满腹苦水的引力子在抱怨：在弱力膜上，引力远要比电磁力、弱力和强力微弱；而在靠近引力膜的时候，引力远要强大得多（其强度几乎与其他力接近）。

[81]同上一章一样，本章的几何是弯曲的，但这里只有一个膜——引力膜。尽管这意味着第五维度是无限的，但在本章中你将看到，在弯曲时空里，这绝对是可行的。

[82]这里我们设想的是一根平直的水管，而不是以前说的洒成圆圈的装置，由此更容易想象一个弯曲时空的情景。

[83]现实生活中有一个这样的例子：科罗拉多河。在这里，水坝和灌溉渠保证了大部分的水都供应美国西北部，而到达墨西哥时，河里只剩下了很少的水量。而在靠近加州湾的地方再建一座水坝（这就如在远离膜的地方再加上一个膜），对于拉斯维加斯得到的水量不会产生影响。

[84]卷曲空间在数学上仍是“平坦的”，这是因为你可以想象将一个卷曲的维度打开、摊平，而对一个球体，你就无法这样做。

[85]根据我们的姓名缩写，这一模型还被称作KR。

[86]流形可以拥有不同数量的洞。例如，球形没有洞，而像炸面圈一样的圆环就有一个洞。

与思想有关
a mind-boggling
business

理论物理学大师
丽莎·兰道尔宇宙三部曲

DARK MATTER AND THE DINOSAURS

THE ASTOUNDING
INTERCONNECTEDNESS OF THE UNIVERSE

向爱因斯坦宣战的理论物理学大师 丽莎·兰道尔 风靡世界之作

暗物质与恐龙

宇宙万物的互联

[美] 丽莎·兰道尔◎著 苟利军 李楠 尔欣中等◎译

LISA RANDALL

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社 于2016年12月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：暗物质与恐龙

著者：[美]丽莎·兰道尔（Lisa Randall）

译者：苟利军，李楠，尔欣中

字数：330000

电子书定价：53.99美元

DARK MATTER AND THE DINOSAURS

Copyright © 2015, Lisa Randall

目 录

版权信息

湛庐文化“科学素养”专家委员会 寄语

科学伴光与电前行，引领你我展翅翱翔

湛庐文化“科学素养”书系 专家委员会

重磅赞誉

推荐序

发现的激情

中文版序

宇宙的故事

引言

暗物质与恐龙灭绝

第一部分 宇宙的历程

01 神秘的暗物质

02 发现暗物质

03 宇宙的答案

04 几乎就是最开始：一个很好的起点

05 星系诞生

第二部分 活跃的太阳系

06 流星体、流星和陨石

07 彗星短暂而壮丽的一生

08 飞出太阳系，哪里才是边际

09 危险地活着

10 陨石坑，天外来客的“名片”

11 灭绝，生命故事的另一面

12 地球霸主恐龙的末日

- 13 宜居带中生命的故事
- 14 天外来客的旅程
- 15 从奥尔特云中逃逸的彗星

第三部分 破解暗物质的秘密

- 16 隐形世界中的神秘物质
- 17 充满荆棘的旅程：在黑暗中探寻
- 18 暗物质的社交天性
- 19 黑暗的速度
- 20 寻找暗物质盘
- 21 暗物质、彗星与恐龙

结语

一个奇妙的宇宙正等着我们去探寻

致谢

译者后记

译者简介

参考文献



你不是一个人在读书！

扫码进入湛庐“趋势与科技”读者群，

与小伙伴“同读共进”！

湛庐文化“科学素养”专家委员会 寄语

科学伴光与电前行，引领你我展翅翱翔

欧阳自远

天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

当雷电第一次掠过富兰克林的风筝到达他的指尖；

当电流第一次流入爱迪生的钨丝电灯照亮整个房间；

当我们第一次从显微镜下观察到美丽的生命；

当我们第一次将望远镜指向苍茫闪耀的星空；

当我们第一次登上月球回望自己的蓝色星球；

当我们第一次用史上最大型的实验装置LHC对撞出“上帝粒子”；

……

回溯科学的整个历程，今时今日的我们，仍旧激情澎湃。

对科学家来说，几个世纪的求索，注定是一条充斥着寂寥、抗争、坚持与荣耀的道路：

我们走过迷茫与谬误，才踟蹰地进入欢呼雀跃的人群；

我们历经挑战与质疑，才渐渐寻获万物的部分答案；

我们失败过、落魄过，才在偶然的一瞬体会到峰回路转的惊喜。

在这泰山般的宇宙中，我们注定如愚公般地“挖山不止”。所以，

不是每一刻，我们都在获得新发现。

但是，我们继续。

不是每一秒，我们都能洞悉万物的本质。

但是，我们继续。

我们日日夜夜地战斗在科学的第一线，在你们日常所不熟悉的粒子世界与茫茫大宇宙中上下求索。但是我们越来越发现，虽这一切与你们相距甚远，但却息息相关。所以，今时今日，我们愿把自己的所知、所感、所想、所为，传递给你们。

我们必须这样做。

所以，我们成立了这个“科学素养”专家委员会。我们有的来自中国科学院国家天文台，有的来自中国科学院高能物理研究所，有的来自国内物理学界知名学府清华大学、北京师范大学与中山大学，有的来自大洋彼岸的顶尖名校加州理工学院。我们汇集到一起，只愿把最前沿的科学成果传递给你们，将科学家真实的科研世界展现在你们面前。

不是每个人都能成为大人物，但是每个人都可以因为科学而成为圈子中最有趣的人。

不是每个人都能够成就恢弘伟业，但是每个人都可以成为孩子眼中最博学的父亲、母亲。

不是每个人都能身兼历史的重任，但是每个人都可以去了解自身被赋予的最伟大的天赋与奇迹。

科学是我们探求真理的向导，也是你们与下一代进步的天梯。

科学，将给予你们无限的未来。这是科学沉淀几个世纪以来，对人类最伟大的回馈。也是我们，这些科学共同体里的成员，今时今日

想要告诉你们的故事。

我们期待，

每一个人都因这套书系，成为有趣而博学的人，成为明灯般指引着孩子前行的父母，成为了解自己、了解物质、生命和宇宙的智者。

同时，我们也期待，

更多的科学家加入我们的队伍，为中国的科普事业共同贡献力量。

同时，我们真诚地祝愿，

科技创新与科学普及双翼齐飞！中华必将腾飞！

欧阳自远

湛庐文化“科学素养”书系 专家委员会

主席

欧阳自远

天体化学与地球化学家，中国月球探测工程首任首席科学家，中国科学院院士，发展中国家科学院院士，国际宇航科学院院士

委员（按拼音排序）

陈学雷

国家杰出青年科学基金获得者，国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

陈雁北

加州理工学院物理学教授

苟利军

中国科学院国家天文台研究员，中国科学院大学教授

李淼

著名理论物理学家，中山大学教授，中山大学天文与空间科学研究院院长、物理与天文学院行政负责人

王青

清华大学物理系高能物理核物理研究所所长，中国物理学会高能物理分会常务理事

张双南

中国科学院高能物理研究所研究员和粒子天体物理中心主任，中国科学院粒子天体物理重点实验室主任，中国科学院国家天文台兼职研究员和空间科学研究部首席科学家

朱进

北京天文馆馆长，《天文爱好者》杂志主编

朱宗宏

北京师范大学天文系教授、博士生导师，教育部“长江学者”特聘教授，北京天文学会理事长

重磅赞誉

韩涛

著名理论物理学家，美国匹兹堡大学物理天文系杰出教授
匹兹堡大学粒子物理、天体物理及宇宙学中心主任

人类真的生活在一个具有多维空间的膜宇宙之上吗？暗物质真的是毁灭“地球霸主”恐龙的“幕后黑手”？发现了“上帝粒子”希格斯玻色子的大型强子对撞机，以及未来的超级对撞机，会为这些玄妙的问题提供深刻的答案吗？听天才理论物理学家丽莎·兰道尔教授用妙趣横生的案例、通俗易懂的语言，对科学求索的真相与未来娓娓道来，让人欲罢不能。这是时下科学研究前沿最振聋发聩的声音！振奋人心，启迪心智！

张双南

中国科学院高能物理研究所和国家天文台双聘研究员
中国科学院粒子天体物理重点实验室主任

我们还没有探测到暗物质，但恐龙的灭绝竟然是暗物质造成的？兰道尔“宇宙三部曲”将告诉读者，想理解地球和人类的现在、历史与未来，我们必须搞清楚物质最深层次的结构和宇宙最大尺度的规律！唉，我真为其他想写类似主题的作家们担心，再写出这么出色的书恐怕很难了。

陈学雷

国家杰出青年科学基金获得者
国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直开展着最前沿的科学研究。在这套科普书中，兰道尔教授介绍了物理学家们是如何研究、探索宇宙之谜的。她并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家们现在正在进行的猜想和探索，使读者真切地欣赏到科学研究的丰富多彩和趣味，体验科学家们在构造假说、探索未知、获得新发现时所体验到的激情。我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。

朱进

北京天文馆馆长

在兰道尔教授的笔下，额外维度、暗物质、暗能量、对撞机，这些科学家的“烧脑伙伴”也变得平易近人起来。这套科普书通俗易懂，与晦涩无缘，揭示了即使是门外汉都读得懂的宇宙真相。

苟利军

中国科学院国家天文台研究员，中国科学院大学教授“
“第十一届文津奖”获奖图书《星际穿越》译者

几千年来，人类一直在试图回答“宇宙是什么”这一古老问题。现代天文观测和研究揭示，宇宙包含了时空和普通物质以及很多神秘“角色”。作为世界知名的粒子物理学家，哈佛大学物理系教授丽莎·兰道尔在她的这套科普书中，以其渊博的知识、广阔的视野、通俗的语言，以及丰富有趣的事例，给我们讲述了宇宙的基本组成和包含万物的时空，非常值得一读。作者大胆推断，地球上恐龙的灭绝与银河系中的某种暗物质有关。如果这能够被证实，将颠覆我们对宇宙神秘物质的现有认识。

吴岩

科幻作家，北京师范大学教授

简明扼要，通俗易懂，内容独创。地球人非读不可！

万维钢（同人于野）

科学作家，畅销书《万万没想到》作者

“得到”App《万维钢·精英日课》专栏作家

过去几十年来，理论物理学中最酷的话题已经从量子力学、相对论和黑洞变成了超弦、希格斯粒子和暗物质。如果说，黑洞让人着迷、量子力学让人困惑、相对论让人脑洞大开，那这些新概念则更难让人理解！不过一旦你理解了，就会获得更大的智力愉悦感。物理学家一直致力于在不用公式的情况下让公众理解物理学，丽莎·兰道尔正是这项事业的新晋翘楚。她用一贯的机智语言告诉我们，这一代的物理学正在发生什么。

郝景芳

2016年雨果奖获得者，《北京折叠》作者

在这个信息爆炸的时代，我们收到的碎片化信息太多，反而难以获得真知。碎片化文章看得再多，也不如读一本真正的好书，尤其是深入浅出、结构恢弘的好书。兰道尔“宇宙三部曲”就是难得一见的、视野辽阔的好书，每一本都选择了令人好奇的话题：宇宙结构、宇宙历史、宇宙物质，并且还与恐龙灭绝这样有趣的话题相结合，更加吸引人，让人读起来手不释卷。而最为难得的是，兰道尔的文笔简洁、优美，你在书中找不到像一般物理学科普图书那种艰深晦涩的语句。她用小说一样的文笔娓娓道来，让你理解人类对宇宙最全面的认知。

基普·索恩

天体物理学巨擘，畅销书《星际穿越》作者

加州理工学院理论物理学教授

《暗物质与恐龙》是科普作品中的杰作：既像一个展示了科学研究本质的侦探故事，又解释了人类的存在如何与充盈宇宙的暗物质那意想不到的性质，联系在了一起。

麦克斯·泰格马克

麻省理工学院物理学家

2003年度世界十大科学成就奖获得者

知名物理学家丽莎·兰道尔为世界上最古老的“谋杀谜案”之一——恐龙的灭绝，带来了新的转折。她以生动的写作手法和通俗易懂的解释方式，令人信服地暗示了一个导致那次袭击的新“嫌犯”：暗物质。

悉达多·穆克吉

哥伦比亚大学医学中心癌症医师和研究员

畅销书《众病之王》作者

只有丽莎·兰道尔可以带领我们经历这样一场激动人心的科学旅行——从恐龙到DNA，到彗星，到暗物质，再到人类的过去和未来。兰道尔的研究是如此彻底，故事是如此强大，她的讲述是如此具有说服力，以至于我都无法把这本书放下。

杰克·霍纳

美国古生物学家，《侏罗纪公园》系列电影技术顾问

丽莎·兰道尔创作了一本集乐趣和洞察力于一体的书，出色地将宇宙和生物这两个主题编织在了一起……最后用一条简单、优雅的理论解释通了大规模生物灭绝事件。这本书对于任何一个对“地球生命是否仍旧安全”感兴趣的人来说，都是必读书目。

蒂莫西·费瑞斯

NASA长期太空探索顾问，《银河系简史》作者

《暗物质与恐龙》是对科学家如何发现人类生存奥秘，以及广阔宇宙之间深层联系的挑衅和揭示。这是一本非常棒的读物。

《纽约时报》

成功的科学写作需要讲述一个完整的“如何”，然后有条不紊地把“为什么”讲解清楚——兰道尔教授就是这样做的！

《科学》

兰道尔以一种有趣且引人入胜的方式，成功地引导读者穿越了宇宙和地球的历史，穿越了从宇宙大爆炸到生命出现的历史。阅读本书，无论是对外行读者，还是对科学家来说，都是一次非常愉快的阅读体验。

《书单》

兰道尔的理论本身就是引人入胜的。在这个理论中，一些尖端学科被应用其中，包括对暗物质大胆地重新定义，它开启了一个对尖端科学的启蒙调查。正如兰道尔在《叩响天堂之门》《弯曲的旅行》中所做的，她给我们带来了令人兴奋的新理念。

《科克斯书评》

以看似闲谈的行文风格，兰道尔为我们打开了一扇令人着迷的窗口，让我们得以洞察新发现的兴奋之情以及新假说检验和构思所需要的严谨。这是来自权威研究人员的一流科学著作。

推荐序

发现的激情

陈学雷

国家杰出青年科学基金获得者

国家天文台研究员及宇宙暗物质与暗能量研究团组首席科学家

拿到这套书的样章，让我想起20多年前（1993年），我作为一名物理学研究生，参加了由李政道先生创办的中国高等科学技术中心组织的一个国际物理学会议。会议日程上列出的报告中有几位大名鼎鼎的学者，他们的名字，我们在粒子物理学教科书中早已熟悉。但当时还有一个我不很熟悉的名字“Lisa Randall”，而且在日程中排在十分显著的位置。会议开始后，我见到了她：一位面容美丽、身材苗条的女子。她看上去似乎比我大不了几岁，却十分高冷。而且，我听说她酷爱攀岩。然而在会议中，无论是演讲、问答还是讨论，她都显得学识渊博、机敏睿智、充满自信，与那些年龄、资历都老得多的学者辩论时，完全不落下风，成为会议的中心人物之一。这完全打破了我那时对女性物理学家的错误刻板印象。诚然，我从小遇到过很多成绩比我更优秀的女同学，但也许是因为女孩子们的谦让、文静和不好争辩，总让我怀疑她们不过是比我更用功、更擅长作业和考试而已。对于她们是否能深刻地思索或者作出创造性的发现，我内心总有一点儿怀疑。在物理学发展史上，女性物理学家，特别是理论物理学家，也确实屈指可数。然而，站在我面前的就是一位活生生的杰出的女性物理学家，这证明之前我错了。当然，自那之后，我有幸遇到过很多优

秀的女性物理学家，其中也包括丽莎·兰道尔教授的一位中国女弟子苏淑芳博士。她们都向我证明了，女性在物理学或者其他科学研究中，完全可以取得毫不逊色于男性的成就。

兰道尔教授先后在麻省理工学院、普林斯顿大学、哈佛大学这几所世界最著名的大学担任理论物理学教授，并一直在进行着最前沿的科学研究。她有许多卓越的成就，其中最著名的是她与桑卓姆合作提出的“额外维度”模型。在这个模型中，我们所熟知的三维空间只是高维空间中的“膜”（参见《弯曲的旅行》一书）。兰道尔教授的这套科普书系，介绍了物理学家是如何研究、探索宇宙之谜的。《叩响天堂之门》一书不仅介绍了大型强子对撞机所进行的研究的意义，也用科学的道理和事实，澄清了人们对科学的各种误解；《弯曲的旅行》一书，重点是对高维空间的探索；《暗物质与恐龙》一书则介绍了作者提出的一种特别的暗物质模型，并就此提出了一个关于恐龙灭绝的有趣假说，借此又阐述了从宇宙起源到暗物质、从太阳系演化到恐龙等多方面的知识。这三部著作的共同特点是，作者并不满足于仅仅介绍那些已经被广泛接受的科学知识，而是着重展示科学家现在正在进行的猜想和探索，当然也清楚地说明了哪些仍仅仅是猜想和假说。这些猜想也许未必都正确，其中许多可能也会被未来的实验和观测所否定。但是，对这些内容的介绍更可以使读者真切地欣赏到科学研究的丰富多彩和趣味，体验到科学家们在构造假说、探索未知、获得新发现时所体验的激情。

我相信，想了解科学探索前沿的读者一定会享受阅读这套书带来的乐趣。我也特别希望，这些书能鼓励那些喜爱科学、希望未来从事科学研究的女孩子们。

中文版序

宇宙的故事

得知我的三本书将在中国出版，我感到十分兴奋。不论在理论物理学还是在实验物理学的舞台上，中国都在扮演着日益重要的角色。

我有一些优秀的中国学生以及博士后，而且我也发现，近年来在中国这片土地上，人们对我研究方向的兴趣正在不断增长。不仅如此，中国的实验物理学也在近期取得了一些重要成果。例如，在大亚湾中微子实验室中对最轻、最重中微子混合的振荡测量，其结果震惊了世人，而且它比人们的预期早了至少一年。现有的暗物质探测器，包括PandaX与CDEX，标定了一些重要的能量范围，并且仍在不断探索，以揭示神秘的暗物质粒子的本质。展望未来，计划在中国建造的最大型的对撞机至关重要，它将成为国际主要的粒子物理学实验装置，并能够胜任探索超越已知领域的重任。

作为一位理论物理学家，我的研究领域涉猎甚广，小到物质的内部结构，大到宇宙、空间的本质。这些研究令人兴奋，然而又很难向他人解释清楚——在没有对应语境的情况下更难说清。这三本书给了我一次机会，不仅可以向世人解释我的研究，还可以同时解释作为我研究基础的量子力学、相对论、粒子物理学与天体物理学等物理学知识。我将乐于讲述一些展现这些领域中研究前沿的宏大故事。

《叩响天堂之门》一书，解释了科学的本质，并强调了尺度的重要性，也即如何在基本粒子、原子、普通物质或是宇宙的尺度上思考

科学问题。《叩响天堂之门》一书也探索了科学的发展历程、什么是“对”与“错”，以及创造力在科学发展中的意义。在这本书中，我还预测了大型强子对撞机（LHC）上的物理结果。大型强子对撞机是建造在日内瓦附近的大型加速器，能让高能质子对撞，以产生新的粒子与新的相互作用，它可以用来研究人类之前所不能及的更小尺度。这本书解释了大型强子对撞机如何运作，以及在这一实验中，科学家正在研究什么以及他们未来将要研究什么。

《弯曲的旅行》一书讲述了我对空间中可能存在的卷曲的额外维度的研究——额外维度是在我们容易观察到的三个维度（左-右、前-后、上-下）之外的某个维度。额外维度可能具有重要意义，它将解释基本粒子的质量，并为它们之间的相互作用引入新的理论可能性。这些卷曲也将容许空间具有一个无穷大的额外维度，它将与我们观测到的一切事物相容。为了讲述这个故事，我回顾了前沿研究中的量子力学、相对论、粒子物理学的基础（既有理论，也有实验），还回顾了弦理论。在这本书中，我会讲述我们是如何把所有研究领域联系在一起，我们是如何得到了这一切，以及我们已经走到了哪一步的大故事。

《暗物质与恐龙》一书，既向外审视宇宙的宏大图景，又向内一窥物质的内部结构。它解释了暗物质的本质及其在宇宙演化中扮演的角色——暗物质是宇宙中捉摸不定的物质，只与引力而不与光相互作用。《暗物质与恐龙》一书也强调了物质的基本性质与我们今日所见的地球、宇宙之间的联系。这本书的内容不仅涵盖了宇宙学，还涉及星系、太阳系和地球之间的相互作用，及其与周边环境的联系。在这一旅程中，我还将解释我对暗物质的新理念：暗物质可能包含了某个小组分，这个组分通过自身的媒介物质——光进行相互作用，而普通物质不与之相互作用。这可能会产生激动人心的结果，包括在银河系

平面上，暗物质盘将形成，其引力效应可能导致巨大的流星体撞击地球，从而最终导致恐龙的灭绝。

《叩响天堂之门》《弯曲的旅行》《暗物质与恐龙》三本书包含了粒子物理学的广阔思想领域，是对我的研究以及更宽泛的粒子物理学和宇宙学的一个简述。这三本书把许多新颖且多元的理念与科学领域结合在一起，给出了对今日科学家工作状态的一个感性认知。

在完成《暗物质与恐龙》这本书之后，我已继续投身于对暗物质的研究中。在许多已有研究的基础上，暗物质已经是一个比较成熟的研究主题了。我们在实验上有了许多直接的结果，也开始着手于更好地理解用来探索宇宙的天文望远镜、人造卫星是如何阐明“暗物质是什么”这一问题的。同时，理论也在不断发展，人们已经超越了对暗物质粒子非常狭隘的假设，并对“暗物质如何相互作用”有了更深刻的想法。我正在思考关于暗物质粒子全部带电（而不只是一小部分带电）的可能性，并自问这一假设可能导出什么结果。现有的研究忽视了某些使这个假设可行的重要结果，或许这可能只是因为：暗物质着实不像粒子物理学家之前所假设的那么无趣。

理解暗物质这种神秘物质的本质，是一个非常令人兴奋的研究主题，毕竟它占据了全部物质能量的85%。我希望中国的读者能享受这一旅程：跟随我一起探索我们是由什么组成的，宇宙中的相互作用是如何发生的，以及我们这些科学家是如何研究宇宙问题的。我确信，你们将会从中学到许多新知识，同时又能提出属于自己的问题和观点。



暗物质是什么？

是它毁灭了“地球霸主”恐龙吗？

扫码关注“庐客汇”，

回复“暗物质与恐龙”，

听兰道尔教授解答

其中的神秘关联。

引言

暗物质与恐龙灭绝

除了一些虚拟游戏以及科幻大片之外，我们很少会同时听到“暗物质”和“恐龙”这两个词。暗物质是宇宙难以捉摸的组成部分，它能够像普通物质一样通过引力相互作用，却不会发射光或者吸收光。天文学家能够探测到它的引力效应，却看不到它。而对于恐龙，我想我无须多言，这种生物在2.31亿年~6600万年前可是陆生脊椎动物中的霸主。

尽管暗物质和恐龙各自都很有趣，但你或许会认为，暗物质这种看不见的物质和恐龙这个很受欢迎的代表性生物是完全不相关的。这样想也非常合理，而且这或许就是事实。但是，按照定义，宇宙是一个单一的实体，在原则上，它的各个组成部分会相互作用，因此，本书探讨了一种假设：我和合作者认为，或许是暗物质最终（间接地）导致了恐龙的灭绝。

古生物学家、地质学家和物理学家已经证明，在6600万年前，一个直径至少为10公里的物体从太空冲向地球，导致了恐龙的灭绝，并杀死了地球上其他3/4的生物。那个天体或许是来自太阳系最外区的一颗彗星，可至今没有人知道为什么这颗彗星的轨道受到了扰动——尽管这颗彗星受到的束缚力很弱，但却很稳定。

我们的理论是，太阳系在通过银河系（由恒星和明亮的尘埃组成的条状带，可以在晴朗的夜晚看到）的中间平面时，遇到了由暗物质

构成的盘面，这些暗物质改变了太阳系远处物体的运行轨道，从而导致了这一灾难性的撞击。银河系附近有大量暗物质围绕，形成了一个巨大的光滑且弥散的球形晕。

导致恐龙灭绝的那种暗物质与宇宙中难以理解的大部分暗物质非常不同。那种暗物质会让暗物质晕保持完整，它与众不同的相互作用也会让它凝结成一个盘状物，而这个盘状物正好就处在银河系的中间。这个很薄的区域很可能非常致密，当太阳系通过它时，由于太阳在上下振动的轨道上穿过银河系，因而暗物质盘的引力影响会特别强烈。它的引力非常强大，从而会改变位于太阳系外边缘的彗星的轨道。在那里，与之竞争的太阳引力却太弱，很难将这些彗星控制在轨道上。脱离原轨道的彗星将会从太阳系中被弹出去，又或者会重新向太阳系的中心区域冲去，这时候就很有可能撞击到地球上。

坦白说，我并不确定这个理论是否正确。它只是一类意想不到的暗物质，会对生物产生可测量的影响（严格来说，它们已经不再是“生物”了）。这本书讲述的正是我们对暗物质及其影响的理论猜想。

尽管这些推测性的想法或许有些挑衅，但它们并不是本书的主要焦点。而这些想法对本书内容的重要性，可以和毁灭恐龙的彗星故事中所涉及的背景，以及支撑它的科学知识的重要性相当。这里所说的科学知识，包括一个很完善的宇宙学框架以及有关太阳系的科学知识。同时，我也感到非常幸运，我所研究的课题时常引导着我思考更宏大的问题，比如宇宙是由什么物质构成的，宇宙中空间和时间的本质是什么，以及宇宙中的一切事物是如何进化到今日之貌的。通过这本书，我希望能和读者分享这些收获。

在即将讲述的科学故事中，我的研究课题让我开始更为深刻地思考宇宙学、天体物理学、地质学，甚至生物学。不过我的焦点依旧是

基础物理学。在做了许多传统的粒子物理学（关于物质基本构成的研究，比如你正在阅读的这本书的纸或电子屏幕的构成）工作之后，我发现，探索那些有关暗物质世界的已知（或者即将知道）内容也会让人耳目一新。当然，探索基本物理过程对于太阳系和地球的意义也同样重要。

《暗物质与恐龙》这本书解释了我们现有的一些知识，包括宇宙、银河系、太阳系，以及是什么催生了地球的宜居环境并最终诞生了生命。我不仅会讨论暗物质和宇宙，还会深入讨论彗星、小行星、生命的产生和灭绝，尤其是那些坠落到地球并杀死恐龙和其他许多生命的天体。我希望这本书能传达出许多不可思议的联系，这些联系让人类得以出现，让我们现在能够更有意义地理解正在发生的一切。当我们思考地球的今日之貌时，我们或许更想了解地球演化的背景。

当我专注于这本书中的基本概念时，我感到了无限敬畏，常常会陶醉其中，不仅是因为我们对所处环境（包括太阳、银河系和宇宙）的现有了解，还因为从我们所处的这个微小的栖息之地——地球，我们是多么期望能更好地了解它。我被许多现象之间的关联所折服，正是因为这些关联的存在，才最终使人类得以出现。需要澄清的是，我的观点不包含任何宗教意味，所以我无须指定一个目的或者意义。然而，我还是不禁将这种强烈的情感归为某种宗教信仰，尤其当我们逐渐理解了宇宙的无限、我们的过去，以及它们如何完美地组合在一起时。在我们处理日常生活中所犯的那些愚蠢之事时，这种强烈的情感给每一个人都提供了某种新视角。

这个新近的研究结果让我以不同的眼光看待这个世界，以及创造了地球和我们所处的这个宇宙。我成长在纽约市的皇后区，看过很多非常壮观的建筑，但很少能亲近自然。我看到的少许具有自然风貌的地方，也是人工公园或草地，几乎没有地方能够保持原本的自然之

貌。然而，当你在沙滩上散步时，你正走在碾碎的生物之上，或至少走在它们的保护壳上。在海滩上或者乡村里，你看到的那些石灰石峭壁，或许也是由生活在几百万年前的生物所构成。山脉起源于相互碰撞的地球板块，而驱动这些板块运动的岩浆是从地心附近放射出的物质。我们的能量来自太阳系的核反应过程，在最初的核反应发生之后，其所产生的能量已经通过多种不同的方式转化并且保存了下来。我们使用的许多资源都是来自外太空的重元素，它们是由小行星或彗星带到地球表面的。一些氨基酸，或者说生命，又或者说生命的种子，也是由小行星带到地球上的。而在这一切发生之前，暗物质坍缩成块，然后吸引了更多物质，最终形成了星系、星系团和类似于太阳系一样的恒星。普通物质虽然很重要，但却不是故事的全部。

尽管我们可能觉得自己生活在一个自成一体的环境的幻象之中，但当太阳在白日升起，月亮和星辰在夜晚出现时，都提醒着我们，地球其实并不孤单。恒星和星云作为更进一步的证据，表明地球存在于一个星系之中，而银河系又处在一个更大的宇宙之中。地球在太阳系内的轨道上运动，四季更迭又告诉我们地球轨道的方位变化；以天和年为单位的时间尺度，也显示了和周围环境的相关性。

不可思议的关联

从本书的研究以及查阅各种文献中，我得到了四个令人鼓舞的经验，在此我想与大家分享。这贴合了我内心对理解宇宙的满足感，理解宇宙的很多方面是如何通过一些不可思议的方式关联起来的。在最基本的层面上，最重要的经验就是：基本粒子物理学、宇宙学以及生物学，是关联在一起的。这些关联不是从某种“新时代”的意义上而言的，而是以一些不可思议的方式关联起来的，这些方式非常值得我们研究。

来自外太空的物体会不时地撞击地球，地球与周围环境的关系“爱恨交织”。行星可以从我们周围的环境中得到一些好处，但其实受到的大部分影响都是致命的。地球所处的位置可以得到合适的温度，对于那些朝向太阳系内运动的小行星和彗星，在它们撞到地球之前，外太阳系区域的行星能够改变它们的轨道。月球和地球之间的距离会让地球轨道变得稳定下来，以防止地球温度发生剧烈变化。同时，外太阳系区域能够屏蔽危险的宇宙射线，防止地球生命受到伤害。撞击地球的流星可能会带来对于生命而言很关键的资源，但它们同时也通过一些有害的方式影响生命的演化轨迹。至少有一个这样的天体在6 600万年前，导致了地球生物发生了一次大规模灭绝，不过话说回来，这次毁灭性的撞击虽然导致了陆栖恐龙灭绝，却为哺乳动物（包括人类）的出现铺平了道路。

第二个经验，也是非常令人印象深刻的一点是，我即将讨论的这些科学发现都是在最近获得的。尽管在任何一个历史时刻人类都可以作出如下声明，但这并不会降低其有效性：在过去的X年，人类的知识已经得到了极大的发展。对于我将描述的研究，这个数字小于50。当我在做研究或阅读别人的研究成果时，我时常惊讶于很多发现是如此之新，并且极具创新性。对于这个世界我们所了解的事物，这些发现时常是令人惊喜的，并且总是有趣的，不过偶尔也是可怕的。可见，科学家在处理这些事物时，人类的聪明才智和执着在不断涌现。本书所呈现的科学认识只是宏大历史框架的冰山一角——是138亿年还是46亿年，取决于你关注的是宇宙还是太阳系。然而，人类揭开这些想法的历史也仅仅100多年而已。

尽管恐龙灭绝发生在6 600万年前，但古生物学家和地质学家直到20世纪七八十年代才推断出了恐龙灭绝的真相。在相关想法首次被提出之后，只花了几十年，科学界就对它们作出了更为全面的评估。这个时间点并非完全是巧合。当宇航员登上月球，近距离观看陨石坑

时，太阳系动态本质的详尽证据才向人类徐徐展开。由此，地外天体导致恐龙灭绝的想法才变得可靠起来。

在过去的50年中，粒子物理学和宇宙学取得了重大进展，这让我们了解了粒子物理学的标准模型，就如我们今天所理解的那样——它描述了物质的基本构成。宇宙中暗物质和暗能量的含量也是在20世纪的最后几十年里才被确认。我们对太阳系的认识也在同一时间框架内发生了变化。直到20世纪90年代，科学家才发现冥王星附近的柯伊伯带（Kuiper Belt），这才知道冥王星并不是在独自运动。行星的数目减少了，但你在小学学过的基础科学知识如今变得更丰富、更复杂了。

第三个经验是对进化发生速度的认识。当物种有充足的时间进化时，“自然选择”将允许物种适应。但是，这种适应不是极端的或快速的，而是非常缓慢的。恐龙没有时间为了一个直径约为10公里的天体要撞击地球而先去做准备，它们无法应对这种突发状况。那些被困在陆地上的恐龙体型太大了，根本无处藏身，因而也无处可逃。

每当新想法或新技术出现时，对灾难性变化与渐进性变化的争论就会成为主要问题。理解最新进展（科学或其他）的关键是这些过程的快慢。我经常听到人们说，某些方面，比如遗传学或互联网，其进展非常巨大，这并不完全正确。对疾病或循环系统的进一步了解可以追溯到几百年前，当时它带来的变化像如今的遗传学带来的变化一样深刻。书写文字的出现以及印刷术的发明，都影响了人们获取知识的方式和思考的方式，其影响之大完全不亚于互联网所带来的影响。

随着以上技术的不断发展，如今，变化的一个非常重要的因素就是其快速性。这个话题不仅仅与科学的发展历程相关，也与环境和社会的变化相关。现如今，虽然流星导致死亡的事情已经无须过分担心，但是环境变化和生物灭绝的速度过快却引起了人们的担心。在许

多方面，这个影响都是非常显著的。本书讨论的主题将帮助我们更好地理解人类是如何进化到现在这种状态，并鼓励我们如何明智地运用现有知识。

第四个重要的经验是有关于那些不可思议的科学，它描述了这个世界的隐藏元素及其发展，并描述了对于这个宇宙，我们可以了解多少。很多人都对多重宇宙非常着迷，尽管我们无法触及其他宇宙。但至少，那些隐藏世界听起来非常有趣，无论是从生物学的角度还是从物理学的角度而言，我们都有机会去探索和学习更多的相关知识。在本书中，我想让大家知道，思考已知和探索未知，是一件令人振奋的事情。

从大爆炸到毁灭性的撞击

这本书从介绍宇宙学开始。宇宙学是一门有关宇宙如何演化到今日之貌的科学。第一部分介绍了宇宙大爆炸理论、宇宙暴胀理论和宇宙构成。这部分也解释了暗物质是什么，我们如何确定它的存在，以及为什么它与宇宙的结构有关。

暗物质占宇宙物质的85%，而普通物质，比如构成恒星、气体以及人体的物质，只占15%。然而，人们主要关注的还是普通物质的存在性和相关性，因为普通物质的相互作用要强得多。

把所有注意力集中在与其影响力很不成比例的一小部分物质上，是很不明智的。我们可以看到和感受到的这15%的物质，也仅仅是其中一部分。我之后将解释暗物质在宇宙（比如，对从宇宙等离子体中产生的星系和星系团）中，以及对于保持其结构稳定性方面的关键作用。

第二部分详细介绍了太阳系。太阳系本身即使不能作为一部百科全书，但完全可以作为一整本书的主题。所以我将集中讨论可能影响恐龙灭绝的那些要素，比如流星、小行星和彗星。这一部分将描述我们已了解的那些天体，它们或者已经撞击过地球，或者被我们预期在未来会撞到地球。同时，我们还会讨论一下那些稀少但是不容易排除的有关恐龙灭绝的证据，以及每隔大约3 000万年就至少发生一次的流星撞击。在这一部分，我们还会讨论生命的形成以及毁灭，并回顾五次主要的大规模生物灭绝，包括恐龙灭绝的那次灾难。

第三部分和最后一部分集成了前两部分的想法，从暗物质模型的讨论开始，对暗物质的一些常见模型做了解释，以让我们了解暗物质是什么。对之前我们提到的有关暗物质的相互作用，我们也讨论了一些相关的新理论。

当前，我们只知道暗物质和普通物质通过引力相互作用。引力的效应非常微小，我们只能察觉到具有巨大质量的物体的引力，比如地球和太阳之间的引力，即便这种力也是非常微弱的。而你用非常小的一块磁铁就可以吸起一个曲别针——磁力很容易把整个地球的引力都比下去。

暗物质也可能会感受到其他力。人们通常假设，我们熟知的事物是独一无二的，这是因为电磁力、弱相互作用力和强相互作用力之间存在相互作用。而我们的新模型对这种假设和偏见提出了挑战。物质间的传统作用力比引力要强很多，这可以解释世界上的许多有趣特性。但是，如果一些暗物质也能够感受到非引力的相互作用，那会怎么样呢？如果这个假设成立，暗物质的力量可能会引出基本粒子和宏观现象之间关联的明显证据，可能比我们已经知道的许多联系都更为深刻。

原则上来说，宇宙中的一切事物都可能存在相互作用，而大多数这种相互作用通常都很微弱，不容易被察觉到，只有那些通过可探测方式影响我们的事物才能够被观测到。如果你有某个东西只能施加或者感受到非常小的作用力，那么即使它就在你的眼前，你也不会察觉到它。这大概就是为什么，虽然单个暗物质粒子或许就在我们周围，但到目前为止还没能被探测到的原因。

本书第三部分展示了我们应该如何更广泛地思考暗物质。比如，为什么暗物质的世界如此简单，而我们的世界却如此复杂？这可能会引导我们思考一些新的可能性。也许一部分暗物质能够感受到它自己的力，愿意的话你可以称它为“暗光”。如果大多数暗物质通常属于相对没有影响的那85%的话，我们可以认为暗物质是一种“向上运动的中产阶级”^[1]，它的相互作用和我们熟悉的普通物质非常类似。这些额外的相互作用将会影响星系的构成，并且因为相互作用，这部分暗物质会影响由普通物质构成的恒星和其他天体的运动。

未来5年里，卫星观测将能够比过去更为详细地测量星系的形状、组成和性质，告诉我们更多有关银河系环境的性质，同时检验我们的猜想是否属实。这种可观察的影响意味着，暗物质和我们的模型将会被科学所证实，即使暗物质不是构成人类的“积木”，也仍然很值得探索。暗物质的影响可能包括流星撞击事件，其中一次撞击就有可能成为恐龙灭绝的原因，就如本书书名所暗示的那样。

把这些现象关联起来的背景和概念，为我们提供了一个有关广阔宇宙的三维图景。我写这本书的目的是分享这些想法，并鼓励读者去探索和欣赏我们这个世界，并不断深入地了解它那非凡的丰富性。

DARK

MATTER

AND

第一部分
宇宙的历程

THE

DINOSAURS

THE ASTOUNDING INTERCONNECTEDNESS OF THE UNIVERSE



神秘的暗物质

DARK MATTER AND THE DINOSAURS

几个世纪以来，物理学教给我们的最大的经验就是，有如此之多的物质是人类看不到的。

我们时常会忽略那些不感兴趣的事物：划过黑暗夜空的流星；丛林漫步时，尾随身后的小动物；路过城镇时，那些宏伟建筑的细节之美……所有这些，即使它们在我们的视野之内，我们也时常会忽略它们。我们的身体是细菌的“殖民地”：10倍于人体细胞的细菌细胞居住在我们体内，帮助我们生存下去。然而，我们几乎不会意识到这些微生物的存在，它们却可以帮我们吸收营养、促进消化。只有当这些细菌使我们生病时，我们才会意识到它们的存在。

为了了解这些事情，你必须深入观察，而且你必须知道如何观察。至少从原则上看，这些现象是可以被看到的。想象一下，在理解某个你无法看到的事物时，会有多困难。暗物质就是其中之一。在宇宙中，我们难以找到它，它和我们理解的普遍物质仅仅有着微不足道的相互作用。在这一章中，我将会列举很多测量证据，天文学家和物理学家通过它们来证明暗物质的存在。我还将从以下几个方面介绍一

下这个难以被找到的物质：它是什么，它为什么看起来如此复杂，而从某些重要的角度看，它为什么又如此简单。

暗物质

在宇宙中难以被找到的东西，它和我们理解的普遍物质仅仅有着微不足道的相互作用。暗物质可以直接通过我们的身体，也存在于外部世界中。我们无法察觉到暗物质。至少在人们目前能够探测到的程度上来看，暗物质不与光相互作用。

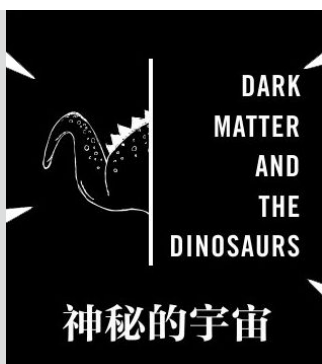
看不见的存在

尽管互联网是一个能够承载数十亿人同时在线的巨大载体，但社交网站上的大多数人并不会直接接触，甚至连间接接触都没有。参与者往往只与志同道合的人做朋友，追踪那些有相似兴趣的人，只关注那些能够代表他们自己个体世界观的新闻媒体。正是因为这种有限制的接触，在线网民被分割成许多不同的、而且不相互接触的人群。在这个人群中，他们之间几乎很难遇到持反对意见的观点，甚至朋友的朋友也不会因为相反意见而争论。所以，大多数网民对持不同观点的陌生群体毫不关心。

人类并不是故步自封的物种，但当谈到暗物质时，我们的确有点愧疚。暗物质并不是普通物质社会网络中的一部分。它存在于我们现在还不知道如何进入的某个“网络聊天室”中。它存在于相同的宇宙中，甚至占据着和可见物质一样的空间。暗物质粒子与普通物质通过一种我们觉察不到的方式相互作用着。就如同那些我们忽略掉的互联网社群一般，除非被告知它就是暗物质，否则，我们在日常生活中并不会感觉到它的存在。

就像我们体内的细菌一样，暗物质正好就是我们眼前众多宇宙中的一个。就像那些微生物一样，它们就在周围。暗物质可以直接通过我们的身体，也存在于外部世界中。然而我们不会注意到它的任何影响，因为它的作用非常弱，因而形成了独特的一类，完全和普通物质分隔开来。

暗物质是很重要的一类。然而，和细菌细胞不同——细菌尽管很多，也仅仅占人体体重的1%或2%；暗物质却占到宇宙物质成分的85%。在你身边的每一立方厘米空气中，就包含有一个质子重量的物质。这个数值的大小依赖于你如何看它。这意味着，如果构成暗物质的粒子的质量和普通物质的粒子质量相当，并且如果这些粒子按照我们所理解的动力学所期望的速度运动的话，那么每一秒钟会有几十亿个物质粒子通过人的身体，而没有人注意到它们。虽然有几十亿个暗物质粒子，但它们对我们的影响依旧微不足道。



我们无法察觉到暗物质。至少在人们目前能够探测到的程度上来看，暗物质不与光相互作用。暗物质不是由像普通物质一样的材料所构成的，不是由原子或者我们熟知的不与光相互作用的基本粒子构成的，而光是我们了解事物本质的主要因素。我和同事最希望解决的谜团就是，暗物质具体是由什么构成的。它包含一类新型粒子吗？如果包含，那么其性质是什么？除了引力之外，它还有任何其他相互作用吗？如果我们走运的话，在现在的试验中暗物质粒子或许会表现出一些异常微小的电磁相互作用力，只是它还太小，我们目前还无法探测到。专门的探测器一直在搜寻（我将会在第三部分解释如何搜寻），但到目前为止，暗物质依旧不为人类所见。以目前探测器的灵敏度，暗物质的效应还没有影响到探测器。

当大量暗物质聚集到一起时，其净引力效应非常巨大，对恒星和周围的星系存在可观测的影响。暗物质会影响宇宙的膨胀，也会影响从遥远天体到达地球的光线的路径，还会影响围绕星系中心

运动的恒星轨道，以及许多其他让我们确信它存在的可观测现象。暗物质的确存在，因为它们具有一些可测量的引力效应。

尽管暗物质不能被看到，也不能被觉察到，它却在宇宙结构的形成过程中扮演着举足轻重的角色，暗物质是这个结构中还未被理解的大多数事物。就如建造金字塔、修高速公路或者组装电器的工人，即使拥有决定权的精英看不到他们，但他们对于文明的发展与进步却极为重要。正如在我们身边其他一些被忽略的人群一般，暗物质对我们的世界来说非常重要。

如果暗物质在早期宇宙中不存在，那么我们现在就不会在这里品头论足，更不要说得到宇宙演化的一致性图景。没有暗物质，宇宙就不会有足够的时间形成现在的结构。暗物质团块作为种子，促进了银河系和其他星系或者星系群的形成。没有星系，就不会有恒星，也不会有太阳系，当然也不会有生命。即使到今天，星系以及星系团能够保持完整还要归功于暗物质的整体作用。如前文所言，如果暗物质盘存在的话，暗物质或许与太阳系的轨迹有关。

但我们并没有直接观察到暗物质。科学家们已经研究了很多形式的物质，对于其组成的研究，都是利用某种形式的光，或者从更为广泛的意义上来说，是利用电磁辐射来观测的。电磁辐射在光学频率上表现为光，在我们所能够看到的有限频率范围之外，它们表现为无线

电波或者紫外辐射等。或者在显微镜下，或者利用雷达装置，或者利用一幅照片上的光学成像，这种效应都可以被观察到。但无论如何，电磁效应总是会涉及其中。尽管很多情况下带电粒子和光会直接作用，但并非所有的相互作用都是直接的。在粒子物理学标准模型中的粒子（也是构成物质最基本的组成部分），它们之间相互作用很强，所以即使光不是它们的“直接朋友”，也至少是所见物质的“朋友的朋友”。

不仅仅是视觉，我们的其他感官（触觉、嗅觉、味觉和听觉）也都依赖于原子的相互作用，这种相互作用依赖于和一些带电粒子的相互作用。而且触觉也依赖于电磁震动和相互作用。人类的感受都是基于一定的电磁相互作用，我们无法通过常规方式直接探测到暗物质。尽管暗物质就在周围，但我们不能够看见或者感受到它。当光照射到暗物质上的时候，它没有任何反应——光仅仅是穿过而已。

从来没有人看到过（或者感觉到或者闻到过）暗物质，所以每次我和人们谈起暗物质的时候，他们对暗物质的存在都很吃惊，并且觉得它很神秘，甚至会觉得这可能是一个错误。通常人们问得最多的问题就是：作为比普通物质总量多5倍的物质，传统的望远镜探测怎么可能探测不到它呢？我个人期望的正好相反，但必须承认，并非所有人都这么认为。如果我们用眼睛看到的物质就是所有存在的物质，那么这对我来说更是不可理解的。为什么我们会有完美的感官，让我们直接感受到所有事物？几个世纪以来，物理学教给我们的最大的经验就是，有如此之多的物质是人类看不到的。从这个角度来看，这个问题就变成了：为什么我们所了解的事物就应该占有那么多的宇宙能量密度？

对一些人来说，尤其是对于暗物质怀疑论者，暗物质听起来非常怪异。但与修改引力定律相比，提出暗物质存在的建议并不草率。尽

管我们的确对暗物质不熟悉，但还是有一个尚可的传统解释，它和现有的所有物理定律完全一致。为什么遵循所有已知物理定律的物质就应该表现得和普通物质一样？简明地说就是，为什么所有的物质都应该与光相互作用？暗物质可以是那些有不同电荷或者根本就没有电荷的物质。因为没有电荷或者没有与带电粒子的相互作用，暗物质就不能吸收或者发射光子。

我对“暗物质”的名字有点想法。我对它的“物质”部分没有问题，暗物质的确是一种物质，这就意味着它能够成团并且施加引力效应，就像其他所有物质一样通过引力相互作用。物理学家和天文学家依赖于这种相互作用的多种方式，探测到了它的存在。但它名字中的“暗”在我看来是不合适的。通常情况下，我们看见暗的东西是因为光被吸收了。另外，这个不吉祥的标签掩盖了它实际的威力，并产生了更大的负面效应。暗物质并不暗，它是透明的。暗的物质吸收光，而透明的物质是“无视”光的。光能够击中暗物质，但不管是物质还是光，都不会因此而改变。

在一次会议中，有很多不同领域的人参加。会上我碰到了一位专注于品牌营销的专家马西莫。当我告诉他我的研究领域时，他难以置信地看着我并问道：“为什么叫暗物质？”他的反对并不是针对科学，而是针对这个名字背后不必要的负面含义，尽管并不是每一个和“暗”相关的名字都具有负面意义，比如“黑暗骑士”（Dark Knight）指的就是一个好人。与电影《黑暗阴影》（*Dark Shadows*）、《黑暗物质》（*His Dark Materials*）、《变形金刚3：月黑之时》（*Transformers: Dark of the Moon*）、达斯·维德（Darth Vader）的《力量的黑暗面》（*Dark Side of the Force*）中使用的“暗”相比，“暗物质”中的“暗”还是非常温和的。尽管我们对黑暗的事物有许多不好的想象，但暗物质和它的名字并没有什么关系。

然而，暗物质的确和那些邪恶之物有着一个共同特征：它藏在我们的视线之外。从这个意义上来说，“暗物质”的名字起得很恰当，因为无论如何加热暗物质，它都不会发光。从这种意义上而言，它的确是暗的，但不是从不透明的意义上而言的，而是从发射光或者反射光的意义上而言的。没有人能够直接看到暗物质，甚至利用显微镜

或者望远镜也不能。就如同许多电影或者小说中的邪恶精灵一般，它的不可见性反而起到了保护性作用。

马西莫觉得“透明物质”这个名字更好，至少不会那么吓人。尽管从物理学的角度来说这是对的，但我并不认同。“暗物质”即使不是我最喜欢的称呼，却也吸引了相当的注意力。无论如何，如果数量不是非常庞大的话，暗物质既不邪恶也不强大。它和普通物质的相互作用非常小，所以找到它非常困难。这也是它为什么如此有趣的部分原因。

黑洞与暗能量，我们居住在一个黑暗年代

黑洞

黑洞是一类太多物质被挤压到一块非常小的区域之内时，所形成的天体。没有任何事物，包括光，能够逃脱其强大引力场的影响。黑洞会吸收光和任何非常靠近它的物质。它既不辐射，也不向外反射光。但是黑洞和暗物质是完全不同的概念，这一点不应该混淆。

通常除了上面提到的听起来不好的影响之外，“暗物质”的名字还会导致一些混淆。比如，许多人都不能区分暗物质和黑洞。黑洞是大部分同类物质被集中挤压到一块非常小的区域之内时，所形成的天体。没有任何事物，包括光，能够逃脱其强大的引力场。除了在“黑色”上类似之外，黑洞与暗物质、黑墨水以及黑色电影没有更多相似之处。暗物质与光不发生作用，黑洞会吸收光和任何非常靠近它的物质。所有进入其中的光都在里面，所以黑洞是黑色的——它既不辐射，也不向外反射光。因为任何形式的物质都能够坍塌成黑洞，所以

暗物质或许与黑洞的形成也有关系。但黑洞和暗物质是完全不同的概念，这一点不应该混淆。

另外一个更深的误解是由“暗物质”这个名字而引起的。宇宙的另外一个成分叫作暗能量，这个名字的选择也很有问题，人们时常将它和暗物质混淆。尽管它又偏离了我们的主题，但对于当今的宇宙学，暗能量是一个非常重要的部分。所以我现在需要再澄清一下。

暗能量不是物质，它仅仅是能量。即使没有真实的粒子或者其他形式的物质在周围，暗能量也是存在的。它并不像普通物质结块成团，它渗透于宇宙的各个角落。暗能量的密度在各个地方都是均匀一致的。它与暗物质非常不同，暗物质能够聚集成物体，在不同地方的密度会不一样。暗物质和普通物质非常相像，它们会被限制在恒星、星系和星系团等之内。而暗能量的分布总是均匀的。

暗能量随时间保持常数，这一点与物质和辐射不同。随着宇宙膨胀，暗能量并不会被稀释，这是它比较稳定的性质。暗能量不会被粒子或者物质所携带，它的密度不随时间发生改变。基于这个原因，物理学家们时常称这种能量为宇宙学常数。

在宇宙演化早期，多数宇宙能量由辐射所携带。但辐射比物质稀释得更快，所以物质最终将超过辐射，成为宇宙能量最大的贡献体。在宇宙演化的晚期，物质和辐射被稀释，但暗能量永远不会被稀释，因此暗能量逐渐占据主导地位，现在大约占宇宙能量密度的70%。

在爱因斯坦提出相对论之前，人们认为只有相对能量，也就是两个不同系统之间的能量差别。但在知道了爱因斯坦的理论之后，我们知道能量的绝对值本身也是有意义的，并能够产生一种引力效果，它能够让宇宙坍缩或者膨胀。量子力学和引力理论表明，暗能量应该存在，并且爱因斯坦的理论告诉我们，它有一些物理效应。所以，对于暗能量的最大疑团并不是它为什么能够存在，而是为什么它的密度如

此之低。知道了它的主导地位，这看起来或许不再是个问题。尽管暗能量现在构成了宇宙能量密度的大多数，但也是在物质和能量被宇宙的膨胀极大稀释之后，并且它也只是在最近才开始与其他类型的能量相竞争。在宇宙早期，相比较更大的辐射和物质的贡献，暗能量的密度比重是极其微小的。在不知道测量结果的情况下，物理学家原本估计的暗能量密度比测量结果大120个数量级，这让人很吃惊。宇宙常数如此之小的问题已经让物理学家困惑了好多年。

暗能量

暗能量不是物质，它仅仅是能量。即使没有真实的粒子或者其他形式的物质在周围，暗能量也是存在的。它并不像普通物质结块成团，它渗透于宇宙的各个角落。暗能量的密度在各个地方都是一样的，而暗能量的分布总是均匀的。

许多天文学家认为，我们现在生活在一个宇宙学复兴的时代。在这个时代，理论与观测已经推进到了这样一个阶段：精确校准的检验将能够帮助我们确认哪个想法在宇宙中被实现了。然而，考虑到暗能量和暗物质的主导地位，以及为什么有如此之多的普通物质能够留存到现在的谜团，物理学家们也开玩笑说，我们居住在一个黑暗时代。

对于每一个宇宙研究者而言，也正是上面这些谜团让这个时代更加激动人心。尽管科学家们在理解黑暗成分时已经取得了很大进展，但依旧有很多问题摆在我们面前，等待着我们去解决。对于像我这样的研究人员来说，这种状况再好不过了。

或许有人会说，那些研究黑暗成分的物理学家正在以一种更为抽象的方式参与一场哥白尼式的革命。不仅地球不是宇宙的中心，而且构成我们的物质也不是组成宇宙能量的核心，甚至对于绝大多数物质

而言也不是。人们研究的第一个宇宙天体就是地球，它是人类最为熟悉的天体；物理学家也曾致力于研究构成我们自身的物质，因为对我们的生活来说，这是最易获取、最明显和最必要的。尽管探索地球上具有地理变化和挑战性的地域并不容易，但是相比较更远的对应天体（比如，太阳系的外边缘区域或更远），全面了解地球仍然是更方便、更容易的。

尽管区分出普通物质的最基本组成也非常困难，对它的研究却比研究透明但不可见的暗物质要容易得多，即使暗物质存在于我们周围。现在的情况正在发生改变。暗物质的研究变得前途无量，因为它能够被传统的粒子物理学解释，并可以被目前正在运行的很多实验探测器探测到。尽管相互作用非常弱，科学家未来10年应该有一个真正的机会，识别并且推断出暗物质的本质。由于暗物质会聚集成星系和其他结构，接下来对于星系和宇宙的观测将允许物理学家和天文学家通过更新的方式来测量它。而且，正如我们即将见证的，或许暗物质甚至可以解释太阳系的一些特殊性，比如与小行星撞击以及生命在地球上的演化进程相关的特性。暗物质在宇宙中并没有被分隔开来（它是真实存在的），所以“进取号”飞船（Starship Enterprise）^[2]不会把我们带到那里。或许，依靠目前已有的理论和技术，暗物质有望成为最后的科研前沿——至少是下一个非常令人振奋的科研前沿。

02

发现暗物质

DARK MATTER AND THE DINOSAURS

有时，当你走过一片森林，一群鸟儿会突然在头顶飞散开，或者一只雄鹿在你面前跑过。但你或许不会遇到惊动这些动物的登山者或者猎人。

当你走在曼哈顿的人行道上，或者驱车在好莱坞的街道上时，你会觉得名人就在附近。即使你并没有直接看到乔治·克鲁尼（George Clooney），但那些手持手机和相机，焦急等待的人群导致交通中断的状况，就已经提醒你：明星就在附近。尽管是间接方式，但通过这个人周围其他人的强大影响，你能够确信这个特别的人就在附近。

有时，当你走过一片森林，一群鸟儿会突然在头顶飞散开，或者一只雄鹿在你面前跑过。但你或许不会遇到惊动这些动物的登山者或者猎人。即便如此，这些动物的反应也能告诉我们狩猎者的存在。

同样，尽管我们没有直接看到暗物质，它仍然会影响周围的环境，就像明星或猎人一样。天文学家利用这些间接的影响来推断暗物质的存在。现在的测量以越来越高的精确度告诉了我们暗物质的能量比重。尽管引力是一种弱力，但只要暗物质的量足够多，它还是会产

生一个可测量的影响。在宇宙中，的确存在许多暗物质。虽然我们还不知道暗物质的真实本质，但接下来将描述的测量结果表明，暗物质是真实存在的，并且是不可或缺的。尽管目前对我们来说，暗物质是不可见的或者不能被直接探测到，但它也并非是完全隐藏起来、不可发现的。

缘起《独狼行动》，暗物质探测简史

弗里茨·兹威基 (Fritz Zwicky) 是一个独立的思考者，他对事物时常有着令人印象深刻的洞察力，不过偶尔也有一些疯狂的想法。他非常清楚自己是一个怪胎，甚至计划写一部名为《独狼行动》 (*Operation Lone*) 的自传。尽管他在1933年得出了20世纪最为惊人的发现之一，但在之后的40年中并没有引起重视。

兹威基在1933年得到的推论的确非常了不起。他观察了后发座星系团 (星系团是引力束缚的巨大星系集合体) 中的星系速度。在星系团中的物质引力和所包含的恒星动能差不多，从而创造了一个稳态的系统。如果包含的物质质量太低，星系团的吸引力将比恒星的动能小，这些星系将会逃离系统。根据对恒星速度的测量，兹威基计算了星系团所需的质量总量，从而产生足够的引力。结果他发现，所需的质量总量是所测量到的发光质量 (产生光线的物质) 总量的400倍。为了解释这一结果，兹威基提出，存在一种额外的物质，并将其命名为“dunkle Materie” (德语对“暗物质”的称呼)。

睿智多产的荷兰天文学家简·奥尔特 (Jan Oort) 比兹威基早一年得到了有关暗物质的相似结论。奥尔特意识到，在临近星系中，如果仅仅把它们的速度归结于发光物质的引力的话，是不会使恒星具有如此快的速度的。奥尔特也推断出应该缺少了某种东西，然而他并没

有去猜想存在一种新的物质形式，而仅仅假设存在一种不发光的普通物质。我接下来会讨论，因为种种原因这一提议被否决的故事。

奥尔特或许也不是第一个获取这个发现的人。我在斯德哥尔摩参加一个宇宙学大会的时候，我的瑞典同事拉斯·伯格斯特龙（Lars Bergstrom）告诉我一个相对不知名的观测项目，这个项目是由瑞典天文学家克努特·伦德马克（Knut Lundmark）完成的。像奥尔特一样，尽管伦德马克没有提出一个全新物质形式的大胆建议，但他对暗物质和可见物质之间比值的测量非常接近真实值，我们现在知道这个值大约是5。

尽管有这些早期的观测，但暗物质在相当长的一段时间里基本上是被完全忽略的。这个想法在20世纪70年代才再次流行起来。天文学家对卫星星系（在大质量星系附近的小质量星系）的运动进行了观测后发现，只有额外的看不见的物质存在时，才能够解释它们的运动。正是类似这样的观测，才让天文学家开始对暗物质进行严肃的探讨。

暗物质的地位真正得到巩固是因为维拉·鲁宾（Vera Rubin）的工作。鲁宾是美国华盛顿州卡耐基研究所的一位天文学家，她当时和天文学家肯特·福特（Kent Ford）一起工作。在从乔治城大学研究生毕业之后，鲁宾决定从仙女座星云（M31）开始测量星系中恒星的角向运动，部分原因也是避免踏入其他科学家过度保护的领地。鲁宾的毕业论文是有关星系速度测量的，从而确认星系团的存在，不过她的论文最初遭到了科学界很多人的拒绝，部分原因是她侵入了其他人的研究领域，于是她改变了自己的研究方向。对于毕业之后的研究方向，鲁宾决定进入一个并不热门的研究领域——恒星的轨道速度。

鲁宾的决定催生了或许是她那个时代最让人兴奋的发现。20世纪70年代，鲁宾和福特发现，恒星的旋转速度在距离星系中心的任何距离上基本上是一样的。也就是说，恒星以恒定的速度旋转，甚至在包含发光物质的区域以外的地方也是一样。唯一可能的解释是，存在一些没有考虑到的物质。正是这些物质的引力作用控制着外围遥远恒星的运动，让它们的运动比预期的要快。如果没有这些额外物质的贡献，恒星就会具有鲁宾和福特所测量的那些速度，那样的话，恒星将飞出星系。这些了

不起的推断发现，为了让恒星保持在它们的轨道上，普通物质仅仅占所需质量的1/6。鲁宾和福特的观测结果给出了那个时代有关暗物质的最强证据，而星系的旋转曲线继续成为一个重要线索。

从20世纪70年代起，关于暗物质存在的证据变得越来越强，它占宇宙净能量的比例被测量得越来越精确。能够让我们获知暗物质信息的动力学效应，包括星系中恒星的旋转。然而，那些测量仅仅适用于像银河系一样的漩涡星系，它们拥有由可见物质构成的盘，并且还有一个向外延伸的旋臂。

另外一个重要的类别是椭圆星系，在这类星系中，发光物质的形状更接近球状。在椭圆星系中，类似于兹威基对于星系团的测量，可以测量速度弥散（星系中恒星间的速度变化大小）。因为这些速度是由星系的质量所决定的，所以它们充当着星系质量测量的代理量。椭圆星系的测量更进一步表明，发光物质不足以解释恒星的动力学测量结果。除此之外，那些未包含在恒星中的星际气体，其动力学测量表明也需要暗物质。因为这些气体到星系中心的距离是可见物质尺度的10倍，所以这些观测表明，暗物质不仅存在，而且它的范围远远延伸到星系的可见部分之外。测量气体温度和密度的X射线确认了这一结果。

引力透镜效应，让暗物质现形

星系团的质量也能够通过光的引力透镜效应测量（见图2-1）。再强调一次，没有人能够看见暗物质，但暗物质能够通过引力影响它周围的物质，甚至是光。比如，根据兹威基对后发星系团的观测结果，暗物质通过他能够探测到的一些方式，影响了星系的运动。尽管暗物质是不可见的，但是它能够通过影响可见物体而被探测到。

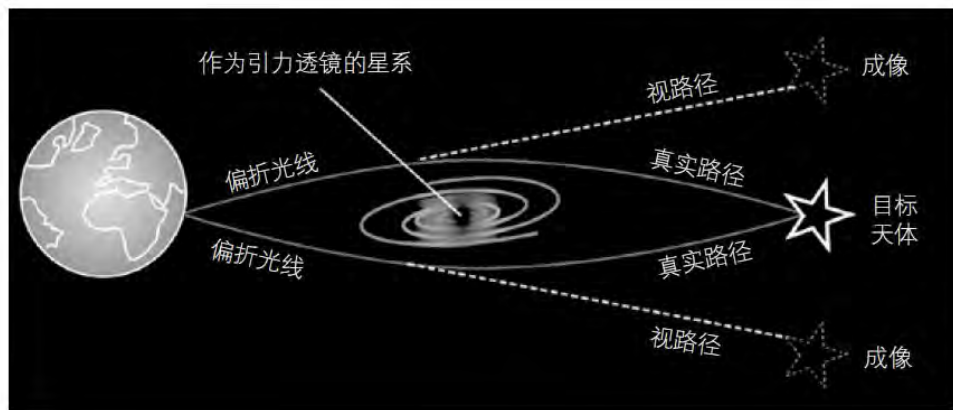


图2-1

类似恒星或者星系的明亮天体发出的光线，在大质量天体（比如星系团）的周围光线会发生弯曲。地球上的观测者将观测到的光线投影成发射源的多个像。

第一个提出了引力透镜效应建议的人正是多才多艺的兹威基。这个提议背后的想法是：位于其他地方的发光物体所产生的光线，会因为暗物质的引力效应改变其传播路径。位于传播路径中的大质量天体（比如星系团）会导致发光物体的光线发生偏折。当星系团足够重时，路径的偏折效应可以被观测到。

偏折的方向取决于光线的初始方向：经过星系团上部的光朝下偏折，右边的光朝左偏折。假设光线按照直线传播，往回追溯这些光线，对于最初产生光线的任何天体，观测将会看到多个像。兹威基意识到，光线和多个像的变化取决于中间星系团所包含的总质量。而通过测量光线和像的变化，我们能够探测到暗物质。发光物体在强引力透镜作用下能够产生多个像；在弱引力透镜情况下，星系的形状会被扭曲，不会产生多个像，因此能够被应用于星系团的边缘，在那里，引力效应不是那么明显。

引力透镜

位于传播路径中的大质量天体（比如星系团），会导致发光物体的光线发生偏折，从而出现多个像。

就像星系团中星系的速度观测结果导致了兹威基的第一个疯狂结论一般，尽管暗物质本身不可见，但通过被偏折的光所产生的可观测效应可以获知星系团的总质量。在首次被提出很多年后，这个神奇的引力透镜效应被人们观测到了。

因此，透镜效应的测量成为暗物质研究中最重要的一种观测。引力透镜效应非常有趣，在一定程度上它是一种直接看到暗物质的方式。位于发射源和观测者之间的暗物质会让光发生偏折。利用恒星或者星系速度测量时，我们需要做一些动力学假设。但透镜效应与动力学假设无关，透镜效应直接测量了发光体和地球之间的质量。位于星系团（或者包含暗物质的天体）之后的某个天体发射出光，星系团会让光线偏折。因为星系物质（包括不发光的暗物质）的引力效应，来自星系背后类星体的光线会产生多个像，这种透镜效应已经被用来测量这些星系中的暗物质。

子弹星系团，暗物质存在的最强证据

透镜效应提供了最有说服力的暗物质证据，这个证据来自合并的星系团，就像现在在物理学家中很出名的子弹星系团（见图2-2）中所发生的那样。子弹星系团由至少两个星系团合并而成。合并前的星系团包含了暗物质和普通物质，也就是X射线辐射出的气体。气体能够感受到电磁相互作用力，这能够非常有效地阻止两个星系团的气体持续穿过彼此，导致的结果是，最开始随着星系团运动的气体被阻挡在了中间部分。另外，就像子弹星系团所表现的那样，暗物质不管是与气体还是其自身，都几乎没有相互作用，所以它可以没有任何阻碍地穿

过，从而在合并星系团的外区形成类似米老鼠耳朵的球形形状。气体就像从不同方向来的车在并道之后导致的交通堵塞，而暗物质就像动作灵敏、可以自由移动的摩托车，能够不受限制地通过。



图2-2

不同星系团合并形成了子弹星系团，气体被困在中间的合并区域。暗物质通过，形成了包含暗物质的球形外围区域。

利用X射线测量可以确认气体存在于中心区域，天文学家也可以利用引力透镜效应，在星系的外部区域找到暗物质。这或许是现有的能够证明暗物质存在的最有力证据。尽管人们还在考虑用修改引力的方式来解释，但如果不存在一些不相互作用的物质，那么要想解释子弹星系团（或者其他类似观测）的奇怪结构，是很困难的。子弹星系团或者类似的星系团，通过最为直接的方式表明暗物质是存在的，它们在星系团合并时会毫无阻碍地通过。

暗物质的总能量密度是多少

以上观测确认了暗物质的存在，却依旧留给我们一个问题：宇宙中暗物质的总能量密度到底是多少？即使知道星系和星系团中包含多少暗物质，我们也不知道它们的总量。事实上，多数暗物质应该在星系团中，因为所有物质的特性会集结成块。因此暗物质也应该存在于引力束缚的结构中，而非弥散地分布在宇宙中，这样包含在星系团中的暗物质质量应该非常近似地等于它的总量。但如果不用做这个假设，也能够测量出暗物质所具有的能量密度，那就更好了。

实际上，有一种更为可靠的测量暗物质总量的方法。暗物质的总量会影响微波背景辐射（CMB）——来自宇宙最早期的辐射遗迹。我们目前已通过精确测量对这种辐射的各种性质有所了解，这些性质在建立正确的宇宙学理论方面扮演着非常关键的角色。暗物质总量的最佳测量手段就来自对微波背景辐射的研究，这是目前已知的用来探测宇宙早期阶段的最干净的方式。

我想提醒读者的是，即使对物理学家来说，这些计算也是非常需要技巧的。然而，分析中所涉及的一些重要概念其实非常容易理解。一个关键的信息就是在最开始，原子（由带正电荷的原子核和带负电的电子构成的电中性束缚态）是不存在的。只有当温度降低到原子的键能之下，电子和原子核才能够稳定地结合成原子。在那个温度之上，辐射光子会将质子和电子分开，也就是将原子拆开。由于早期的这些带电粒子的存在，弥漫于宇宙各处的辐射在最开始时不能自由穿行，会被这些带电粒子所散射。

随着宇宙的冷却，在某个特定温度（再复合温度）条件下，带电粒子将复合形成中性原子。缺少了带电粒子，光子就能不受阻碍地运动。结果是，从这一刻之后，带电粒子的运动状态不再是自由运动，

而是结合成了原子。因为有了带电粒子的散射，再复合之后所辐射出来的光子就能直接被望远镜观测到。所以当观测微波背景辐射时，我们实际上是在回看宇宙早期的状态。

从测量的角度来看，这非常棒。这在宇宙的一生中也非常早——大约在宇宙大爆炸之后的38万年，那时宇宙结构还没有形成。就像宇宙诞生之初所呈现出的那样，它基本上是均匀的和各向同性的。也就是说，不管你研究哪一块天空或者选择哪一个方向，温度几乎是一样的。但万分之一程度的微小扰动还是微微降低了这种均匀性。这些扰动的测量结果包含非常多的有关宇宙成分和之后的演化信息。这些结果帮助我们推断出宇宙的膨胀历史和其他一些性质。通过这些性质，我们又可以了解宇宙过去和现在所包含的辐射、物质和能量的含量，从而对宇宙的性质和成分有更详细的认识。

为了理解为什么远古的辐射包含那么丰富的信息，第二件需要对早期宇宙理解的事情是，在再复合的时候，也就是中性原子最终形成的时候，宇宙中的物质和辐射开始产生振荡。在大家所了解的“声学振荡”（acoustic oscillations）中，物质的引力将物质向里拉，而辐射的压力驱使着物质向外运动。这些力相互竞争，让处于坍缩的物质收缩和膨胀，从而产生振荡。暗物质的量决定了将物质向内拉的引力势的强度，从而来抵制辐射向外的压力。这种相互作用帮助形成了振荡，也允许天文学家测量当时暗物质的能量密度。在一种更为细微的效应中，在物质开始坍缩（这个发生在物质中的能量密度，超过了辐射中的能量密度）和再复合的时间之间——也就是开始振荡的时候，暗物质能够影响时间的流逝。

宇宙饼图

宇宙饼图里包含了很多信息。即使不知道细节，我们也清楚这些测量结果是非常精确的。这些数据让我们可以非常准确地确定很多宇宙学参数的值，包括暗物质的能量密度值。这些测量不仅确认了暗物质和暗能量的存在，也限制了它们在宇宙中的能量比例。

暗物质所含的能量比例大约是26%，普通物质大约为5%，暗能量大约为69%（见图2-3）。普通物质的大多数能量包含在原子中，这也是为什么在宇宙饼图中既可以使用“原子”，也可以使用“普通物质”。另外，暗物质的能量是普通物质所含能量的5倍，意味着它占宇宙物质总能量的85%。让人欣慰的是，利用宇宙微波背景辐射测量所得到的暗物质比重结果，与之前星系团测量得到的结果是一致的，由此进一步确定了宇宙微波背景辐射测量结果的可靠性。

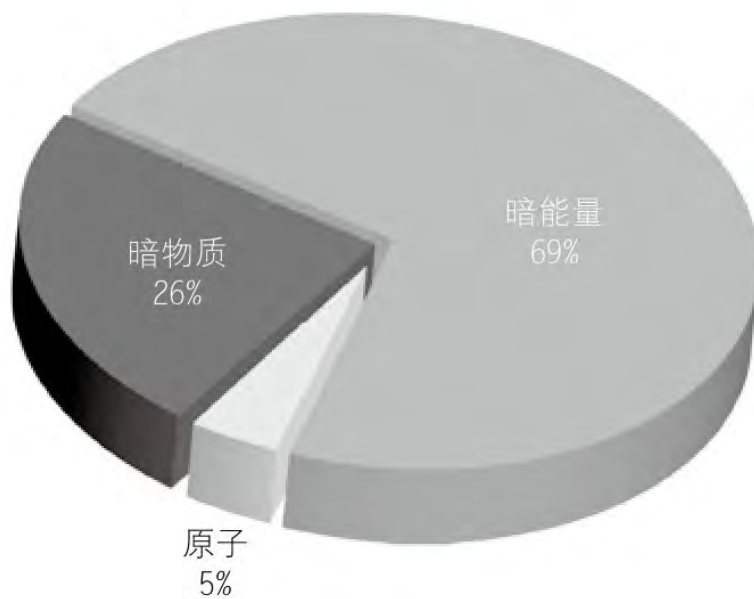


图2-3

普通物质（原子）、暗物质和暗能量的比例图。请注意，暗物质是总能量密度的26%，但占据了85%的物质能量，因为物质仅仅包括了原子和暗物质的贡献，而不包括暗能量的贡献。

微波背景辐射的测量也确认了暗能量的存在。暗物质和普通物质会通过不同方式对从宇宙大爆炸时期产生并留存下来的微波背景辐射产生扰动，而背景辐射的测量数据确认了暗物质的存在及其所占的比例。并且，暗能量（存在于宇宙中却不存在于任何物质之上的神秘能量）也影响着这些扰动。

暗能量的真正发现来自超新星的测量，这是由两个不同的研究团队作出的：一个由索尔·珀尔马特（Saul Perlmutter）领导，另外一个由亚当·里斯（Adam Riess）和布赖恩·施密特（Brian Schmidt）领导。尽管暗物质才是我们真正关心的，暗能量的发现看似是一个小的“插曲”，但暗能量也是非常有趣和重要的，所以这个短暂的“插曲”很值得我们进一步了解。

“标准烛光”，照亮暗能量的发现之路

在暗能量的发现过程当中，Ia型超新星扮演着非常重要的角色。在热核聚变燃烧耗尽其核心区所有的氢和氦之后，某些恒星在其演化的最后阶段，便会成为看似平淡无奇的白矮星。而Ia型超新星正是来自白矮星核的爆炸。如果质量高于一定数值，白矮星就会变得不稳定从而爆炸。就像一个石油大国，将所有资源都输出了，结果发现自己的大量国民非常不满，他们随时可能发动革命。白矮星只需要吸收一定量的物质，就会让它自身处于爆炸的边缘。^[3]因为能够发生爆炸的所有白矮星具有同等质量，所以所有的Ia型超新星具有同样的亮度，天体物理学家称它们为“标准烛光”（standard candles）^[4]。

因为Ia型超新星光度的一致性，并且因为它们很亮，即使在很远的地方，也可以容易地看到，所以Ia型超新星是特别有用的宇宙膨胀

速率探测器。除此之外，作为标准烛光，Ia型超新星的视亮度变化只是因为它们到地球的距离不同而已。

一旦天文学家测量了星系的退行速度和星系的亮度，他们就能够决定星系距地球的距离，以及星系所暗含的宇宙膨胀速率。有了这些信息之后，天文学家就可以确定宇宙随时间膨胀的规律。

利用这个想法，两个超新星研究团队在1998年测量了许多Ia型超新星所在的星系红移，从而发现了暗能量。红移是退行物体所发出的光频率的改变，就像当救护车呼啸着离我们远去的时候，它警笛的音调会慢慢降低一样，红移能够告诉我们一个事物远离某个光源或者声源的速度。当研究人员弄清所研究超新星的红移和亮度时，他们就能够测量宇宙的膨胀速率。

让研究者感到非常诧异的是，他们发现超新星比预期的要暗一些。这表明：测量得到的超新星到地球的距离，要比利用当时传统的宇宙膨胀速率假设所得到的预期距离远一些。这一观测结果产生了一个很了不起的结论：某种能量源一直在加速宇宙的膨胀。暗能量正好符合要求，随着时间推移，它的引力效应让宇宙膨胀的速率越来越大。宇宙微波背景辐射的测量结果和超新星的测量结果共同确认了暗能量的存在。

暗物质终曲

现在所有的测量结果都很好符合预估结果，所以大多数宇宙学家都在谈论一个 Λ CDM ^[5] 的宇宙模型。 Λ 过去被用来指暗能量，我们现在知道它是存在的。利用宇宙饼图中暗能量、暗物质和普通物质的含量，目前所有的测量和模型预测结果都很符合预期。

微波背景辐射的密度扰动很微小，但却表现出了非常多的扰动现象。我们目前对宇宙微波背景辐射已经做了精确的测量，从而也确定了许多宇宙学参数，包括普通物质、暗物质和暗能量的能量密度，以及宇宙的年龄和形状。我们将在第5章讨论来自威尔金森微波各向异性探测器（Wilkinson Microwave Anisotropy Probe, WMAP）和普朗克卫星的新数据。这些数据和观测数据（比如来自Ia型超新星的研究数据）的高度一致，是对宇宙学模型的一种重要验证。

我必须说明的是，还有一项非常重要的证据表明了暗物质的存在，这项证据对我们来说是最为重要的，即结构（比如星系）的存在。如果没有暗物质，这些结构将没有充足的时间形成。

为了理解暗物质在这一重要过程中所扮演的举足轻重的角色，我们需要对宇宙早期的历史有一些了解。在讲述结构形成之前，我们要首先了解一下宇宙学——它是一门关于宇宙随时间如何演化的学科。

03

宇宙的答案

DARK MATTER AND THE DINOSAURS

就像一张脸一样，宇宙同时具有美和潜在的秩序。

有很多次，当我告诉人们我从事“cosmology”（宇宙学）的研究时，他们会错误地把我当作“cosmetologist”（美容师），这让我觉得非常有趣，如果我真的去做美容师的话，将是多么不适合。但这个错误的的确促使我去查了一下这两个词，如果不仔细听的话，它们听起来确实非常相似。这两个词都起源于希腊语“kosmos”的拉丁语版本，所以经常有人将两个词混淆。公元前6世纪，毕达哥拉斯或许是第一个将“kosmos”用于描述宇宙的人。公元1200年后，“cosmos”主要意为“秩序，好的秩序，有秩序的安排”。不过，这个单词一直到19世纪中期才开始流行起来，这是因为德国科学家和探险家亚历山大·冯·洪堡（Alexander von Humboldt）作了一系列报告之后，把它整理制作成了一个名叫《宇宙》（*Kosmos*）的小册子。这个小册子影响了很多读者，包括作家爱默生、梭罗、爱伦·坡、惠特曼。读到这里，你或许会明白天文学家、科普作家卡尔·萨根（Carl Sagan）重启科普电视片《宇宙》的原因。

“Cosmetic”（化妆品）这个单词可以追溯到17世纪40年代的法语单词“cosmétique”，这个词又起源于希腊语“kosme-tikos”，意为“擅长装饰或者安排”。在线词典给出了双层含义，我猜测只有洛杉矶的居民才能够完全理解^[6]，除了“宇宙，世界”的含义之外，“kosmos”还有一个很重要的含义，与“饰品，一个女人的服饰或装饰”有关。无论如何，我碰到的那个令人尴尬的困惑，并不完全是一个巧合。“Cosmology”（宇宙学）和“cosmetology”（美容学）都源于“kosmos”。就像一张脸一样，宇宙同时具有美和潜在的秩序。

宇宙学是关于宇宙演化的科学，已经发展成为一门独立的学科。它已经进入了一个实验和理论均有革命性进展的时代，所产生的见解要比大多数人在30年前的预期更广泛、更细致。改进后的技术与根植于广义相对论及粒子物理学的理论相结合，提供了一个关于宇宙早期阶段的详细图景，以及它如何演化到今日之宇宙。下一章我将会解释，20世纪的这些进步如何深刻影响了我们对宇宙历史的认识。在探寻这些非凡的成就之前，为了解释在一些有关人类最古老和最基本的问题上，科学能够告诉和不能够告诉我们什么，我要先暂时变身哲学家。

没有答案的问题

宇宙学是关于大的方面的追问：宇宙如何开始，又如何演化成为现在的状态。在科学变革之前，人们试图利用仅有的方法，比如哲学和有限的观测数据，来回答这样的问题。一些想法被证明是正确的，但也有很多想法是错的，这都是很正常的。

到现在也是如此，尽管我们取得了许多进步，当思考一些未解之谜时，人们还是情不自禁地求助于哲学。事实上，这也迫使我们面对哲学和科学的差别。科学关注那些至少原则上我们能够通过实验或者观测来验证或者排除的想法。至少对于一个科学家而言，哲学关注的那些问题，我们永远不可能给出确定的答案。技术有时会滞后，但我们相信至少在原则问题上，科学的提议是能够被验证或者证伪的。

这让科学家们处于一种困境。宇宙肯定超过人类能够观测的区域。如果光速的确是有限的，而且如果宇宙仅仅存在有限的时间，不管技术有多进步，那么我们也只能看到有限的空间区域，看到那些光触及到的区域，或者其他以光速传播的事物所能到达的区域。只有从那些区域，信号才可能在宇宙产生的时间内抵达地球。任何比这更远的事物，处于物理学家所说的宇宙视界（cosmic horizon）之外，我们目前所做的任何观测都无法触及。

这意味着，从本质上来说，科学知识并不适合于视界之外。对于那些视界之外的猜想，没有人能够通过实验证实或者证伪。根据我们对科学的定义，在那些遥远的区域，哲学至上。但这并不意味着拥有好奇心的科学家不用思考应用于那些遥远区域的物理原则和过程。其实许多人还是会这样去做的。我不想忽视这些探索，因为这些探索既有趣又深刻。但由于这些知识的局限性，这些内容并不是那么可信，即使是科学家的答案，也不比其他人更可靠。然而，因为我时常被人问起，所以利用这一章，我将对那些时常需要被解答的几个问题，谈一谈我的看法。

我时常听到的一个问题是：在这个世界上，为什么有“存在事物”，而不是“空无一物”呢？尽管没有人知道真正的答案，但我会给出我的两个回答。首先，也是不可否认的，



如果是“空无一物”的话，你不可能坐在这里提出问题，我也不可能在这里解答问题。其次，我认为“存在事物”是更为可能的。毕竟，“空无一物”是非常特殊的。如果你有一个数字序列的话，在所有可以被选择的具有无限数目的数字当中，“零”仅仅是这当中无限小的一个点。“空无一物”是如此特殊，你根本不会期望用它来表述宇宙的状态，这无须一个潜在的理由。但是即使潜在的理由也不是“空无一物”。你至少需要物理定律来解释一件非随机事件的发生。有起因就意味着一定有某种事物的存在，尽管它听起来有些可笑，但我深信这一点。你或许不能总是找到你想找的东西，但你总会无意地找到某些其他东西。

当物理学家思考构成我们自身的物质时，就会有一个科学而非哲学的问题冒出来。在宇宙中，为什么会有这么多物质（比如，质子、中子和电子）组成我们自身？尽管对普通物质已经有了很多了解，但我们并不完全明白为什么还有那么多普通物质依旧存在。包含在普通物质中的能量总量是一个悬而未决的问题，我们至今不太明白为什么会有这么多的普通物质在宇宙中保存了下来。

这个问题可以归结为：**为什么物质和反物质存在的数量不相等？**反物质是质量相同，但和普通物质电荷相反的物质。物理理论告诉我们，任何一个物质粒子，必然存在着一个对应的反物质粒子。比如，如果电子的电荷为-1，那么必须有一个叫作正电子的反物质粒子存在，与电子相比，它具有同等质量但相反的电荷+1。**为了避免任何不必要的混淆，我要明确指出，反物质不是暗物质。**反物质带有像普通物质一样的同种类型的电荷，所以它会和光子相互作用。唯一的差别是，反物质的电荷与其相关物质的电荷是相反的。

反物质携带着与普通物质相反的电荷，两者的净电荷为0。既然物质和反物质在一起时不带有任何电荷，根据电荷守恒和公式 $E=mc^2$ ，我们知道，当物质和暗物质相遇时，就会消失变成纯粹的能量——这个能量也是不带电荷的。

我们本来期望，在宇宙冷却时，所有已知的普通物质最终将会和反物质一起湮灭。这意味着，物质和反物质合并变成纯粹的能量，从而消失。但是我们现在还坐在这里讨论问题，很明显这不是事实。普通物质被保留了下来，在图2-3中你看到它占了宇宙能量的5%，那么宇宙中物质的总量一定比反物质的总量要多。对于宇宙和我们自己而言，一个很关键的特点是，普通物质能够以足够的数量保留下来，从而创造出动物、城市和恒星，这和前文提到的标准热期望不同。不过这只有在普通物质主导反物质的情况下才能发生，也就是说有一个物质和反物质的不对称性。如果总量总是一直相等的话，那么物质和反物质总是能够找到彼此，然后湮灭并且消失。

为了能够让物质保留至今，物质和反物质之间的非对称性必须在宇宙早期的某个时候就已经建立。物理学家已经提出了许多可能的情形，从而创造出这种不平衡，但我们现在还是不知道哪一个想法是正确的。这个非对称性依旧是宇宙学中没有解决的重要问题之一。这意味着，我们不仅不理解暗物质，也不完全理解普通物质。在宇宙演化的早期，一定发生了某种特别的事情，从而可以解释为什么只占宇宙构成5%的物质可以保留至今。

第二个目前没有回答的问题是：大爆炸过程中究竟发生了什么？科学家和大众媒体经常用“大爆炸”代指宇宙时间小于 10^{-43} 秒和尺度小于 10^{-33} 厘米，甚至利用非常漂亮的多色图来演示大爆炸。但“大爆炸”很容易误导人，这个我会在第4章深入讨论。天文学家弗

雷德·霍伊尔 (Fred Hoyle) 倾向于一个静态宇宙，1949年，他在BBC电台节目中发明了这个术语，用来贬低这个他不相信的理论。

无论你对大爆炸宇宙学的态度如何，它都非常成功地描述了我们所知道的宇宙在开始的几分之一秒后的演化，而在更早的时刻所发生的事情，没有人知道。对于大爆炸或者之前的可靠表征则需要一门叫作量子引力的理论。在这个与最早时间相关的微小距离尺度上，量子力学和引力理论都很重要，但并没有人找到可行的理论，能够应用于这个极小的距离尺度范围。只有当我们对这个微小距离尺度上的物理过程有了更多了解之后，我们才能够对宇宙的最初状态了解更多。不过即使到了那个时候，通过观测去检验结论可能还是行不通。

我还时常会听到另外一个更加不可能回答的问题就是：大爆炸之前发生了什么？回答这个问题估计比理解大爆炸需要知道的知识更多。我们现在不知道大爆炸时到底发生了什么、在此之前发生了什么。但在你失望之前，我还是要告诉你，即便发现所有的答案，都仍然不会令人满意。或许宇宙已经存在了无限长的时间，或许它开始于某个特定的时刻。这两个答案似乎看起来都不让人满意，但至少是一些选择。

更进一步来说，如果这个宇宙一直存在，并且大爆炸仅仅是其中一部分，或者我们的宇宙就是所有的一切，或者其他宇宙也是从其自己的大爆炸而来。在多重宇宙中，除了我们自己的宇宙，还有很多其他宇宙。在这种情形下，将会有许多其他膨胀区域——每一个都构成了它自己的宇宙。

这个推理给我们提供了三个选择：宇宙开始于大爆炸；它一直存在，但最终还是经历了大爆炸理论所预言的膨胀阶段；我们的宇宙只是众多宇宙中的一个，而其中的每一个宇宙都产生于已经存在的一个宇宙或者多重宇宙。这涵盖了所有的可能性。不过对我来说，最后一

个可能性看起来似乎是最有可能的，因为它无须假设地球，甚至我们已知的宇宙是特殊的，这个推理自从哥白尼时代就一直被使用。依照我的思考方式，这一选择也意味着，宇宙的空间尺度很可能是无限的，而不是有限的，演化的宇宙也不可能有一个时间起点和终点，尽管我们的宇宙确实有可能“有始有终”。在这三个不能完全理解的可能性中，多重宇宙出现但最终消失的可能性，或许是最不能令人满意的一个。

上述问题促使我做了最后一个哲学思考：多重宇宙是否存在？现存的物理理论表明，多重宇宙是非常有可能存在的，尤其考虑到目前很多量子引力理论表明，存在很多可能的解。不知道那些计算是不是经得起推敲，不过我敢打赌，其他无法探知的宇宙是应该是存在的。为什么要否定它们呢？仅根据目前已知的物理定律和科技水平，就认为它们不存在，完全是一种短视。在我们世界里，没有事物与多重宇宙的存在是相违背的。

这并不意味着我们会知道什么。如果没有事物比光运动得快的话，任何在宇宙视界之外的区域都将超出人类观测的极限。但从原则上讲，其他区域可能包含其他宇宙，且完全与我们的宇宙分隔开来。在某些情况下，一些来自其他宇宙的信号可能会被探测到：随着时间演化，不同的宇宙会发生接触。但这是非常不可能的，因为其他宇宙通常是无法接触的。

对于我的那些忠实的老读者，我想说句题外话。在我们讨论多重宇宙时，我并不是指《弯曲的旅行》一书中所描述的多个维度。在宇宙的视界之内，也许还存在着额外维度，我们所能观测到的三维空间包括左-右、上-下和前-后。尽管没有人看到那个额外维度，但它也是可能存在的。从理论上来说，一个被高维空间所分隔开来的宇宙也是可能存在的。这种类型的宇宙通常被称为膜世界（braneworld）。读

过《弯曲的旅行》的读者知道，最让我感兴趣的是，膜世界很有可能有一些可观测的结果，因为它们无须很远。然而，当大家讨论更为广义的多重宇宙时，会涉及许多独立宇宙，即使通过引力它们也不会作用，膜世界并不是通常大家所指的这种独立宇宙。多重宇宙通常是非常遥远的，如果一个物体以其中的一个宇宙中以光速飞行，那么在我们的宇宙寿命之内，也不会抵达我们这里。

大众对多重宇宙的想法有着浓厚的兴趣。最近我和一位朋友聊天，在提到多重宇宙的想法时，他表现得非常兴奋，但他不明白为什么我没有像他那样觉得多重宇宙有趣。对我而言，第一个原因就是上面说过的：在所有的可能性中，我们永远也不会肯定地知道，我们是否居住在一个多重宇宙中。即使其他宇宙的确存在，它们也是不可探知的。尽管我的朋友有些失望，但兴趣依然。我怀疑他和其他人一样，对存在多重宇宙感兴趣只是因为，他认为另一个自己就生活在某个遥远的宇宙中。必须要指出，我并不赞同这种观点。如果其他宇宙存在，它们很有可能跟我们的宇宙完全不同。它们甚至不会包含类似于我们这个宇宙的物质和力。如果那里有生命的话，我们很有可能不能识别他们。即使它离我们并不遥远，也仍然不能探测到它。将创造人类的无数因素汇聚在一起更是不太可能。与许多宇宙一起，一个更大的宇宙也是有可能存在的。在我给朋友解释完之后，他开始明白我的观点。

事实上，即使多重宇宙的设想是正确的，大多数其他宇宙也将是不稳定的，它们或许会坍缩或者爆炸，无论是哪种情形，几乎在瞬间它们就会变得什么都没有。仅仅只有一些，像我们自己的宇宙，或许会持续很长的时间，形成一定的结构，甚至演化出生命。尽管哥白尼的观点很深刻，但我们这个宇宙看起来的确有一些特殊的性质——允许星系、太阳系的存在和生命的产生。一些人试图假设有多个宇宙存在，其中至少有一个宇宙会具有我们这个宇宙的特殊性质，从而允许生命的存在。许多这样思考的人尝试人择原理（anthropic reasoning），这个原理试图从生命（或者至少能够支持生命的星系）的重要性上来判断，是不是我们这个宇宙所必须具备的特殊性质。这里的问题是，我们不知道哪一类性质是人择决定的，哪个是建立在基本物理定

律之上的，或者哪些性质对于生命是必不可少的，哪些是我们主观认为对生命必不可少的。人择原理或许在某些情形之下是正确的，但问题是，我们不知道如何验证这些想法。在所有的可能性中，如果有一个更好更有预见性的想法出现，我们将会把上面那些想法排除。

上文所讨论的想法都只是推断而已。尽管它们很有趣，但我们不会得到答案，至少在短时间内不会。在我的研究中，我倾向于认为物质的多重宇宙“社区”就在这里，而且我们有希望理解它们。尽管我使用了形象化的术语，但距离真相仍然遥远。一个暗物质的宇宙就在我们眼前，然而我们通常与它没有相互作用，也不知道它到底是什么。理论和实验物理学家目前正在推进对“暗宇宙”的了解，我们或许很快就会知道答案。我也相信，这样的发现是值得等待的。

04

几乎就是最开始：一个很好的起点

BARBARA WALTER AND THE UNIVERSE

如果宇宙正在膨胀，那么它会膨胀到什么里面去？

一位非常有趣并且坦率的俄罗斯理论物理学家在准备下一周的一个学术报告，当他在喝咖啡时和大家聊起这个的时候，让大家很是吃惊。物理类的学术报告通常是一个总结性的报告，面向学生、博士后、教授和所有具有物理学背景的人。但这位个性的物理学家打算在报告上“谈论宇宙学”。当被别人指出这个话题可能有点太宽泛，毕竟宇宙学是整个领域时，他认为在宇宙学中只有几个想法和几个值得测量的量。他自认为在一个小时的报告中，自己完全能够覆盖所有这些内容，包括他自己的贡献。

我会让你自己判断他对宇宙学的极端观点是否正确。我想强调的一点是，我对此观点很是怀疑，许多问题还有待探索和理解。事实上，宇宙早期演化的优美之处就在于，许多方面它简单得令人惊讶。天文学家和物理学家通过观察今天的这同一片天空，能够反推得到宇宙数十亿年前的组成和活动。在本章中，我们将探讨人类在理解宇宙历史方面所取得的惊人进步，这些都是过去一个世纪的理论 and 观测带给我们的。

宇宙大爆炸理论

我们没有工具来可靠地描述宇宙的原初。尽管不知道宇宙如何开始，但这并不意味着我们知之甚少。宇宙的原初并没有已知的理论能够对其进行描述，在宇宙爆炸之后的极短时间开始，宇宙的演化才遵循已有的物理定律。利用相对论方程以及简化的宇宙成分的假设，物理学家可以确定大约 10^{-36} 秒之后宇宙演化的行为，也只有在这个时刻之后，用来描述宇宙膨胀的大爆炸理论才能够适用。早期的宇宙充满了物质和辐射，它们是均匀分布并且各向同性的，也就是说在各个区域和各个方向上都是一样的，因此仅仅需要几个量就完全足够描述宇宙的早期物理性质。这种特性使得宇宙的早期演化很简单，可以预测，还很容易理解。

宇宙大爆炸理论的核心是宇宙的膨胀。在20世纪20年代和30年代，俄罗斯气象学家亚历山大·弗里德曼（Alexander Friedmann）、比利时牧师和物理学家乔治斯·勒马（Georges Lemaître）、美国数学家和物理学家霍华德·罗伯森（Howard Robertson），以及英国数学家亚瑟·沃克（Arthur Walker），后两人一起工作，分别得到了爱因斯坦广义相对论方程的解，推断出随着时间流逝宇宙必须膨胀（或收缩）。他们甚至还计算出了空间膨胀速率在物质和辐射的引力作用下如何变化，而物质和辐射的密度随着宇宙的演化也是变化的。

宇宙本来很可能就是无限的，因此宇宙膨胀也许是一个奇怪的概念，但其实是空间本身在膨胀而已，也就是说像星系之类的天体之间的距离会随着时间的流逝而增加。我经常被问：“如果宇宙正在膨胀，那么它会膨胀到什么里面去？”答案是，它没有膨胀到任何东西中去，空间本身也在膨胀。如果你想象宇宙是一个气球的表面，气球本身在变大（见图4-1）。如果你在气球的表面标记了两点，那这两点的距离会变得越来越远，正如膨胀宇宙中的星系在彼此退后一样。我们的类比并不完美，因为气球的表面仅仅只是二维的，而它实际上是膨胀到了一个三维的空间。如果想象气球的表面是所有的一切，也就

是空间本身，那么就可以做这个类比了。如果这是真的，即使没有任何事物可以膨胀到里面去，标记的点之间的距离仍然会继续变大。

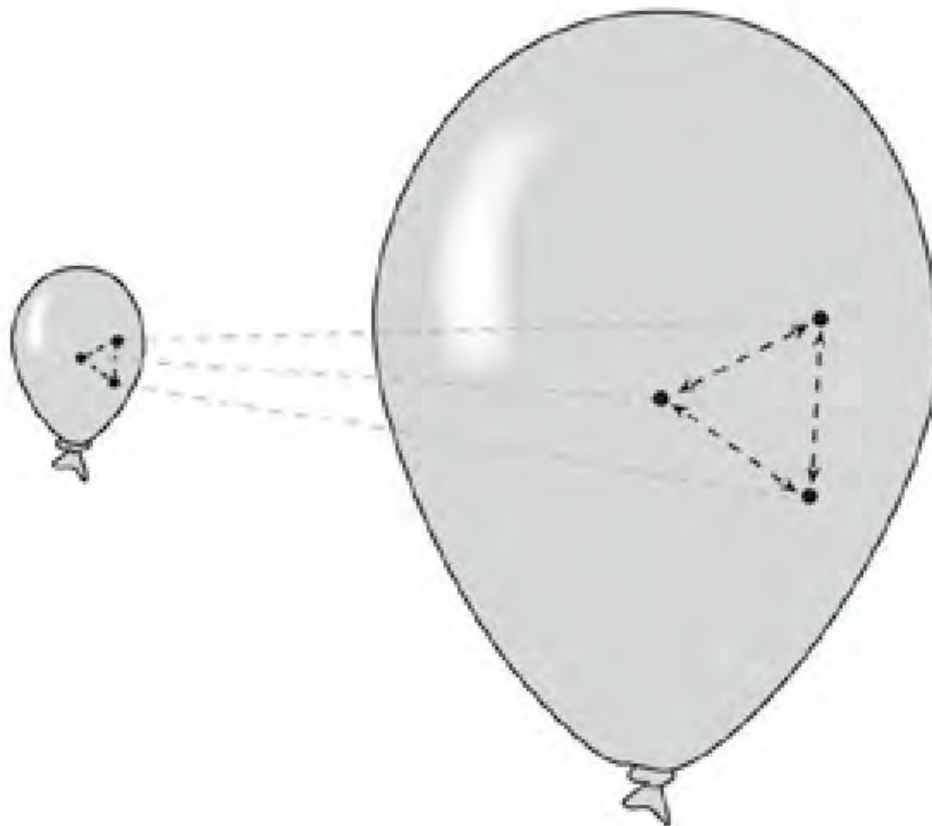


图4-1

随着宇宙的膨胀，星系相互远离对方，这和随着气球膨胀，上面两个点之间的距离逐渐变远的过程非常像。

宇宙的年龄谜题

继续上述类比，只是标记点之间的空间在膨胀，而不是点本身。即使在一个膨胀的宇宙中，恒星、行星或者被很强的引力或者其他力紧紧约束在一起的天体，都不会经历让星系彼此分开的膨胀。原子（包含一个原子核和一些电子）由于电磁力的作用，彼此靠近，体积

不会变得更大。相对致密的强束缚结构，比如星系或我们的身体，它们的密度超过宇宙平均密度万亿倍，但仍然不会发生膨胀。驱动膨胀的力也会作用于这些致密的束缚系统，但因为力的贡献非常强大，我们的身体和星系不会随着宇宙的膨胀而增长；或者即使会膨胀，那么膨胀的量也可以忽略不计，我们永远不会注意或测量到它的效应。物体的大小保持不变，只是随着空间的变大，物体之间的距离变得越来越远。

众所周知，爱因斯坦是从相对论方程中得到宇宙膨胀结论的第一人。然而，因为他所得到的结果是在宇宙膨胀被发现之前，所以他对自己的观点也不是特别自信，更没有宣传。在试图解决他的理论预言和静态宇宙之间矛盾的过程中，爱因斯坦引进了一个新的能量。在他看来，这个新的能量来源能够阻止预测中的宇宙膨胀。埃德温·哈勃（Edwin Hubble）在1929年证明了爱因斯坦的解决方案是错的。他发现，宇宙实际上是膨胀的，随着时间增长，星系彼此互相远离（但让人难以置信的是，哈勃是一个不相信任何特定理论的观测者，他并没有接受对他发现结果的这种解释）。爱因斯坦因此欣然放弃了他做的“傻事”，并且将其称为自己“最大的错误”（这也许是后人杜撰的）。

然而，爱因斯坦提出的那种能量确实存在。最近的测量显示，需要一种新的能量，我们现在称之为“暗能量”——虽然不是能够抑制宇宙膨胀的那种类型，却是精确解释观测所必须的，而且是一种完全相反的效果：宇宙在加速膨胀。我认为，爱因斯坦之所以真的认为是他的错误（如果他真的那么说过），是因为他没有意识到最初膨胀预言的正确性和意义，而这本来会成为他的重要预言。

公平地说，在哈勃提出他的结果之前，我们对于宇宙的认知非常少。哈罗·沙普利（Harlow Shapley）测定了银河系的大小，直径为3

0万光年，但他相信银河系包含了宇宙的一切事物。20世纪20年代，哈勃意识到这一说法并不对，因为他发现许多星云实际上包含很多星系，距离我们数百万光年，而沙普利原来认为那是一大团尘埃云（这也是一个名副其实的平常名字）。20年代后期，哈勃获得了一个让他更为出名的发现：星系红移（redshift of galaxies），即光的频率的改变能够告诉科学家，宇宙在膨胀。就像一个行驶中的救护车的警报声变低会告诉人们它已远离一样，星系的红移说明其他星系在远离地球，表明地球处在一个星系互相远离的宇宙当中。

哈勃常数

又称哈勃定律，是指河外星系退行速度与距离的比值，是一个常数，通常用H表示，单位是公里/（秒·百万秒差距），也是目前学界认为的宇宙膨胀速度。

如今，我们常会提起哈勃常数（Hubble constant），这是宇宙目前膨胀的速率。现在来看，它是一个常数，在空间的各个地方，它的值都是相同的。但实际上哈勃参数不是一个恒定值，它会随着时间发生变化。在早期的宇宙中，当事物更为致密、引力效应更为强烈的时候，宇宙的膨胀要比现在快许多。

直到现在，我们对哈勃常数的“测量”仍然有很大的不确定性，这意味着我们不能精确地确定宇宙的年龄。宇宙的寿命取决于哈勃常数的倒数，所以如果这个测量有2倍的不确定，那么年龄也有2倍的不确定。

记得我还是个小孩的时候，在报纸上看到一些测量已经引起宇宙年龄的发生改变。因为不知道那些数字代表了膨胀速度的测量结果，所以我当时非常惊讶。像宇宙年龄这样重要的事情怎么能够被随意地改变呢？原来即使不知道宇宙的确切年龄，我们也可以在一个定量的水平

上，理解许多有关宇宙的进化结果。也许，更精确的宇宙年龄估计会促使我们更好地理解宇宙的成分和其中的基本物理过程。

无论如何，这种不确定性现在得到了更好的控制。当时还在卡耐基天文台的温迪·弗里德曼（Wendy Freedman）和她的合作者测量了膨胀速率，并且最终平息了争论。事实上，因为哈勃参数的值对于宇宙学非常重要，所以共同努力确保了最大可能的精度。利用哈勃太空望远镜，天文学家得到的测量值为72公里/秒/Mpc（意味着在百万秒差距距离上的事物，以72公里/秒的速度后退），其精确度为11%，与哈勃之前最初的不准确测量值500公里/秒/Mpc相去甚远。

Mpc（megaparsec）即100万个秒差距，而秒差距，就像许多天文单位，是早期测量的一个传统。它是“秒视差”（parallax second）的缩写，与物体在天空中张开的角度有关，这就是为什么它有一个角度单位。就像许多因为历史缘故而保留下来的测量单位一样，许多天文学家仍然使用这个单位，但不少人不太喜欢用秒差距做单位。将其转化成或许我们稍微熟悉的距离单位，一个秒差距大约是3.3光年。这个神秘的单位和更容易被解释的物理量大小相当，这仅仅是一个偶然的巧合。

对于哈勃参数，哈勃望远镜的结果有10%~15%的不确定性，而非2倍的不确定，从而使其结果更为精确。最近，微波背景辐射数据的研究给出了更好的结果。我们现在可以把宇宙年龄的不确定限制在几亿年之内，而且测量精度一直在提高。当我写《弯曲的旅行》时，宇宙年龄还是137亿年，但现在我们相信宇宙更老一些，即从大爆炸以来大约是138亿年。请注意，这不仅仅是不断变化的哈勃参数导致了结果的改进，我在前文提到的暗能量的发现也能对其有所解释，因为宇宙的年龄依赖于这两个量。

大爆炸演化的预言

根据大爆炸理论，宇宙起源于138亿年前的一个炙热并且致密的火球，它由许多相互作用，并且温度高于一兆兆摄氏度（ 10^{24}C ）的粒子组成。根据爱因斯坦的理论，所有已知的（也可能有未知的）粒子在以接近光速的速度向周围四处运动，不断相互作用、湮灭，又从能量中不断产生。所有强烈相互作用的各种物质有着一个共同的温度。

物理学家称充满早期宇宙那炙热且致密的气体为辐射（radiation）。出于宇宙学的研究目的，辐射被定义为任何以相对论速度运动的事物，意味着以光速或者接近光速的速度运动。要称得上辐射，物体必须具备足够多的动能，甚至要远远超过其质量所储存的能量。早期宇宙非常炙热，能量非常高，以至于由基本粒子组成的气体很容易满足这一标准。

此时的宇宙只有基本粒子，而没有原子（由原子核和电子束缚在一起）或质子（由更基本的所谓夸克组成）。在如此多的热量和能量面前，没有什么能被困在一个被束缚的物体中。

随着空间的膨胀，弥漫在宇宙四处的辐射和粒子变得稀薄、冷却下来。它们表现得就像是被困在一个气球中的热气体，随着气球的膨胀变得越来越不那么致密。每个能量成分的引力效应影响膨胀的方式不同，所以研究宇宙膨胀随时间的变化，能够让天文学家区分出辐射、物质和暗能量的不同贡献。物质和辐射会随着膨胀而被稀释，非常像警报声远离我们而去时，它的声音也随之变小，辐射会红移到更低的能量，所以辐射会比物质稀释得更快。然而另一方面，暗能量根本不会被稀释。

随着宇宙的冷却，当温度和能量密度不足以产生某个特定粒子时，尤其当一个粒子的动能不超过 mc^2 的时候（其中 m 是特定粒子的质

量， c 是光速），一些显著事件就会发生。对处在冷却中的宇宙，大质量粒子会一个接着一个地变重，通过与反粒子相结合，大质量粒子会发生湮灭，转变成能量，从而加热其余质量比较轻的粒子。这些大质量粒子逐渐脱耦，从而最终消失。

即使宇宙的成分发生了改变，也是直到宇宙大爆炸发生的几分钟之后，才会看到一些可观测效应。因此，我们将跳到宇宙成分发生巨大改变的时候，至少可以去验证它们。哈勃膨胀就是对宇宙大爆炸理论的一个验证。另外两个涉及宇宙成分的重大测量结果增强了物理学家的信心，从而相信这个理论的正确性。在宇宙极早期会形成不同类型的原子核，我们首先考虑这些原子核相对丰度的预测结果，这与已经观测到的密度非常吻合。

在大爆炸发生几分钟之后，质子和中子停止了单独飞行。温度降到足够低，使得这些粒子被强相互作用力束缚在一起成为原子核。也就是从那个时候起，最开始让中子和质子数目保持相等的物质相互作用不再起作用；而中子依旧能够通过弱相互作用力衰变成质子，所以它们的相对数目发生了变化。

因为中子衰变发生得足够慢，所以相当大一部分的中子能够存活足够长的时间，并和当时的质子一起形成原子核。氦、氘和锂原子核就这样产生了，然后这些元素（包括氢原子，当氦产生时，它的密度就会被消耗）作为宇宙遗迹的含量就可以确定了。不同元素的相对含量取决于质子和中子的相对数目，同时也取决于相对宇宙的膨胀速度所需物理过程发生的快慢。所以核合成理论（nucleosynthesis，这个过程现在是已知的）的预言结果检验了原子核物理理论，以及大爆炸膨胀的一些细节。观测与预测的结果惊人地一致，这是对大爆炸理论和核物理的一个重要确认。

这些测量结果不仅验证了现有的理论，也对新理论给出了限制。这是因为，当核丰度确立的时候，膨胀速度主要取决于我们已知物质类型所携带的能量。当时存在的任何新物质不可能贡献很多能量，否则膨胀速度将会过快。当我们在推测宇宙中是否存在物质时，这个限制对我和我的同事们来说非常重要。只有少量的新物质能够存在于平衡态中，并且它的温度与核合成中已知物质的温度相同。

这些预测的成功也告诉我们，即使在今天，普通物质的含量也不会远大于观察到的含量。过多的普通物质和核物理的预测结果将会与宇宙中观测到的重元素丰度不一致。与前一章描述的测量结果（它告诉我们发光物质不足以解释观测结果）相结合来看，核合成理论成功的预言结果告诉我们，普通物质不能解释所有宇宙中观测到的物质，这也在很大程度上排除了一个可能性：人们看不见一种物质，只是因为它不燃烧或反射得不够。如果发光物质中有更多普通物质，除非有一些新的成分，否则核物理预言的结果将不再适用。如果普通物质在核合成的过程中没有被隐藏起来的话，那么我们可以断定，暗物质肯定是存在的。

就宇宙学预言的详细检验而言，宇宙演化过程中最为重要的一个里程碑发生在比较晚的时候，大约在大爆炸之后的38万年。宇宙最初充满了带电和不带电的粒子。在这个时候，宇宙已经足够冷，此时带正电的原子核和带负电的电子结合形成原子。自此，宇宙开始包含了中性物质（一种不带电荷的物质）。

光子是传播电磁力的粒子，而带电粒子变为中性原子是一个实质性的变化。在带电物质不会让光子发生偏折的情况下，光子可以顺畅地穿越宇宙。这意味着，早期宇宙的辐射和光可以直接到达地球，而与宇宙后来可能发生的更为复杂的演化没有关系。我们今天所看到的宇宙背景辐射是宇宙演化了38万年时所产生的辐射。

这种辐射与宇宙开始大爆炸膨胀之后立即出现的辐射是一样的，但它现在正处于一个更低的温度。光子已经冷却，但它们并没有消失。今天的背景辐射温度是2.73开尔文^[7]，这是极其冷的温度。辐射温度也仅仅比零度高几度而已，这个零度也被称为绝对零度，是任何事物所能承受最低温度的极限。

从某种意义上来说，探测到这种辐射也是对大爆炸理论最确切的验证，并且可能是最有说服力的证据，这表明方程式是正确的。出生于德国的天文学家阿尔诺·彭齐亚斯（Arno Penzias）和美国人罗伯特·威尔逊（Robert Wilson）在使用新泽西贝尔实验室的望远镜时，于1963年意外地发现了这个宇宙微波背景辐射。彭齐亚斯和威尔逊当时并不是为了寻找宇宙的遗迹，他们感兴趣的是将无线电天线用于天文学研究。当然，附属于一家电话公司的贝尔实验室对无线电波也很感兴趣。

当彭齐亚斯和威尔逊试图校准他们的望远镜时，他们记录了一个均匀的背景噪声（犹如静态的），它来自四面八方，而且不随季节变化。这个噪声一直不消失，所以他们知道不能忽视它。由于这个噪声没有一个方向的倾向性，所以它不可能来自附近的纽约^[8]、太阳或前一年的核武器测试。在清理完望远镜中鸽子的粪便之后，他们得出的结论是：这个辐射也不可能来自鸽子的“白色介电材料”，彭齐亚斯礼貌地这么称呼这些粪便。

威尔逊告诉我，现在看来他们发现这个辐射的时间是多么幸运。他们当时并不知道关于大爆炸的任何事情。但就在附近，普林斯顿大学的理论物理学家罗伯特·迪克（Robert Dicke）和吉姆·皮布尔斯（Jim Peebles）却是知道的。当普林斯顿的物理学家发现，他们已经被贝尔实验室的科学家赶超的时候，他们当时还正处在设计实验来测量辐射遗迹的过程中。他们早已经认识到这种遗迹辐射对大爆炸理论的重要意义，但是贝尔实验室的人却还没有意识到。对于彭齐亚斯和威尔逊来说幸运的是，麻省理工学院的天文学家伯尼·伯克（Bernie Burke），既知道普林斯顿大学的研究结果，也知道彭齐亚斯和威尔逊的神秘发现。威尔森向我把伯克描述成他的早期“私人互联网”，而伯克所做的是把两者结合，让两方的相关人员接触。经过咨询理论物理学家迪克之后，彭齐

亚斯和威尔逊意识到了他们发现的重要性及价值。这个背景辐射的发现，使得贝尔实验室的这两位物理学家在1978年获得了诺贝尔奖。与更早的哈勃膨胀发现一起，这个发现确认了曾经预言宇宙在一直冷却和膨胀的大爆炸理论。

这是“科学在行动”很好的例子。这一研究是为一个特定的科学目的而做的，但带来了辅助性的技术和好处。天文学家开始并不是在寻找他们所发现的东西，但职业素养让他们没有忽视那些发现。尽管这项研究在开始时是在寻找相对很小的发现，却无意中催生了一个意义深远的发现。他们之所以能够获得这个发现，也是因为其他人也同时在以更大的图景思考着同样的问题。贝尔实验室的科学家们的发现是偶然的，但这个发现却永远地改变了宇宙学这门科学。

在这个发现之后的几十年内，这种辐射促进了宇宙学的重大发展，其详细测量帮助验证了宇宙暴胀理论的预言结果，而宇宙暴胀是发生在宇宙极早期的一个爆炸阶段。

宇宙暴胀

“科学的突破是如何发生的”这个问题还存有争议。是逐渐发生，还是突然发生，或者无论如何都会发生？这个问题的答案就像我们最初对宇宙膨胀的无知一样。在理解当今世界的变化速度这个问题上，考虑技术进步的影响或者环境变化的影响非常有用，但人们时常忽视这个重要因素的相关性。

关于变化速度的争论，在19世纪许多有关达尔文进化论的核心冲突中得到了很好的体现。我们将在第11章看到，争论使得查尔斯·莱尔（Charles Lyell）所支持的地质学中的渐进主义，以及他的助手达尔文所支持的突发地质变化的观点，形成了鲜明对比。突发地质变化的观点是由法国人乔治斯·居维叶（Georges Cuvier）提出的。居维

叶也意识到了另外一种极端的变化，表明新物种不仅会出现，就像达尔文所展示的那样，而且它们也会因为灭绝而消失。

关于变化速度的争论也是我们理解宇宙发展的关键。对于宇宙而言，第一个让人诧异的是，它一直都在演化。当大爆炸理论在20世纪初被提出时，它与神学上所青睐的静态宇宙非常不同，而静态宇宙是当时大多数人所接受的一种宇宙观。但另外一个，也是后来让人惊奇的理论认为，宇宙在最早期经历了一个爆炸性的膨胀阶段，也就是宇宙暴胀阶段。正如地球上的生命一样，无论是渐进的进程还是灾难性的进程，都在宇宙的历史中发挥了重要作用。对于宇宙而言，“灾难”是暴胀过程；而对于“灾难”而言，我所指的是这个阶段发生得非常突然和迅速。暴胀破坏了宇宙中最初存在的那些成分，但在暴胀结束之时，它也创造了充满宇宙的很多物质。

宇宙暴胀

该理论认为，宇宙初期经历了一个爆炸性的膨胀阶段，其速度非常快，使得宇宙的尺度在极短的时间内增大了几十个数量级。

迄今为止，我所讲述的是一个标准的大爆炸理论历史，它描述了一个不断膨胀、冷却、老化的宇宙。它是非常成功的，但并非全部。宇宙暴胀发生在标准的大爆炸演化之前。尽管我不能告诉你在宇宙的最开始发生了什么，但我可以比较确定地说，在其演化的最早期，或许早至 10^{-36} 秒的时候，这个被称为暴胀的有趣阶段就已经发生了（见图4-2）。在暴胀阶段，宇宙膨胀的速率远远要比它在大爆炸演化阶段的膨胀速度快得多，以至于在暴胀阶段，宇宙尺度在不断地呈指数级增加。比如，指数膨胀意味着当宇宙年龄是暴胀开始年龄60倍的时候，宇宙的尺度已经增加了一兆兆倍（ 10^{24} ）以上；而如果没有暴胀的话，宇宙尺度将只增加8倍左右。

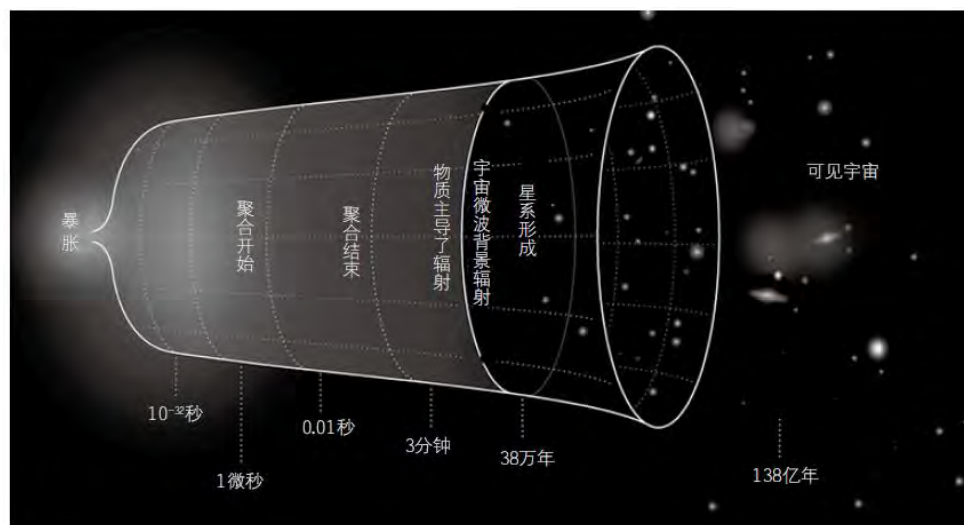


图4-2

暴胀和大爆炸演化的宇宙历史，包括原子核的形成、结构开始形成、在天空留下印记的宇宙微波背景辐射，以及现代宇宙（星系和星系团已经成形）。

也仅仅只花了几分之一秒，暴胀就结束了，留下了一个巨大的、平滑且平坦的均匀宇宙，然后就进入到了宇宙正常演化中，而其之后的演化就如传统的大爆炸理论所预测的那样。在一定意义上，暴胀就是“爆炸”，就像刚才描述的那样，它让宇宙演化变得更为平滑和更为缓慢。暴胀稀释掉了最初物质和辐射，因为快速的冷却使得温度降到了非常接近于零度。只有等到暴胀结束，驱动暴胀的能量转化成数量极其众多的基本粒子时，热的物质才会被重新引进来。当暴胀结束时，传统而缓慢的膨胀就开始了。从这个阶段之后，旧的大爆炸宇宙学就可以适用了。

尽管大爆炸理论很成功，但它依旧有几个悬而未决的问题，所以物理学家艾伦·古斯（Alan Guth）创立了暴胀理论。如果我们的宇宙成长于一个极度小的区域，那么为什么会包含那么多东西？为什么宇宙已经存在了那么长时间？基于引力理论，你可能会期待一个包含有这么多东西的宇宙已经膨胀成虚无，或者是已经很快坍缩。尽管它包含着巨大的物质和能量，宇宙三个无限大的空间维度是非常接近平坦

的，而宇宙的进化也足够缓慢，慢到能够让我们庆祝它138亿年的存在。

退耦

即防止前后电路网络电流大小变化时，在供电电路中所形成的电流冲击对网络的正常工作产生影响。

最初，大爆炸宇宙学当中的一个主要遗漏是，没能解释宇宙为什么如此均匀。当我们现在所观察到的宇宙辐射产生时，宇宙大约只有现在尺寸的1%，这意味着光可以走过的距离会小很多。然而，当观测者现在看到来自不同天空区域的辐射时，辐射看起来是相同的，这就意味着温度和密度的变化非常微小。按照最初的大爆炸理论，这一点非常让人困惑。当宇宙辐射和带电物质退耦的那个时间，宇宙年龄很小，因而光没有足够的时间来穿过哪怕是天空尺寸的1%。也就是说，如果时间倒流，你提出一个问题：存在于不同天空中的辐射，是否能够接收来自彼此的任何信号，或者发射任何信号给对方？答案是否定的。如果不同的区域从来不会相互交流，那么为什么它们看起来是一样的呢？这就好比1 000个陌生人住在有着不同商店的不同地方，同时阅读着不同的杂志，然而你却和他们穿着一致地走进剧院。如果你从未与他们接触过或者没有共享过信息，结果你们都穿得一样，那么这将是一个惊人的巧合。太空的均匀性更是让人吃惊，因为它的不均匀性在万分之一以内。现在看来，在宇宙开始的时候，有超过10万块的区域，而且这些区域间当时没有任何的交流。

由于大爆炸理论上的这些缺陷，古斯在1980年提出的想法看起来就很有吸引力。他提出，存在一个更早的时期，在这个时期之内，宇宙的膨胀极其迅速。而在标准的大爆炸理论中，宇宙在平稳地增长。在暴胀期间，宇宙经历了一个爆炸性膨胀的阶段。根据宇宙暴胀理

论，极早期的宇宙在极其短暂的一段时间之内，从一个极小的区域经过一个指数级变化增长到巨大区域。一束光线本来可以穿过的区域，在尺寸上增加了大约一兆兆倍（ 10^{24} ）。这取决于暴胀何时开始和它持续了多长时间，光线穿过的最初区域大小为 10^{-29} 米，在暴胀结束后，膨胀到了至少1毫米大，比一粒沙子大一点。从一定意义上而言，在一粒沙子里，或者如果你能够测量宇宙在那个时候的可见尺寸的话，至少在沙粒这个尺寸上，就像威廉·布莱克（William Blake）让你相信的那样，你的确会发现一个宇宙。^[9]

暴胀宇宙的极速膨胀解释了宇宙的巨大、均匀和平直的特点。宇宙是巨大的，它在极其短的时间里呈指数级增长，变得非常大。相比于经典大爆炸理论下的慢速膨胀，一个呈指数级膨胀的宇宙覆盖了远远大很多的区域。宇宙是均匀的，在暴胀期间，急速的膨胀能够平滑掉时空结构中的不均匀性，就像拉伸衣袖来消除折痕一样。一块微小区域中的所有东西本来都是非常靠近，并且通过辐射彼此交流，因为宇宙的暴胀，这块小的区域最后膨胀变成了我们今天所看到的宇宙。

暴胀也解释了宇宙的平直性。从动力学的角度看，宇宙的平直意味着宇宙的整体密度位于临界值上，从而宇宙可以在非常长的时间里保持这种状态。任何更大的能量密度都会导致正的空间曲率，通常一个圆球会具有这样的曲率，它会使宇宙迅速坍缩。任何小一些的密度都会导致宇宙膨胀速度非常快，结构就永远不会合并形成。从技术上讲，我讲得稍微有一点夸张。如果有一个极其微小的曲率，宇宙也还是可以持续这么长时间，但是，如果没有暴胀来判断这个值的话，曲率将会小得让你想象不到。

在暴胀情形之下，宇宙目前是如此巨大和平直，这缘于它在极早期的增长。想象一下，你可以把气球吹得像你想要的那么大。如果你把注意力集中在气球上的某一区域，当气球变得越来越大时，气球表

面会变平。同样，人们原本以为地球是平的，因为他们只看到一个巨大球面上的一小块区域。同样的事情对宇宙也是一样的：当宇宙膨胀时，它就会变平；不同的是，它的膨胀超过了一兆兆倍。

宇宙的极端平直是对暴胀的主要证据。这并不让人惊讶，因为平直性问题就是暴胀理论需要解释的几个问题之一。在暴胀理论被提出时，人们认为的宇宙要比简单预期所建议的宇宙平直许多，但人类没有必要的精确度来检验暴胀理论的极端预言结果。现在，在1%的精确度上，宇宙已经被测量是平直的。如果这不是真的，那么暴胀将已经被排除在外了。

20世纪80年代，我还是个研究生。那时候，暴胀被认为是一个有趣的想法，但是大多数的粒子物理学家都不是很重视它。从粒子物理学的角度来看，一个长期的指数级膨胀的环境似乎是非常不太可能的。事实上，它们今天还是不大可能。暴胀本来应该是解决宇宙膨胀初始条件的自然性的。但如果暴胀本身是不自然的，那么这个问题还没有得到真正解决。暴胀如何发生（包括它潜在的物理模型），仍然是一种推测。在20世纪80年代困扰我们构建模型的问题现在仍然值得关注。像斯坦福大学俄裔物理学家安德烈·林德（Andrei Linde），是第一批从事于暴胀模型的人之一，他认为暴胀肯定是对的。尽管暴胀想法的最早提出仅仅是因为没有人找到有关宇宙尺寸、平直和均匀性问题的其他答案，但暴胀却能够一下子解决这所有问题。

考虑到新近宇宙微波背景辐射的详细测量结果，大多数物理学家现在都支持暴胀。尽管我们还需要确定暴胀的理论基础，并且暴胀发生在很久以前，它还是可以产生可检验的预言结果，这使得大多数的人认为暴胀（或者非常类似暴胀的过程）发生过。最精确的观测集中在由彭齐亚斯和威尔逊所发现的2.73开尔文的宇宙背景辐射的细节上。美国国家航空航天局的宇宙背景探测器（COBE）测量过相同的辐

射，更全面，而且其频率范围更大上，确认了它在整个天空的高度一致性。

宇宙背景探测器最为壮观的发现是，早期宇宙并不是完全均匀的，这赢得了几乎所有对暴胀持怀疑态度的人。总体而言，暴胀使整个宇宙极其均匀。但暴胀又引进了非常微小的不均匀性，从而偏离完美的均匀性。量子力学告诉我们，暴胀结束的确切时间是不确定的，这意味着，在天空的不同区域，暴胀会在不同的时间结束。这些微小的量子效应会在辐射上留下偏离完美均匀性的印记。尽管很小，但它们却像当你把鹅卵石扔进池塘时，在水里的扰动一样真实存在。

在过去几十年中，最令人兴奋的发现之一无疑就是宇宙背景探测器发现宇宙的量子涨落，它产生于宇宙大约还是一粒沙大小的时代，并最终成为你、我、星系和宇宙中所有结构的起源。宇宙最初的不均匀性产生于暴胀结束的时候。尽管它们开始时的尺度极小，之后却被宇宙的膨胀拉伸到了可以形成星系种子的尺度，或者可测量结构的尺度（第5章我会详细解释）。

一旦这些密度扰动被发现，就像知道了温度和物质密度的微小偏离一样，对它们进行详细研究也仅仅是时间问题。从2001年开始，以更高的精确度以及在更小的角度尺度上，威尔金森微波各向异性探测器测得了密度扰动。和南极望远镜一起，威尔金森微波各向异性探测器观测到了辐射的密度涟漪或者扰动，而这个辐射密度中包含了最开始被创造出来的复杂信息。这些测量的细节证实了宇宙的平直性，确定了暗物质的总量，并验证了早期指数膨胀的预测结果。威尔金森微波各向异性探测器最惊人的一个结果就是，它通过实验证实了暴胀图像。

2009年5月，欧洲航天局发射了它自己的卫星——普朗克卫星，更为细致地研究了扰动。普朗克卫星的结果提高了大多数已知物理量的

精度，并且帮助巩固了人类对早期宇宙的认识。普朗克卫星最为重要的成就之一就是，它确定了一个对暴胀动力学可以作出限制的额外量，而正是暴胀动力学驱动了暴胀般的膨胀。正如宇宙大都是均匀的，但有一些小的违反均匀性的扰动。尽管天空中的扰动振幅大多是空间独立的，但也表现出对尺度的微小依赖。对于尺度的依赖反映了宇宙在暴胀结束时能量密度的变化。在这个对暴胀动力学令人印象深刻的确认中，威尔金森微波各向异性探测器和测量精度更高的普朗克卫星测量了这个尺度依赖关系，确定了早期暴胀阶段逐渐停止，并测量了限制暴胀动力学的数值。

虽然我们的理解并不完整，但宇宙学家已经确认暴胀和之后的大爆炸是宇宙历史中的一部分。因为早期宇宙有高度的均匀性，相对比较容易研究，所以我们可以建立详细的理论。未来，方程可以被求解，数据可以很容易地被评测。

然而，在数十亿年前，当宇宙结构开始形成时，宇宙从一个相对简单的系统变为一个更为复杂的系统，所以宇宙学在解释宇宙后来的演变时，面临着更大的困难。当恒星、星系和星系团这样的结构形成以后，宇宙成分的分布变得更难预测、更难解释。

但无论如何，仍然有大量信息包含在不断演化的宇宙结构中，而这些信息最终能够被观测、模型和计算机的综合运用所揭示。正如我们将在本书后文所看到的那样，对这一结构的测量和预测将会教给我们很多东西，包括暗物质与我们这个世界的相关性。现在，让我们先来探讨一下这个结构最初是如何产生的。

05

星系诞生

DARK MATTER AND THE DINOSAURS

然而，无论如何，不管比值如何变化，只要有光的地方，就会有暗物质存在。

你或许还记得在慕尼黑我和马西莫在晚餐时的对话，马西莫是一个品牌营销专家，他反对“暗物质”这个名字。在同一个晚宴上，马西莫向我介绍了另外一个参会者马特。马特问我人类驾驭这种尚未被理解的物质的能力有多大。作为一个游戏设计者，他问这个问题是可以理解的。不久后，一位从事剧本创作的朋友又问了我同样的问题。这也不足为奇，因为她在自己的科幻小说中对此做了预测。

以上这些询问仅仅代表了一个非常美好的愿望，我把它再次归因于“暗物质”这个名字选择得不好上。在我们周围的宇宙环境中，暗物质既不是一种有威胁的物质，也不是一种取之不尽、用之不竭的战略资源。鉴于已知物质和暗物质之间的影响极其微弱，没有人可以将暗物质收集起来放在地下室或车库里。利用我们的双手和普通物质制成的工具，我们既不能制造暗物质导弹，也不能制造暗物质陷阱，因为找到暗物质非常困难，而且驾驭它又完全是另外一回事。即使我们

能够找到一种方法把暗物质控制起来，它也不会以任何明显的方式影响我们，因为它只通过引力或通过一些其他力与普通物质相互作用，这些作用力因为太弱而不能被探测到，即使是非常灵敏的探测器也不行。若非存在巨大天文尺度的物体，暗物质对地球的影响将会更加难以被感知到。这也是它如此难以被找到的原因。

事实是，宇宙含有大量暗物质。弥散于宇宙四处的暗物质坍缩并且聚集在一起，创造了星系和星系团，并继而形成了恒星。

虽然暗物质并没有通过任何可以观测到的方式直接影响人类或实验室里的实验，但它的引力作用对宇宙结构的形成却至关重要。

因为大量暗物质聚集在这些巨大的坍缩区域中，普通物质也存在于此，因而暗物质将继续影响恒星的运动和星系的轨迹。我们将很快会看到，一种非常规类型的暗物质会坍缩得更为致密，从而会影响太阳系的轨道。尽管人类不能驾驭暗物质，但更为强大的宇宙却可以。本章将解释暗物质在宇宙演化中，以及在已知的、有限寿命的星系形成中的关键作用。

蛋与鸡

结构形成理论认为，宇宙暴胀最终形成了极度（但并非完全）枯燥而均匀的天空，恒星和星系就是从中发展而来的。本书多次提到的“结构形成的一致性图景”是一个相对较新的进展，但这个理论现在已经牢牢地扎根于宇宙学（例如以暴胀理论作为补充的大爆炸理论），以及有着更好测量的物质组成（比如暗物质）之上。这让我们能够解释，早期宇宙的炙热、无序和未分化区域，如何能够形成我们当今所看到的星系和恒星。

最初，宇宙是炙热的、致密的，并且基本上是均匀的，也就是说在太空的每一点都是一样的；它也是各向同性的，这意味着它在各个方向上都是相同的。尽管粒子之间会相互作用、出现，然后消失，但粒子的密度和行为在各个地方都是一样的。这个图像当然会非常不同于你看到的宇宙照片，也不同于当你仰望星空时看到的美丽夜空的图景。

宇宙现在不再是均匀的了。星系、星系团以及恒星挣脱了空间的膨胀，突显出它们在整个天空分布的不均匀性。这样的结构是我们这个世界中所有事物的核心。如果没有那些致密的恒星系统，就不会有这个世界的一切。这些恒星系统对重元素以及所有包括生命在内的神奇事物的形成，都至关重要。对生命而言，它至少要产生于一个恒星聚集的环境之中。

宇宙的可见结构产生于气体和恒星系统。这些恒星的集合有着各种尺寸和形状。双星——一个恒星绕着另外一个旋转，构成一个恒星系统，就如星系一样，不过星系是由数十万颗甚至万亿颗恒星所构成的。星系团的恒星数量又是星系的上千倍，它们也是恒星系统。

为了了解所涉及的物体类型，先来考虑一下宇宙所含物体的典型质量和大小。天文尺寸一般用秒差距或光年来衡量，天文质量通常以太阳质量来衡量，即等价于多少个太阳。星系的尺寸变化很大，小至1000万太阳质量的矮星系，大至100万亿太阳质量的最大星系。银河系是一个相对较小也比较典型的星系，大约为1万亿太阳质量，即其总质量，其中包含了占主导地位的暗物质。大多数星系的直径在几千到几十万光年。星系团的质量在百万亿到千万亿太阳质量之间，典型的直径在500万~5000万光年。星系团包含了多至大约1000个星系，而超星系团包含的星系数目更有星系团的10倍之多。

尽管这些物体至今仍存在，但早期的宇宙并不包含它们。早期的宇宙是极为致密的，恒星或星系还没有形成。对于恒星和星系而言，它们的密度要远远低于早期宇宙的密度。只有在宇宙冷却到某个温度之下，即宇宙的平均密度比最终形成天体的密度还要低时，恒星系统才有可能形成。结构形成也要等到宇宙中的物质能量比辐射能量高的时候才开始。请注意，我使用的是宇宙学中的辐射定义，它包含任何事物，包括诸如光子的粒子，也包含接近光速运动的粒子。在炙热的早期宇宙，几乎所有的事物都满足这个条件，因为温度非常之高，使得辐射在宇宙能量中占主导地位。

随着宇宙的膨胀，辐射和物质都被稀释了，它们的能量密度也同样如此。因为红移效应会降低辐射的能量，所以辐射能量将降低得更快，对于物质，在等待了10万年之后，终于走到聚光灯下，最终主宰了宇宙的能量。在这个具有里程碑意义的时刻，物质超过了辐射，成为宇宙能量的主要贡献者。

理解结构最初如何增长的一个很好的起点，大约就是从宇宙开始演化的10万年之后，这个时候物质开始占据主导地位。这个时间要比原初扰动开始增长的时间晚一些，但又比我们所观测到的宇宙微波背景的形成时间早一些。对于宇宙学而言，物质主导的意义非常重大，运动速度比较慢的物质携带有比辐射少得多的压力，因此会以不同的方式影响宇宙膨胀。当物质主导时，宇宙的膨胀速度就会发生变化。但对于结构形成更为重要的是，小而紧凑的结构可能就在当时开始增长。对于那些以光速或者接近光速运动的辐射，因为减速不够（速度仍然很快），就不会被那些弱引力束缚系统限制住。但辐射会抹平扰动，就像风会抹去沙子留在海滩上的涟漪。另外一方面，物质会减速并聚集成团——只有缓慢运动的物质才能够充分坍缩形成结构。这也是为什么宇宙学家有时说暗物质是冷的，这意味着它不是热的，并且不是以相对论速度运动的，从而与辐射表现出很大的差别。

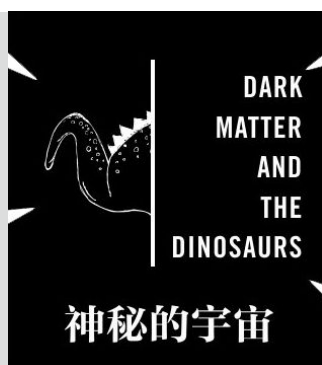
当物质在宇宙能量密度中占主导地位之后，密度扰动的区域会诱发物质坍缩，从而形成结构增长的种子。这些密度扰动的区域中，一些区域密度稍高，另外一些区域的密度会相对较低，而这些结构是在暴胀阶段结束时形成的。这些扰动随后增长，最终经过放大，将最初的均匀宇宙转化成我们看到的差异化很大的天空。小于万分之一的微小密度变化就足以从一个几乎均匀的宇宙中创造出结构，因为宇宙是平直的，这意味着存在一个临界能量密度，刚好位于快速坍缩和快速膨胀的边界上。临界密度创造了一个最佳点，从而宇宙能够缓慢膨胀并且持续足够长的时间，使结构能够得以形成。在这个微妙确定的环境中，即使很小的密度扰动也能导致某一区域中的物质坍缩，从而结构开始形成。

当这个坍缩导致结构形成开始时，两个相互竞争的力都有贡献。引力将物质向里拉，而尽管辐射不占主导地位，但它会将物质向外推。一旦超过某个阈值，平衡就会被破坏，这个阈值被称为金斯质量（Jeans mass）。在这块区域之内，向外的辐射压不能和向内的引力相平衡，从而导致区域内的气体坍缩，物质和增长于引力势中的天体变成了形成发光星系和恒星的种子。

高密度区域会比低密度区域产生更多吸引力，从而让本来致密的区域越来越致密，并进一步消耗掉周围区域中已经很稀疏的气体。因为富含物质的区域变得更加富有，而物质贫乏的区域变得更加贫乏，所以宇宙变得更加凹凸不平。这种物质聚集会一直持续，从而产生引力束缚的天体，并以正反馈的方式持续坍缩。恒星、星系和星系团都是在这个时候产生的，它们都是起源于引力效应作用于最初的暴胀结束之时，微小的量子力学波动。

具有吸引力的势阱能够吸引物质并使其坍缩。最初产生这种势阱的大都是暗物质，而非普通物质，一是因为它与辐射没有相互作用，

二是因为其总量更大。我们能看到恒星和星系，是因为它们能够发光。但实际上，最初是暗物质将可见物质吸引到这些致密区域的，从而才可能产生星系及恒星。当一块足够大的区域坍缩时，暗物质会形成一个球形晕。在这个晕内，由普通物质组成的气体会冷却，朝中心坍缩，最终碎裂形成恒星。



相比较仅仅有普通物质存在的区域，暗物质的存在会让这些区域坍缩得更快，因为更大的物质总能量密度会让物质比辐射更快地占据主导地位。除此之外，暗物质之所以很重要还有另外一个原因，即电磁辐射在开始时，会阻止普通物质在大约1%星系大小的尺度上形成结构。只有搭上暗物质这个便车，宇宙中星系大小的天体和恒星的种子才会有时间形成。如果没有暗物质导致的坍缩，就不会有目前恒星的星族和分布。

因此，是暗物质开启了坍缩，从而形成了结构。不仅是因为它更多，更是因为暗物质不受光的影响，所以电磁辐射不能将它们像普通物质一样驱散。暗物质从而建立了物质分布中的原初扰动，在物质和辐射退耦之后，普通物质会根据原初扰动作出反应。暗物质有效地给了普通物质一个很好的开始，为星系和恒星系统的形成铺平了道路。因为暗物质不与辐射作用，即使当普通物质不能坍缩时，它依然可以坍缩，形成基板，把质子和电子带到坍缩区域中来。

暗物质和普通物质同时坍缩成可见天体，比如星系和恒星，这对结构形成非常重要，同样对观测也很重要。尽管我们只能够直接

看到普通物质，但我们非常确信暗物质和普通物质存在于同一个星系中。由于普通物质依赖暗物质的原初扰动来产生形成结构的种子，所以普通物质多数情况下位于包含大量暗物质的结构中。因此，从某种意义上来说，“在路灯下寻找暗物质”的说法是恰当的。

暗物质在今天仍然发挥着重要作用。它不仅对吸引力的大小有贡献，防止恒星飞走，同时也把一些被超新星喷发出去的物质吸引回星系里。暗物质有助于把重元素保留在星系中，而这些重元素对于恒星的进一步形成和生命的出现都很关键。

尽管物理学家可以在理论的基础上预测早期的结构形成，但目前没有观测者可以目睹最初结构形成过程中宇宙相变的细节。望远镜探测到的光，有时甚至能够让我们看到数十亿年前形成的最早星系，但通常都是最近发射出来的。另外一方面，可观察的宇宙微波背景辐射来自宇宙还是充满辐射的那个时间，坍塌的引力束缚的天体尚未形成。背景辐射把早期密度波动的痕迹印在了38万年之后的宇宙进化中，但等到恒星和星系存在并产生可观测的光，又过去了大约5亿年。

在再合并之后（此时中性原子形成，微波背景辐射产生）和发光天体出现之前的中间一段时间，是一个非常黑暗的时代，目前的观测仪器无法对其进行观测。天体没有光发出，是因为恒星尚未形成。在早期，微波背景辐射会与无处不在的带电物质相互作用，但从此时起微波背景辐射也不再会照亮天空。这是用传统望远镜观测不到的（见图5-1），但这正是宇宙的原初之“汤”转化为丰富和复杂的宇宙最初结构的时代，这些最初的复杂结构演化成了我们现在所看到的一切。

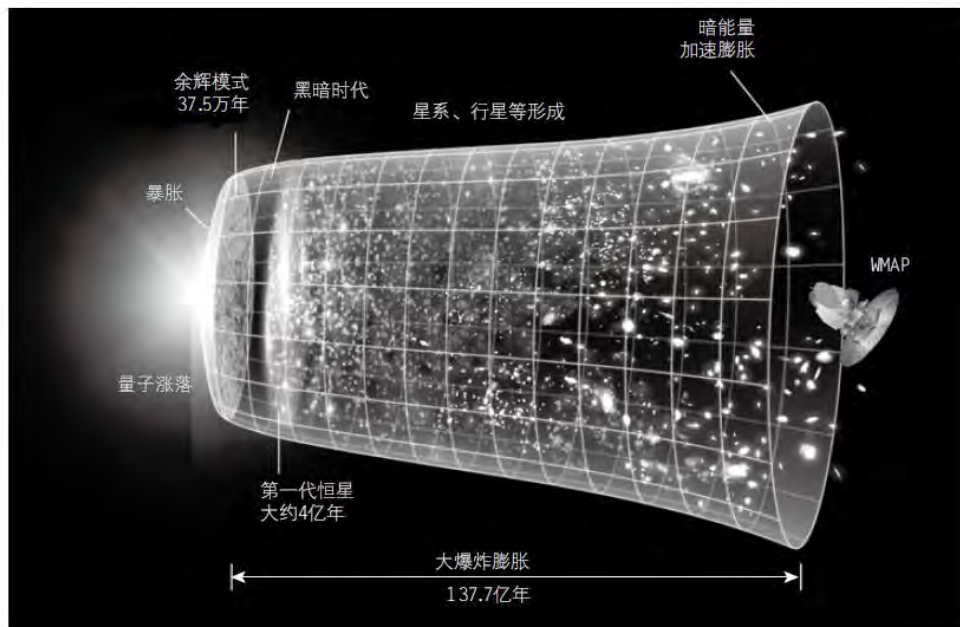


图5-1

在我们能够看到宇宙微波背景辐射之后，一个黑暗时代随之而来，结构开始形成，第一代恒星出现和爆发，星系和其他结构也随之形成，暗能量逐步主宰宇宙膨胀。（感谢NASA提供图片）

哈佛大学天体物理学家阿维·勒布（Avi Loeb）做了一个类比，目前的技术无法观察到最早恒星的形成，就如同我们不能目睹一只小鸡如何最初从鸡蛋中孵出来一般。一个鸡蛋包含了像汤一样的糊状结构。让一只母鸡孵蛋的话，从鸡蛋中就会孵化出一只小鸡，并最终成长为一只发育成熟的鸡。从破碎的鸡蛋中，除了蛋黄和蛋白，别无他物，但它们包含了可以孵出小鸡的“种子”。这个转变发生在蛋壳之内，如果不借助特殊工具，没有人可以看到内部究竟发生了什么。

我们需要新技术来探索最早的结构形成。没有人能看到宇宙进化时的黑暗时期，尽管已有一些理论提议。然而，就像一个鸡蛋，密度扰动包含了后续结构的种子。但是，不像“是鸡生蛋，还是蛋生鸡”的窘境，对于宇宙进化我们知道谁先谁后。

等级结构

上述结构形成的图景包含了大量与坍缩相关的物理知识，它是建立在各自的扰动过程之上的，每一个扰动都成为一个星系的种子，而每一个星系都会独立演化。进一步的研究表明，巨型恒星最先形成，但它们或是很快以超新星爆发掉，将第一代重元素释放到宇宙中，或是成了黑洞。这些重元素在宇宙之后的发展中扮演着重要角色。也只有金属（天文学家称这些重元素为金属）存在之后，更小的恒星（比如太阳）才能够在一些更冷更致密的区域形成，我们现在所能看到的结构才能够被产生。

但在恒星形成之前，星系就已经产生。事实上，星系是最先出现的复杂结构。每一个星系看起来都自成一体，但是，就像我们很快看到的，它们又都是联系在一起的，所以星系在很多方面都是宇宙的基石。一旦形成，星系可以合并成更大的结构，比如星系团。在足够的坍缩之后，恒星会在最密集的区域之内形成。我们今天所见结构的形成是开始于星系的。

然而，这个星系单独形成的图景是简化了的模型。在现实中，就如这个图景让你相信的一样，星系并不是孤岛。与星系的碰撞和合并对于它们的发展也至关重要。星系的形成是有等级的，较小的星系首先形成，更大的结构在之后形成。甚至那些似乎是孤立的星系其实也是被更大的暗晕所包围，而这些暗晕与其他星系的暗物质晕是接触在一起的。因为星系占据了一个相对比较大的空间比，大约为1%，所以星系比恒星要碰撞得更为频繁，而恒星所占的体积比差不多为千万亿分之一。通过合并以及其他引力相互作用，星系持续地互相影响着。星系持续吸引气体、恒星和暗物质进入，从而进一步演化。

有了这个更深层次的知识，让我们重新审视在结构形成过程中发生了什么。为了更好地理解这个过程，“富人更富、穷人更穷”的比喻非常贴切。就像当今世界的社会状况一样，穷人不仅变得更穷，也

会变得越来越多。事实上，在一些激烈的争论中，我有时会听到有关人性的灾难性预言，富人将会被挤到很小的区域里，随着越来越多贫穷阶层的增加，富人会被挤到城市的最边缘。在这种并不是很受欢迎的情形中，富人将居住在城镇的郊区，就像当我访问位于南非德班地区时，在白人居住的郊区看到的一样。邻近的城镇也将会经历类似的现象。一旦过分向外延伸，邻近城镇会相互影响，在相交的区域只留下富人。这些富有的、被隔离的人群可能会在商业和安全系统方面投资，但所有这些快速发展都会被留在某些节点当中，而这些节点是社会不同特权阶层相互交汇而形成的。

虽然这并不是一个很有吸引力的社会图景，但这和宇宙结构形成的演化方式非常相似。稀薄地区的膨胀比宇宙整体要迅速得多，而稠密地区扩大得更慢一些。结果是，稀薄区域把稠密区域挤到了一边，使得它们存在于从开始就一直膨胀的低密度区域的边缘区域。弥散区域中的气体被耗尽，最终演变成空洞，这些低密度区域就像牧羊人驱赶羊群一样，把物质驱赶到高密度区域，从而使自己的体积有所增长。

等级模型

在宇宙结构形成的演化过程中，物质的聚合方式。宇宙物质先是形成高密度纤维状结构，之后是形成宇宙网，其中，最致密的物质位于纤维相交的节点上。这些节点成为形成星系的基础。这一过程不断从小尺度到大尺度重复着，也即产生自下而上的等级模型。

当这样的高密度区域相交时，高密度区域的纤维状结构就会形成，其吸引力会把剩余的大量物质吸引过来。越来越多的物质被限制在一个宇宙网上，这个宇宙网是由很薄但是致密的片状结构和围绕着

片状结构的空洞区域构成的。宇宙网因此也变成了一个由纤维状结构组成的网络，在这个网络中，最致密的物质位于纤维相交的节点上。所以这一过程并不是一个简单的球形坍缩，物质会先沿着片坍缩成纤维状，而纤维相交形成节点（见图5-2）。然后这些节点成为形成星系的种子。这一过程随着时间持续进行着。小尺度结构形成后，这个模式在不断增大的尺度上继续重复。这就产生了自下而上的等级模型，在这个模型中，小的结构比大的结构更早形成，也就意味着小的星系先形成。

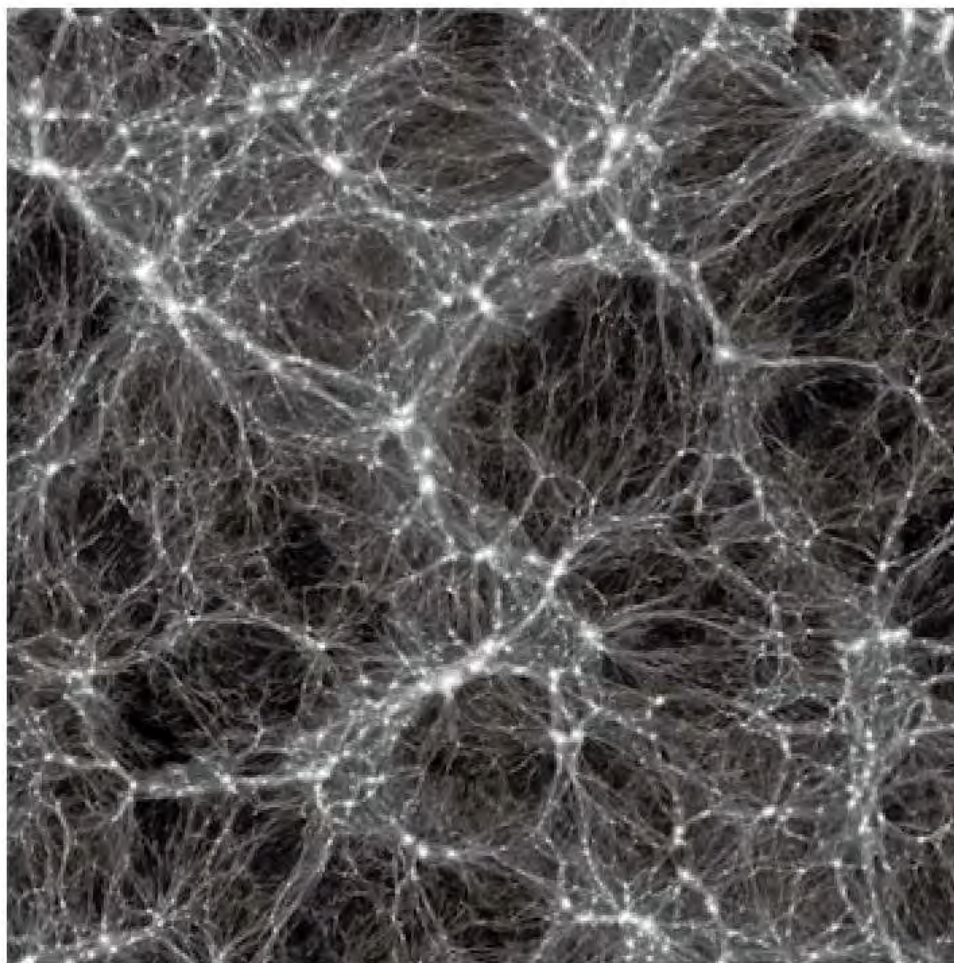


图5-2

由物质构成的宇宙网的模拟图。暗物质构成的纤维状结构相交形成节点，而这些纤维状结构又包含了暗的、相对空的空洞区域。非常亮的区域是星系团所在的地方，

位于节点上。这是贝内迪克特·迪默（Benedikt Diemer）和菲利普·曼斯菲尔德（Philip Mansfeld）通过一个厚为18Mpc、边长为179Mpc的数值模拟切片，利用了克勒（Kaehler）、哈恩（Hahn）和阿贝尔（Abel）在2012年研究出来的可视化算法，制作出来的暗物质密度投影图像。

数值模拟在最大尺度上证实了这些预言，暗物质正确地解释了宇宙中结构的密度和形状。较小尺度上的差异可能会对这一理论的进一步改进提供线索，但我们将把这些不完全确定的预言、观测和模型留到以后讨论。有些模型或许可以解释它们。

因为普通物质和暗物质同步坍缩，来自星系的辐射会追踪暗物质较多的区域。就像世界各地的灯光绘制出了各个城市所在的位置一样，宇宙中最亮的区域展示了星系中最为密集的区域，具有最多数量的恒星。辐射的强弱反映了整体的质量密度，就像世界上的光地图反映了人口密度一样。

然而要记住，就像光分布一样，与真实人口分布相比，我们所看到的部分可能会有变化。暗物质和发光物质之间的比值依赖于星系的类型，例如矮星系、正常星系或星系团。无论如何，不管比值如何变化，只要有光的地方，就会有暗物质存在。对于验证结构形成理论而言，这是一个很有价值的观测工具。

银河系，我们的保护伞

在结束这一章以及本书第一部分之前，现在让我们转向我们最为熟知的星系——银河系，以及我们最为喜欢的恒星——太阳，看一下普通物质在它们中的分布和影响。我们的星系因为具有乳白色的光带而被命名为“银河系”，这个光带在晴朗干燥的夜晚清晰可见。而我们所看到的光是银河平面上的众多暗星的光。

银河系位于一个被称为本星系群（Local Group）的星系群中。这是一个引力束缚的星系系统，它的密度要比宇宙的平均密度高。银河系和仙女座星系主导了这个星系群的质量，但也有几十个小星系属于这个群，绝大多数是这两个大星系的卫星。本星系群的引力束缚作用能够防止因为宇宙膨胀而导致的银河系和仙女座星系彼此远离。它们的距离实际上是在收缩的，在大约40亿年之后，它们将会碰撞并且合并。

银河系有一个由气体和恒星构成的盘，直径大约13万光年，在垂直方向上大约有2 000光年，这个平地状的结构使其具有了独特的形状。银河系盘包含了很多恒星和被称为星际介质（interstellar medium）的物质。星际气体由氢原子构成的气体和小的固体尘埃颗粒组成，它的总质量大约为恒星总质量的1/10。现实中，在银河系的中心附近聚集着银河系大多数恒星，但我们并没有真的看到更亮的光聚集于此，这是因为星际尘埃挡住了光。然而，天文学家在红外波段看到了星系中心，这是因为尘埃没有吸收这种频率较低的光。银河系的中心也包含着一个大约为400万太阳质量的黑洞，有时也被称为人马座A*。

中心黑洞和暗物质是完全不同的东西。然而，暗物质确实存在于一个巨大的球形晕中，直径大约是65万光年。就尺寸和质量而言，银河系中的最大组成部分，是一个重约一万亿太阳质量的近球形结构，它包括银盘区域。就像所有星系一样，暗物质首先凝聚，然后吸引可见的普通物质（见图5-3）。

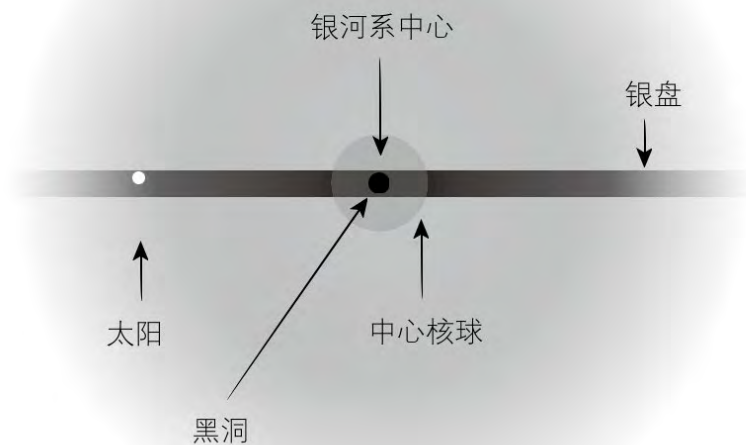


图5-3

银河系和它的中心核球、黑洞以及围绕在其周围的暗物质晕。太阳的位置也做了标记。

我还没有描述盘形成的过程与形成的原因，以及为什么这对“暗物质盘”这一提法以及小行星天体很重要（后文我将详细论述）。普通物质很有趣，因为与暗物质相比，它在星系中有着非常不同的分布。暗物质形成了一个弥散的球形晕，而普通物质可以坍缩成一个盘，比如银河系平面上熟悉的恒星盘。

普通物质与电磁辐射的相互作用导致了这一坍缩。普通物质和暗物质的一个重要区别是，普通物质可以产生辐射。如果没有导致冷却的辐射，普通物质将继续和暗物质一样保持弥散状态。实际上，因为普通物质的能量只有暗物质的1/5左右，它的致密程度会比暗物质更

低。然而，普通物质与光子的相互作用使得能量能够耗散掉，从而冷却，以致它能坍缩成为一个更集中的区域，也就是一个盘。通过光子辐射导致的能量损失很类似于蒸发的过程，水的汽化能够从你的皮肤上带走热能。但与辐射能量的物质不同，当你出汗并体温回归正常时，你是不会坍缩的。然而，因为普通物质会释放能量，气体会坍缩，聚集在一个较小的塌陷区域中，在这块区域中，它的密度比暗物质密度要高。

普通物质位于一个圆盘而不是一个小球体内，是因为物质的净旋转，而转动是从气体云中继承而来，气体云是从它们形成之初获得角动量（转动能量）。冷却降低了一个方向的抵抗力，但在另外两个方向上的坍缩是不允许的，或者至少因为所包含气体的旋转导致的离心力减缓了坍缩。如果没有摩擦力或者其他力作用其上，在圆轨道上运动的一块大理石将会永远保持转动。类似地，一旦物质开始旋转，它将会保持角动量，直到一些力矩作用其上，或者角动量随着能量一起被耗散掉。

由于角动量守恒，气体区域在径向（由旋转所定义）不能像在竖直方向上一样有效地坍缩。虽然物质在平行于旋转轴的方向上或许会坍缩，但它不会在径向坍缩，除非角动量通过某种方式被移除。这种差异性的坍缩导致了银河系的相对扁平的盘状结构，我们才会看到它伸展着穿过天空。上述原因也是导致大多数旋涡星系产生盘状结构的主因。

太阳和太阳系

尽管暗物质在银河系的净质量中占主导地位，但集中在银盘上的普通物质主导了在银河系盘面上的各种物理过程。普通物质在结构形

成开始的作用很有限，但因为其高密度以及核力和电磁力的相互作用，所以普通物质对于很多包括恒星形成的重要物理过程，都至关重要。

恒星是以核聚变为燃料，炙热、致密、被引力束缚的气体球，在星系的气体致密区域形成。当气体围绕星系中心旋转的时候，它就会碎裂成一些更为致密的团块，这些团块会进一步坍缩。在晕中，那些坍缩到非常高密度的气体就会形成恒星。

其中的一个气体球——太阳，就是诞生于45.6亿年前，它当时是一个充满活力的系统，引力、气体压、磁场和角动量在其中都起着作用。包含几乎和太阳系一样古老物质的陨石已经被发现，而且许多博物馆都有展品。太阳的位置非常靠近银盘的中间平面，位于距离中心大约27 000光年的地方，在径向上至少要比银河系中3/4的恒星要远。

像银河系中的千亿颗恒星一样，太阳以每秒220公里的速度绕银河系转动。按照这种速度，它需要大约2.4亿年才能完成一圈。由于银河系的盘面不到100亿岁，在这段时间内，其中的恒星转了不到50圈。对于系统而言，这些时间足够让系统均匀化，表现出一些稳定的特点。但实际上，并不需要这么长时间。

过去几十年中，很多科学知识都得到了快速发展，太阳系和它的形成就是其中之一。像大多数恒星一样，太阳和太阳系出现了一大团分子云气体，太阳周围的天体运动很快，碰撞发生得很频繁。大约在10万年之后，系统坍缩形成一个原恒星（protostar，这时核聚变还没有发生）和一个原行星盘（protostar disk，这个盘最终会演变成太阳系内的行星和天体）。大约在5 000万年之后，气体氢开始发生聚核反应，我们认为这就是太阳的开始。太阳吞噬了绝大多数的星云质量，但剩余的一些物质将会保留在太阳周围的盘上，这个盘最终会形

成行星和太阳系天体，比如彗星和小行星。一旦太阳产生的能量能够阻止引力收缩，太阳系就诞生了。

真正使我和我的合作者们感到惊讶的是，分子和重元素在冷却气体中发挥着关键作用，从而足以形成大多数恒星。重元素不仅仅对核燃烧很重要，它们对通过散射让物质冷却到某一温度也至关重要，这时，燃烧甚至变成了一种选择。形成类似太阳大小的恒星需要极端冷的温度，大约几十开尔文，过高温度的气体永远不会变得足够致密，所以无法点燃核燃烧。而另外一个神奇的关联存在于基本物理过程和宇宙的本质之间：如果没有重元素和普通物质经历的分子冷却，产生太阳的气体永远不会冷却下来。

在我开始新研究之后（与我之前的粒子物理学研究相比，新研究更多的是集中在对天文系统的细节研究上），我才真正理解了宇宙动态系统的美丽和一致性。星系形成、恒星产生、产生于这些恒星的重元素，以及恒星喷发出来的气体，对恒星的进一步形成都有贡献。尽管在人类时间尺度上看起来一切似乎都是静态的，但宇宙和其中的一切都从未真正平静过。不仅恒星是不断演化的，而且星系也是如此。

本书第二部分将集中在太阳系，并讨论小行星、彗星，它们的影响，以及生命的出现和消失。我们将会看到，在与人类直接相关的环境中，同样的作用和变化模式也是成立的。

DARK

MATTER

AND

第二部分
活跃的太阳系

THE

DINOSAURS

THE ASTOUNDING INTERCONNECTEDNESS OF THE UNIVERSE



06

流星体、流星和陨石

DARK MATTER AND THE DINOSAURS

天空中出现的壮丽光亮，源自太空中飞过的尘埃或卵石，浪漫而神秘。

当我在科罗拉多州大章克辛附近的沙漠观光时，有人借给我一副夜视眼镜，这让我非常高兴。这种特别设计的眼镜非常强大，能使光增强很多倍，通常对人眼来说太暗的东西也可以被看到，因此美国法律禁止将它们出口。军方用这些眼镜找寻敌军，而山民则用它们来寻找夜行动物。

我对那些功用没有兴趣，我只想利用它来观察天空，希望看到一些不用辅助手段永远也注意不到的暗弱天体。在晴朗的夜空中，最吸引我的就是在大气中频繁出现的流星。在仅仅几分钟的时间里，有大概5~10颗流星在我的视野里飞驰而过。我很幸运，因为我看向天空的时候正好有一场流星雨，闪光经过视野的频率比平时更多。然而，即使没有流星雨来增强频率，微小的沙粒也总是在大气层中持续燃烧。

这些尘埃粒子造成的流星让人特别激动。天空中出现的壮丽光亮，源自太空中飞过的尘埃或卵石，浪漫而神秘。当然，前提是它们不会冲到地面——没人愿意被高速坠落的石头砸中，且我们也肯定不

希望有大石头砸中地球。幸运的是，虽然在极少数情况下一些大个头会击中或靠近地球而造成伤害，但大多数靠近地球物体并不值得担忧。每天有上百万的小流星体，携带大约有50吨的天外物质进入地球的大气层，然而我们并没有受到什么明显影响。

第一部分以暗物质和整个宇宙为重点，只在最后简略介绍了银河系和太阳系；第二部分将集中讲述太阳系，尤其是可能与暗物质盘的存在相关的东西。这一部分将探索能够到达地球或者地球附近的天体，以及已经对地球生物造成关键影响的天文现象。本章讨论行星、小行星、流星、流星体、陨石以及各种容易混淆、不断变化的天文术语。第7章将转向另一种地球束缚轨道天体——彗星，及其“前身”所处的太阳系的更远地带。

模糊的边界

我和我的合作者大多是理论物理学家，这意味着我们研究的是基本粒子（即组成物质的基本成分）的性质。而天文学家主要研究天空中最大的物体，他们考察这些东西是什么，以及它们是如何由基本物质合并而成，并演化成今日之貌。众所周知，粒子物理学家总是在创造一些新奇的术语，或者用人名来命名尚未被发现的、有些纯粹是假想的对象，比如“夸克”“希格斯玻色子”“轴子”。和大多数天文学名字相比（它们经常被粒子物理学家笑话），我们的命名系统就显得非常有条理。天文学命名要考虑历史背景，而不是基于已知科学的解释说明。现在的天文学命名惯例和计量单位是非常晦涩并且不直观的。那些术语通常和事物发现时的已知知识或者仅仅是猜测的东西相关，而与人们目前对事物的理解无关。

例如，你可能以为星族I指的是宇宙中的第一代恒星，但星族I被用于指代一种稍后形成的恒星，星族II被用于指代另外一种。而当一种短暂存在的最早期形成的恒星被提出时，它们被叫作星族III。另一个同样令人混淆的例子是“行星状星云”这个称呼，它是红巨星演化的最终阶段，与行星一点关系都没有。这个让人困惑的名字起源于18世纪末期，天文学家威廉·赫舍尔（William Herschel）在望远镜里第一次看到这类天体的时候，误认为是行星^[10]，所以将其命名为行星状星云。

天体物理中有一些让人困惑不已的术语。人类在长达几个世纪的时间里一直进行天文学观测，并依据当时的观测得出结论，同时给出命名。而要在很久以后，能够正确解释这些术语本质的理论才会出现。只在极少数情况下，在一开始就有人理解了全部图像，但通常都是在一段时间之后才出现的。因为没有更好的理解，命名不可能建立在正确有条理的原则之上。

对行星、小行星和流星的术语命名也不例外。最初的分类非常宽泛，包含非常多不同类型的对象。只有从新天体的发现表明最初的命名不恰当之后，人们才意识到这个问题。虽然如此，最初的命名通常都会保留下来，不过具体的定义会随着时间发生改变。我总是对更改名字很谨慎。在商业或者政治活动中，改变称谓经常被用来转移对真正问题的注意力。然而，大多数天文术语的演化反映了真实的科学进步。目前，术语所代表内容的扩大是一个令人激动的现象，反映了我们对太阳系的理理解取得的重大进步。

行星的等级世界

最开始，“行星”这个术语是被广泛应用的。当古希腊人第一次发明这个词时，他们并不清楚天空中大多数物体之间的区别。当时的科学家需要更加精密的测量工具，才有可能了解到天空中看上去一样光亮的点是不一样的。有一件事情希腊天文学家可以观测到，那就是有些天体是在运动的，因此他们造出了一个单独的词“*asteres planetai*”（游走的恒星）来称呼它们。但是最初的定义不仅包括行星，还包括太阳和月亮。

后续的发现要求使用更精确的术语。虽然起初含义广泛，但随着时间推移，“行星”这个术语变得越来越严格。它的意思一开始是指肉眼可见的五大行星（不包括地球，在以地球为中心的模型中地球算不上行星），后来也包括其他由望远镜发现的行星。

现在我们知道，行星是在太阳诞生之后形成的。尘埃颗粒不断地积聚当时与之碰撞的大量物质，然后增长到与今天差不多的状态。这一过程大概需要几百万年、几千万年的时间。从天文学的角度看，这只是短暂的一瞬间。

行星的成分和状态取决于它的温度。对于小行星和彗星来说温度同样重要。你可以想象到，对于靠近太阳的行星，吸积到此行星上的物质，比吸积到离太阳更远的行星上的物质温度高得多。较高的温度使得水和甲烷处在气体状态，这个范围延伸到4倍日地距离的地方。因此最初在这里几乎没有这类物质凝结。此外，太阳发射出带电粒子，把其附近区域中的氢和氦冲走。因此，只有结实的物质，比如铁、镍、铝以及硅酸盐，这些在这样的高温下也不会融化的物质，才能凝结到内侧的行星上去。

正是这些物质组成了靠近太阳的4颗类地行星——水星、金星、地球和火星。这些元素相对罕见，因此这些靠近中心的行星需要一定时

间长大。它们能达到今天的体积，碰撞和并合必不可少，但是和靠外的行星相比，它们的体积仍然很小（见图6-1）。

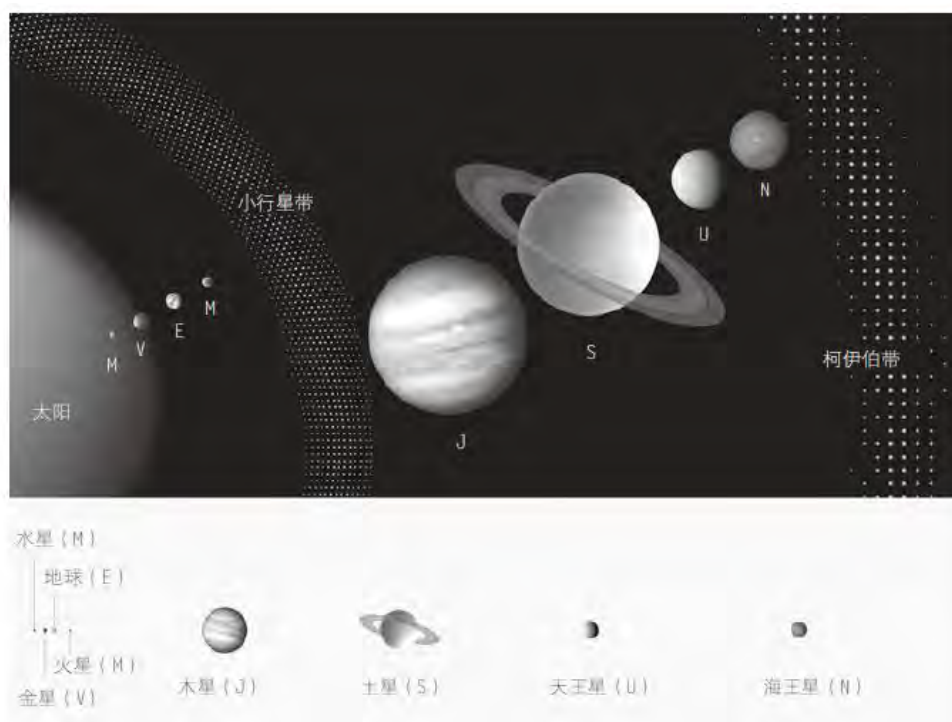


图6-1

4个靠近太阳的岩石组成的行星，以及4个靠外的体积较大的气态行星，相对大小如图所示。小行星带和柯伊伯带也在图中标出了。下方的图例给出了行星的名字以及它们在太阳系中的相对位置。

远离太阳的地方，在火星和木星的轨道之间有一个边界。在这个边界之外，水和甲烷这样易挥发的化合物冻结成晶体。这个外部区域的行星的增长更为高效，因为和构成类地行星的物质相比，构成它们的物质的丰度要大得多。这其中包括了氢，当行星足够快速地形成的时候，能够大量积聚氢。这4个气态巨行星（包括木星、土星、天王星和海王星）的总质量是太阳系质量（不包括太阳本身的质量）的99%。而木星，这个最接近物质能否积聚的分界线的行星，又占了其中的大部分。

在过去的20多年里，在太阳系的外部区域，发现了更多类行星物体，更不用说那些围绕在其他恒星轨道上的物体了。行星不再是一个简单的类别。这其中的成员彼此之间差异巨大，小的可以比月亮还小，大的几乎可以达到恒星核燃烧的尺寸。此前发生过多次对定义的修订。比如，谷神星（Ceres）被发现之后在长达50年时间里都被认为是行星，然后被重新归类为一个小行星。然而，不久之前的一次讨论引发了很多人的关注。

你也许记得当时的新闻故事，关于冥王星是否继续有资格作为一个行星。天文学家们现在仍在私下讨论这个问题，有时候甚至展开投票，想要让冥王星恢复原来的身份。近期的一些发现引发了最初的讨论。辩论愈演愈烈，但争论并非完全出乎意料，因为自从20世纪20年代，当冥王星首次被发现之时，人们就觉得它有些怪异。冥王星的轨道比其他行星的轨道偏心率更大，也就是轨道被拉长得更多。而它的轨道倾角（相对于太阳系平面的角度）比其他行星大得多。与太阳系里的其他距离太阳较远的行星（那些被称作气态和冰态的行星）相比，冥王星非常小。在行星王国里，它明显是个异类。

然而，70年后，有几个类似的天体在附近轨道被发现，表明古怪的冥王星其实并没有那么特别，本来对于它是行星的身份而言，也没有必要非要把它单独挑出来。对于改变冥王星分类的理由，简而言之，类似于制定许多随意的规则时所采用的理由一样：“如果我们让你加入，那么我们不得不让所有其他人也加入。”这是个偷懒的理由，为了避免更加微妙的分类而设定，不过它很难令人满意或者具有说服力。但是和冥王星大小以及轨道都相似的天体也被发现了。如果冥王星继续作为行星，那么另一个相似的在2005年被发现的叫作阋神星（Eris）的天体也应该是行星。还有好几个其他天体也很有可能如此。阋神星尤其麻烦，测量发现它比冥王星还要重27个百分点。之后有了更多类似的发现，某个人（或者某个组织）将不得不规定行星身

份质量的下限。但是如果把冥王星降级，问题就可以解决了。这就是国际天文学联合会（IAU）2006年在布拉格大会上作出的决定。他们利用了人们在此类情形下的通常做法：他们改变了准入规则。

所以，现在行星的定义为，由于自身引力成为圆球，并且“清扫了邻近区域”，在其附近不存在围绕太阳运动的更小天体的物体。这意味着，对于像冥王星和阋神星一样的天体，是附近有独立轨道的一些天体构成的天体带的一部分，所以它们就不再属于行星了。而像水星和木星这样的天体，近似球形，并且在它们的轨道上是孤立的，虽然互相之间差异巨大，它们仍是合格的行星。

这意味着，虽然我们大多数人都出生在一个太阳系有九大行星的时代，但现在，只有八大行星了。你也许会觉得这很让人沮丧，但这可能比不上美国在1984年进入大学的人所经历的。因为1984年7月17日法律上的一个改变，他们从合法饮酒的年龄被降级为未达法定饮酒年龄^[11]。2006年，当国际天文学联合会改变行星准入门槛的时候，冥王星被降级了。

有趣的是，最初对于阋神星和冥王星的相对大小的估计，后来被证明是个误导。虽然阋神星被认为比冥王星大，但误差非常大，天文学家不得不等待更详细的观测结果来核实这个说法。2015年7月，“新视野”号（New Horizons）宇宙飞船近距离飞越冥王星，拍下了壮观的图像，得到更加详细的信息，结果表明其实冥王星的体积更大。如果从一开始就没有这个不确定性，也许冥王星现在依然位于大行星的精英之列。

作为安慰奖，在“行星”被重新定义的同一次会议上，国际天文学联合会发明了术语“矮行星”（dwarf planet），来称呼和冥王星一样掉入小行星和行星裂缝之间的天体。冥王星成为这个新创建俱乐部的首位成员，也是模范代表。“矮行星”这个特殊的名字一直都有

争议。不像矮恒星其实就是恒星，矮行星并不是行星。这个名字的产生是由于最开始区分就不清楚。其他被提议的名字更荒谬，比如“微行星”（planetoid）、“子行星”（subplanet）。

和行星一样，矮行星围绕太阳运动，而不是像月亮一样绕着另一个行星运动。矮行星和小行星也不一样，小行星只是形状任意的岩石。按照定义，矮行星比小行星大，必须质量足够大，能够在其自身引力作用下成为近似的圆球形。但矮行星和真正的行星不一样，它们没有独立的轨道。很多其他天体也在附近有绕转。就是因为这个不孤立因素，它们没有清空近邻区域，从而被排除在行星身份之外。一位天体物理学同事开玩笑说，行星就像资深的教授，把附近的轨道都清空了；矮行星更像博士后，独立工作，但是他们的办公室挨着研究生；研究生就像小行星一样，还没有成型。

目前，矮行星的成员还有限。冥王星和谷神星是唯一被确认的矮行星，它们是小行星带上最大的，但又是已知矮行星中最小的天体。谷神星是内太阳系的唯一一个矮行星。在遥远的区域，妊神星（Haumea）、鸟神星（Makemake）和阇神星也被官方认定为矮行星，因为它们足够大，同时它们几乎可以肯定是圆的，虽然它们的形状还需要进一步进行可靠观测。其他的候选者也可能满足要求，比如神秘的塞德娜（Sedna），我们只有在完成更好的探测之后才能确定。然而，很多天文学家认为更多，也许数量多达100个或者200个的矮行星还存在于遥远的柯伊伯带中（这点我们稍后会讲到）。柯伊伯带很可能是上述物体起源的地方，也非常可能是有待发现的更多类似种类的源头。

小行星，太阳系最资深的漂流者

与“行星”和“矮行星”不同，“小行星”这个专业名词的定义仍然模糊，只是在口头上这样称呼它，天文学界还从未正式地定义它。直到19世纪中叶，“小行星”和“行星”这两个词还是可以交替使用的，并且通常认为是同义词。今天我们用“小行星”这个词的时候，通常指的是比流星体大，但是比行星小的物体，它在太阳系内圈，直径从几十米到几公里不等。《纽约客》杂志撰稿人乔纳森·布利策（Jonathan Blitzer）曾经对它们这样描述：“小行星是太阳系最资深的漂流者，有着岩石般的身体，围绕着太阳运转，从太阳系形成之时被遗留。其太小不能成为行星，但又并非小到可以被遗忘，它们能够揭示出很多关于宇宙原初历史的事情。”

与矮行星不同，小行星的形状通常是不规则的（见图6-2）。因为观测到的小行星旋转速率的上限很低，让科学家们认为，大部分小行星并不是结合紧密的天体，仅仅是岩石碎屑的聚集物。如果旋转速度更快，碎石块就会飞散。探测器已经到访过一些小行星，以及小行星卫星的观测都支持这一推测，证明小行星的密度较低。

小行星的数目非常多，也许有几十亿个，而它们的成分千差万别。它们中的大部分，要么是S型的，由普通硅酸盐岩石组成，大部分在火星附近发现；要么是C型的，富含碳元素，距离上更靠近木星。当人们研究太阳系中生命的起源时，特别注意了后一种小行星，因为碳元素对于生命至关重要。有趣的是，对于陨石所做的实验研究发现，一些小行星含有微量的氨基酸，由此使得它们更加受关注。下一章我们会看到，彗星也是如此，让它们成为另一个考虑生命起源时的重要主题（稍后我会讲述）。对生命而言，水很重要。一些小行星包含水，而彗星中普遍含有更多的水。主要由铁和镍组成的金属小行星上也有水。尽管至少有一个相对仔细研究的小行星，它有一个镍铁核以及一个玄武岩的壳层，但还是比较罕见的，只占小行星总数的百分之几。

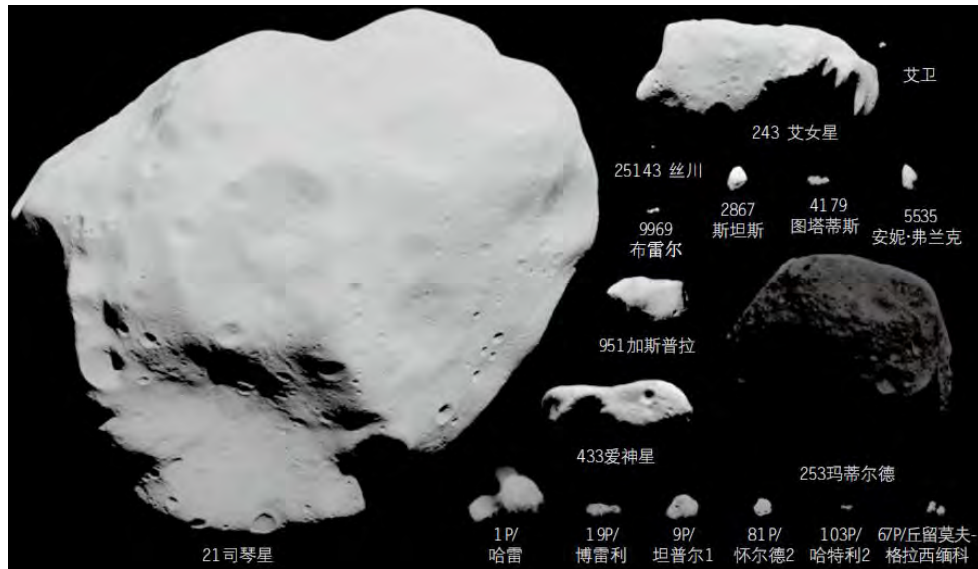


图6-2

截止到2014年8月，被探测器访问过的小行星和彗星的图像。它们的直径从几百米到100公里不等。（图像由艾米丽·拉克达沃拉[Emily Lakdawalla]合成。数据来自NASA/JPL/JHUAPL/UMD/JAXA/ESA/OSIRIS团队/俄罗斯科学院/中国国家航天局。图像由艾米丽·拉克达沃拉、丹尼尔·马查歇克[Daniel Machacek]、特德·斯特里克[Ted Stryk]、戈登·乌加尔科维奇[Gordan Ugarkovic]处理）

和行星不一样，小行星极少是孤立的。它们在太阳系中特定的区域做轨道运动，和附近许多其他的小行星一起绕转。大多数小行星位于小行星带。这一区域从火星一直延伸到包含木星轨道的位置，涵盖了从岩石类行星区域外边缘一直到遥远的气态行星（见图6-3）。小行星带延展宽度大约为2天文单位^[12]至4天文单位，距离太阳大约2.5亿公里~6亿公里远。在主带之外，另一类小行星——特洛伊（Trojans）族小行星，被一个更大的行星或者一个行星的卫星束缚，保证了其稳定性。

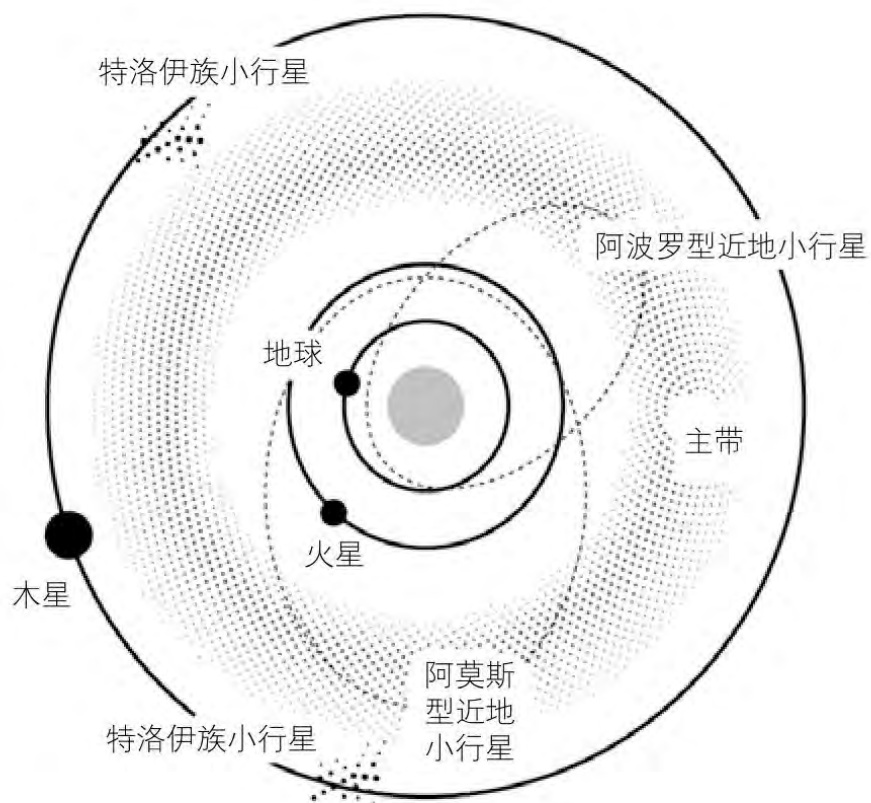


图6-3

火星和木星之间的主小行星带以及特洛伊族小行星，这是一个阿波罗型小行星和一个阿莫斯型小行星的例子。

小行星的分布

自21世纪以来，天文学家开始了解早期太阳系的行星迁移。关于小行星带形成的科学研究取得了很大进展。现在我们知道，在行星开始形成之后的几百万年时间里，由太阳发射出的带电粒子，把盘上剩余的大部分气体和尘埃都吹走了。行星的形成就此终止，然而太阳系的形成仍在继续。在此时间之后，行星会迁移，有时是破坏性的，把剩余的物体全都弹到太阳系外，或者把较小的物体移到周围。在过去几十年里，行星科学中很重要的一项研究进展是，理解并认识到这种

行星迁移对太阳系形成的作用。气态行星迁移得最显著，并且影响了小行星和彗星的形成。内部的行星也会向内迁移，但幅度很小，因此很可能起的作用也比较小。很有可能，当几个外部的行星向外迁移，而木星同时向内迁移的时候，大量小行星被抛到了太阳系内圈，开始了被称为晚期重大的撞击事件，它发生在大约40亿年以前（太阳系形成之后的大约5亿年）。月球和水星上大量由这些冲撞造成的陨石坑为这一事件提供了证据。

天文学家认为小行星是原行星盘（行星形成之前存在的盘）的遗迹。小行星带最初可能有远大于今天的质量，而大部分质量在太阳系充满活力的早期丢失掉了。在它们并合之前，木星将很多原本在这个区域的物体散射开来。这也许解释了这个区域行星缺失的原因，因为丢失了太多的物质。虽然这里有几十万个直径大于1 000米的物体，但这些物体今天的总质量也只有月亮的1/25，而谷神星自己的质量就占据了其中的1/3。把谷神星和另外3个质量最大的小行星加起来，就占了这些物体总质量的1/2，剩下的则由上千万个小质量物体构成。除了几十万个或者100万个直径大于1 000米的小行星之外，还有更多更小的小行星。虽然很难被识别，它们的数目在小规模地迅速增加。所遵循的经验规律是：尺寸小1/10，物体的数目大约增加100倍。

在把星子（planetesimal）从小行星带弹射出去的过程中，木星可能也把一些携带水的天体扔向了地球。虽然我们对地球上水的起源所知甚少，由于最初水比较容易在太阳系的外部冷的区域积聚到足够数量，这些木星在早期也许对地球获得丰富的水供给起到了一定的作用。有趣的是，在早期撞击时期结束“不久之后”（在地质学的尺度上），大约是38亿年以前，生命开始出现。幸运的是，从那时到今天，虽然小型撞击偶尔还是会发生，但小行星和彗星早已没有早期那么危险了。

和行星一样，最初发现的小行星被分配了各种符号。到1855年，这些符号达到20个的时候，它们的名字还带有神话色彩，而近期的发现则被分配了一些别出心裁的名字，包括流行文化偶像，比如“詹姆士·邦德”“柴郡猫”，甚至还有发现者亲戚的名字。看着这些代表单个小行星的符号，让我感觉像在看一块象形文字的匾牌（见表6-1）。正如我的合作者所说，这些名字非常像那些“曾被称作普林斯（Prince）的艺人”的名字，原名非常难读。这个比喻特别恰当，因为像普林斯一样，这些天体现在有更容易发音的名字。

表6-1 不同小行星及表示符号

小行星	符号	年份
1 谷神星		谷神的镰刀，像反着的字母 C 1801
2 智神星		智慧女神（雅典娜）的长矛 1801
3 婚神星		天后，婚神星权杖上的一颗恒星 1804
4 灶神星		灶神的祭坛和神圣之火 1807
5 义神星		一个天平，或者倒置的锚，公正的象征 1845
6 韶神星		韶华女神赫柏的杯子 1847
7 虹神星		彩虹（彩虹女神）和一颗恒星 1847
8 花神星		一朵花（花神）（特指英格兰玫瑰） 1847
9 颖神星		智慧之眼和一颗恒星 1848
10 健神星		健神星的蛇和一颗星，或者医神的蛇杖 1849

早期科学家获得发现的方式可能和古代埃及人区别不大，都是在一种神秘的状态下发现的。这并不代表古人没有努力地找寻规则。宇宙十分复杂，要弄明白宇宙所含物体的本质，时间、专注和技术都必不可少。在观测能力有限时，因为一个物体的尺寸、成分和位置的差

别，很难判断一个天体是不是更暗或更亮，更大或是更小。只有随着时间推移和更好测量工具的出现，才会带来真正的科学理解。

在行星最初被发现时，没人知道有小行星。小行星和行星一样，不会发出可见光。行星、小行星和流星体只有在反射太阳光时才是明亮的。然而，找寻小行星困难得多，因为它们的尺寸要小很多，更加暗淡因，从而更难以被看到。彗星有明亮的长尾巴，流星相对比较近、比较明亮，小行星则没有这些很明显的特征。因此，在以前发现它们很困难——现在也没有容易到哪儿去。

的确，人类花了几千年的时间才意识到小行星的存在。如果没有特别灵敏的工具，发现这些暗弱天体的唯一途径就是长时间地盯着它们看，当然如果事先知道往哪里看也会有些帮助。最初的尝试依靠后一种方式。天文学家并不知道最佳的目标位置，但他们采用了一个启发式的方式，认为这样可能会帮助指导他们的研究。提丢斯-波得定则（Titius-Bode law）所给出的位置似乎和已知行星的位置相符合，同时能够预言其他行星的位置。1781年，天王星在被这个定律所预言的位置被发现了，这似乎是一个很大的成功。即便如此，并没有一个真正的理论来判断此“定则”的正确性。但不管怎样，海王星的位置并没有遵从定则的预言。

尽管预言的位置很任意，但以18世纪的技术水平，寻找行星所采用的技术方法却相当可靠（请注意，当时还没有小行星被发现）。观测者比较不同夜晚的星空图，寻找位置发生改变的天体。近距离的行星运动很明显，而遥远的恒星看上去是不动的。利用这一方法（同时以提丢斯-波得定则做指导），西西里岛巴勒莫天文台的创建者和台长、天主教牧师朱塞普·皮亚齐（Giuseppe Piazzi）在1801年的新年第一天，在火星和木星之间发现了一个物体。数学家卡尔·高斯（Carl F. Gauss）之后计算出了这个天体到地球的距离。

现在我们知道，他们所发现的天体——谷神星，并不是一个行星，而是第一个被发现的小行星，它位于我们现在所知的火星和木星之间的小行星带上。伴随相继的几个这类发现，天文学家威廉·赫歇尔爵士（Sir William Herschel）提议用一个单独的名词“小行星”来描述它们。这个名字源自希腊词语“asteroeidēs”，意为“恒星形状的”，因为和行星相比，它们更像一个点状物。谷神星，现在我们知道它近于球形，直径大约为1 000公里，甚至比其他小行星更特别，它后来被证明是第一个被发现的矮行星。

人类最初对小行星所知甚少，直到科技和航天项目取得一定发展，更多这类物体被更好地观测到。而这一领域的研究人员所作出的出色进展令人惊叹。和发现小行星一样令人兴奋，观测和探索它们更加有趣。正在进行的航天计划已经设计了好几个近期的任务，都和这一目的相关。当近距离的图像首次揭示出小行星的不规则形态时，这些更加直接的探索将极大地改进早在20世纪70年代开始但没有细节的观测结果。

过去，其他值得关注的小行星任务之一是包括近地小行星探测计划“会合-苏梅克号”（NEAR Shoemaker），它是第一个小行星探测器。2001年，它给第一个靠近地球的小行星爱神星拍了照片，甚至登陆了上去。日本的“隼鸟号”（Hayabusa）项目在2010年带回了石质小行星样本。而近期，日本的科学家发射了更加野心勃勃的“隼鸟2号”，它将登陆一个小行星，并部署三个探测车，任务结束时它将会收集到更多样本。美国国家航空航天局将要发射源光谱释义资源安全风化层辨认探测器（OSIRIS-Rex），它应该会带回一些小行星的碳质样本。^[13]

在近期新闻中更加出名的是欧洲的罗塞塔（Rosetta）探测器。它先是飞过了直径大约为100公里的鲁特西亚和直径大约为4.6公里的斯坦斯两个小行星，并且获得了它们的详细信息。之后更为出名的事件是罗塞塔探测器与彗星会合。美国国家航空航天局的黎明号（Dawn）

探测器最近也有消息。它已经访问了灶神星（Vesta），并且现在已经到达了矮行星谷神星。

计划在未来对小行星进行雄心勃勃的开采行动现在也在考虑当中，虽然这不一定是最经济的路线，但很有可能也会使相当多的小行星被发现。目前正在进行的以小行星偏转为设计目标的探测器，例如美国国家航空航天局正在发展的小行星转向任务（ARM），也可能会获得同样的发现。目前的美国太空项目以小行星为重点：小行星与行星相比，没有那么有趣，但是通常更容易到达，而且它们很可能告诉我们更多关于太阳系的新信息。

流星体、流星和陨石

下面我们将从小行星转向更小的天体——流星体（meteoroid）。对流星体的研究是以听起来有些尴尬的术语“meteoritics”（陨星学，而不是“meteorology”[气象学]，meteoritics意指对陨石的研究），后者是一个听起来更合理的名字。这个源于希腊语“meteoreon”（在高空中）和“logos”（知识），在天文学采用它之前，气象研究已经使用它了。古代希腊人认为对天气的研究符合“meteorology”一词，即对天空中物体的研究。

流星体

在行星际空间运动的固态物体，尺寸比小行星小很多，同时又比一个原子大很多。

对于流星体的标准定义，直到1961年才由国际天文学联合会确定：在行星际空间运动的固态物体，尺寸比小行星小很多，同时又比

一个原子大很多。虽然与“气象学”相比，从天文学角度来看更加合适，但这个定义仍然不是很具体。1995年，两位科学家建议将尺寸限制在100微米~10米之间。但当直径小于10米的小行星被发现时，陨石协会的科学家们建议将尺寸范围改为10微米~1米之间，1米大约是观测到的小行星的最小尺寸。但这一改变建议一直没有被官方采纳。在本书中我将“流星体”来指代天空中那些中等尺寸的物体，但对于尺寸更小的物体，我将用更加精确的名字“微流星体”或者“宇宙尘埃”来指代它们。

和小行星一样，流星体的本质千差万别，这源于它们在太阳系中差别巨大的起源。有些是雪球一样的物体，密度只有冰的1/4；有些是致密的、富含镍和铁的岩石；有些碳含量很丰富。

虽然口语中使用的“流星”经常包含了创造它的流星体或者微流星体，对这个词的正确使用对应着这个术语的希腊语词根，其意为“悬浮在空气中”，并且仅指代我们在天空中看到的東西。流星是当流星体或者微流星体进入地球大气层时，所产生的肉眼可见的一道光。虽然定义如此，不过大多数人，甚至记者，在讲到流星掉向地球时都并不准确，就像1979年的电影《流星》(*Meteor*)的名字一样^[14]，但公平地说，它有它的娱乐重要性。

有趣的是，“流星”这个词的早期定义和天气相关，最初包含所有的大气现象，比如冰雹或者台风。风被称作“空气流星”；雨、雪以及冰雹被称作“水流星”；光的现象，比如彩虹和极光被称作“光流星”；而闪电以及我们现在所说的流星被称作“火流星”。这些术语是古时候的遗迹，那时没人知道那些东西都有多高，也不知道天气特征和天文学特征的缘起有着非常大的区别。“气象学”作为专有名词也许并不完全是误导，因为天气的确与地球在太阳系中的位置相

关，当然和最初想象的非常不同。幸运的是，尽管“气象学家”早期对“流星”有误解，不过这个词后来再没有这样被误用过。

人们很容易看见流星，因为引起这个现象的物体进入大气层后被加热，会产生让人眼看得到的光的物质，而且流星体的运动速度很快，产生的光呈弧形。虽然很多流星随机发生，但是流星雨的发生却更加规律，产生于地球经过彗星的碎片。当然在太阳光没有掩盖流星时，流星在晚上更容易被看到。这里没有哲学思维实验中所讨论的森林里树木倒下的困局^[15]。流星的存在并不依赖于观测者实际的观察，而且原则上它所产生的一道道光，必须要在可见光范围之内才能够被看到。

大多数流星产生于尘埃大小或者石子尺寸的物体。每天有数百万个这类物体进入地球的大气层。因为大多数流星体在海拔50公里的高空破碎，流星通常发生在海平面以上75~100公里叫作中间层的地方。尽管精确的速度依赖于物体的具体性质，以及相对于地球的角度，但总体来说，其速度在每秒几十公里这个量级上。流星的轨道能帮助识别产生它的流星体来自哪里，而流星发出的可见光的谱线，以及它对射电信号的影响，都能帮助科学家确定流星体的成分。

能够穿过大气层并且击中地球的流星体称为陨石。陨石是地外物体击中地球之后，分解、融化并且部分蒸发之后留在地球上的石块。陨石又是另外一个确凿的提示，表明地球本质上是宇宙环境的一部分。如果幸运的话，你能在流星体撞击点的附近找到一颗陨石，但你更可能会在实验室、博物馆或者那些足够痴迷的、幸运的或者富人的房子里看到它们。梵蒂冈天文台博物馆有着相当棒的收藏，与史密森美国自然历史博物馆一样，后者有着最大的收藏量。一位三星上将曾告诉我，美国国防部也有不少很棒的各种陨石。他们的收藏与导弹防御相关联，但遗憾的是，他们关于流星体撞击的数据仍然保密。那些

能够接触到的陨石研究已经告诉了我们很多有关太阳系以及其起源的事情。

陨石也能源于彗星，彗星是太阳系外圈的物体（我将在下一章详述）。太阳系内圈和外圈的物体是非常不一样的，只有在木星轨道以内的物体才能被叫作小行星或者次行星（minor planets）。次行星和小行星不同，它的名字听上去有些轻视的意味，但却是官方术语。彗星和小行星的区别看上去很明显，比如彗星有明显的尾巴，但二者之间的区别要细微得多。彗星通常有着更加拉长的轨道，但一些小行星也有着相似的椭圆轨道，也许是因为它们最初就是彗星。此外，含水的小行星并不一定与在太阳系外部形成的彗星的种群不同。小行星的多样成分也表明，被归类为小行星的种群和被归类为彗星的种群是有一些重叠的。

由于小行星和彗星的区别过于模糊，因此在2006年，国际天文学联合会定义了一个专有名词“太阳系小天体”（small Solar System body），这个定义同时包含两者，但不包含矮行星。矮行星本来也可以包含在太阳系小天体中，但由于它们的尺寸更大、形状更圆，也表明引力越强，越有可能成为固态物体，因而国际天文学联合会没有将矮行星归到太阳系小行星类别里，而是用它们自己的名字区分开来。总的说来，国际天文学联合会更喜欢用“太阳系小天体”这个术语，而不是“次行星”，因为小行星带中的有些物体带有彗核的特征。一个包含了小行星和彗星两个类别的名称，虽然提供了较少的信息，但避免了错误。即便如此，小行星大体上要更加岩石化一些，而彗星通常包含更多的易挥发物质，因此大多数天文学家还是坚持它们之间的差别。

当我在本书其余部分提到撞击地球的大物体时，这些复杂的命名让我陷入两难。在天空中就烧尽的小物体是流星体或者微流星体。更

大的物体，源自小行星或者彗星，偶尔可以到达地球或者只到达大气层。但我们只有在观察了轨迹，知道了它的起源之后，才能知道到底是哪一种。我们需要一个能够同时指代二者的词语。严格说来，“太阳系小天体”这个烦琐的词汇符合了这个要求，但这个术语极少被用于描述来自天空飞入地球大气层的物体（尤其那些接近或者击中地球的物体）。看一下很多新闻中的标题，你会发现人们更经常使用“流星”“流星体”，甚至“陨石”这些名词，严格说来，如果物体的直径大于一米，这些词都是错的。现在似乎还没有一个通用的并且足够明确的口语词汇（虽然“撞击天体”和“火流星”有时会被使用），在本书中，当提到从地球以外进入大气层或者击中地球的物体时，我将把这类物体叫作“流星体”（这些物体是最没有攻击性的一类）。这可以算是对术语的轻微滥用，因为这个词通常仅指代尺寸较小的物体，但联系上下文你就会明白我的意思。



耶稣诞生像中的光的余辉毫无疑问是一颗彗星，并很有可能是哈雷彗星。

如果你有机会前往直意大利城市帕多瓦，一定要参观一下斯科洛文尼教堂（Scrovegni Chapel）。这个从14世纪早期完好保存到现在的瑰宝里，收藏了由文艺复兴早期的艺术家乔托（Giotto）所画的一组宏伟壁画。在那里，有我最喜欢的、也是我所有物理学同事都珍爱的一幅壁画，那就是《博士来拜》（*The Adoration of the Magi*, 见图7-1）。这幅画描绘了一颗耀眼的彗星掠过耶稣诞生时的场景。也许，正如艺术史学家罗伯特·奥尔森（Roberta Olson）所说，彗星取代的是大卫王之星（Star of David）这个人们更为熟悉的构图元素。而就在这幅画完成的几年之前，人们就曾目睹了这个异常明亮的壮丽天体。独立于寓言的意图之外，耶稣诞生像中的光的余辉毫无疑问是一颗彗星，并很有可能是哈雷彗星，在地球上那一块区域的任何人都应该会看到。1301年9月和10月，巨大的彗尾扩散在天空中很大的范围内，形成一幅壮观的景象，尤其是在那样一个还没有电灯的时代。

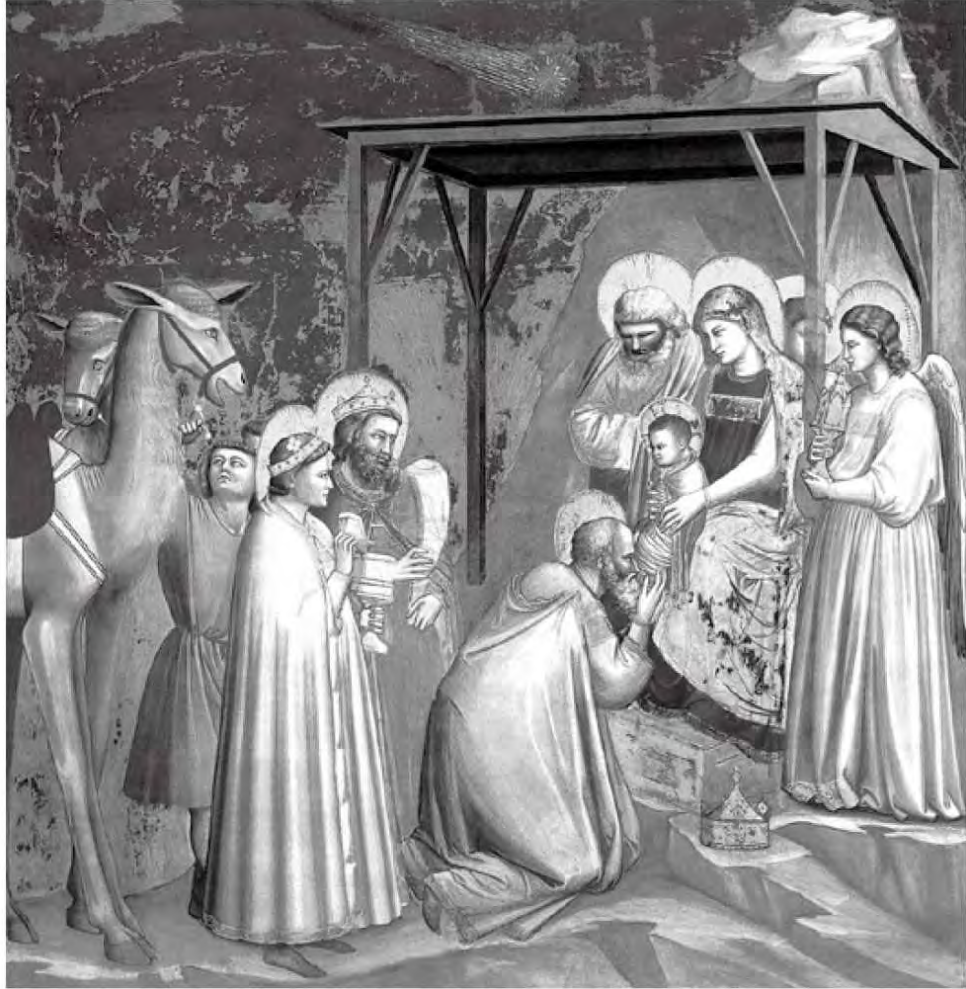


图7-1

乔托绘制的《博士来拜》，耶稣诞生像之上有一颗彗星。

我想，在14世纪早期，意大利人向天空凝视并欣赏到的天体奇观，也同样让今天的我们惊叹。古希腊和古代中国文明的证据表明，至少在意大利人开始观测的两千年前，人们就已经在观测和欣赏彗星了。亚里士多德甚至试图探究彗星的本质，并将其解释为上层大气中的一种现象，干燥、炎热的物质会在那里燃烧。

自古希腊时代以来，我们对彗星的研究已经取得了很大进展。基于数学模型和更好的观测数据，最新研究告诉我们，彗星是冷的，没

有任何东西在燃烧。一旦足够接近太阳，它们所包含的挥发性物质将很容易转化为蒸气或者液态水。

在前文，我们已经探索了来自太阳系的小行星及其本质，接下来让我们转向彗星。彗星源于被称为散盘（scattered disk）的遥远区域，跨越了柯伊伯带和位于太阳系外部区域的奥尔特云。彗星也存在于其他恒星系，但这里我们的重点将放在那些我们了解最多、源于太阳系的彗星。

长发彗星的内核

尽管我们现在知道，彗星来自遥远的太空区域，只有极少数彗星能够沿着接近地球的轨道运动，然而第谷·布拉赫（Tycho Brache）早在16世纪就得出结论：彗星位于地球大气层以外。这是早期科学理解上的一个重要里程碑。第谷整合了不同地方观察者的观测结果，得到了1577年大彗星的视差，从而确定了彗星到地球的距离至少是月地距离的4倍。实际距离肯定是被低估了，但在当时对于彗星的研究来说这是一个巨大的飞跃。

当牛顿意识到彗星沿着倾斜轨道运动时，他得到了另外一个重要的推论。用他的引力平方反比定律，也就是说对于距离两倍远的天体，引力的强度变为原来的 $1/4$ ，牛顿证明了天空中的物体必须遵循椭圆、抛物线或双曲线轨道。当牛顿用一个抛物线拟合1680年大彗星的路径点时，抛物线完美地把所有观测数据点都连接了起来。这表明，人们一直看到并且相信是不同种类的物体，实际上沿着一个单一轨道运动，是同一个物体的足迹。尽管彗星的轨迹实际上是一个拉长的椭圆，但路径的形状与抛物线非常接近，因此牛顿对于单一运动物体的推论仍然正确。

在早期时候，被发现的第一批彗星以出现年份被命名。但到了20世纪初，命名的惯例有所改变，改为以预言轨道的人的名字命名。例如，德国天文学家约翰·恩克（Johann F. Encke）、业余天文学家威廉·冯·比拉（Wilhelm von Biela），两位都有以其名字命名的彗星。

虽然早在20世纪之前就被辨认出来，哈雷彗星也是以对其轨道有着最好理解、预言了其回归的人的名字命名的。1705年，运用牛顿定律并考虑了木星和土星的扰动，作为牛顿的朋友以及出版商，埃德蒙·哈雷预测，一颗已经在1378年、1456年、1531年、1607年、1682年出现过的彗星，会在1758—1759年之间再次出现。哈雷第一个提出了彗星的周期运动，而且他是正确的。三位法国数学家做了更加精确的计算，预测的日期精确到了1759年的一个月內。今天我们可以做类似的计算：到2061年，人类才能再次一睹哈雷彗星的归来之貌。

到20世纪后期，彗星的命名惯例再次改变：以发现者的名字命名。然而，当彗星的发现成为基于先进观测工具的集体努力时，彗星就以观测仪器的名字来命名。目前被发现的彗星大约有5 000颗，但对彗星总数的合理估计认为，它至少是这个数目的1 000倍，真实的数目也可能会更多，或许高达一万亿。

了解彗星的性质和成分需要对物质的状态有所了解。物质的相态包括固体、液体和气体，对于水而言就是冰、水和水蒸汽。每一个相态下，原子的排列方式不同。固体冰的结构最有规律，而气态水蒸气最随机。当相变发生，液体转换成气体（当水沸腾时），或固体转变成液体（当冰块融化时），物质成分并不会发生改变，因为所有的原子和分子都没有发生变化。但这种物质的性质变得非常不同。物质的形态取决于它的温度和成分。对于任何特定的物体，温度和成分决定了它的沸点和熔点。

听说最近有人试图利用物质的不同状态，把一瓶水带过机场的安全检查，我觉得非常有趣。他把水冻成冰，并认为他瓶子里的固态冰没有违抗任何液体禁令。不过，交通安全管理局的人并不这么认为。如果那位管理局的人学过物理学，他可能会说服这位乘客，只有那些在标准温度和压力下呈固态的物质是允许的。然而，我敢肯定他不会这么说^[16]。

熔点和沸点是所有结构的关键，因为它们确定了材料所处的状态。一些元素，例如氢和氦，具有极低的沸点和熔点。例如，氦在绝对零度以上4度就变成液体。行星科学家称不论物质实际处于什么状态，熔点低于100开尔文的这些元素为气体。而那些具有低熔点，但并非如气体熔点那样低的物质，再一次被行星科学家称为冰。当然物质是否真的是冰也同样取决于实际的温度。这就是为什么木星和土星被称为气态巨行星，而天王星和海王星有时被称为冰态巨行星。在这两种情况中，它们的内部实际上都是热的、稠密的流体。

气体（在行星科学家所使用的意义上）是挥发物的一种，是具有低沸点的元素和化合物，例如氮、氢、二氧化碳、氨、甲烷、二氧化硫和水，这些物质可能存在于行星里或者大气里。具有低熔点的材料将更容易变成气体。你也许见过由冷液氮制成的冰激凌（不少餐厅和快餐车上都会卖这种冰激凌，不过我不大喜欢它们的口味）。如果你看到过任何这种例子，你会注意到，氮原子在常温下是如何很容易地以气体状态逃逸。

月球含有很少的低挥发性物质，主要是由硅酸盐组成，还有少量氢、氮或碳。彗星含有丰富的挥发物，从而形成了明显的尾巴。彗星起源于远远超出木星轨道的太阳系外部区域。那里，水和甲烷仍然处于寒冷和冰冻状态。在这些远离太阳的非常寒冷的区域，冰不能转变为气态，冰依然是冰。只有当彗星经过内太阳系时，它们接收到更多

太阳热量时，彗星里挥发性的物质蒸发，和灰尘一起，在彗核的周围形成一个被叫作彗发（coma）的大气。这一圈大气比彗核大很多，跨越数千甚至数百万公里，有时甚至会增长到太阳的大小。较大尺寸的尘埃颗粒会留在彗发中，而较轻的尘埃则被太阳的辐射和带电粒子的辐射推送至彗星尾部。彗星包括彗发、它围绕着的彗核以及吹散的尾巴。

流星雨是由彗星在其身后留下的固态碎片形成的，它也是彗星的壮观证据。在一颗彗星穿越过地球轨道之后，一些被抛出的物质会出现地球的轨道上。然后地球会定期经过这些碎片，从而创造出非常壮观的周期性流星雨。斯威夫特-塔特尔（Swift-Tuttle）彗星的碎片是英仙座流星雨的起源，发生在每年8月初。我曾无意间在美国阿斯彭某个晴朗的夜空看到过，阿斯彭有一个经常举办暑期研讨班的物理中心。另一个例子是猎户座流星雨，发生在每年10月，是由哈雷彗星散落的碎片产生的。

流星雨

在一颗彗星穿越过地球轨道之后，一些被抛出的物质会出现地球的轨道上。然后，地球会定期经过这些碎片，从而创造出非常壮观的周期性流星雨。

彗星是我们可以用肉眼看到的最壮观的天文现象之一。大部分彗星都是非常暗的，但那些诸如哈雷彗星一样不用望远镜都可以看到的明亮彗星，也许会时不时地（比如10年内）经过地球附近。彗星围绕太阳运动，它的明亮离子尾巴和分离的尘埃尾巴通常会指向不同的方向。这些明亮的尘埃和气体的痕迹是彗星名字的由来，这个源自希腊的词语意思为“留着长头发”。尘埃彗尾一般遵循彗星的路径，而离子彗尾指向远离太阳的方向。当太阳的紫外辐射撞击彗发的时候形成

离子彗尾，将电子从一些原子中撕离。这些电离粒子在所谓的磁层中会形成磁场。

太阳风在彗星的这种外观中起着重要作用。太阳辐射会产生光子，让地球上的我们感受到热和光。而太阳发出的带电粒子，也就是电子和质子，构成了太阳风。20世纪50年代，德国科学家路德维希·比尔曼（Ludwig Biermann）以及另一名独立进行研究的德国人保罗·阿纳特（Paul Ahnert），进行了仔细观测，发现一个彗星的明亮离子彗尾始终指向远离太阳的方向。比尔曼当时就提出，是太阳产生的粒子“推动”了彗尾，使其指向那个方向。用比喻来说，是太阳风把离子彗尾“吹”到了那里。通过对这个过程的研究，科学家们学到了有关彗星和太阳两者的知识，并让我对这个神秘名字的由来豁然开朗。

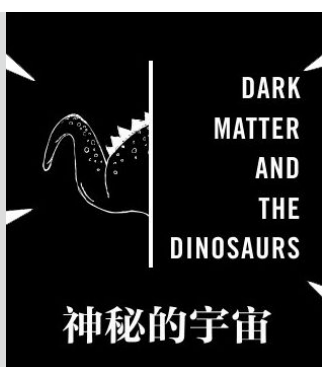
彗尾能够扩展数千万公里。彗核的大小当然要小得多，但和典型的小行星相比较还是相当大的。彗核没有足够的引力使它们变成圆形，因此其形状很不规则，整体的尺寸从几百米到几十公里不等。这有可能是观察偏差，因为更大的更容易被发现，但是应用足够灵敏的设备寻找更小彗星的搜寻工作，到目前为止还是没有什么进展。

就可见度而言，彗星有彗发和彗尾是一件很好的事情。因为彗核的反射率非常低，这使得它极难被看到。观看那些不发光的物体（比如你和我）最常用的方法就是通过反射光。举一个著名的例子，哈雷彗星核的反射率大约只有1/25，这和沥青或木炭的反射率差不多，而后两者十分暗淡。其实彗核反射得更少。事实上彗星表面似乎是太阳系中最暗的^[17]。易挥发、较轻的化合物在太阳的热量消失之后，留下的是颜色更深、更大的有机化合物。这个暗的物质吸收光线，从而加热冰状物，挥发出去的气体就形成了彗尾。木炭和彗星之间反射率的相似并非巧合，因为焦油沥青也是由石油中大的有机分子构成的。现在，想象在天空中数十亿公里远的地方有一大块沥青。如果不付出大

量努力把它们搜寻出来，那么这些很暗的天体无疑将会默默无闻地消失掉。

彗星在太阳系外部区域的时候，很暗并且是冰冻的，光学辐射是极其微弱的。在彗星接近太阳之前，观察它们的唯一方法是通过它们发出的红外光。当它们进入内太阳系时，彗发和彗尾形成，使得彗星较为容易被看到。灰尘会反射阳光，离子会使气体发光，从而产生我们更容易观察到的亮光。即便如此，大多数彗星都只有用望远镜才能看到。

和彗星本身相比，彗星的具体化学成分更难被观测出来。地球上发现的陨石给我们提供了一些线索，这些陨石会运送彗星上的一些实体物质到地球上。科学家们也已经注意到彗星有各种颜色，并且观察了它们的一些谱线。利用这些稀有线索，科学家们得到的结论是：彗核由水冰、灰尘、卵石状岩石和包括二氧化碳、一氧化碳、甲烷和氨的冷冻气体组成。彗核表面看起来是岩石状的，而冰藏在地表以下一点的地方。



虽然牛顿那个时代的天文观测条件很有限，他还是在17世纪就非常准确地解释了彗星。虽然牛顿误认为彗星是紧凑、坚实的固体，但他认识到彗星的尾巴是已经被太阳加热的稀薄蒸汽流。在认识彗星的成分方面，哲学家伊曼努尔·康德在1755年做了一项更好的工作。他推测，彗星的彗尾是由易挥发性物质蒸发而形成的。在20世纪50年代，哈佛大学天文系的弗雷德·惠普尔（Fred Whipple）曾发现过6颗彗星，辨认出彗星中的主要成分为冰，而灰尘和岩石是次要的组成成分，从而产生了“脏雪球”模型（dirty snowball），他也

因此而出名。虽然到目前为止，我们并不完全清楚彗星的组成，但我们的认识不断在更新。

对于彗星而言，一个更加让人感兴趣的特点是，它包含有机化合物，例如甲醇、氰化氢、甲醛、乙醇和乙烷，以及长链碳氢化合物和氨基酸（氨基酸是生命的前身）。甚至地球上陨石中被发现含有的DNA和RNA成分，据推测也可能来自小行星或者彗星。那些携带水和氨基酸并且定期撞击地球的天体，值得我们关注。

彗星迷人的结构和与生命的可能关联，使其成为许多太空任务的目标。第一个研究彗星的空间探测器飞过慧尾以及彗核的表面，收集并分析了灰尘颗粒，也拍了照片，但是距离太远且分辨率不够高，因此不能提供很详细的信息。1985年，美国国家航空航天局的国际彗星探险者任务在得到欧洲的一些资助之后重新定向，成为第一个接近彗尾的空间探测器，不过是在距离3 000公里的地方。不久之后，由俄罗斯发射的两次维加（Vega）太空飞行任务、日本的彗星号（Suisei）探测器任务，以及欧洲的乔托号（正是以《博士来拜》的作者乔托之名命名）宇宙飞船共同组成的哈雷舰队（Halley Armada），试图更好地研究彗核和彗发。值得注意的是，乔托超越了所有其他任务，该飞船接近了哈雷彗星的核心，距离不到600公里。

更近的一些试图直接探索彗星及其构成的任务，结果也很好。2004年初，星尘号（Stardust）空间探测器收集并分析了来自怀尔德2号彗星的彗发尘埃颗粒，并于2006年把样本送回地球进行研究。样本显示，彗星并非主要由星际间普通物质组成，这和对形成于遥远奥尔特云的物体的预期不同。彗星主要由来自太阳系内被加热的物质组成。科学家发现彗星含有铁和硫化铜，而如果没有液态水，这些是无法形成的。这意味着，彗星的温度最初比较温暖，也即形成于接近太阳的

地方。这一结果进一步表明，彗星和小行星的成分并非总是像科学家原本预想的那样不同。

Deep Impact ^[18]不仅是一部电影的名字，也是一个太空探测器的名字。在2005年它发送了一个撞击器，希望进入到坦普尔1号彗星的内部。此探测器的设计目的就是研究彗星的内部，并拍摄撞击坑，虽然撞击产生的尘埃云使图像有一些模糊。探测器发现了结晶物质，而结晶物质需要在比彗星目前的温度更极端的温度下才能形成。这表明，要么这些物质是从内太阳系区域进入彗星的，要么是因为彗星最初形成的区域离它现在所处位置很遥远。

最近的彗星探测更加令人兴奋。欧洲航天局有一个非常出色的进展。他们于2004年发射了名为罗塞塔的探测器。这个飞船围绕67P/丘留莫夫-格拉西缅科彗星运行。随后他们又将一个名为“菲莱”（Philae）的直接探测器降落在此彗星表面，目的是近距离研究彗核的组成和内部区域。2014年11月，菲莱号发生了“大事件”：它的确着陆了，但并不如计划般顺利，而是在反弹了几次之后，降落到了一个不太稳定的位置。这一事件，毫无疑问是紧张刺激的，但已经完成了相当一部分科学目标。虽然打钻任务没有成功，对于菲莱号探测器而言，即使降落在了计划之外的位置，也缺少准备就绪的附着机制，但它还是比以往任何时候都更详细地研究了彗星的形状和大气。

罗塞塔现在围绕彗星运行，随着彗星进入内太阳系，探测器将继续绕转。这整个任务已经是一个颇为壮观的成就。考虑到它的发射是在莱特兄弟第一架飞机首飞之后的100年内发生的，因此更加令人印象深刻。

短周期和长周期彗星

即使有了很大的进展，但除了更好地确定了它们的构成成分，有关彗星的许多有趣问题依然没有得到解答。所以天文学家想更深入了

解彗星的轨道以及彗星形成的方式。我们并不指望得到一个统一的解释，因为证据表明彗星的种类区别很大。根据绕太阳一周所需要的旅行时间，彗星分为短周期和长周期两种类型。区分短周期和长周期的分界线被定为200年，但总的说来周期从几年到几百万年不等。

彗星起源于海王星之外，而这些海外天体（指位于海王星轨道之外的天体）的聚集地位于距离太阳不等的不同轨道带上。其中靠里的区域产生了短周期彗星，即柯伊伯带和散盘；靠外的区域是假设的奥尔特云，产生长周期彗星。天体物理学家提出还有一个额外区域，这个区域位于散盘和奥尔特云之间，即分离天体（detached objects，但在这里我们不会着重讨论它）。

在很大程度上，对彗星起源的内部区域和外部区域的分类，与彗星的轨道周期有些重叠。我们最常看到的彗星是短周期彗星，如哈雷彗星。它们在一定的周期内回归，并常被人看到。短周期彗星来自靠近地球的区域，而长周期彗星主要是从遥远的地方到达地球。我们偶尔也会看到长周期彗星，但只有当它们进入内太阳系的时候才会被看到，这可能是由遥远的奥尔特云受到扰动而引起的。太阳的引力对这些彗星的束缚力很弱。因此，即使是小的扰动也能够使物体偏离它们的轨道，向内并朝太阳飞来。甚至是类似哈雷的短周期彗星，也有可能最初是从一个更遥远的长周期轨道被踢出，成为内太阳系中的一个运行周期较短的彗星。

短周期彗星可分为两个子类：轨道周期大于20年的哈雷族彗星，以及轨道周期较短的木星族彗星。在短周期轨道上可能也有一些小行星或者休眠彗星或者死彗星，但极少有小行星的轨道周期大于20年。长周期彗星的轨道椭率更大，意味着与短周期相比，它们的轨道被拉得更长。这是有道理的，因为彗星只有在太阳附近的时候才能被我们看到。尽管短周期彗星在相对靠近太阳附近的区域绕转，而那些既能

被看到又有很长周期的彗星就应该有如下轨道：轨道向内延伸靠近太阳，但之后又远远向外延伸，形成很长的路径，从而旅行一圈需要很长的时间。长周期彗星的轨道似乎也更接近黄道面（黄道面是行星运动的平面），并且都在朝同一个大方向绕行。

一旦进入内太阳系，这些物体的命运将取决于可能的进一步扰动。木星是已知的相对最大的本地微扰源，因为它的质量是所有其他行星总质量的两倍多。所以内太阳系的新彗星就可能会进入一个新的轨道运行，或者它们只出现一次，之后就会被踢出太阳系或与行星发生碰撞。一个例子是苏梅克-列维（Shoemaker-Levy）彗星，在1994年，这颗彗星撞毁在木星上的事件非常出名，场面也是非常壮观。

柯伊伯带和散盘

现在让我们考虑这样的区域，它们包含了冰冷的小太阳系天体。如果这些小天体受到扰动进入内太阳系，就将变成彗星。我们的第一个主题将是柯伊伯带（见图7-2）。它本身并不是短周期彗星的储备地（散盘才是），但是它是散盘的一个重要标志。

对我来说，柯伊伯带最有趣的地方是，早在20世纪40年代和50年代它就被预言存在，但直到最近才被发现。1992年，天文学家决定必须修正我们之前对太阳系的理解，以加快对柯伊伯带的探索。而在当时，很多人都学过有关太阳系的知识，并且认为它们有着坚实的基础。即使你从来没有听说过柯伊伯带，但你可能熟悉其中的几个天体，它包括三个矮行星，其中就有以前被认为是行星的冥王星。虽然离柯伊伯带很远，海王星的卫星海卫一崔顿（Triton）和土星的卫星菲比（Phoebe）的大小和成分也表明，它们曾经也是在这个区域内形成的，之后被经过的行星拉离。

一个天文单位大约是1.5亿公里，是地球和太阳之间的大致距离。柯伊伯带的位置与太阳的距离比日地距离远30多倍，大约为30~55天文单位。它含有大量的小行星，其中大部分是位于柯伊伯带上，也就是距离太阳大约42~48天文单位。在垂直方向上，这个区域延伸在黄道面之外大约10°之内，虽然平均位置仅具有几度的倾角。其厚度使得它更加像甜甜圈的形状，而不是带形。即便如此，它那颇有误导性的名字还是被保留了下来。

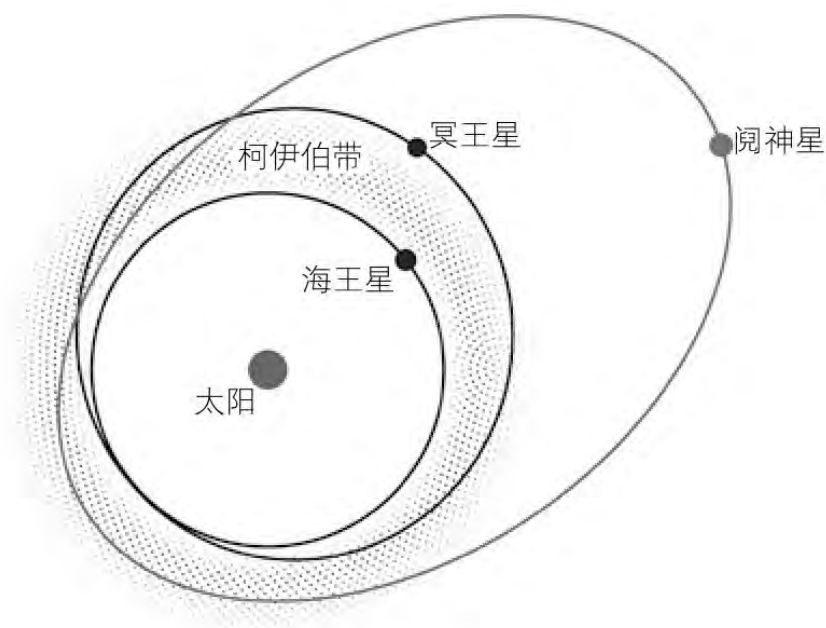


图7-2

柯伊伯带，位于海王星之外，包含作为其最大天体的冥王星。散盘位于柯伊伯带稍微向外，包含质量更重的天体阋神星。

说柯伊伯带的名字有误导性还有另外一个原因。之前对柯伊伯带性质的猜测有很多并且十分多样化，因此，到底是谁对这个概念的提出功劳最大，我们并不十分清楚。在20世纪20年代它被发现之后不久，许多天文学家都怀疑冥王星并不孤单。早在1930年，科学家们就引入了各种有关海外天体的假说，但天文学家肯尼斯·埃奇沃斯（Ken

neth Edgeworth) 在这方面可能最应该受到赞誉。1943年，他认为，在早期太阳系海王星以外区域的物质过于稀疏，无法形成行星，因而会形成一组较小的物体。他继续推测，有时这些对象的其中一个会进入内太阳系，成为一颗彗星。

科学家们现在所青睐的情形非常类似于埃奇沃斯所提出来的理论。初期太阳的盘凝结成比行星更小的物体，有时也被称为星子。杰拉德·柯伊伯 (Gerard Kuiper, 柯伊伯带因他而得名) 在相对较晚的时候，也就是在1951年，提出了他的假设。另外，柯伊伯的答案当时也并不是完全正确。他认为，柯伊伯带是一个短暂的结构，到今天应该已经消失了，这是因为他认为冥王星要比我们现在所知道的实际大小要大，并因此能够像其他行星一样清空它附近的区域。由于冥王星其实比柯伊伯所预期的小很多，所以这个现象并没有发生。因此，那些位于冥王星所在区域的许多其他天体和柯伊伯带的确存活下来了。

埃奇沃斯的功劳与柯伊伯一起通过“埃奇沃斯-柯伊伯带”的命名被认可，但这个名字太长了，人们常用其简称。就像“瑞典国家银行纪念阿尔弗雷德·诺贝尔经济学奖”常被称为“诺贝尔经济学奖”，因为全称太复杂了。可能你很少听到这个天文学术语的全称，虽然它恰当地承认了埃奇沃斯的贡献。

根据埃奇沃斯和柯伊伯的建议，科学家意识到，彗星本身能够给我们提供柯伊伯带存在的线索。20世纪70年代发现了太多的短周期彗星，无法全部用奥尔特云 (奥尔特云是很多距离遥远彗星的储备库，我们稍后讨论) 来解释。短周期彗星出现在太阳系的平面附近，而不是像奥尔特云的彗星一样，围绕太阳呈球状分布。在此结果的基础上，天文学家胡里奥·费尔南德斯 (Julio Fernandez) 认为，彗星可能由位于一个带状区域来解释，这个区域就是我们现在知道的柯伊伯带所在地。

尽管有所有这些猜测，但发现它还是要求观测灵敏度达到必要的精度。因为要找到小的、远的、又不发光的天体非常不易，直到1992年和1993年年初，柯伊伯带内才首次发现除了冥王星之外的天体。珍妮·刘（Jane Luu）和大卫·朱维特（David Jewitt）开始他们的研究时，朱维特还是麻省理工学院的教授，而珍妮·刘还是个学生。他们在亚利桑那州基特峰国家天文台（KPN0）和智利托洛洛山美洲际天文台（CTIO）进行了探测。在朱维特搬到夏威夷大学之后，他们可以使用学校的一台直径为2.24米的望远镜继续研究。这个望远镜位于休眠火山莫纳克亚山（Mauna Kea）的山顶，这里是一个美丽的景点，有着非常晴朗的天空（如果你来夏威夷的大岛上，这里非常值得游览）。在5年的寻找之后，他们发现了两个柯伊伯带天体，第一个在1992年夏天发现，另一个在次年年初。从那时起，许多其他类似天体被发现，尽管它们几乎可以肯定仅代表所有存在天体的非常小的一部分。我们现在知道，柯伊伯带包含千余“居民”，它们被称为柯伊伯带天体（KBOs）。但是计算认为，那里可能有多达10万个直径大于100公里的天体。

尽管冥王星失去了行星的地位，它仍然很特殊，这就是为什么它早于柯伊伯带中其他天体被发现。根据我们现在了解的其附近物体的质量，冥王星比后来大家所预期的要大。^[19]这一个天体似乎占据柯伊伯带的总质量的百分之几，并且极有可能是那里同类天体中最大的一个。事实上，柯伊伯带的总质量比较低，这对它起源的研究也是一个有趣线索。尽管估算的质量范围是地球质量的4%~10%，然而太阳系的形成模型给出的柯伊伯带的质量大约是地球质量的30倍。如果它的质量一直这么低，那么直径大于100公里的物体就不会加入进去的，这就和冥王星的存在是矛盾的。这告诉我们超过99%的预测的质量，现在都已经不在那里了。要么柯伊伯带天体形成于距离太阳更近的其他的地方，要么有什么东西让大部分质量疏散到其他地方去了。

和冥王星具有相似轨道的许多其他天体被称为冥族小天体（plutinos），它们位于距离太阳略小于40天文单位的地方，其高度偏心轨道意味着它们的距离随轨道变化。冥族小天体是共振柯伊伯带天体，其运行轨道相对于海王星有一个固定的比例。例如冥族小天体，在海王

星绕太阳运行三圈的时间里可以绕太阳运行两圈。这一固定比例防止天体过于靠近海王星，因此它们能逃脱海王星更强的引力场，否则它们将被踢出该区域。有趣的是，国际天文学联合会要求冥族小天体要像冥王星一样，也必须以地狱诸神的名字命名。我们知道至少存在1 000个这样的天体。不过，鉴于目前人类有限的观测巡天能力，科学家认为真实存在的应该更多。

然而，柯伊伯带占主导地位的类群并不包括冥族小天体，而是属于经典柯伊伯带天体。通过观测巡天已经发现很多这样的天体。Pan-STARRS巡天项目现在所有的时间都用于搜寻太阳系内明显移动的任何物体，很可能会找到更多。经典柯伊伯带中的天体具有稳定轨道，不会被海王星扰动——即使没有共振轨道以使它保持在一个固定远的距离。这些颜色偏红的经典天体的相当一部分都有很圆的轨道。第二大种类轨道更偏心并且更倾斜，最大达到约 30° ，但这类天体数目少得多。这使得柯伊伯带内剩余一些相对空旷的不稳定区域，只包含近期才进入的一些天体。

曾经位于柯伊伯带内的天体，有可能是很多我们观察到的彗星的前身，或者至少与这些彗星有关。因此，它们的成分基本上和彗星一样就并不令人惊讶了。它们大部分由冰状物质构成，包含甲烷、氨和水。物质以冰态而不是气体存在是由于柯伊伯带的位置以及由位置导致的低温（大约50开尔文，这比水的冰点要低 200°C ）导致的。新视野号探测器将收集到许多有关冥王星和柯伊伯带的信息。在科学家完成这些数据分析之后，我们就应该能够知道得更多。

柯伊伯带内的轨道是稳定的，因此彗星不是源自于那里。柯伊伯带的永久居民不会向太阳靠近。相反，短周期彗星是从散盘产生的，散盘是一个相对空旷的包含冰状的小行星的区域，与柯伊伯带重叠，但是延伸得更远，距离太阳一直到100天文单位或者更远的地方。散盘

包含的天体的轨道会受海王星扰动而变得不稳定。散盘天体与柯伊伯特天体的差别在于，前者天体具有更大的偏心率、更大的位置范围，以及更大的倾角，最高达到约 30° 。同时，散盘天体也更不稳定。散盘天体具有中等到高的偏心率，这意味着它们的轨道是拉伸的形状，而不是圆形。它们的偏心率非常高，即使那些最大轨道非常远离海王星的天体，当它们的轨道足够接近海王星时，就会受到其引力场的影响。这就是为什么海王星的影响有时能将散盘天体发送到内太阳系里面。在那里，这些天体被加热，释放出气体和尘埃，从而被识别为彗星。

阋神星，作为已知的大小与冥王星相当的一个小行星，位于柯伊伯特之外的散盘中，并且是被人类识别出的第一个散盘天体。

为了找到它，在莫纳克亚山上天文学家使用了电荷耦合器件（CCD，数码相机所用技术的高级版本），以及更好的计算机处理方法。这些技术使得更加遥远的天体也能够被观察到，并促成了阋神星在1996年被发现。几年后天文学家又发现三个散盘天体。另外一个散盘天体，毫无诗意地被命名为1995TL8（48639），它是在1995年被发现的，但在之后才被分类为散盘天体的。自那之后又有几百个散盘天体被发现。它们的总数可能和柯伊伯特天体数量相当，但因为距离更远，观察它们更为困难。

柯伊伯特天体和散盘天体包含类似的化合物。像其他类海王星天体一样，散盘天体密度很低，并主要由类似水和甲烷的冰冻易挥发物组成。许多人认为，柯伊伯特天体和散盘天体产生于同一区域。但由于引力的相互作用，主要是与海王星的作用，将其中一些天体送到柯伊伯特稳定轨道。还有一些向内进入被称为半人马的天体区域。这一区域位于木星和海王星的轨道之间。剩余的天体则由引力相互作用，发送到散盘的不稳定轨道上去。

几乎可以肯定的是，柯伊伯特和散盘的大部分结构取决于靠外行星的引力作用。在某个时刻，木星看来是向内并朝向太阳系的中心漂

移，而土星、天王星和海王星则向外移动。木星和土星利用彼此来稳定自己的轨道。木星围绕太阳的速度正好是土星的两倍。但是，这两个行星使天王星和海王星变得不稳定，使它们进入不同的轨道。海王星的轨道偏心率更大，在更远的地方运行。在到达其最终目的地的途中，海王星可能将很多星子散射到更偏心的轨道，并将其他许多星子散射进入更内部的轨道。后者在内部轨道上可能被木星的影响再次散射或者弹出。这会使柯伊伯带不到百分之一的天体幸存下来，而大多数已经被散射出去。

一个很有竞争性的提议是，柯伊伯带首先形成，而散盘中的天体来自柯伊伯带。在这个提议中，在许多方面与前面一个提议类似，海王星和外部行星将一些物体散射进入偏心和倾斜的轨道，要么向着内太阳系区域，要么向外进入太阳系更远的地方。然后一些从柯伊伯带向外散射的物体就变成散盘天体。另一些可能成为半人马天体。这能解决一个谜题，半人马天体有着不稳定的轨道，并只能在它们的领地停留几百万年。但它是如何能存在到今天的呢？柯伊伯带可能为它们提供了补充。彗星，也同样只有有限的（但光荣的）生命。太阳的热量逐渐侵蚀它们，使它们易于挥发的表面不断蒸发。如果没有新天体的持续加入，我们周围的彗星将不复存在。

奥尔特云

散盘是短周期彗星的储备库，而奥尔特云则是长周期彗星的假定源头，是一个巨大的球状分布的由冰冷微行星构成的“云”，也许包含一万亿小的行星（见图7-3）。奥尔特云是以荷兰天文学家简·奥尔特名字命名的。有几个重要成就都是他的功劳，至少有两个物理专有名词带有他的名字。奥尔特一个最著名的成就是在1932年建立的一种方法，可以测量星系中的物质总量，包括暗物质的质量。

奥爾特的另一個重要工作成就是對現在被稱為奧爾特雲的推斷。20世紀30年代，天文學家恩斯特·奧匹克（Ernst J. Öpik）首次提出這樣一個作為長周期彗星起源的雲的存在。到1950年，奧爾特根據理論和觀測的雙重證據，推測出這樣一個由許多遠距天體組成的球形雲的存在。首先，他觀測到來自各個方向的長周期彗星有着非常大的軌道，這表明彗星的起源地要比柯伊伯帶遠得多。奧爾特進一步認識到，如果彗星一直在其目前的軌道上，它們的軌道不可能存在足夠長的時間，以致到今天還能被觀測到。因為彗星的軌道是不穩定的，所以行星的擾動會導致它們最終與太陽或某顆行星碰撞，或者被完全送出太陽系。此外，在彗星消失之前不可能永遠持續擴散氣體，彗星因為太多次靠近太陽飛過，就會失去活力，不再產生氣體。奧爾特的假設是，我們現在所說的奧爾特雲補充了新鮮彗星的供應，因此今天仍然可以觀察到這些新的彗星。

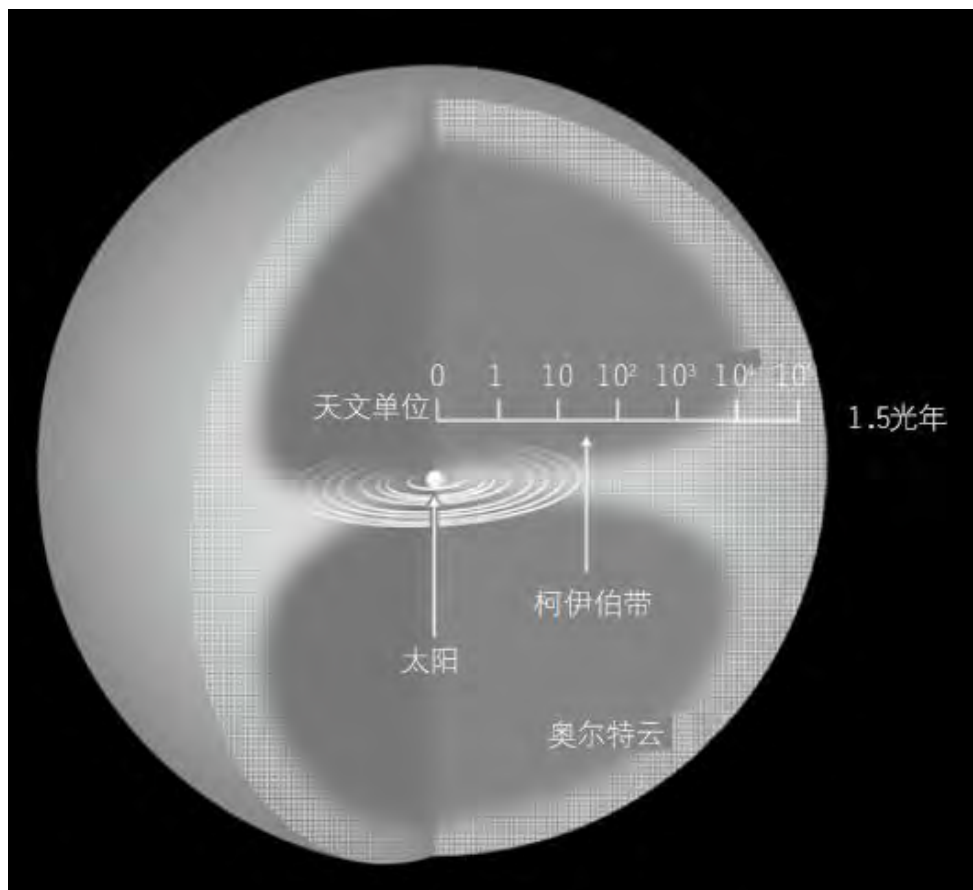


图7-3

太阳系遥远区域的奥尔特云位于行星和柯伊伯带的领地之外，延伸范围在1000~50000天文单位之间。

到奥尔特云的可能距离是非常远的。目地距离为1个天文单位。海王星，作为最遥远的行星，到太阳的距离是30天文单位。天文学家认为，奥尔特云与太阳的距离或许从接近1000天文单位延伸直到超过50000天文单位，比我们迄今考虑过的任何天体都要远很多。奥尔特云离太阳的距离，达到了太阳到与其距离最近的恒星——比邻星（Proxima Centauri）的相当一部分，后者大约是27万天文单位（4.2光年）。从奥尔特云发出的光几乎需要一年时间才能到达地球。

在太阳系遥远边缘的物体有着微弱的引力束缚能，这能够解释为什么它们很容易受小引力的扰动影响，从而导致我们可以观察到彗

星。轻轻的扰动就可以将它们推出原来的轨道，进入内太阳系，从而产生长周期彗星。当某个行星进一步改变了它们向内的轨迹，对这些弱束缚物体的扰动就可能产生短周期彗星。所以奥尔特云可能是造成所有长周期彗星的原因，如最近观察到的海尔-波普彗星（Hale-Bopp）。奥尔特云甚至也是造成一些短周期彗星的原因，哈雷彗星也许是其中之一。此外，尽管大多数木星族短周期彗星可能来自散盘，但其中一些彗星含有的碳和氮同位素的比值，类似于那些来自奥尔特云中的长周期彗星，这表明那里也是一个起源地。最后，甚至更具有破坏性的可能性是，在奥尔特云中受扰动的天体可以进入内太阳系，并和某个行星碰撞，也许是地球，从而形成了彗星撞击。稍后我会再讨论这个有趣的可能性。

长周期彗星提供了奥尔特云“居民”的性质线索。与其他彗星一样，它们含有水、甲烷、乙烷和一氧化碳。但奥尔特云中的某些成分可能是岩石质的，在成分上更类似于小行星。虽然被称为“云”，彗星的储备地似乎也有结构，包括一个甜甜圈似的环形内部区域（有时也称为希尔斯云，这是出于对杰克·希尔斯[Jack Hills]的尊重，他在1981年提出了这个单独的内部区域），以及一个由彗核构成的球形外部区域，它延伸到非常远的地方。

尽管尺寸巨大，外奥尔特云的总质量可能仅仅是地球质量的5倍。然而，它最有可能包含了数十亿个直径大约为200公里的天体，以及上万亿个直径大于1 000米的天体。模型表明，奥尔特云的内部区域一直延伸至大约2万天文单位，可能包含了比这个数目更多倍的物体。这个内云可能为外奥尔特云提供了天体来源，以补偿弱束缚的外奥尔特云的损失。如果没有内云，整个奥尔特云也无法存活下来。

由于奥尔特云过于遥远，我们无法去现场查看这些冰冷的天体。观察太阳系遥远边界的小天体非常困难，因为它们反射的光极少。奥

奥尔特云距离太阳比柯伊伯带远1 000倍，对于像奥尔特云这样距离极远的物体，目前的观测条件不可能看到。所以奥尔特云仍然是一个假想，因为没有人观测得到它的结构，或者它所包含的物体。尽管如此，奥尔特云还是被认为是太阳系的一个相当确定的组成部分。来自天空各个方向的长周期彗星的轨迹令人信服地证明了它的存在，以及彗星在这个遥远地区的起源。

奥尔特云可能是从原行星盘产生的。原行星盘最终导致了太阳系许多结构的出现。彗星碰撞、银河潮汐，以及与其他恒星的相互作用，都可能对奥尔特云的形成作出了贡献，尤其是在过去当这些相互作用被认为更加频繁时。在宇宙充满活力的早期，在接近太阳的地方形成的物体可能受到气态巨行星的影响从而向外移动，并创造了奥尔特云，或者奥尔特云中的物体族群可能是从散盘里的不稳定物体形成的。

当然，我们目前还不知道所有答案，但最新的观测工作和理论研究，让我们获得了更多有关太阳系外边缘的知识。也许我们不应该感到惊讶，因为那将是一个迷人的、充满活力的地方。



如果穆罕默德不走向大山，那么大山就会来到穆罕默德面前。

1977年，美国国家航空航天局发射了旅行者1号（Voyager I）太空探测器，用来执行一个为期4年的土星和木星探测计划。几十年过去了，它以惊人的持久力和稳定性飞行了超过125个天文单位的距离（在这个距离上，即使是光信号也需要大半天的时间才能传到地球），而且还在继续航行。旅行者1号太空飞船以及它所得到的测量成果，是以前所有航天器都未曾有过的。虽然它那基于八声道磁带的的数据收集系统一路上经常出错，它的摄像头也已经不能正常工作，其设备内存的大小只是现在普通智能手机的百万分之一，但这个太空飞船却仍然可以运行，而且它是人造天体中飞离地球和太阳最远的一个。

旅行者1号虽然已经退役，却忽然又在2013年成为新闻热点，因为美国国家航空航天局宣布，旅行者1号于2012年8月25日进入了星际空间。当新闻报道说旅行者1号已经到达了太阳系的边界时，引起了科学界内外很多人的激烈讨论。Twitter上的争论尤其激烈和有趣：为旅行者1号飞出太阳系而欢欣雀跃的推文以及愤怒地斥责人们不该再这么说的推文交相出现。我花了一些时间才发现，人们并不是在反对有关旅

行者1号的新闻，而是在质疑这条新闻的准确性。到底哪里才是太阳系的边界呢？

例如，我们现在已经探索到了奥尔特云，这是一个合理的太阳系边界候选者，但是旅行者1号和其他太空飞船却从未到达过如此远的地方。既然暗物质与流星雨的联系依赖于奥尔特云及其附近的区域，那就让我们先简单地讨论一下，旅行者1号飞出太阳系到底是什么意思。哪里是太阳系的边界？为什么它的边界如此难定义？

旅行者1号飞出太阳系了吗

太阳系只占了整个可见宇宙的很小一部分，但它也是非常非常大的。据最可靠的测量结果来看，它包含了奥尔特云，而奥尔特云至少延伸到5万倍日地距离（即5万天文单位）的地方，也就是比1光年的距离还要远。为了更确切地认识这个距离有多远，让我们想象一下目前最先进的宇宙飞船飞到太阳系外边缘需要多长时间。宇宙飞船飞行的速度和地球绕太阳运转的速度差不多，也就是说它在一年里飞过的距离大约就是地球绕太阳一周的轨道长度。这样算的话，宇宙飞船将需要大约八九千年的时间才能飞行5万天文单位，也就是到离太阳系最近的恒星1/5的路程。但是太阳系到底有多少个天文单位大呢？

目前占主导地位的有两个定义，而它们给出了不同的答案，而且第二个定义给出的结果非常模糊。在第一个定义中，太阳系就是以太阳的引力为主导的区域，其他恒星的引力在这一区域范围可以忽略。从这个定义来看，旅行者1号现在还在太阳系里面。确实，因为奥尔特云一直被认为是太阳系的一部分，而旅行者1号现在还没有进入奥尔特云，并且根据现在的估计，它要至少再过300年才能到达奥尔特云，然

后再过3万年才能穿过去。所以根据这个定义，旅行者1号已经不在太阳系里了这个说法就很难被接受。

然而，我们并不清楚太阳引力的“势力范围”，因此第一个定义其实也比较模糊。所以我们有了第二个定义：星际空间的起点就是太阳风所携带的磁场的终点——大约150亿公里远，或者说100个天文单位远。在这么远的距离上，无线电波也需要大约一天的时间才能到达我们。但是这比奥尔特云要近一些。

我们在第9章介绍过太阳风，它由太阳发出的带电粒子（也就是电子和质子）组成。这些粒子携带着磁场以每秒400公里的速度流向星际空间。星际空间的定义就是恒星之间的区域，而这些区域并不是真空的。它包含有冷的氢气体、尘埃、电离的气体，以及一些从爆炸了的恒星或者除太阳以外的恒星吹出的恒星风带来的其他物质。最终，太阳风和这些星际介质相遇。在它们相遇的区域造成一个空腔，我们叫它为日球层（heliosphere），而这两个区域的边界我们称其为日球顶层（heliopause）。因为太阳系的运动，这个边界并不是球形的，而是泪滴形的。

有些科学家认为，日球顶层就是太阳系和星际空间的边界。那么如果旅行者1号碰到的来自太阳系内的带电粒子越来越少，而来自太阳系外的带电粒子越来越多的话，就说明旅行者1号已经到了日球层的边界进入了外太空。这两种不同来源的粒子是可以区分的，因为它们具有不同的能量。高能的带电粒子起源于太阳系外的超新星所辐射出的宇宙射线。2012年8月，旅行者1号的数据表明，它碰到的这种高能粒子的数目突然增加，而同时低能量的粒子探测数目也有显著减少。既然这些低能粒子是来自太阳的，而较高能的粒子来自星际介质，那么关于这两种粒子的探测数据都表明旅行者1号已经离开了日球层。

然而，最初关于日球顶层的定义还有另外一个限制条件，那就是磁场的强度和方向也应该变得和日球层外面的一样。这个定义其实是随时间变化的，它依赖于太阳的“天气”——即在给定时刻太阳风的行为。结果是，虽然测到的带电粒子数据满足了飞出太阳系的条件，但是并不能满足更严格的磁场限制条件：旅行者1号并没有发现磁场有变化。

所以，虽然在2012年8月25日旅行者1号周围等离子体环境出现了变化，但是直到2013年3月，关于旅行者1号是否已经进入了星际空间的问题仍然饱受争议。尽管如此，美国国家航空航天局在2013年9月12日还是宣布它进入了星际空间。科学家们认为，磁场的改变其实并不是必要条件。他们决定用一个并不十分强的限制条件，也就是电子密度增强近100倍，这种情况在日球层外是应该碰得到的。

所以根据第一种以太阳引力为基础的定义，旅行者1号还处在太阳系内，而且在很长一段时间内它仍然会在太阳系里面。然而，根据第二种修正定义，旅行者1号已经进入了星际空间。关于旅行者1号是否已经离开了太阳系的问题，似乎取决于你使用哪一种定义。

还有一个有趣的补充信息：旅行者1号携带了一个镀金磁盘唱片，唱片中包含了给外星人看的有关人类社会的各种信息——如果外星人恰好能够捡到这个唱片的话。唱片中有吉米·卡特（Jimmy Carter，也就是当时的美国总统）用英语说的问候语，还有用49种其他语言说的问候语，鲸鱼的声音，以及查克·贝里（Chuck Berry）创作的经典曲目《Johnny B. Goode》（查克·贝里当时甚至还去了发射现场）。这种外星人在几百年后还可以播放我们的录音的想法还是挺让我吃惊的，因为这种事情并不太可能发生：外星人的个头会不会和我们差不多，他们有没有必要的播放设备——要知道，这种设备对在地球上的大部分人来说都是不容易找到的。而且，是不是也要考虑一下翻译的问题，或者他们能听到的音域和人类也许不一样的问题。不过提前多想一些也没什么坏处。这个唱片至少有一个好结果。它是这张唱片的创意导演安·德鲁彦（Ann Druyan）与天文学家、天体物理学家卡尔·萨根（Carl Sagan）合作的

原因。即便外星人很可能并不能理解这张唱片，但它至少成就了一个美好的爱情故事^[20]。

现在，我先将外星人的故事放在一边，并将话题转到我们更加确定的那个来自外太空的撞击上，也就是流星撞击地球或者冲入大气层的事件。如果穆罕默德不走向大山，那么大山就会来到穆罕默德面前^[21]。也就是说，虽然没有人能在近期到奥尔特云中去，但那些来自奥尔特云的类太阳系小天体确实偶尔会飞向地球。



09

危险地活着

DARK MATTER AND THE DINOSAURS

如果我们眼睁睁地看着一个天体在危险的近地轨道上运行几年，却没有能力改变人类命运的话，那么我们都会觉得自己是个大傻瓜。

某天，我趁着哈佛大学的春日假期去了趟科罗拉多看朋友，并在那里做了些工作，顺便滑了滑雪。洛基山是一个非常适合人们坐下来思考的地方，那里的夜晚和白天一样绚烂，似乎充满灵性。在晴朗而干燥的夜晚，天空被灿烂的星光照亮，不时会有一颗流星划过长空。这些流星就是很久以前在地球轨迹上解体的小天体。一天晚上，我和朋友站在我住的房子外面，惊奇地看着这美丽非凡的天空，天空中满是各种发光星体。在我看到了两颗流星之后，我和朋友又同时发现了一颗持续数秒的大流星。

虽然我是个物理学家，但我在看到这种壮观的景象时也常会停止思考，只享受眼前的美景。但这一次，我仔细思考了一下那个大流星是什么物体以及它的轨迹意味着什么。这颗流星——一个45亿年长的故事的高潮部分，只闪耀了几秒钟的时间，却也意味着这颗明亮的流星大约滑行了50公里~100公里后才气化并消失的。这颗流星距离我们

的高度也差不多是这个距离，所以我们看到它在天空中的轨迹是一个大的圆弧。流星不只具有美丽的外表，而且我们还可以理解它的一些本质。当我说这颗像尘埃或者石子大小的物体迅速划过苍穹的样子看起来非常美妙时，我那个不是物理学家的朋友感到非常惊讶，他说他本来以为这个物体的直径都快2 000米那么长了。

我们的对话很快从安静地享受美丽的夜空转到了思考一个直径约2 000米长的大物体猛冲向地球时所能造成的损害。这么一个危险的大家伙撞向地球的概率是很小的，而它撞向有人的区域并造成很大伤害的概率就更小了。虽然如此，从月球表面来推测（地球上存留下来的陨石坑太少了，我们无法据此精确估算），在地球的生命过程中，有数百万个直径大小从1 000米~1 000公里的物体曾经撞击过地球。但是大部分撞击都发生在几十亿年前一个叫作后期重轰炸（Late Heavy Bombardment）的时期。虽然被称作“后期重轰炸期”，但是这个时期其实是在太阳系刚刚形成后不久还没有演化成稳定状态的时期。

大流星撞击地球的频率自从重轰炸时期过后就已经很低了，而这是生命存活的必要条件。即便2013年2月被行车记录仪拍下来，并在Youtube上疯狂传播的在西伯利亚发生的撞击——点亮了半边天的车里雅宾斯克（Chelyabinsk）流星，直径也不过只有20米左右。最近唯一一个可以和我朋友想象的那种大撞击相似的事件发生在1994年，当时约2 000米大小的苏梅克-列维彗星碎块撞向木星。最初的彗星要更大一些，在它碎成小块之前有几公里大。这次撞击之后，我们可以观测到木星表面上有个大约和地球同尺寸的黑云，这就是约2 000米大的碎块所能造成的损害。20米是很大，但是直径将近2 000米又完全是另外一回事。

流星带来的并不全是毁灭。很多散落在地球上的流星和微流星也带来了不少好东西。陨



石——流星在地球上残留的碎块，可能是形成生命所必须的氨基酸以及生命存活必须的水的来源。毫无疑问，我们在矿场里挖出来的大部分金属也都来自地外撞击。我们甚至也可以这么认为：如果不是流星撞击（我将在第12章详细讲述）毁灭了地球上的恐龙的话，哺乳动物不会迅速统治地球，人类也不会出现——当然，也有人认为这并不一定是个好事。

6 600万年前的物种大灭绝只是将地球生命与太阳系其他部分联系起来的其中一个故事。这本书是关于抽象物质的，比如我所研究的暗物质，但是它也涉及地球与其周围宇宙环境的关系。下面我将开始讲述我们知道的一些撞击过地球的流星、彗星，以及它们在地球上留下的疤痕。我也会谈及在未来有可能撞到地球的物体，以及我们如何阻止这些破坏性的不速之客。

当陨石飞向地球

像来自太空的物体撞击地球这种奇怪现象听起来有些不可思议，确实，科学界最初是几乎不承认这类说法的。虽然古人相信外太空的物体会到达地球表面——即便是现代，一些村庄里的人也这么认为，但直到19世纪，受到良好教育的人也会对此表示怀疑。有些没上过学的牧羊人看到过有东西从天空坠落，但这些见证者没有可信度，因为很多和他们背景相似的人也曾经报告过一些想象出来的场景。即便科学家最后承认会有东西掉落到地球上，但他们最初不相信这些石头是从太空来的。他们倾向于地球来源说，比如在火山爆发中喷出的东西再重新掉落到地球上。

1794年6月在锡耶纳学院发生的一次偶然的陨石飞落事件，才让外太空的陨石理论开始慢慢成立。当时很多受过教育的意大利人和英国游客都在现场见证了这起事件。这个戏剧性的场景始于一朵喷着烟雾和火花的高高的黑云，这黑云慢慢地移动着，伴随着红色的闪电，随后就下起了石头雨。当时在锡耶纳的阿比·索尔丹尼（Abbe A. Soldani）觉得这些掉落的物质非常有意思，于是他收集了目击者们的描述，并且寄了一份掉落物质的样本给一位在那不勒斯的化学家古列尔摩·汤姆森（Guglielmo Thomson，真名为威廉·汤姆森，曾因为和他的小男仆发生了不轨行为而羞愧地逃离牛津大学）。

汤姆森经过仔细研究发现，这块物体是来自地球之外的。这比当时盛传的来自月亮或是雷电击中的尘埃的解释要合理得多，也比另一种认为其来自当时正处于活动期的维苏威火山（Vesuvius）的说法也更合理一些。维苏威火山在事发之前18小时碰巧刚好有一次喷发，所以将它当作天降石头的来源也情有可原。然而，维苏威火山距事发地320多公里，并且方向也不对，所以这种解释行不通。

化学家爱德华·霍华德（Edward Howard）在法国贵族科学家寇穆特（Jacques-Louis, Comte de Bournon，他在法国大革命时期被驱逐到了伦敦）的帮助下，最终推断出这些石头来自流星。霍华德和寇穆特分析了一块落在印度贝那勒斯（Benares）附近的陨石。他们发现，镍元素在这块陨石中的含量比在地球表面以及高压下形成的石头状物质中要高很多。汤姆森、霍华德以及寇穆特所做的化学分析正是德国科学家恩斯特·克拉德尼（Ernst F.F. Chladni）曾经建议的，他曾经想用这种分析，来证实他的猜想：这些物体撞击地球的速度太快了，因而和其他解释都不相符。其实锡耶纳的陨石雨只发生在克拉德尼的书《论铁的起源》（*On the Origin of Ironmasses*）发表两个月之后，而这本书最初所受到的评论都是负面的，直到柏林的报纸最终在两年后开始报道锡耶纳陨石雨的时候才有所改观。

在英格兰更为传播的是英格兰皇家学会会员爱德华·金（Edward King）在事发当年所写的一本小书。金的书总结了锡耶纳事件以及克拉德尼书里的内容。1795年12月13日，当一个重达25千克的石头落在约克郡的伍尔德村（Wold Cottage）时，陨石理论在英格兰被确立了下来。随着刚从炼金术中分离出来不久的化学方法越来越为人们所称道，再加上这么多一手的证据，陨石最终在19世纪被人们承认。自那之后，更多被承认的地外天体落到了地球上。

更近期的“拜访”事件

流星和陨石绝对是博人眼球的新闻标题。虽然人们会很热衷于追捧这些壮观的事件，但是我们不应该忘记：我们现在所生活的太阳系是处于平衡态的，也就是说基本不会有非常剧烈的分裂瓦解事件发生。几乎所有的流星都非常小，因而它们在地球大气层顶端就会将其固态物质蒸发殆尽。大一些的天体则很少到达地球。但小天体确实会来拜访我们，而且它们无时无刻不在造访。大多数时候进入大气层的是一些微流星体，这些小颗粒太小了甚至都烧不起来。虽然和微流星体比起来少了很多，毫米级的天体闯入地球的频率也很高——大约每30秒一个，而它们会被烧掉并且不会造成什么大影响。两三厘米大小的天体会在大气层中燃烧掉一部分，因此它们残留的碎块可能会落到地面，但是这些碎块太小了，也不会造成很大影响。

但每过几千年，就会出现一个由大天体在大气层底端引起的爆炸现象。历史上有记录的最大规模的这种爆炸1908年发生在俄国通古斯（Tunguska）。即便在没有表面接触的情况下，大气层中的大爆炸也可以在地球上造成可观的影响。造成这次大爆炸的那颗小行星或者彗星——我们无法确切知道是哪个，在西伯利亚森林中的通古斯河附近的天空中爆炸。这颗将近50米大小的火流星，一个来自太空并在大气

层中瓦解的天体相当于10兆吨~15兆吨TNT炸药，是广岛原子弹爆炸能量的1 000多倍，不过比现今爆炸的最大的核弹能量要小一些。这次爆炸摧毁了2 000平方公里的森林，产生的冲击波相当于里氏5.0级地震。值得注意的是，在原爆点上的树都得以保持直立，而它们周围的树却都被压倒了。这些仍然保持直立的树木所占区域的大小，以及人们并没有发现陨石坑，表明这颗天体很可能是在距离地面6公里~10公里的空中爆炸分解的。

我们对于风险的估算一直在变，部分是因为对通古斯天体的大小估算一直在变——估算的大小从30米~70米不等。在这个大小区间内的天体撞上地球的频率大概是几百年一次到两千年一次。即便如此，大部分击中或者接近地球的流星们都比较接近无人区，因为人口密集的区域和地球的总面积相比，毕竟是少数。

通古斯流星也不例外。它在西伯利亚一个未开发地区的上空爆炸，离此最近的一个贸易点在70公里之外，而离最近的村庄尼志讷-卡勒林斯科（Nizhne-Karelinisk）就更远了。即便如此，冲击波还是足够将这个并不近的村庄里的窗户击碎，将行人冲倒。村民们都被迫转身以避开这天空中足以致盲的强闪。科学家们在爆炸发生20年后回到这里，发现一些当地牧民受到了噪声和冲击波的伤害，其中有两人还在这次事件中丧生。而它对动物界的影响是毁灭性的，爆炸导致的大火烧死了大约1 000头驯鹿。

通古斯事件甚至影响到了更大的区域。在距离事发地点方圆相当于整个法国面积的区域内，都可以听到爆炸的声音，而且全球气压也因此受到了影响。爆炸产生的波动环绕了地球三次。在通古斯大爆炸之后还发生过一次更大的、也是我们研究得更好的希克苏鲁伯大撞击（Chicxulub impact，也就是毁灭恐龙的那次撞击事件），这类事件

产生了很多破坏性后果：风、火、气候变化，以及大气层中一半的臭氧层消失。

由于这颗流星是在一个偏远无人区爆炸的，并且当时通信不发达，因此大部分人直到几十年后的一次调查揭示了它的全部影响范围之后，才注意到这个巨大的冲击波。通古斯非常遥远，而且当时正处于第一次世界大战和俄国革命期间，因此这次事件就更难被人知晓了。如果这次小行星撞击仅仅提前或者拖后一小时，那么它将可能碰到一个大的人口聚集区，那样的话，大气层里的爆炸或者海洋产生的海啸很可能造成几千人丧生。如果这些变成现实，那么这次撞击将不仅会改变地球表面的自然环境，还会改变20世纪的历史——之后的政治和科学发展将是完全不同的走势。

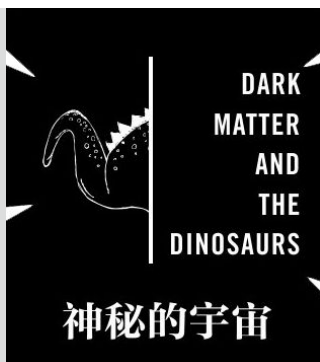
在通古斯大爆炸之后的100年里，还发生过几起虽然较小却仍有新闻价值的天外来客造访地球的事件。虽然记录并不完整，但是1930年在巴西亚马孙河上空爆炸的一颗火流星应该是其中较大的一个。它释放出的净能量比通古斯事件要小，估计大约为通古斯事件的 $1/100 \sim 1/2$ 。即便如此，这颗流星的质量也在1 000吨以上，甚至可能重达25 000吨——大约是10万吨TNT当量。我们对流星撞地球的概率估计不是很确定，但10米~30米的物体撞击地球的概率大约是10年一次到几百个世纪一次。我们对撞击概率的估计强烈依赖于天体的精确大小。大小上变化两倍可能会导致撞击概率变化高达10倍。

两年之后，一颗和亚马孙河上空的那颗差不多大的火流星在西班牙上空15公里处爆炸，释放出的能量大约是20万吨TNT当量。在接下来的50多年里又发生相当于多次火流星爆炸，但它们都不如巴西的这颗大（我不在这里一一列举了）。一个值得注意的例子是1979年发生在南大西洋和印度洋之间的维拉事件。维拉事件是根据发现它的美国维

拉号防御卫星命名的。虽然最初大家以为它很可能是流星，但现在都认为它是地球上的一次核爆炸。

当然，一些探测器也探测到过真的火流星。美国国防部的红外探测器和能源部的可见光探测器于1994年2月1日探测到了，在太平洋马歇尔群岛附近的一个5米~15米宽的流星信号。在距离事发地点几百公里处的密克罗尼西亚科斯雷岛（Kosrae）海岸附近的两个渔民也看到了这次爆炸。另一个更近期的10米级天体爆炸于2002年发生在希腊和利比亚之间的地中海上空，释放出的能量大约是2.5万吨TNT当量。当然更近期的一个事件于2009年10月8号发生在印度尼西亚的波讷（Bone）。这颗很可能是产生于一个直径10米的天体，释放出的能量高达5万吨TNT当量。

偏离轨道的彗星或者小行星都有可能变成流星。遥远彗星的轨迹很难预测，但是足够大的小行星在远未到达地球之前就可以被探测到。2008年撞在苏丹的一个小行星就是一个很好的例子。2008年10月6日，科学家们通过计算发现他们刚刚发现的这颗小行星第二天早晨就会撞上地球。而且它确实撞上来了。这次碰撞并不严重，而且附近也没有人居住。但是它确实显示出有些碰撞是可以预测的——即便我们能提前多久知道发生时间与探测灵敏度有关，而探测灵敏度又与撞击天体的大小和速度有关。



最近一次有新闻价值的事件是发生在2013年2月15日的车里雅宾斯克流星。这次事件不仅有很多现场照片，而且很多人都对它有很深的印象。这颗火流星在俄罗斯乌拉尔区域南部的上空爆炸，产生的能量大约是50万吨TNT当量。大部分能量被大气层吸收掉了，但是部分能量随着一个冲击波在几分钟后击中了地球表面。这次事件是由一个直径15米~20米大的小行星

引发的，它大约有1.3万吨重，估计下降的速度是18公里/秒，这个速度是声速的60倍。人们不仅看到了这次爆炸，也感受到了它传来的热浪。

大约有1 500人因此受伤，但大部分都是因为次级效应造成的，比如被击碎的玻璃划伤。受伤的人基本都是跑去窗边看那耀眼闪光的人。闪光的传播速度是光速，因此人们首先通过光发现了这次奇异事件。不幸的是——这也是很好的恐怖片素材，天空中的光先将人们吸引到了危险的地方，然后冲击波撞击，从而造成了很大伤害。

在这个流星撞上地球的时候，新闻曾报道说有另一个小行星正在接近地球。而这颗车里雅宾斯克流星并没有被提前探测到。新闻报道中的那个30米的天体，在16小时后到达最接近地球的地方，但后来并没有进入地球大气层。很多人曾推测这两颗小行星可能拥有共同的起源，但是后来的研究表明并非如此。

近地天体，最经常的撞击

就像2013年2月那个预测会撞地球的小行星一样，很多非常靠近地球却最终没有撞到地球的天体吸引了人们的大量注意力。另一些天体确实飞到了地球上，但是这些撞上来的天体中，绝大多数都是无害的。即便如此，之前的撞击已经影响了地球上的地貌和生态，而且以后很可能也还会造成影响。随着人们对小行星的理解日益增进，对其潜在危害（很可能有些夸大）的认知不断提高，搜寻有可能穿过地球轨道的小行星的任务会被人们重视起来。



最经常的撞击——虽然并不一定是最大的，来自近地天体（NEOs），这些非常靠近地球 的物体，它们与太阳的最近距离不超过日地距离的30%。大约有一万个近地小行星（NEAs）以及少数彗星满足这个条件。一些跟踪范围内的大流星也算是近地天体——严格来讲，一些绕太阳转的宇宙飞船也算是其中一种。

近地小行星可以分成很多类（见图9-1）。进入地球势力范围，和地球靠得很近，但是和地球轨道没有相交的一类天体叫阿莫斯（Amors）——以1932年飞近地球达160万公里（0.11天文单位）的小行星名字命名。它们虽然现在并不会穿过地球的轨道，但是对地球也有潜在的威胁：木星或者火星对它们造成的扰动有可能增加其轨道的偏心率，因此最终它们还是有可能穿过地球轨道。阿波罗（Appollos）也是以一个小行星的名字命名的，这是一类在径向方向穿过地球轨道，但轨道面却在地球椭圆轨道（太阳在天球上的轨迹就标示着地球的轨道）之上或者之下的天体，它们一般也不会与地球轨道相交。然而它们的轨道会随着时间变化，于是也有可能偏离到危险区域以内。另一类天体被称作阿特恩（Atens），和阿波罗相似，它们和阿波罗的不同之处在于其轨道区域比地球的还小。阿特恩天体也是以一个这种小行星的名字命名的。最后一类近地小行星叫阿提拉斯（Atiras），这种小行星的轨道完全在地球轨道圈以内。它们很难被找到，因此已知的只有那么几个。



图9-1

4类近地小行星。阿莫斯型小行星的轨道处于地球和火星之间。阿波罗型小行星和阿特恩型小行星的轨道会穿过地球轨道，但是在一部分轨道时间内有可能向外延展。阿波罗型小行星的轨道的半长轴大于地球轨道的半长轴，而阿特恩型小行星的则比地球的要小。阿提拉斯型小行星的轨道完全在地球轨道圈以内。

近地小行星在地质时间尺度和宇宙学时间尺度上并不一直存在。它们只能存在几百万年，在那之后就会被扔到太阳系之外，或者与太阳或太阳系内的一颗行星相撞。也就是说，我们需要源源不断的新的的小行星来填满地球附近的轨道。而这些小行星的来源很可能是受到木星扰动的小行星带。

大部分近地小行星都是石质小行星，不过也有很多碳质小行星。只有阿莫斯型小行星会超过10公里宽，而它们在现阶段并不会穿过地球轨道。但是有很多阿波罗型小行星的大小超过了5公里，这足以造成很大破坏——如果它的轨道不幸与地球相交的话。最大一个近地小行星有32公里宽，名字叫作嘎尼米德（Ganymed），这是特洛伊王子的德语拼法，在英语中为“Ganymede”。嘎尼米第是木星的一个卫星，是完全不同的天体，但它是同类（也就是太阳系里的卫星）里最大的一个。

近地小行星还包含有另外一个研究领域，是在最近50年内成熟起来的。早些时候，基本没有人认真考虑过外来天体撞地球的问题。而

现在，世界各地都有人在尽力收集近地小行星的信息并跟踪其轨迹。即便在前段时间，我在访问加纳利群岛参观特那利菲（Tenerife）望远镜的时候，发现他们的研究所所长带领着12个学生也在利用望远镜数据寻找近地小行星。这个又小又老的望远镜并不是当前最先进的，但是这帮兴趣盎然的学生以及他们对搜寻方法的热爱，都给我留下了很深的印象。

如今更先进的望远镜利用电荷耦合器件（charge-coupled devices, CCD）来搜寻小行星。CCD是一种利用半导体将光子转化成电子的装置，它可以将光子碰到的地方用电信号标识出来。自动读出系统也提高了小行星的发现速率。哈佛大学史密森天体物理中心的国际天文学联合会小天体中心的网站<http://www.minorplanetcenter.net/>会及时报道被发现的小行星、彗星以及近地天体的最新数目。

由于众所周知的原因，离地球轨道最近的轨道受到了最多的注意。美国和欧盟为了更好地搜寻这种小行星，合作建立了一个叫太空卫士（Spaceguard）的公司——为了纪念亚瑟·克拉克（Arthur C. Clarke）的科幻小说《与拉玛相会》（*Rendezvous with Rama*）。太空卫士的第一个任务由1992年的美国国会报告决定，即在10年内搜寻并分类归档大小在1000米以上的离地球最近的天体。1000米已经很大了，比能造成危害的最小天体还要大。之所以选择这个尺度是因为1000米大的天体更容易被找到，而且这种大小足以造成全球性的破坏。幸运的是，在我们所知的1000米大小的天体中，大部分都处在火星和木星之间的小行星带上。在它们改变轨道变成近地天体之前，是不可能对地球造成威胁的。

在仔细综合了观测数据、投影轨道以及计算机模拟之后，天文学家到2009年时几乎按时地完成了太空卫士所设定的目标——搜寻到了大部分1 000米大小级别的近地天体。最新的结果显示，大约有940个近地小行星的尺度是1 000米甚至更大。一个由美国国家科学院召集的委员会认为，即便考虑各种误差，这个数字也是相当准确的，总数应

该在1100以内。这些搜寻工作还认证了10万个小行星以及大约1万个小于1 000米的近地小行星。

大部分作为太空卫士搜寻目标的大型近地小行星来自小行星带的内边缘和中心地带。美国国家科学院的委员会认为，其中大约20%的近地小行星离地球的距离少于0.05个天文单位。他们将这些危险天体称作“有潜在危险的近地天体”。他们同时认为，所有这些天体都不会在一个世纪内对地球造成威胁，这当然是个好消息。这个结果其实也不太令人吃惊，因为1 000米大小的天体撞击地球的概率本来预期就不会超过几十万年一次。

实际上，只有一个已知的近地天体可能会在不久的将来击中地球并造成破坏。但是它撞上来的概率只有0.3%，而且这在2880年之前也不可能发生。即使考虑了所有的误差之后，我们基本上还是非常安全的，至少现在看来是这样。有些天文学家提出对另一个小行星的担忧。这颗小行星被冠以一个魔鬼的名字——阿波菲斯（Apophis）^[22]，有300米宽。天文学家预测，它在2029年最接近地球的时候不会与地球相撞，但是有可能会在2036年或者2037年返回并撞上地球。这是一种被称作“重力锁眼”（gravitational keyhole）的机制，在这种机制作用下，这颗小行星可能会被引导至地球。然而，进一步的计算表明，这只是一场虚惊。阿波菲斯和别的已知天体在可预见的未来都不会撞到我们。

在大松一口气之前，我们需要记住，还有很多较小的天体也会造访地球。虽然它们比太空卫士所关注的公里级别的天体要小，也不会造成那么大的破坏，但它们造访的频率却更高。所以美国国会在2005年延长了太空卫士项目的期限，并鼓励它去跟踪、归档以及描述至少90%的大小在140米以上的、具有潜在危险的近地天体。科学家们几乎肯

定不会发现什么对地球带来灾难性影响的目标，但是做这么一个星表仍然还是很值得的。

我们应该担心吗

很明显，小行星有时候会飞得离地球很近。小行星与地球的相撞毫无疑问会发生，但是相撞的频率和强度一直是个有争论的话题。是否会有东西在我们关心的时间尺度上撞到地球并造成破坏？人们并没有完全确定的答案。

我们应该担心吗？这与时间尺度、花费、我们的焦虑阈值、社会认知，以及我们自以为的控制能力都有关系。本书主要关心在百万年甚至十亿年的时间尺度上发生的事情。我在后文会描述我研究过的一个模型，它可用来解释具有300万~350万年周期性的大的（大约几公里大小）流星撞击。这些时间尺度不在人类的担心范围之内，人们有更加紧迫的其他事情需要担忧。

虽然接下来的话题有些偏离了本书的主题，但是在讨论了很多流星撞击地球的内容之后，不介绍一下科学家们关于它们对世界潜在影响的结论也是说不过去的。这个话题在新闻和谈话中出现过很多次，所以我们在这讨论一下最新的估计也无妨。这与政府部门也是相关的，特别是当他们考虑小行星的探测和使小行星轨道偏转时很重要。

根据美国国会2008年颁布的《强化财政预算法案》（*Consolidated Appropriation Act*），美国国家航空航天局请美国国家科学院的国家研究理事会（NRC）来研究近地天体。研究目的并不是要回答抽象的撞击问题，而是为了评估偏离轨道的小行星造成的威胁，以及是否可以实施某些措施来减轻这种威胁。

参与这项研究的科学家们将注意力放在了较小的近地天体上，这些近地天体撞击地球的概率更高，而且是有可能被转向的。在短周期轨道上的彗星与小行星的轨道相似，因此它们也可以用相似的方法被探测到。而长周期的彗星基本上不可能提前看到，它们也不太可能在地球的轨道平面上——它们来自各个方向，因此探测这种彗星更加困难。无论如何，虽然一些最近观测到的事件可能来源于彗星，但是彗星是很少到达地球附近的。我们也基本上不可能提前足够长的时间发现长周期彗星并作出有效反应，即便以后我们的技术发展到可以使小行星轨道偏折的地步，也无法做到。现在人们基本上也没有办法来做一个完整的危险的长周期彗星列表，所以现今的巡天项目只关心小行星和短周期彗星。

长周期彗星，至少是那些来自太阳系外缘的彗星，是我们之后所关心的目标。来自太阳系外缘的天体相比于内太阳系的天体所受的引力束缚要弱得多，因此，引力或者其他什么造成的扰动会更容易将它们移出原来的轨道，并送到内太阳系或者踢出太阳系。虽然这并不是美国国家科学院减灾研究的课题，但它们仍然可以作为科学研究的目標。

科学家的结论

2010年，美国国家科学院在一篇题为《保卫地球：近地天体巡天以及危险应对策略》（*Defending Planet Earth: Near-Earth Object Surveys and Hazard Mitigation Strategies*）的文件中发布了他们关于小行星及其所造成的威胁的研究结果。我将展示一些此文件中非常有意思的结论，以及其中的一些总结性图表，最后我会简单解释这些结论和图表的意义。

当你看到这些数字的时候，请记住乘上人口众多的城市区域的低密集度百分比——全球城市绘制计划（GUMP）对此给出的估计值大约是3%。虽然人们不愿看到任何破坏，但对城市区域的破坏是最恐怖的。城市所占区域的低密集度百分比告诉我们，小天体撞击人口密集区并造成严重破坏的概率是它们造访地球概率的1/30。例如，如果一个5~10米的物体撞击地球的概率是100年一次的话，那么它撞到城市的概率大约是3 000年一次。

同时需要注意，基本上在所有的估算方面都存在误差，科学家们在最好的情况下也只能将误差降到10倍以内。很多新闻中提到的遥远天体的威胁从未实现过的原因之一就是，即便对某些特定种类特定大小的天体来讲，轨道测量上的一个微小误差都将大大改变对它撞击地球概率的计算结果。即便是对已知巨大天体所能造成的影响和破坏，我们也不是完全清楚。虽然这些数字有如此大的误差，美国国家科学院的研究结果还是很可信、很有用的。接下来，让我们暂且容忍一下那些不确定性，先来看看2010年这些神奇的统计结果吧。

我最喜欢的一张表是表9-1。根据这些结果，每年由于小行星造成的死亡人数为平均91人。虽然小行星造成的死亡人数远比大部分灾难造成的死亡人数要少——其死亡率大约和轮椅相关的致命事故率相当（表中没有列出），表中的数字91仍高得不可思议。在各方面误差都相当大的情况下，这个数字也精确得有些可笑。显然，小行星造成的91条人命并非每年都有。实际上，我们在历史记录中只能找到很少几起此类死亡事件。表9-1中给出的高数字很有迷惑性，因为它包含了预测中的大量撞击事件，而这种事件很少发生。图9-2可能对我们的理解更有帮助。

表9-1 世界范围内多种原因下年平均死亡人数预测

原因	预计年死亡人数（人）
鲨鱼袭击	3~7
小行星撞击	91
地震	3.6 万
疟疾	100 万
交通事故	120 万
空气污染	200 万
艾滋病	210 万
烟草	500 万

注：这是美国国家科学院给出的关于全球各种致命事件每年造成的平均死亡人数的统计。统计结果是基于数据、模型以及推测得出的。

从图9-2中可以看出，表9-1中提到的死亡人数大部分都来自较大天体的撞击，即图9-2中1~10公里之间的那个高峰，而这种大天体撞击据预测是极少发生的，堪称小行星撞击事件中的“凤毛麟角”。如果我们只关注小于10米的小行星，那么每年的死亡人数就都在个位数以下，而这仍然是个不小的数字。那么，不同大小的天体撞击地球的概率分布到底是多少呢？让我们来看图9-3（这张图信息有些多，我会慢慢解释）。图9-3很好地总结了我们现有的知识。

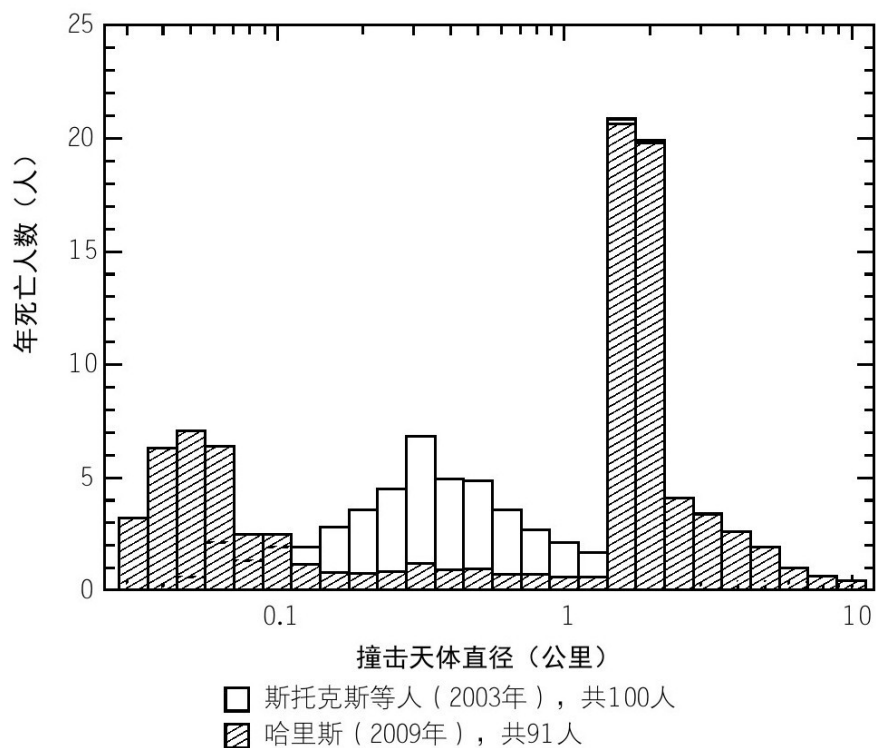


图9-2

美国国家科学院给出的由于小行星造成的年平均致死亡人数。数据是根据太空卫士一个完整度为85%的巡天项目得出的。该图利用了2009年的近地天体大小分布数据，并包含了新的海啸和空中爆炸所造成的威胁。2003年的估算结果也显示在图中，以便比较。

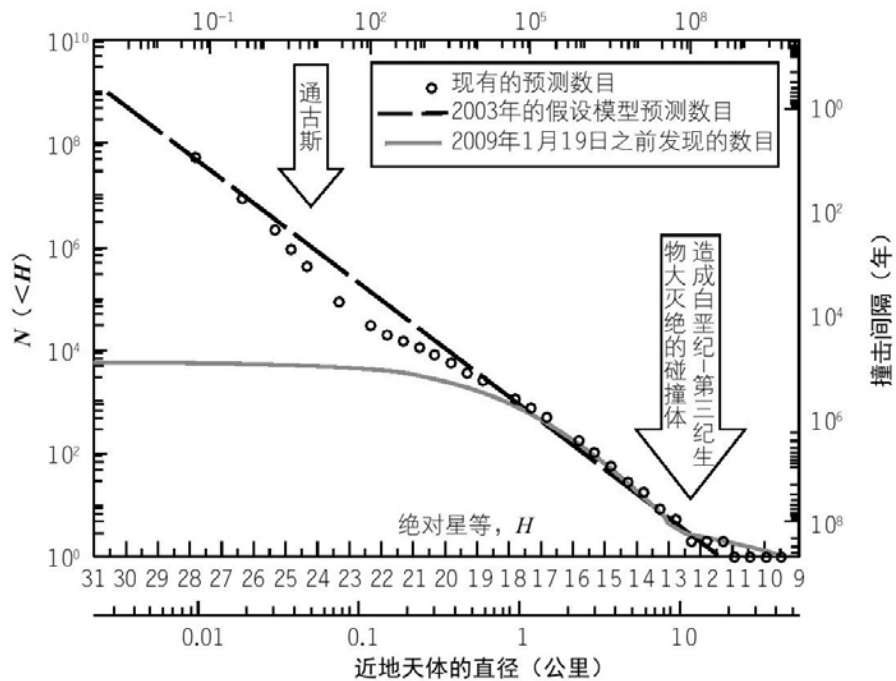


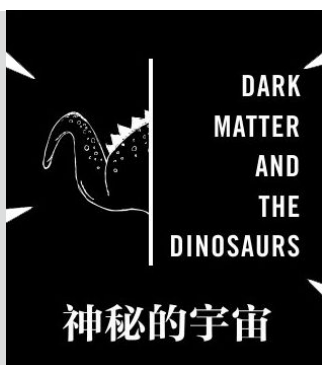
图9-3

撞击的数目（左侧纵坐标）和大致的撞击间隔时间（右侧纵坐标）与近地天体的直径的关系。上面的坐标给出了假设天体在撞击时的速度是每秒20公里的情况下，所能释放出的能量（以TNT当量为单位）。下面的横坐标上同时标出了天体的星等：虚线是2003年的估算，圆圈是新的估算结果，实线表示2009年以前发现的小行星数目。

图9-3虽然不太容易读懂但确实包含了很多信息。图中标度都是对数尺度，意味着天体大小的变化对应的时间尺度变化可能比你想象的要大得多。比如，10米大的天体可能每10年造访地球一次，而25米大的天体可能每200年才会撞击地球一次。这也就意味着，测量值的微小变化可能会对预测值造成很大的改变。

图9-3中上边框的标度显示一个给定大小的天体在每秒20公里的速度下所能释放的能量，单位是兆吨。比如，一个25米大的天体能释放出的能量是1兆吨。图9-3还能告诉我们不同大小的天体的数目会有多少，以及它们有多亮——这关系到探测以及追踪它们的难易程度。虽

然较小的小行星数量非常多，但是它们非常小也非常暗，使得此类小行星更难被探测到。



举例来说，一个500米大小的天体撞击地球的频率估计为10万年一次，1 000米大的天体大约为50万年一次，而一个5 000米大小的天体要2 000万年一次。图9-3也同时告诉我们，一个可以灭绝恐龙的大约10公里大的天体撞击大约1 000万年~1亿年才会发生一次。

如果你只对撞击的频率感兴趣，那么图9-4将给你更清晰的答案。注意，图9-4中的纵坐标上面年数最小、下面年数最大，因此大撞击发生的频率远小于小撞击。注意，纵坐标是以指数形式增长的，例如， 10^0 是1， 10^1 是10， 10^2 是100。

最后，为了解释不同大小的天体所能造成的危害，我需要再展示最后一张来自美国国家科学院研究报告中的表格（见表9-2）。表9-2告诉我们，一个直径几公里的物体会影响整个地球。大流星撞击比自然灾害要少见得多，因此它们并不会造成迫在眉睫的威胁。但是一旦发生，它们所释放出的能量和严重程度将是毁灭性的。表9-2也显示了，比如，一个300米的物理撞击地球的频率可能是几十万年一次。这种撞击会使大气中硫的成分提高到喀拉喀托火山（Krakatau）周围物质成分的程度，这样会破坏地球上很多地方的生命或者农作物。当然，这种灾难具体所能造成的破坏依赖于天体的大小和所撞击的地点。

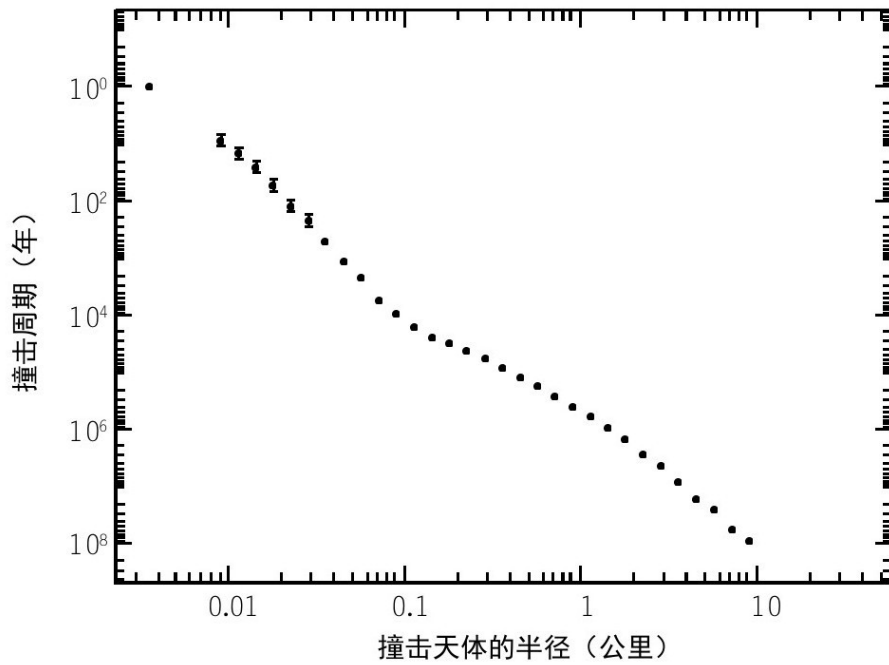


图9-4

不同大小（直径3米~9 000米）的近地天体撞击地球的平均时间间隔。

表9-2 不同大小近地天体的平均撞击间隔和撞击能量的估算值

事件类型	直径	撞击能量 (兆吨)	撞击间隔 (年)
空中爆炸	25 米	1	200
地方性影响	50 米	10	2 000
地区性影响	140 米	300	3 万
大陆性影响	300 米	2 000	10 万
全球性灾害下限	600 米	20 000	20 万
可能的全球性灾害	1 000 米	100 000	70 万
全球性灾变上限	5 000 米	1 000 万	3 000 万
大规模物种灭绝	10 000 米	1 亿	1 亿

注：需要注意的是，这些值取决于撞击天体的速度以及物理化学特性。

保护地球，让小行星转向

那么，我们对此该做何种结论呢？首先，所有这些天体都在同一个空间中沿一定的轨道运行，这是一件非常令人着迷的事情。我们认为地球是特别的，当然我们会想要保护它。但是在更大的图景上来看，地球只是内太阳系的一颗普通行星，围绕着一个普通的恒星旋转。其次，虽然我们知道周围的邻居们离我们非常近，但是这些小行星并不是人类最大的威胁。小行星撞击可能发生甚至可能造成一些破坏，但是它对人类实际上并没有眼前的威胁。

即便如此，我们也会问，如果危险的事情真的发生了，我们应该怎么办呢？如果我们眼睁睁地看着一个天体在危险的近地轨道上运行几年，却无能力改变人类的命运的话，那么我们都会觉得自己是大傻瓜。没有近在眼前的威胁并不意味着，我们可以对可能发生的流星造成的破坏漠不关心甚至放任自流。

当然有一些人已经想到了这个问题，并且提出了很多应对来自太空危险天体的提议——虽然现在还没有实际的设备。两个应对的基本措施就是摧毁它或者使其偏移轨道。仅仅摧毁并不一定是个很好的办法。如果你把一个正要撞向地球的天体炸成很多飞向同一方向的小石块，反而会增大撞击的概率。虽然每个小块的破坏力比较小，但如果有办法能把撞击数目也减少的话就更好了。

因此，将小行星偏移轨道的方法似乎更加合理。最有效的偏移方法是使飞来的天体增速或者减速，而不是侧推。地球其实是很小的，而且围绕太阳转动的速度很快，大约每秒30公里。根据飞来天体的方向，将它的轨道改变一下，只须让它提前或者拖后7分钟——这是地球在公转轨道上移动一个地球半径的时间，这样做就可以避免一次撞击而将其变成一次激动人心却无害的擦肩而过。这并不是一个很大的轨

道改变。如果我们能提前足够的时间（哪怕只是几年）探测到这种危险天体，即使是速度上的一个很小改变，就足够达到目的了。

无论是摧毁或者偏移的方法都无法将我们从一个直径大于几公里的、可以造成全球性破坏的天体手里拯救出来。幸运的是，这种撞击在接下来的100万年里估计不会发生。如果碰上小一点的天体，原则上我们是自救的，最有效的偏移方法是核爆炸，它能阻止一个最大直径为1 000米的天体的撞击。然而，法律是不允许在太空进行核爆的，所以现在这种技术还没有被开发出来。另一种可能的方法（虽然并没有核爆那么有力）是用一个物体与飞来的小天体撞击，这样此物体的动能（也即运动的能量）就会转给小行星。如果提前量足够，特别是如果多次撞击小行星，将会使直径几百公里^[23]大的天体转向。其他可以使小行星发生偏移的方法可以是太阳能板、将卫星作为引力拖船、火箭发动机等任何可能产生足够推力的物体。这种技术最终可能会对100米大小的天体有效，但是需要有几十年的提前量。所有这些方法（以及小行星本身）都需要更多的研究，因此现在就下定论哪种方法更有效还为时尚早。

这些提议虽然有趣也值得仔细考虑，但是现在只是关于未来的一些可能性。并没有现存的技术能实现它们。然而，有一个叫小行星撞击及偏移评估（Asteroid Impact and Deflection Assessment）的计划——其主要目标是测试动力学撞击小行星的可能性，已经做了很多相关工作了。还有另一个相关项目，叫小行星重定向任务（Asteroid Redirect Mission），它的目标是将一个小行星或者一块物体改道成绕月亮旋转，并计划派一个宇航员去造访这颗小行星。然而所有这些计划都没有实际展开建造工作。

有些人会反对在地面上建造反小行星的设备，因为这可能会在更大的意义上造成伤害。比如，有人担心这种技术会被应用到军事上而

不是用来保卫地球，不过我觉得这种事情不太可能发生，因为要制成相关设备需要很长的研制时间。还有人提出：如果发现一个会与地球交会的小行星，但是如果我们没有足够长的时间或者足够先进的技术去阻止它，那么小行星的发现可能会引发心理学和社会学上的问题。这种想法让我相当吃惊，因为它看起来就像拖延战术，而这种拖延却阻止了很多有益措施的发展。

暂且不管这些无谓的担心，我们总是可以问，是否需要做提前准备？如果需要的话，什么时候开始？这其实是一个成本和可利用资源的问题。国际航天学会（IAA）召开了很多会议，就是为了解决这些问题并讨论最佳的战略方法。

一个曾参加过2013年在美国亚利桑那州弗拉格斯塔夫（Flagstaff）召开的行星防御会议（PDC）的同事告诉我，会议给参会者布置了一项小作业，让每个人想象一个假想的要撞向地球的小行星，然后给出自己能想到的最好的处理方案。具体的问题包括：

- 如何处理小行星的尺寸和轨道参数的不确定性，并不断地修正它们？
- 应该什么时候作出应对措施？
- 什么时候应该给总统打电话（当然因为这是在美国开的会）？
- 什么时候应该疏散区域内人员？
- 什么时候应该发射核导弹去阻止相撞？

这些问题——虽然在某种程度上在我看来非常有娱乐性，清楚地显示出：即便是熟悉相关知识且对此感兴趣的天文学家也会在如何看待及应对天外来客的问题上存在很大分歧。

希望我已经说服了你，虽然存在小行星撞击地球的潜在危险，但是这种危险并非迫在眉睫。虽然存在某个小行星撞到地球并将一个大型人口密集区夷为平地的可能性，但是这种可能性在可预见的未来是极为微小的。我身体中的“科学家部分”在尽力做着搜集和理解各种天体轨道的事情。而我身体中的“极客部分”则认为，一个可以将危险的近地天体送出地球轨道的宇宙飞船非常酷。但是，说真的，没人确切知道该如何实施。

对于社会的终极问题是：在穷尽了这么多科学和技术上的努力之后，我们该重视什么？我们学到了什么？除此之外还有什么其他益处？当你要作出某种选择的时候，你可以用一些基本事实作为考量的基础。现有数据是有帮助的，但是它们并不完备。而在制定政策方针的时候，我们需要综合考虑专家们的猜想、实际情况以及伦理道德。我认为，即便在没有任何威胁的情况下，这方面的科学问题也是很有趣的，值得投入相应资金来寻找更多的小行星并且做更多的研究。但是，只有时间才能告诉我们，这个社会以及私人企业的最终决策会是什么。



陨石坑，天外来客的“名片”

ORR MATTER AND THE DINOSAURS

撞击坑是一颗飞驰的流星体坠落到地球上之后留下的非凡“名片”。

最近一次去希腊的时候，我在那里遇到的当地人偶尔会让我觉得自愧不如：他们有着令人印象深刻的英语词汇量，有时他们用的词语是连我这个以英语为母语的人在使用的时候都会犹豫的。当有人用“eponymous”（与作品同名的）一词的时候，我又谈起这件事，我的对话搭档提醒我这个词源于希腊语。当然，英语里的很多词都源于希腊语。

“Crater”（物体坠落等造成的坑、火山口）一词就是其中之一。古希腊人虽然是葡萄酒的极大爱好者，显然也欣赏节制的美德。除非在狂欢的时候，平时葡萄酒要与3倍的水相混合，而“krater”（调酒缸）就是用来混合水和酒的容器。调酒缸有一个大大的圆形开口，形状和地球、月球上也叫这个名字的巨大张口区域有点类似。但是具有类似名字的地质特征可以横跨200公里，而其周边受到影响的地区甚至可以更大。

地球上的一些环形山是由火山喷发形成的，没有任何外来力量的帮助。例如，在加那利群岛的特内里费岛（Tenerife），你可以在泰德（El Teide）火山的巨大熔岩区看到几个令人难以置信的环形山。这是地球表面之下的混乱动荡偶尔冒泡出来的证据。我也是从这了解

到，“caldera”（火山口）一词在西班牙语中意为“大锅”。我发现，我们对火山凹陷所使用的词语和“crater”这个词有着相似的起源。而另一方面，撞击坑^[24]的形成地点比较孤立，而且更值得注意的是，它们仅由地球之外的力量造成的。

撞击坑（陨石坑）

撞击坑是飞驰的流星体坠落到地球平面时造成的坑体。理解撞击坑的形成、形状和特性，可以帮助我们确定不同大小的石块撞击地球的频率，并可以更好地理解流星体在生物灭绝中可能起到的作用。

大多数流星撞击，包括所有大的撞击，都发生在有人类可以勘察之前，更不用说有人来记录它们了。撞击坑是一颗飞驰的流星体坠落到地球上之后留下的非凡“名片”。撞击坑或凹陷以及其周围的物质，往往是唯一的证据，记录着这些意外访客到达时所造成的巨大破坏。残骸中掩埋的疤痕、岩石的类型和化学丰度，给我们提供了关于这些久远事件的最可靠的信息。

撞击坑为地球与其外围环境，即太阳系的最后关联，提供了非凡的证据。理解撞击坑的形成、形状和特性，可以帮助我们确定不同大小的石块撞击地球的频率，并可以更好地理解流星体在生物灭绝中可能起到的作用。

在这一章，我将解释这些令人惊叹的撞击坑最初形成的原因和方式，以及它们和地面上由火山造成的凹陷的区别；我还将讨论一系列对象，它们以足够强的力量击中地球，并造成持久的影响。^[25]

流星的痕迹

在深入了解撞击坑的形成以及地球上撞击坑的完整列表之前，让我们花点时间再回想一下最早的那个发现，它将天外物体和地球表面联系在了一起（见图10-1）。虽然流星陨石坑这个名字有点不准确——请记住，“流星”是指在空气中的光痕，而流星陨石坑是由流星体造成的，按照定义，流星陨石坑都是撞击坑。这个特殊的陨石坑位于亚利桑那州弗拉格斯塔夫附近。

按照流星命名的惯例，它的名字与附近的一个邮局相关。这个邮局是由西奥多·罗斯福在1906年建立的。当时他的朋友，同时也是采矿工程师和商人的丹尼尔·巴林杰（Daniel Barringer），开始调查这个神秘大坑的成分和来源。地质学家最初怀疑巴林杰的提议，但最终巴林杰证明大坑起源于一个流星体。为表彰他的贡献，人们将这个凹坑称为巴林杰陨石坑。

虽然还存在更大的撞击结构，这个陨石坑是美国规模最大的陨石坑之一。它的直径长达1 200米左右、深170米、边缘的环状轮廓高约45米。这个陨石大约有5万年的历史，你甚至可以在地球表面上看到它。你可以在美国一睹其貌，因为这个坑已经被私有化了。巴林杰家族通过巴林杰陨石坑公司拥有这个陨石坑，并向参观者收取16美元的门票。该所有权在1903年开始受到保护，当时，巴林杰和数学家、物理学家本杰明·蒂尔曼（Benjamin Tilghman）一起声明了所有权，这一声明在不久之后获得了总统的签署。而持有股份的标准钢铁公司（SIC）得到许可，拥有约259万平方米土地的使用权，从而可以在该区域进行开采矿产。



图10-1

位于亚利桑那的直径超过1 000米的巴林杰陨石坑。（航拍图由D. 罗迪提供）

由于是私人财产，陨石坑不能成为国家公园系统的一部分。只有联邦政府所有的土地可以作为国家纪念场所，所以巴林杰陨石坑仅仅是一个国家自然地标。其好处是，当政府关闭的时候它不会被关闭——2013年，当地政府就倒闭了。私人所有权的另一个好处是，因为巴林杰家族对靠陨石坑获益，因此会对陨石坑多加保护，而它确实被认为是世界上保存最完好的流星撞击地点——当然，这也和陨石坑的形成年数很短有关。

与这个陨石坑相关的陨石被称为暗黑陨石（Diablo meteorite），以鬼镇坎宁迪亚布洛（Canyon Diablo，暗黑峡谷）命名，它位于同名峡谷的沿线。一些直径50米的流星体，几乎是由纯的铁和镍组成的，大概以每秒13公里的速度撞击到地面。撞击产生的能量相当于至少两个百万吨级的TNT炸药。这是车里雅宾斯克事件所释放能量的几倍，相当于一颗氢弹的能量。大部分初始物体被蒸发，碎片难寻。一

些好不容易找到的碎片被陈列在当地的博物馆里，有的甚至在被出售。

由于缺乏足够的碎片做研究，起初人们很难确定这个大坑是由地球之外的物体而非火山爆发造成的。19世纪，来自欧洲的移民第一次偶然发现它时，以为这是个火山口。在当时这个假设很合理，因为外星球的解释太奇怪。此外，旧金山火山区就向西65公里左右的地方，这一点相当有误导性。

有一个关于科学出错，后来才得以修正的富于启发性的故事。美国地质调查局首席地质学家格罗夫·吉尔伯特（Grove Gilbert）在1891年作出官方定论：这是一个火山。他从费城矿物经销商阿瑟·富特（Arthur Foote）那里听说了这个大坑。富特对1887年牧羊人在附近发现的铁很感兴趣。他辨别出了金属的外星起源，并且到现场去看了看他还能挖掘出些其他什么东西。除了铁之外，富特还发现了微小的钻石。这些在撞击的时候已经形成，但富特并不知道这个，错误地认为撞击地球的物体和月亮一样大。富特还犯了一个错，他没有将大坑和他正在调查的陨石材料联系起来。虽然他知道在地上的材料来自外星球，但在他的脑海中，附近的大坑是一个独立的现象，是由火山活动形成的。

另一方面，吉尔伯特从富特那里了解到这个大坑，他也是第一个提出大坑的起源于流星体的人之一。但是在吉尔伯特企图科学地认证他的推断时，也得出了错误的结论。由于当时没有人明白撞击坑的形态，他错误地排除了自己的撞击假设，因为在环形边缘物体的质量和大坑里丢失的物体质量不一致。另外，大坑的形状是圆形而不是椭圆形——如果撞击来自一个特定的方向，他预测撞击坑应该是椭圆的。此外，没有人发现铁含量的任何磁性差异的证据，以表明撞击物是来自外星球。由于缺乏是流星体的证据，吉尔伯特被他自己的方法误导

而得出了错误的结论：大坑是火山活动而不是由撞击形成的。我将很快提到，他的方法忽视了一些撞击坑形成的微妙因素。

大坑的起源在1905年最终被正确确定下来。巴林杰和蒂尔曼在《费城自然科学院学报》发表了几篇优秀的论文，证明了该流星陨石坑确实来自一次外星撞击。他们的证据包括翻倒的边缘地层（有人告诉我这看上去相当壮观）；还包括沉积物中的氧化镍。然而，撞击坑周围的30吨氧化铁陨石碎片导致巴林杰犯下了一个昂贵的错误。巴林杰认为，剩余大部分的铁被埋在了地下，于是他花费了27年的时间挖掘寻找。如果真的有所发现，这将会是巴林杰的另一个财源。1894年，他从也是位于亚利桑那州的联邦银矿赚得1 500万美元（相当于今天的10亿多美元）。

陨石比巴林杰预想的要小，大多数陨石在撞击的时候被高温烧掉蒸发了。所以，巴林杰没有挣到钱，即使在挖掘完成之后，他也没有成功说服很多人相信大坑的起源。在卸任陨石坑探采公司主席职位几个月之后，巴林杰因心脏病发作去世。巴林杰和他的公司在勘探陨石坑上损失了60万美元，但至少，巴林杰活了足够长的时间以维护他的假说。

由于行星科学的发展，人们最终开始更全面地了解陨石坑的形成，巴林杰的推论得到更多科学家的认同。最后的证实发生在1960年。一个在科学理解撞击上的关键人物——尤金·苏梅克（Eugene Shoemaker），在撞击坑中发现了二氧化硅的罕见形式。这种形式，只能在内含石英的岩石受到由撞击压力产生的严重冲击时才会产生。除了核爆炸——而在5万年以前这不太可能出现，而流星体的撞击是唯一可能的已知原因。

苏梅克仔细绘制了撞击坑的地图，展示了这个大坑和内华达州的核爆炸大坑之间的地质学相似性。他的分析使地球上的外星撞击概念

合理化，成为地球科学的一个里程碑，同时吸纳了地球与其宇宙环境相互作用的显著性。

撞击坑的形成

我对攀岩的喜欢很大一部分来自调查岩石的材料、质地和密度所得到的快乐，我可以通过近距离检查岩石表面，以确定最安全和最有效的路线。但是，埋藏在岩石中的真正宝藏是它们悠久的历史。与它们表现出的板块运动的证据一起，岩石的形态和成分为地质学家提供了可以评估的信息宝库。古生物学家也从地球的嵌入化石和其地形中学到很多。

岩层的形成在讲述着一个又一个故事。在这一方面，一些地方尤为壮观。

不久前，我访问了西班牙毕尔巴鄂市巴斯克地区（Basque country）的大学。我很幸运，有一个研究物理学的同事告诉了我关于苏玛亚（Zumaia）附近小镇复理层地质公园的事情。地质公园是一个很不错的生态旅游地。它以露出地表的让人难以置信的石灰岩为特色。这些石灰岩代表了几百万年的地质历史。这个地方非常迷人，因为它不仅对那里的地质宝藏的利用提供了可持续发展的经济，还提供了多样化的科学活动和发现。

当我参观地质公园时，那里的科学负责人向我指出，跨度6 000万年的岩层沿垂直的山崖随时可见。山崖位于迷人的海岸边（见图10-2）。他将悬崖描述为一本展开的书，每一页都在同一时间可见。K-T分界线（白垩纪-第三纪分界线，现在被称为K-Pg[白垩纪-古近纪]分界线，我将在稍后讲到）分离开含有化石的白色岩石层和上方没有化石的灰色岩石层。这条标志着最后一次大灭绝的线被完好地保存在巴斯克地区这个安静的地方。



图10-2

复理层地质公园的岩石上那看得见的6 000万年历史。摄于西班牙苏玛亚附近的Itzurun海滩。（感谢乔恩·尤勒斯提拉[Joh Urrestilla]提供图片）

这样壮观的岩石层并不是了解过去的唯一途径。撞击坑，这些在地球表面上最显著的结构，形成了非常不同的信息源泉。尽管我们对于流星体是如何以及何时撞击的了解有限，但科学家们对于撞击坑的地质学了解了很多。大坑的形状、岩石形态和成分提供的线索有助于将撞击坑和火山口或其他圆形凹陷区分开来。而且，由于撞击坑独特外观和成分可以在很大程度上从其来源得到理解，流星体造成的地面崩塌处的凹陷和特殊的岩石类型告诉了我们很多关于最初形成陨石坑的事件信息。

如果不是已经被极不成功的军事政策败坏了，“冲击与震慑”很可能是对撞击坑的形成的最中肯描述。撞击坑是地外物体撞击地球，撞击能量足够大以至于形成了一个冲击波，冲击形成一个圆形的坑。

冲击波，而不是直接的撞击，使得撞击坑保持了圆形的形状。如果是更直接的挖掘，将会生成一个有方向偏向的凹陷，反映出撞击物的最初方向，而不是看起来四处都一样的大坑。这是误导了吉尔伯特对巴林杰陨石坑分析的虚假论点。但是，撞击坑不能被简单地理解为撞击物对岩石的向下冲击推动。撞击坑是这样形成的：当撞击物以非常大的力量向下冲击地球时，被压缩的区域像活塞一样，会迅速减压以释放应力，从最初的冲击反弹，并喷射出物质。通过冲击波的半球模式的压力释放是实际创建火山口的爆炸。这个地表以下的爆炸产生了撞击坑独特的圆形形状。

形成撞击坑的物体通常撞击地面的速度是地球逃逸速度倍，也就是11公里/秒，最为典型的大约20~25公里/秒。对于较大的物体，这一速度是声速的很多倍，保证了巨大的动能被释放，因为动能不仅随质量增长，同时还会随速度的平方增长。对坚固的岩石的一次撞击，可以与一次核爆炸产生的影响相当，产生的冲击波压缩同时压缩来自太空的物体以及地球的表面。撞击释放的冲击加热它所遇到的物质，并且几乎会一直熔化并蒸发进入的流星体，如果流星体足够大，也能熔化并蒸发掉目标区域。

不断扩大的超声波产生的压力远远超过当地物质的承受强度。这催生了罕见的结晶结构，如冲击石英，这些结晶体结构只在撞击坑和核爆炸的冲击区域才会被发现（见图10-3）。其他特征性质包括岩石上破碎的锥体，它们是锥形的结构，其顶点指向碰撞点（见图10-4）。碎裂锥体也是一个明确的证据，再一次表明一次高压力的事件只能由撞击或核活动解释。碎裂锥有趣的地方在于，它们的大小从几毫米到几米的范围变化，从而提供了物质的宏观尺度效应。和晶体变形以及岩石熔化的证据一起，震裂锥帮助人们区分出了真正代表撞击事件的撞击坑。

还有一些具有撞击特征的岩石是在高温下形成的。这些被称为玻璃陨石和冲击熔融球粒等玻璃材料的物质，起源于熔岩。由于它们是在高温下（并不见得是在高压下）产生的，可以想见，它们也能起源于火山，而火山也是于大坑形成另一个主要原因。但是撞击坑通常具有不同的化学组合物，包括金属和其他物质，如镍、铂、铱和钴，这些是地球表面上罕见的。这些额外的线索帮助确证了陨石坑源于撞击。

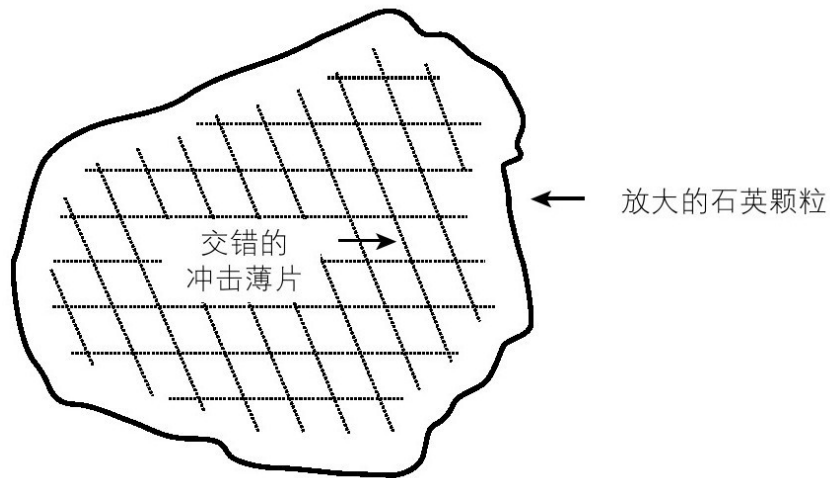


图10-3

冲击石英中独特的交叉变形模式，表明它产生于陨石的强力撞击。

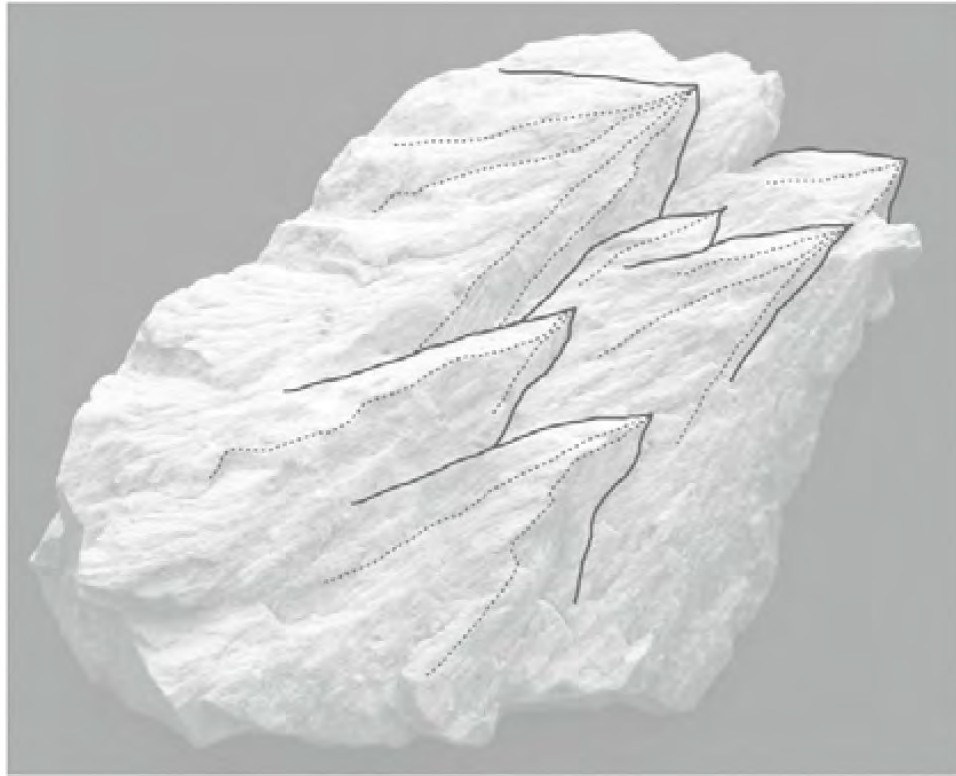


图10-4

在同一块岩石上，多次出现了不同尺寸的、明显的锥形形状。它是岩石结构在高压条件下形成的宏观迹象。

撞击物体的化学成分也能有其他的鲜明特点。例如，特定的同位素——也就是具有相同的电荷但不同中子数目的原子，大都产生于地球之外。由于大部分原始物质都被蒸发了，因此这个方法只对剩余物质的一小部分有用。

冲击角砾岩对区分陨石坑也很有用，它由通过细粒度基岩材料连接在一起的岩石碎片组成——再一次表明了撞击将原本在那里的物体击碎了。受冲击熔化的玻璃也很有趣，它们的形成既需要高压也需要高温，其不寻常的高密度有助于识别它们。另一个显著的特征存在于撞击坑底部或中心岩层的岩脉中，它们由玻璃颗粒组成，并在撞击坑底部组合成复杂结构。

这些与众不同的冲击和熔化的特点是证实撞击事件的关键，因为它们在其他条件下无法形成。然而，找到拥有这些特点的岩石并不容易，因为它们可能深埋于岩石碎片并且可能熔化了。虽然如此，陨石比比皆是，许多自然历史博物馆都有展示。我喜欢在纽约的美国自然历史博物馆展出的2米多高、34吨重的陨石阿尼吉托（Ahnighito），它是目前在展出的最大陨石。这块巨大的陨石是后来收购的，加入到博物馆自从1869年创立以来已收藏的陨石藏品中。

物质材料有助于识别陨石坑，而陨石坑独特的形状也能够帮助辨别。陨石坑是中心区域下陷而低于周围地面，而大多数火山口由喷发产生，所以高于周围平面。陨石坑的环状边缘也会抬升起来，这对于火山口也是不典型的。

另一种识别特征是倒置地层（inverted stratigraphy），也就是翻转的边沿地层。这是由于中心的物质被挖掘出来之后“翻转”到撞击坑外部导致的，它类似一叠大煎饼的边缘。在地球表面或在任何行星或月亮上的大致圆形的凹陷，都有着抬升的边缘和倒置的地层。这也提供了明确的证据：一个大质量的物体以巨大的速度撞击到表面上。

虽然区分陨石坑的物质大多数是在突发冲击波释放的过程中成形的，但陨石坑的形状也依赖于后续的形成历史。最初天体击中目标时，撞击天体会减速、被撞击物质会加速。撞击、压缩、减压以及冲击波的外流，都在零点几秒之内发生。一旦冲击波过去，变化就发生得较为缓慢了。被击中的加速物质——由初始的激波加速，在激波消失之后仍然继续移动，其运动速度是亚音速的。即使如此，陨石坑继续形成，其边缘上升，更多的物质被喷射出来。然而，陨石坑尚未稳定，重力会使其崩塌。对于小陨石坑，边缘落下了一点，岩屑向下冲向陨石坑的四壁，而熔化的物质侵入陨石坑更深的部分。最终的结果

仍然是碗状的，看起来很像是初步形成时的样子，但是大小很可能小很多。例如，巴林杰陨石坑是其原始大小的一半。此后，角砾、熔化的和喷出的岩石填充了空洞。一个简单的陨石坑如图10-5所示。

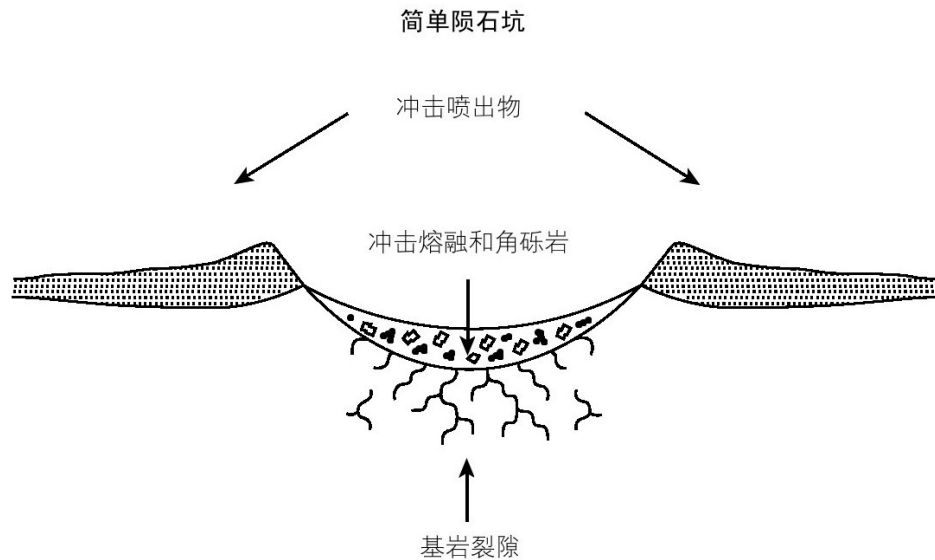


图10-5

由撞击形成的简单陨石坑有一个被掏空的碗状中心区域。此区域被相对比较平的角砾覆盖，并有着鲜明的上升边缘。

更大的冲击不仅会改变物质的位置并喷出物质，也会使被击中的原来地面的一部分蒸发。这些熔化的物质能够覆盖空腔的内部，而汽化的物质通常会扩展出去，造出一个蘑菇云的效果。大多数粗糙的物质将在几个撞击坑半径之内降下来，但一些更细小的颗粒物质可以消散到全球范围。

当撞击体的直径大于1 000米时，形成的撞击坑直径将达到20公里或更大。在这种情况下，撞击体实质上在大气中造出了一个洞，而喷出物填充这个空洞——先向上运动，之后下降到一片较广的区域上。最热的物质能够上升至平流层以上，而汽化物质的火球可以被广泛扩

散，就像由于白垩纪-第三纪撞击沉积，在世界各地发生的富铀黏土情况一样（我们后文将提到）。

更大的冲击形成了一个复杂的陨石坑（见图10-6），在最初的陨石坑形成之后，坑内经过了更大的变化。坑的中央区域上升，而边缘部分坍塌，因为冲击波在地里传播的过程中，不会与不均匀的岩石相互作用，生成一种新的和冲击波传播方向相反的波，并且将冲击波“卸掉”。这个变稀薄的波将深处的物质拉到较浅的地方，在大陨石坑下面留下变薄的地层。这一切发生的速度之快令人惊讶。几公里深的大坑能在数秒内生成，而峰顶可以在几分钟内上升到几公里的高度。

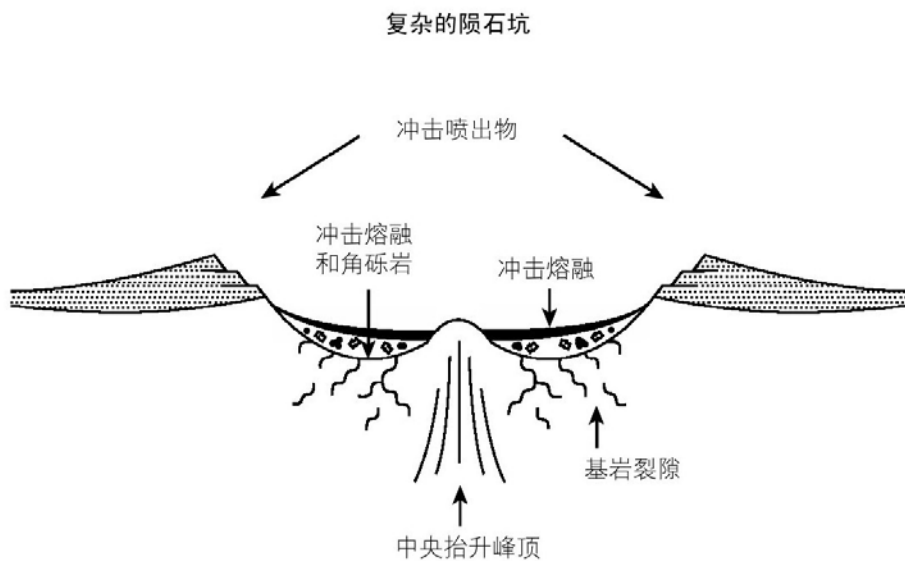


图10-6

一个复杂的陨石坑，它和简单陨石坑一样有一个升起的边缘，但是边缘具有阶梯结构。其内部有一个隆起的区域，并且有更大量的坍塌物质。

复杂陨石坑和由小型冲击形成的简单陨石坑有着不同的外观。精确的陨石坑形状取决于坑的大小。当陨石坑的层状沉积岩直径大于2 0

00米，或者更强的火成岩或变质结晶岩的直径大于4 000米时，它一般具有中央隆起区、广阔的平坦坑面和阶梯状的墙壁。这是在最初的压缩、挖掘、改变和坍塌之后留下的。

当直径超过12公里时，在陨石坑中心可能会升起一个完整的平台或者圆环状。所有这些线索对于“挖掘”（它有双重含义）过去都是关键信息。20世纪80年代，这些鲜明的特点有助于认证和白垩纪-第三纪灭绝有关的尤卡坦（Yucantan）陨石坑（详见第12章）。

地球上的“疤痕”

在过去的半个世纪，许多陨石坑已经被发现。通过研究它们的化学成分以及陨石坑疤痕——被破坏最厉害的陨石坑仍然留有可以辨别的痕迹，我们可以填补地球的“访客”名单，而“来宾留言本”就是地球撞击数据库。

地球撞击数据库当然包含了一些你可以在互联网上找到的引人入胜的列表，如果仔细查看，你会发现包含很多造访过地球的地外天体。它们撞击过地球，并且留下了巨大的陨石坑，从而让人类可以发现并辨认出它们。但这不是撞击事件的完整列表，因为地球上很多非常古老的陨石坑被地质活动抹平了，我们现在看到的大部分陨石坑来自相对近期的，也是频率较低的撞击。

大多数撞击很可能发生在39亿年以前太阳系的早期阶段，当时由行星形成剩余的物质被吸引至太空，四处移动。但是，地球、火星、金星和其他更多的地质活跃天体往往随着时间的推移丢失了陨石坑的遗迹，这就是为什么地质活动被动的月亮上陨石坑更加明显的原因。

即使是最近的撞击的遗迹也大部分丢失了。虽然较小的撞击经常发生，但它们并不会留下明显的疤痕，至少不能长期留下。事实上，小陨石坑甚至比你预期的会更加少见，因为地球有着稠密的大气层。像在金星和土卫六上一样，大气层保护着我们，让我们不用遭受许多小撞击，像经常发生在水星和月亮上的那样，那里的大气层保护不了它们。^[26]



较大的撞击很少发生，这对于地球生命的稳定性来说，是相当幸运的事情。每隔几十万到100万年，一次足够猛烈、足以产生一个直径达20公里宽的陨石坑的撞击可能会发生，并导致全球性的损害。然而，即使这样的概率也没有在地球撞击数据库中表现出来。如果你仔细查看，你只会发现43个这种陨石坑的遗迹；而在过去的5亿年里，只产生了34个；在过去2.5亿年内，又产生了26个（见表10-1）；总数只有大约200个。

表10-1 从地球撞击数据库获得的地球上已知的直径大于20公里、形成时间小于2.5亿年的陨石坑名单

陨石坑名	年龄 (百万年)	直径 (公里)
圣马丁 (Saint Martin)	220 ± 32	40
马尼夸根 (Manicouagan)	214 ± 1	85
罗什舒阿尔 (Rochechouart)	201 ± 2	23
欧宝龙 (Obolon')	169 ± 7	20
普切日 (Puchezh-Katunki)	167 ± 3	40
莫罗昆 (Morokweng)	145.0 ± 0.8	70
戈斯峭壁 (Gosses Bluff)	142.5 ± 0.8	22
雷神 (Mjølfnir)	142.0 ± 2.6	40
艾伯特王子 (Tunnunik [Prince Albert])	>130, <450	25
图库努卡 (Tookoonoka)	128 +/-5	55
卡斯威尔 (Carswell)	115 ± 10	39
斯蒂恩河 (Steen River)	91 ± 7	25
拉帕加维 (Lappajärvi)	76.20 ± 0.29	23
曼森 (Manson)	74.1 ± 0.1	35

续前表

陨石坑名	年龄 (百万年)	直径 (公里)
卡拉 (Kara)	70.3 ± 2.2	65
希克苏鲁伯 (Chicxulub)	65.17 ± 0.64	24
波泰士 (Boltysk)	66 +/-0.03	150
蒙塔格奈 (Montagnais)	50.50 ± 0.76	45
卡缅斯克 (Kamensk)	49.0 ± 0.2	25
洛甘察 (Logancha)	40 ± 20	20
霍顿 (Houghton)	39	23
米斯塔斯汀湖 (Mistastin)	36.4 ± 4	28
波皮盖 (Popigai)	35.7 ± 0.2	90
切萨皮克湾 (Chesapeake Bay)	35.3 ± 0.1	40
里斯 (Ries)	15.1 ± 0.1	24
卡拉库尔 (Kara-Kul)	<5	52

注：大小表示陨石坑自身从边沿到另一边沿的直径，它比受影响的冲击区域要小。

有几个因素导致了陨石坑记录的缺乏。第一个相关因素是，地球表面70%被海洋覆盖。这不仅因为海面下的陨石坑难于被发现，而且因为在一开始海水就能对陨石坑的形成进行干扰。此外，海洋底部的地质活动可能会消除所有实际上已经形成的疤痕，除了最近期形成的那些。海洋底部的证据每两亿年就会在很大程度上被消除，因为板块构造用一种传送带似的展开和俯冲过程改变了海底的结构，这一过程可以在这个时间尺度上覆盖任何之前存在陨石坑的证据。

即使在地面上，地质活动（例如风或水的侵蚀）也会破坏陨石坑存在的证据。这也是为什么大多数陨石坑在大陆更稳定的内部区域被发现（并且更有可能在地质活动较少的行星上保留下来，如金星）。当然，即使不像海平面以下4 000米的地方一样难以到达，流星体也可能在陆地上比较不易进入的地区着陆。最后，人类的活动可以通过改变地球的表面掩盖陨石坑遗迹。因此，从某些方面来说，陨石坑的名单有现在这么多已经很让人惊讶了。

一些个例大多因为是相对近期的事件（在地质时间尺度上来看），而相对有名。在过去100万年里，产生了两个直径10公里宽的陨石坑，一个在哈萨克斯坦、一个在加纳。另外两个知名陨石坑在南非和加拿大——弗里德堡（Vredefort）和萨德伯里（Sudbury）。这些陨石坑甚至比由造成了白垩纪-第三纪灭绝的撞击事件产生的希克苏鲁伯陨石坑更大，但是它成形的时间在更加遥远的过去，大约20亿年前。加拿大萨德伯里矿被创建起来是为了挖掘镍和铜，因为当产生陨石坑的巨大物体撞击并熔融壳层的时候，它们出现并沉积在了那里。在萨德伯里的撞击体并没有直接带来大部分金属，而是由于撞击熔化了一片如海洋般巨大的地壳，而被熔化的地壳花了很长时间才能结

晶。这留下了足够的时间使已经存在于地壳的少量镍和铜沉降到冲击熔体池的底部。然后这些金属再进一步由热冲击熔片产生的热液活动浓缩，生成了经济上可行的可开采的矿石。

萨德伯里矿在粒子物理学界相当有名，因为那里有一个地下实验室。这里虽然仍然是一个活跃的矿藏，同时也是一个活跃的物理实验场所。萨德伯里实验室的位置在地下2 000米的深处，以保护里面的探测器不受宇宙射线的影响，这使得萨德伯里实验室成为研究太阳发出的中微子的理想地点。例如，它在1999—2006年间进行了相关研究。对于搜索暗物质，这里也是一个很好的地方，这也是目前其内部安置的几项实验的目的。

不过，大多数撞击故事并不那么美好。很快我将会提到发生在墨西哥的希克苏鲁伯撞击事件，这一撞击造成了巨大的破坏力。正是这个流星体在6 600万年前引发了白垩纪-第三纪灭绝事件。不过在讲述这个令人难以置信的故事之前，让我们首先回顾一下过去5亿年间导致大规模生物灭绝事件的“大故事”。这些故事将告诉我们：地球上的生命既脆弱又稳定。



一些证据表明一些物种过去已经从地球上完全消失了。

达尔文的自然选择学说对生命如何进化给出了著名的解释。当有物种无法成功竞争并适应周围环境的变化，或者无法进入其他合适的栖息地以至于灭亡后，新的物种就会出现。尽管达尔文的进化论能够合理解释很多生物学现象，却不能解释我们所知道的生命的全部。这其中所缺少的最关键元素是生命的起源。

达尔文帮助我们了解到，一旦生命出现，一些生命形式是如何让位给其他形式的。尽管进化原理发挥着作用，达尔文的理论并没有解释生命最初是如何形成的。关于这一主题有许多相关文章和书籍，但是生命的起源依然是最难以解决的科学问题之一，无论这个问题是地球上生命的起源还是包含生命的宇宙的起源。关于生物进化后期阶段的理论与科学的实验给出的结果契合得很好。即使这些实验不都是实验室里的受控实验测试，至少也是能够通过研究化石记录或者丰富而古老的宇宙数据进行的检验。然而，最初状态几乎是无法实现的。倾向于理论研究的科学家们喜欢解释起源的问题，其实更合适的说法是猜测起源之前发生了什么。而一些实验导向的生物学家可能会试图复

制早期太阳系形成生命的基本过程。尽管这个领域里有些新的进展，生命的起源仍然非常难以确定，至少目前来看仍是如此。

我们在本章的重点是从另一个角度阐述生命的故事。和生命的起源一样，这个角度所反映的图景并不完全包含在达尔文有关自然选择的最初理论中。但这个图景具有和生物演化后期的理论一样的优点，即可被观测验证。生命故事中的这一重要组成部分涉及生命如何应对根本性的改变，其中包括大规模的物种灭绝，即许多物种大约在同一时期灭亡，没有留下直系的后代。

达尔文理论的核心部分是渐进（gradualism）的概念，达尔文认为，变化是在许多代的时间中缓慢积累而发生的。达尔文的理论不适用于根本性的变化，当然他也没有预想过由外星入侵引起的变化。达尔文脑海中的图像基于缓慢的演变，而环境的灾难可能是非常突然的。与达尔文最初的设想相比，当前的进化理论考虑了更加快速的变化。普林斯顿的生物学家彼得和罗斯玛丽·格兰特（Rosemary Grant），跟随达尔文的脚步对加拉帕戈斯群岛（Galapagos Islands）的雀类进行了研究，并得到了一个著名的发现：生活在加拉帕戈斯群岛上的雀类的喙对降水的改变适应得很快，适应的时间如此之短，以至于格兰特他们在连续的访问中都看到了相应的改变。但是，灾难通常发生得非常快，并且同时会带来突发的严重后果，这使得许多物种无法生存下来。



恐龙的确在不断适应着周围的环境，并且作为一个种群，它们已经生存了数百万年。在其他情况下，恐龙几乎肯定会存活更长的时间。但它们无法适应之前从来没有经历过的环境条件，我们将很快看到：这样的环境改变起源于一个来自外太空的物体。

现在关于生物进化的研究让我们认识到，相对于任何形式的环境变化，除了最缓慢的环境变化，适应几乎总是一个很慢的过程。“适应”似乎只在孤立的环境中产生具有独特属性的物种。对环境变化的首选反应经常是迁徙到一个有着更合适环境的新地方，当然只有当这种环境是可以到达的情况下才行。当一个物种无法适应或迁移到一个合适的栖息地，它就没什么希望了。在迅速变化的环境中，人们会很好地考虑到这一点。尽管技术进步了，但在评估今天不断变化的环境对地缘政治可能带来的影响时，这一教训可能是值得吸取的。

与宇宙的故事类似，我们对物种灭绝的故事感兴趣，是因为这个星球上的生命和地球、太阳，甚至可能和银河系的环境之间都有联系。人们很容易忘记我们的存在与许多让生命形成以及消亡的偶发事件息息相关。本章将讨论灭绝的概念、它的起因以及五次最大规模的灭绝事件。在这五次事件中，在几百万年的时间框架内，有1/2~3/4的物种灭绝了（古生物学家还没有统一的定义）。本章还将讨论很可能正在进行的第六次大规模物种灭绝。

物种灭绝从两个角度将我们的星球和气象事件联系起来：天气和太空。更好地了解其中的联系是具有挑战性的，但可能依然在我们的掌握之内。这门科学对于作为一个物种的人类来说很重要，即使这些故事展现的时间尺度比大多数人考虑的要长得多。

生存和死亡

在地球的历史中，简单生命出现得比较早。地球表面上最古老的岩石中包含着生命的证据，在距今大约35亿年前的化石里，即地球形成之后大约10亿年，这个时间点上，来自太空的小行星和彗星对地球进行的轰炸刚刚结束。有氧光合作用在大约10亿年后出现，同时出现

的还有最有可能引发物种灭绝的大气。这样的大气也导致了多细胞藻类的出现。大约5亿年以后才开始了“无聊的10亿年”，在这期间没有什么激进的演化——至少就我们所知是这样的。这段漫长而宁静的时光在寒武纪初的时候突然结束了，在大约5.4亿年前，复杂的生命爆发式地出现。

我们对从寒武纪的生物多样化时期到今天的这段时间内的生物进化过程的了解比较详细。这一时间段被称为显生宙（Phanerozoic eon）。化石记录包含的印记正好从这一时期初开始，这时许多硬壳类动物首先出现，并产生了一个确定的、持久的记录，大多数动物和植物的生命也随后出现了。化石所留存的地区非常广泛，包括加拿大落基山脉的伯吉斯页岩（Burgess Shale）、中国的三峡、东北西伯利亚以及纳米比亚。所有化石都包含着各类生命在地球的不同位置爆发性出现的证据，例如澳大利亚的早期伊迪卡拉（Ediacara）化石群、纳米比亚的那抹型（Nama-type）化石群、纽芬兰的阿维隆型（Avalon-type）化石群，以及一些来自俄罗斯西北部白海地区（White Sea）的化石群。后面的这些区域包含了一些已知的最早的复杂生命，它们来自紧接寒武纪的生命大爆发。

化石记录，除了告诉我们生命的分布，还提供了物种消失的信息，让我们得以洞悉不同形式的生命消失（没有留下任何后代）的时间。虽然大部分化石记录的灭绝是很久远的，但是人类对灭绝的了解却是相对较新的。19世纪初期法国的博物学家和贵族乔治·居维叶（Georges Cuvier）意识到：一些证据表明一些物种过去已经从地球上完全消失了。在居维叶之前，当人们发现远古动物的骨头时，他们总是试图将这些骨头与现有的物种联系起来，这当然是合理的第一猜测。毕竟，虽然猛犸象、乳齿象与大象是不同的物种，但区别并没有那么大，以至于你一开始很可能会混淆它们，或者会将它们的残骸联系起来。居维叶找出了其中的区别，他证明了乳齿象和猛犸象不是任

何当时存在的动物的祖先。接着，他找出了许多其他已经灭绝的物种。

尽管灭绝的想法现在已经深入人心，但是这种认为整个物种可能不可逆转地消失的想法最初遇到了很多阻力。灭绝的概念和当年主流的看法一定是非常难以调和的，至少不比今天人为造成的气候变化更容易。英国地质学家查尔斯·莱尔、达尔文和居维叶虽然都有助于推进这个想法被接受，但不一定是刻意为之，而且肯定是出于不同的角度。

居维叶和其他人的观点不同，他认为化石记录中的根本转变是全球范围的大灾难的后果。对其观点的强有力支持来自对岩石的观测，因为在化石种类急剧变化的时间点的岩石会展示出灾难性事件发生过的迹象。然而，居维叶也没有得到完整的图像。他过分执着于自己的观点，认为所有已灭绝物种的消亡都是由于灾难性事件造成的，并且他从不认为逐渐的变化也可能作出贡献。居维叶拒绝接受达尔文的进化理论，也不接受物种会在缓慢、持续的过程中灭绝。

凭心而论，即使是现在，人们在看到戏剧性变化的景观时也会感到困惑。这些景观并不总是反映塑造出它们的缓慢过程。在科罗拉多西南区的一个活动上，一位发言人在开车去会场的路上评论道，他想象的戏剧性隆起创造了路两侧那令人眼花缭乱的砂岩峭壁。我提醒他，这个过程跨越了数百万年，尽管时有时无，但这些过程并没有像他说的那么剧烈。

在居维叶提出猜想的时候，大部分科学人员则犯了相反的错误，即认为灾难性变化不会起到任何作用。如果几个世纪之前，灭绝的概念是一颗难以下咽的药丸，那么灾难性改变的想法则显得更不可思议。达尔文是理解渐进式改变的科学家之一，但忽视了对于居维叶而言最关键的观点。达尔文认为，任何与渐进论相矛盾的证据只是地质和化石记录不足的标志。虽然他接受演化，但他假设演化总是发生得

很慢，无法通过各种研究观察到。达尔文在思维上跟随颇有影响力的莱尔。莱尔在19世纪下半叶仍然主张，所有的改变都是平稳和渐进的，所以任何支持突变论的所谓证据都是不完整数据带来的错觉，要么是由地质记录的欠缺引起，要么是由风化造成。莱尔在一定程度上又是被身为苏格兰医生、化学品制造商、农学家和地质学家的詹姆斯·赫顿（James Hutton）启发的，后者认为地球只是通过微小的改变在进行变化，不过在很长一段时间积累后产生了显著的影响。

这些科学家的想法对很多过程来说的确是正确的，比如生物学和地质学。雨和风慢慢蚕食山脉，而山脉本身的形成会经过数百万年的时间，是由缓慢的板块运动所引发的逐渐抬升的结果。但我们现在知道，逐步和快速变化都在塑造着地球的形态，虽然即使最剧烈的变化从人的角度来看仍然是相对缓慢的。这也是这些变化如此难以理解的原因之一。

然而事后看来，我们可以回过头来说，戏剧性变化的证据应该是显而易见的。甚至早在19世纪40年代，科学家们已经发现化石记录中的巨大空白，这表明灾难性的事件确实发生过。研究沉积岩记录的古生物学家辨认出这样的事件，因为他们注意到许多化石种类突然在岩石层上的一个边界消失了，而越过这个边界之后，新物种的证据开始出现。这并不是说证据总是十分明确的，因为许多现象可能导致沉积停止，然后再重新开始。但随着对相应的灾难性事件的辨认，以及仔细地断定年代，以确定较早和较晚的岩石层沉积的相对时间，古生物学家可以解决很多困惑。随着时间的推移，快速变化的证据逐渐强大到无法反驳。

克服不确定性

证实假设或证伪预言，是一项异常艰苦的工作，试图重建历史事件的科学家们必须艰苦地工作。即使有丰富的化石记录，但时间或空间在分辨率上的不确定性会导致非常不同的假设和结论。为了了解一些持续进行的科学辩论的原因，也为了欣赏已经克服了这些障碍的地质学家和古生物学家的智慧和有条不紊，让我们简单地了解其中的一些难题：可靠地断定灭绝是如何迅速且广泛地发生的，以及如何确定其中的根本原因。

第一个障碍是评估物种灭绝速率的困难性。在任何给定时间对地球上存在的物种进行准确地计数是很困难的，因为科学家需要找到、识别并区分每一个存在的哺乳动物、爬行动物、鱼类、昆虫和植物的类型。这同样适用于对今天存在的物种进行计数的工作，虽然原则上说这是最容易得到的。在《生命的未来》（*The Future of Life*）中，生物学家爱德华·威尔逊（E. O. Wilson）感叹，每年都出现太多关于新物种出现的发现，博物学家难以把它们都写成论文。

已被编目的现存物种的数目介于100万~200万之间。对现存物种总数最好的估计在800万~1 000万，但比这高达5倍的估计也是可能存在的。毫无疑问，确定过去的灭绝速度比确定目前现有物种的数量更具挑战性，因为前者不仅要考虑时间上的间隔，辨识多年以前的生命，还要考虑地质事件及其影响。毕竟，和现在相比，过去物种的数量以及它们消失的速率都更难测量。

在辨认大规模物种灭绝时，一个令人困惑的技术细节是：相关数据可以根据具体的定义有所不同。我将主要讨论物种的数目，而科学家们经常倾向于计算属的数目。通过后者，他们可能找到比较有用的分组。相关的生物种类对研究进化和灭绝都非常重要，而我对其的理解能力很大程度上得力于很久以前为高中考试所做的准备。那时我仅是靠死记硬背记住了“界-门-纲-目-科-属-种”。尽管很少再用到这

方面的知识，但我永远不会忘记这些名词。这些对你而言也许并不熟悉的等级，代表着生命的特定形式是如何密切相关的。

在评估是否发生了大规模灭绝时，采取不同的分类方式会得到不同的结果。例如，考虑这样一种情况：每个属中一半以上的物种都被消灭了。但在一个给定的属中只要有一个物种存留，这个属就算幸存下来。这种情况下，按照物种计数的规定，由于超过一半的物种都被淘汰了，即灭绝事件发生了；若按照属计数的规定，灭绝并没有发生，因为属的数目没有改变。这个例子说明了物种灭绝定义的模糊性，这与用一个随机的比例来精确划分大灭绝一样模糊，例如有人说50%，而有人说75%。这并不是说生物大灭绝可以被忽略，只是还没有理想的方式来定义它们。

除了术语的问题，古生物学家的的工作还会受到实质性问题的干扰。显然，识别和理解被打乱的化石记录是必不可少的工作。如果某些种或属在相邻的岩层中留下化石，但在其上面的岩层中没有出现，这可能是灭绝事件的信号。但化石仅能在沉积岩中被发现。住在火山或其他非沉积岩环境中的稀有物种通常不会留下任何痕迹。研究寒武纪（约5.4亿年前）之前的古老生命形式时的一大障碍是没有坚硬的身体成分，这使得辨认更早期的化石沉积层的工作非常具有挑战性。

更近期的记录也很复杂。即使化石的确形成了，但沉积和侵蚀速率对于理解化石的含义至关重要，而这些速率差异很大，对化石的解释可能被混淆。在陆地上，沉积是偶然发生的，侵蚀作用却是稳定存在的；在海洋环境中，沉积是稳定发生的，而侵蚀则是偶然发生的。这使得海洋中的记录比陆地上的更全面、更完整。这些因素意味着，化石记录只有部分幸存下来，然而这些记录即使存在，也可能很难被发现和辨别。古生物学家依然成功地获得了一些化石记录，因为虽然找到任何单个化石的概率很低，但鉴于在一段足够长的时期内有相当

多物种的个体化石被保存下来，所以沉积记录中依然有着丰富的化石信息。

这样的化石可能整齐地保存下一个完整个体的印记，但更多的时候，它们都只是局部的记录，作为证据，这些记录被嵌入岩石中，很容易被掩盖掉。因为通常只有一个物种的坚硬部分会形成化石，有辨识度的身体部位经常会缺失，这导致不同的物种容易被混淆。即使我们精通辨识化石，但可能在它们被发现之前，地球上的风化作用以及其他过程已经把许多相关的印记给隐匿或者毁损了。

除此之外，模糊效应（即西格诺尔-利普斯效应）会混淆对化石的解释。这种现象因菲尔·西格诺尔（Phil Signor）和杰尔·利普斯（Jere Lipps）而得名，它和一个颇为直观的想法联系在一起。模糊效应认为：一个物种最后的化石将位于不同的地方以及不同的地质时期，这使灭绝往往看起来不那么突兀，比它的实际情况更加渐进。根据西格诺尔和利普斯的理论，依据空间上延展区域出现的最后残余化石的深度的变化，并不能果断地确定灭绝是否是渐进式的或突然出现的。这种模糊性会使一个给定的灭绝事件的诱发原因难以被查明。

研究人员往往更喜欢海洋生物的化石，因为它们通常保存得较好。在19世纪，蛤、菊石、珊瑚以及其他大型物种的化石是最容易获得的。而在20世纪，利用更先进的工具，地质学家开始利用微化石来获得更详细的信息，例如，单细胞有孔虫这种非常丰富且分布广泛的生物的化石，它们被保存在水下和隆起的石灰岩中。

确定灭绝事件时的另一项考虑是：化石记录和绝对年龄都相当重要。化石记录与它们被发现处的地质构造相结合，可以帮助估算出相对年龄。由于不同的时间段居住着不同的物种，出现的化石种类能帮助我们确定它们所形成的相对时间。但往往找到一个边界岩石层的绝对年龄（而不只是相对年龄）是非常困难的，这还需要和化石记录相

独立的确定形成年代的方法。以此为目的地质学家们经常使用的一种方法是同位素分析（isotopic analysis）。通过分析同位素含量，科学家能够确定一个原子的不同同位素（在其中，质子数是相同的，但中子数不同）的比例。如果知道一种同位素衰变到另一种同位素需要多长时间，并且还知道初始情况，你就可以通过一种类型的原子所保留的多少来确定一个物体的年龄。

“碳定年”（carbon dating）也许是这种方法中最有名的例子。它被用来确定古老的有机材料的年龄，并且非常精确。然而，由于碳同位素的半衰期不是很长，所以这种方法只对年龄小于5万年的物体有效。这使得这种方法不足以用来测定大多数显生宙的古老岩石。取而代之的是更长半衰期的同位素。

不过，当应用上述方法来测定古老岩石的年龄时，同位素分析会变得更加困难。通常说来，只对现存的相关同位素的含量进行追踪，并以此来决定岩石的年龄并非总是足够精确。例如，钾衰变为氩是一个重要的定年龄的过程。但是在岩石中的氩气会逃逸到大气中，这使被测岩石看上去比实际更年轻。或者，有些氩气可以在岩石形成的时候被捕获并储存下来，导致岩石里有更多的氩含量，从而表现为这个岩石是在更古老的时候形成的。在过去的几十年中，研究方法不断得到改进，各种元素的交叉相关使得进行更好的研究成为可能，甚至微量元素的详细探查也变得更容易了。近期，利用激光从氩晶体中去除气体来确定流星体和白垩纪-古近纪灭绝事件的年龄的工作确实提供了一个惊人的准确例子。我将在后面的章节中提到这个故事。

磁信息也被用来帮助证实绝对年龄。这种方法依赖于地磁的倒转，最初被用于测量与恐龙灭绝相关的岩石的年龄。但由于地壳是由移动中的板块构成的，磁场的方向随着时间会发生变化，这使得磁场最初的方向难以被重建，影响了结果的可靠性。也许这是个好事儿，因

为它的不足加速了另一种方法的出现，这个新方法是由地质学家沃尔特·阿尔瓦雷斯（Walter Alvarez）和他的父亲、物理学家路易斯·阿尔瓦雷斯（Luis Alvarez）建立起来的，并引出了流星体的假设（我将很快对这种方法进行解释）。

为什么大多数生命消失了

地质学家和古生物学家的辛勤工作毫无疑问地揭示了：过去发生了惊人的变化，导致这个星球上的大多数生命都被消灭了。一旦这个论点被建立起来，问题就变成了：这些变化发生的方式和原因是什么？近年来，我们已经经历过一些破坏性的风暴和灾害，但没有任何事件本身会强大到足以消灭地球上的一半的物种。当然，对人类影响的累积效应的最终结果尚未确定。但是，是什么促成了过去那些改变了世界的灾难呢？

在介绍那些可以触发灭绝事件的灾难性事件之前，让我们首先考虑一下其他因素，它记载着可能发挥作用的环境因素。温度和降水的变化在任何情况下，都是两个重要的贡献者。从广义上讲，当天气模式变化时，那些已经适应当地环境的物种不一定能跟得上这个变化。

随着北极冰的融化，适用于特殊物种的环境能够随着变化的温度剧烈变化，以至于那些不能以足够快的速度适应这种显著变化的物种不得不转移到另一个合适的栖息地，或者死亡。气候变化的影响不那么直接，当然，其中最显著的是海平面的变化，它可以破坏稳定的海洋环境并淹没曾经适宜居住的陆地，把陆地环境转变为海洋环境，从而消灭了一些陆生物种。

海洋的变暖也会影响降水模式，再次影响物种的生存机会。在较短的时间尺度上，寄生虫或疾病也可能导致物种的灭绝，因为气候变

化会加剧它们的危险。此外，一个物种所依赖的食物可能会死光，引发食物链的多米诺骨牌效应。

在海洋中，与氧气耗尽一样，酸度的变化是更进一步的潜在杀机。最终，屏障的形成可导致分离的、脆弱的种群，或者屏障的移除能够造成物种的入侵或不同种群的均质化，这两者都能使一个物种灭亡。任何灭绝触发器都会导致至少发一次我刚才所描述的灾难，而大多情况下会导致几个灾难组合性地发生。

为什么会发生这些变化？是由什么环境变化引发的？关于这个问题有两种不同的主流观点。一种观点认为，灾变是渐变的一一这种观点常和地球相关的现象联系起来，例如火山和板块运动。火山喷发的烟尘能够遮挡阳光，显著地改变大气成分从而影响温度。但它引发的结果需要很长时间才会导致生物的灭绝。而板块运动会影响栖息环境，是物种逐渐消亡的另一个论点。随着海洋环境的变化，板块运动能够改变气候和陆地范围，这两点都会导致地球上生物的戏剧性变化。当然火山爆发或者板块运动中的某一个和生物灭绝是相关的，但很可能两者都相关，因为它们趋向于同时发生。

此外还有些“大事件”。另一方的观点认为，突发事件是造成物种灭绝的主因，包含外来天体引发的灾难，例如大彗星的撞击，也包含地球上突然发生的事件。地球自身引发灾难的依据来自那些已知的、可能是突发的能量加速释放。例如，我们知道火山喷发有不同的间歇，但是在西伯利亚和南印度的德干高原，却有大面积的玄武岩层，被称为暗色岩。暗色岩包含的岩层占了高原很大一部分，是由大量火山喷发出来的岩浆扩散形成的。这两块暗色岩是火山喷发率异常的标志。尽管经受了风化的侵蚀，但就算今天，西伯利亚的暗色岩面积依然超过一百万平方公里，体积则达到几十万立方公里。

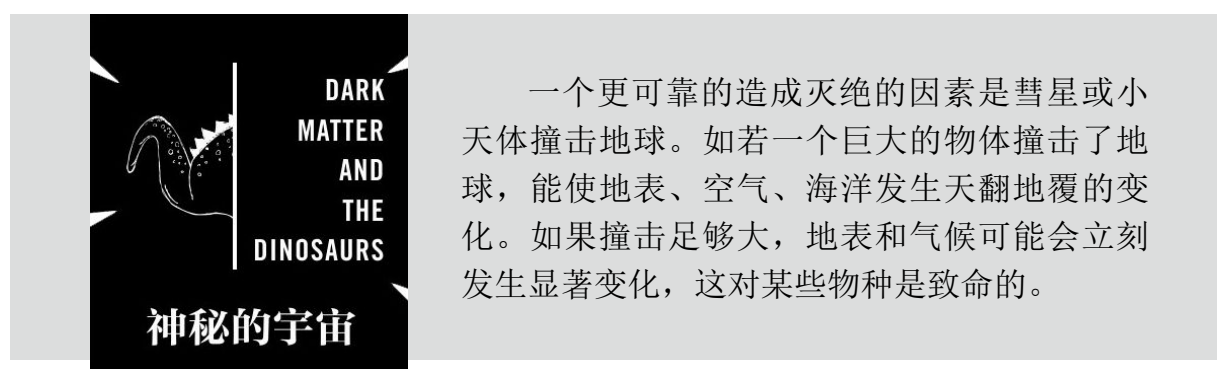
那种能形成暗色岩的火山密集地喷发时，会造成严重的破坏。你大概还记得火山灰浓密到影响飞机飞行的新闻，也就是2010年4月冰岛的埃亚菲亚德拉火山喷发。更猛烈的火山活动能造成更可观的全球性影响，例如对全球气候的影响。喷发物中含有大量二氧化硫，这会增加大气上层的水蒸气含量，加剧温室效应从而造成短时期的全球变暖。但长远看来，这些火山会使全球变冷。这是因为二氧化硫会和水形成硫酸并凝聚形成硫酸盐气溶胶，而把太阳光反射回空气中，从而使大气底层变冷。（这一过程很有效，科学家们甚至在研究向大气中注入硫来应对气候变化。）硫酸盐气溶胶还能破坏大气中的臭氧，还会形成酸雨。该过程的进一步反馈机制（包括已知和未知的）可能产生更持久的气候现象。

仅仅有火山还不够解释所有的物种灭绝。足以毁灭地球主要生物的事件是非常罕见的。更奇异的想法是，宇宙事件引发了快速发生的大灾难。地轴和轨道的变化可能引发了一些气候变化，例如发生在几万年或几十万年的时间尺度上的冰河期，但这些地球活动似乎不能解释低频率的大型灭绝事件。

宇宙射线、超新星和其他太空现象，也被认为是长时标现象的“嫌疑犯”。宇宙射线能从几个方面影响云层覆盖。一是电离对流层中的原子，使水滴能够凝结。这可以加快云的形成，从而影响天气。然而这个理论并不站得住脚。首先，我们不知道宇宙射线和其他电离源比有多重要。其次，原子核（即使形成后）必须要凝结到足够大才能形成云。最后，云的作用还不清楚：它们可能通过反射太阳光来使地球降温；而另一方面，它们还可能再辐射一些能量以使地球升温。无论如何，宇宙射线和气候的关系不足以解释造成灭绝的短时间剧烈天气变化。

超新星也被认为是潜在的因素。提出者认为，超新星会释放高能的X射线和宇宙射线。这些辐射原则上可以破坏细胞和遗传物质，从而杀死生物。辐射还可以耗尽臭氧层，导致二氧化氮形成，这会吸收太阳光而导致全球降温。

尽管有这些潜在的威胁，超新星似乎也不能解释物种灭绝，原因你肯定想得到：太阳系附近的超新星数量不够多。即使当地球穿过银河系旋臂的时候，因为恒星数密度增大了，所以遇到超新星的概率会增加，但是超新星临近地球的可能性还是不够解释灭绝事件。同样，伽玛射线暴（gamma ray bursts）的数量也是不够的。根据有些估计，银河系中的伽马射线暴每10亿年左右才发生一次。



事实上，大多数灾难电影的情节（除了生化危机那一类电影）都是伴随着一个大型撞击而来的。撞击会产生冲击波、火灾、地震和海啸。灰尘还会飞入大气，短时间内会阻止光合作用，终结大多数动物的食物来源。撞击还会导致气候的改变：开始变热，然后变冷，再然后又变热。变冷是由于大气中残留的硫酸盐和尘埃；而后期的变热可能是因为有毒气体和吸热气体会引发全球变暖。有一颗彗星肯定是引起了一次大灭绝事件，我在下一节会详细讲述，这次灾难是显生宙时期的五大灭绝事件之一。

五次物种大灭绝事件

1982年，芝加哥大学的古生物学家杰克·塞科斯基（Jack Sepkoski）和大卫·劳普（David M. Raup）用他们对这一领域的所有既有数据的开拓性分析，给古生物学带来了一场革命。许多观测中的不足使他们基于数值计算的数据导向类研究变得很困难，他们需要作出许多关于如何取舍数据和如何处理数据的决定。然而他们意识到：只要有足够的数据点可用，统计类方法对不完美或者不完整的数据而言依然有效，而且事实也确实如此。尽管劳普和塞科斯基在1982年发表的论文不是最早量化分析化石记录的文章，但是这项工作改变了研究物种灭绝的大方向，因为在他们之前的文章都是依赖于对小范围数据的研究。

在他们的工作中，芝加哥的古生物学家们辨认出了5次大型的生物灭绝事件（见图11-1），还有大概20次小一些的，即那些只有大概20%的生物灭绝的事件。因为进化动力学的显著区别以及更早期的可信证据的匮乏，劳普和塞科斯基集中研究了过去5.4亿年的生命及其毁灭。在寒武纪大爆炸之前，生命形式肯定已经出现过和灭亡过了。但不清晰的化石记录使得人们对更早期的物种计数无从下手。

塞科斯基和劳普辨认出的最古老的灭绝事件是奥陶纪-志留纪物种大灭绝，应该发生在4.5亿~4亿年前之间。本质上来说，那时生物都生活在海洋中，所以所有灭绝的生物都是水生生物。这次灭绝是在350万年里分两个阶段发生的，在这次第二大的灭绝事件中，85%的生物都消亡了。开始的起因似乎是温度降低，大量冰川形成，海平面显著下降。这是因为水都结成了冰——相反如果大量冰川融化，那么海平面就会上升。第二波灭绝的高峰大概是在后来的一段温暖时期里，那些已经适应了寒冷的动物被消灭了。适应温暖的动物，例如热带浮游生物、潜水海百合（海星和海胆的祖先）、三叶虫、甲壳鱼类和珊瑚最先灭亡，然后是已经适应了寒冷的珊瑚、三叶虫和腕足类走向灭亡。

下一个大灭绝事件持续的时间长一些——大约2 000万年，开始于3.8亿年前的泥盆纪晚期，即泥盆纪到石炭纪的过渡期。似乎有可能存在3~7次的灭绝高峰（具体数字不确定），每一个高峰持续几百万年。这次灭绝也严重地打击了海洋生物，可观数量的海洋物种因此灭亡。尽管死亡也蔓延到了陆地上，陆地上的昆虫、植物和早期的雏形两栖动物存活了下来。古生物学家认为这次事件的一个显著的特征是：这次灭绝主要是因为物种的低形成速率无法平衡物种的消失速率。而其实，这期间的物种消失速率并不比通常快多少。

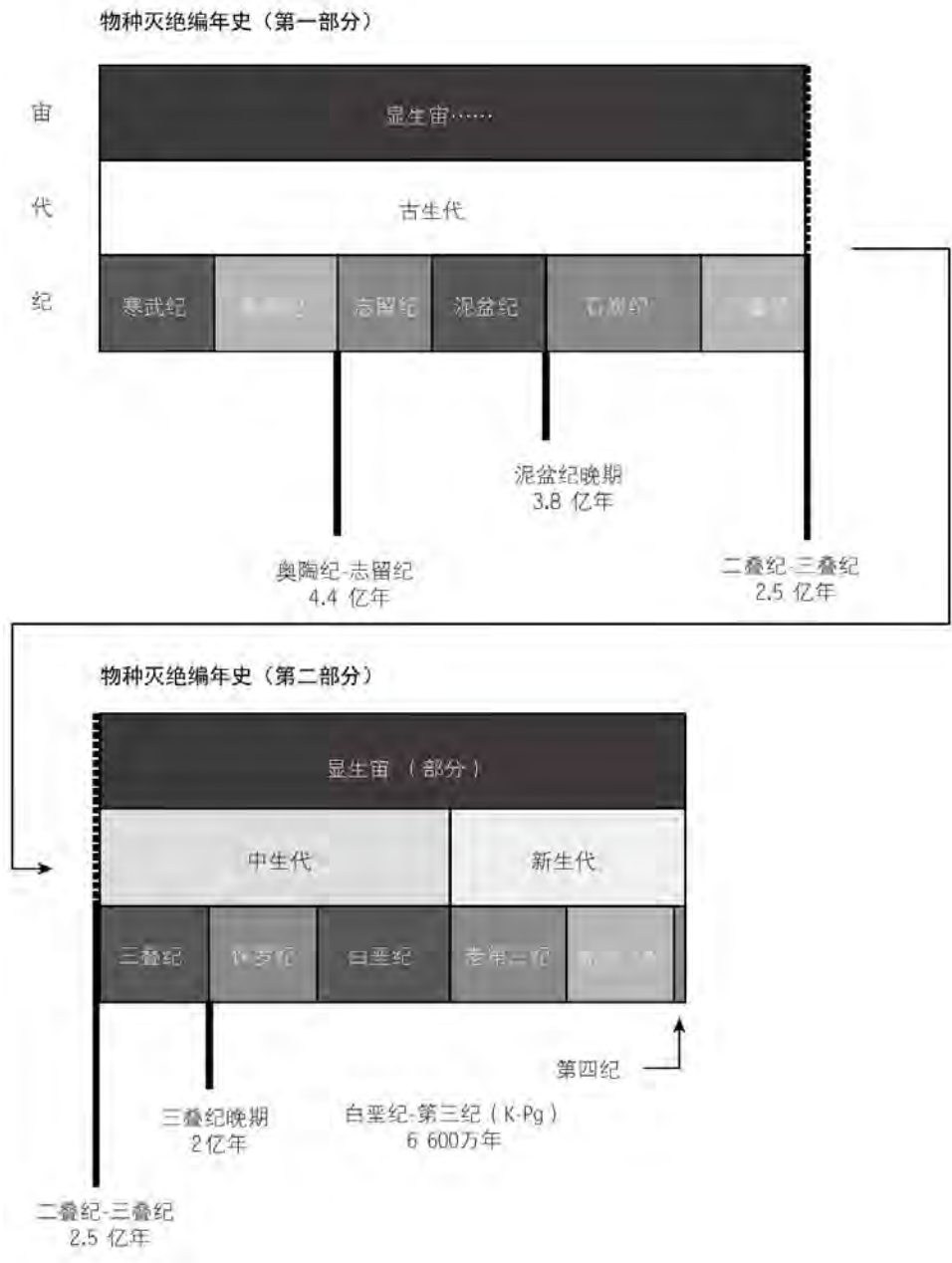


图11-1

五次大灭绝的界限：奥陶纪-志留纪大概是4.4亿年前；泥盆纪后期，大约是3.8亿年前；二叠纪-三叠纪，大约是2.5亿年前；三叠纪结束，2亿年前；白垩纪-第三纪，6 600万年前。此外还有显生宙的各个时期。

从地球上消失的物种比例来看，2.5亿年前的二叠纪-三叠纪灭绝事件是已知的破坏性最大的一次。泥盆纪物种灭绝事件之后，生物（包括两栖类和爬行类）在陆地上和海洋中繁荣了很长一段时间。但是生物在这一次灭绝中走向衰落，至少90%（甚至更多）陆地和海洋物种灭亡了。灭亡的物种包括表层浮游生物以及海底生物（例如苔藓虫类、珊瑚），一些贝类，还有三叶虫——那些从前两次大灭绝中都存活下来的三叶虫。陆地上，连昆虫都灭绝了——这是仅有的一次让它们也吃了很大苦头的大灭绝事件。此外，很大一部分两栖动物消失了，而爬行动物——从上一次大灭绝后才出现的物种，也损失了大部分成员。

这次灭绝的原因还存在争议，但大范围的气候变化，以及大气和海洋化学成分的变化肯定起了很大作用。尽管起因和机制还不清楚，但温度升高了大约8℃，这似乎和西伯利亚大量火山喷发有一定的关系，因为西伯利亚的暗色岩释放出了大量二氧化碳和甲烷。二叠纪-三叠纪灭绝这个已知的最大一次灭绝事件，肯定部分受到了火山喷发的气体影响：加热了地球，蒸发了海水，减少了氧气，污染了大气。直到今天，在大量风化侵蚀后，西伯利亚暗色岩还有100万平方公里的覆盖面积和几十万立方公里的体积。在当时，暗色岩的覆盖面积估计相当于今天俄罗斯的面积大小。

尽管生物几乎消亡殆尽，但“我之毒药，你之美味”。蕨类和蘑菇取代了早期的生物群落，新植物最终出现了。在这个时期以后，地球不再由类似哺乳动物的爬行动物完全主导，但现代哺乳动物是从它们进化而来的。祖龙的出现是另一个显著的结果，并最终成就了恐龙的霸主地位。

前不久，一个朋友非常自豪地给我看了一个保存极好的（而且也很可爱）约2米长的化石，她告诉我这是3亿年前的恐龙化石。如果她早一年给我看，我只会简单地赞叹其细节。但是根据我当时的研究，我知道

她所说的不可能是对的，因为恐龙是在三叠纪才出现的，比她所说的事情晚了2.5亿年。如果那块化石肯定这是恐龙化石，那么我认为那块化石可能没有那么古老。但实际上我们都搞错了：那块化石的确是3亿年前的，但不是恐龙的——它是中龙属化石，一个已经灭绝了的爬行动物。恐龙化石很古老，但仍然没有朋友完美保存的那块化石古老。

由于这一次大灭绝的惨烈程度非常之高，地球上的生物没能很快恢复过来。我们是通过黑页岩知道的：在沉积岩灭绝的界限上足足好几米厚，代表着生命产生的白色石灰岩消失了很长一段时间。无论怎样，至少500万年后，新的软体动物、鱼类、昆虫、植物、两栖动物、爬行动物、早期哺乳动物以及恐龙出现了。但是这次生物繁荣持续了四五千万年后被第四次大灭绝打断了，也就是大约2亿年前。

在三叠纪末侏罗纪前的灭绝中，大约75%的物种走向了死亡。起因不是很明确，但较低的海平面和火山喷发并最终诞生了大西洋的地质活动，可能起到了关键作用。海洋中的大多数大型脊椎食虫动物死亡了，海绵动物、珊瑚、腕足类、鹦鹉螺类和菊石类也遭受到了严重的打击。这次灭绝也导致了大多数类哺乳动物的灭亡，也包括许多大型的两栖类和非恐龙的祖龙。

陆地上，竞争对手的消失使得恐龙走向巅峰。灭绝事件使一些生物走向末路，但也重组了生物进化的条件。之后的侏罗纪时期因为书和电影出了名，尽管电影《侏罗纪公园》中所展示的生物并不都生活在那个时代。但是侏罗纪的确是恐龙走向繁荣的时期——在侏罗纪晚期，恐龙已经在陆地生态系统中占了主导地位。会飞的爬行动物、鳄鱼、海龟和蜥蜴也大量繁殖；哺乳动物也在进化，尽管它们的命运是等待时间线上的下一次大灭绝。

最近的一次大灭绝估计也是最有名的一次。它发生在白垩纪和第三纪的分界线上。这次事件最早被称为白垩纪-第三纪物种灭绝事件（K-T），但现在官方将其改名为白垩纪-古近纪物种灭绝事件（K-P

g)，它发生在6 600万年前。这也是我们熟知的灭绝了恐龙的那次事件。

恐龙并不是唯一灭亡的物种。那个时期里大约75%的物种和50%的种属消失了，这包括许多爬行动物、哺乳动物、植物和海洋生物。在沉积记录中很常见的微生物海洋化石尤其重要，因为数目庞大的此类化石对当时所发生的事情作出了详细的记载。每厘米的海洋沉积物可以映射出一万年的活动，这给出了发生在海洋里的事件的细致图像。国际海洋钻探计划（ODP）用好10倍的精确度检测了一些核心海域的海底沉积物。精确的海洋微生物化石尺度帮助科学家们确定了浮游生物、珊瑚、多骨鱼、菊石、大部分海龟和许多鳄鱼种类的灭亡。

这次大灭绝发生后，哺乳动物成为地球上的主要角色。这由多方面的因素造成，首先这肯定与陆地恐龙的消失相关。如果霸占了主要资源的恐龙没有消失，大型哺乳动物（像人类）可能永远没有出头之日。有一个推测：在希克苏鲁伯陨石撞击事件之前恐龙比哺乳动物更占优势是因为它们会大量产蛋；而哺乳动物的后代较少，而且体型越大的哺乳动物生产的频率越低。恐龙可能只是因为数量多，所以在与其他大型动物的竞争中取得了优势。

因为五次大灭绝中的最后一次不仅是最近发生的一次灭绝事件，同时也是大型哺乳动物登上生物历史舞台的主因，所以科学家们对它的研究最仔细。寻找一个正确的理论来解释全球陆地海洋生物的消失是一个有趣的故事，我们把它放在下一章。几乎可以肯定的是，有一颗巨大的彗星在6 600万年前撞到了地球上。虽然那是很久以前的事了，但是和地球漫长的40亿年历史相比，它只不过是时间长河的一小段。我发现一个引人注目的事实：这个来自太空的给地球带来了严重后果的撞击事件，对地球的影响要比我们想象的还要持久。

第六次物种大灭绝？

灾难也许就近在眼前。如果我就这样结束这一章，不说一说这最后一点非常令人不安的想法，那真是我的责任了。

许多科学家今天认为，我们当下正在经历第六次大灭绝，这次物种灭绝完全是人类一手造成的。为了明确地建立这个论点，科学家们需要确定现存的物种数量和它们消失的速率——但这两点都很困难。尽管还没有定论，但我们已经确定的数字显示了可怕的趋势。证据显示了一个明显高于正常的物种消失速率，这是把当前的物种消失速率和之前的物种消失速率比较后得到的。根据估算的物种消亡速率基线，平均每一年就会有一个物种灭绝。这个基线的估计也许不是很精确，但事实上，如今，物种的灭绝速率大概是这个平均值的几百倍。

如果只看鸟类、两栖类和哺乳类的测量结果，更是令人如坐针毡。哺乳动物只占了全部物种的很小一部分，但是它们却是被研究得最细致的一类。在过去500年里，物种总数只有不到6 000种的哺乳动物中有80个物种消失了。

过去的500年里，哺乳动物灭绝的速率大概是正常值的16倍，在20世纪，这个值被提高了32倍。20世纪，两栖动物灭亡的速率比过去高了100倍，此外，当前有41%的两栖动物也面临着绝种的威胁。而在同一时间段内，鸟类灭绝的速率也比平均值高出了大约20倍。

这些数值和一次灭绝事件的结果相吻合。根据加州大学伯克利分校的生物学家安东尼·巴诺斯基（Anthony Barnosky）等人观察到的结果，现在发生的环境变化也是如此，和二叠纪-三叠纪灭绝发生时的情况可怕地相似。那时二氧化碳的含量升高——温度也是，海水呈现

酸性，海洋环境中无氧的死亡区域增多。不可思议的是，那个时期温度和pH值（测量酸碱度的值）的比率与今天的值相当。

我们几乎可以肯定地把物种多样性的消失归咎于人类的影响，人类从许多方面改变了地球和其生命形式。例如当欧洲人到达北美洲后，80%的大型动物走向了死亡——很大一部分是被直接屠杀的。人类还以其他方式破坏着生态环境。其中一项罪责便是污染，还有地貌的改变，包括滥伐森林和过度捕捞；另一个因素是气候的改变，这是由温度和海平面的变化造成的。干旱、火灾、洪水和暴风，以及气候变暖和酸性海洋都与物种的存活息息相关。人类对栖息地的破坏在某种程度上促进了物种的入侵，从全球角度看，使物种群体更均匀地分布了，却这使得所有疾病或寄生虫都变得更加危险。在条件允许时，物种会迁徙到新的栖息地，但是如果栖息地被破坏，那些隐藏的居民们也会跟着遭殃。由于所有这些破坏性的影响，地球上的生物正面临着一个迫在眉睫的危机的说法绝对不是危言耸听。

巴诺斯基给出了一个有趣的理论，惊人的人口增长导致了当前的物种群体危机，因为人类对能源的消耗直接造成了物种危机。假设资源的分布是公平的，假设大型哺乳动物的大小和分布的范围也是合理的，每天从太阳传播过来的能量能够支撑的动物和物种数量是有限的。50 000~10 000年前，当人类出现在这颗星球上时就开始大量地霸占这个星球上的资源，这使得大型动物的物种数量从大约350种降到了175种。后来，哺乳动物的数量缓慢地恢复到其之前的水平，但是在大约300年前又开始极速下降——基本上是从工业革命促使人类开采地球储备的能源开始，那些储存了几百万年的化石燃料就是没有使用的能源。在这些储备能量的帮助下，尽管物种数量在不断减少，但人类和大型家畜的数量却随着城市化进程的发展呈现出爆炸式的增长。

一些乐观主义者尽管承认存在这个不安的趋势，但却认为：通过设计或复制DNA，我们也许可以创造或复兴一些物种来补偿那些已经消失的物种，从而避免一次大灭绝（这里的灭绝是根据物种或属的消失比例来定义的）。但恢复那些消失物种的真实面貌是非常具有挑战性的，因为它们的DNA并没有被很好地保存下来，而且重建过去那些物种的生存环境也不太可能了。此外，我们制造新物种并使其存活下来的速率似乎不太可能赶上目前全球损失物种的步伐。无论怎样，灭绝只是一个词。评估灭绝是否发生了也只是基于一个数字，这个数字不可能代表灭绝所产生的巨大变化，而事态正朝着这个方向发展（我们不得不承认）。

从技术上讲，另一个避免灭绝的方式是在物种数量消耗到一半之前，逆转这个趋势。例如，当物种数量降低到一定程度时，也许那些在多样的生物环境中无法竞争存活下来的物种反而会存活下来。这种“乐观”的情况其实只是一种推测，而且这个方式其实仅仅是挽回了生物种类的大量损失，而最后还要有一个稳定的环境才行。

这种变化也许最终对将来的物种是有益的。毕竟，即使是二叠纪-三叠纪大灭绝也留下了一些完好无损的生物。例如，从恐龙的观点来看，这还是好事一桩呢。然而，这并不能消除物种灭绝所造成的生命的损失，而且在生物恢复期间（同样遭受着痛苦）也会损失很多物种。因为即使在这一时期，生命依然遭受着匮乏和混乱所造成的痛苦。尽管从全局来看，我们当前引发的变化所造成的后果可能最终是有益的，但对地球上的某些已经进化了并适应了其环境条件的物种来说，却并非如此。

即使有新物种出现，或者最终条件有所改善，一个彻底改变了的世界对我们人类这个物种来说似乎不见得是什么好事。让人类对生物多样性的丧失负责，也许是一种误会，因为这也确实伤害到了我们自

己，例如这一过程会使我们失去食品和药品，失去清洁的空气和水。生命的进化有着非常微妙的平衡机制。我们并不清楚其中有多少可以被改变，而且改变之后又不会导致这颗星球上的生态系统和生命产生戏剧化的变更。你可能会想：我们会相当自私地考虑人类自己的命运，特别是当这许多损失几乎是不可避免的时候。不像6 600万年前那些生命的命运被一颗脱轨的小行星或彗星所左右，人类今天应该有预测即将发生的事情。

A stylized graphic featuring the number '12' in a large, bold, sans-serif font. The number '2' is partially obscured by a silhouette of a dinosaur, possibly a T-Rex, which is depicted in profile. Above the dinosaur, a trail of small circles and dots suggests a meteorite streaking across the sky.

地球霸主恐龙的末日

DARK MATTER AND THE DINOSAURS

一颗直径为10公里~15公里的流星具有巨大的破坏性——无论是对环境还是生命，都是如此。

每个人都喜欢恐龙。无论它们以骨骼、化石，甚至以塑料模型的形式出现，不管老幼都为之着迷。孩子们喜欢这些来自过去的生物，他们用积木拼出恐龙的形状，并能够记住绝大多数成年人都不会拼读的名字。任何展出恐龙的博物馆都会出现众多的观众，包括小孩子以及家长。自然博物馆的馆长们深知这些奇异的古代爬行动物的吸引力。纽约的美国自然历史博物馆将其最主要的展品定为霸王龙（意为“蜥蜴之王”）和雷龙的巨大骨架，并且在入口处布置了相关的模型来迎接观众的到来。

恐龙受欢迎程度的进一步证据是，恐龙在流行文化中的明星角色。从《摩登原始人》（*The Flintstones*）里的迪诺（其实在陆地上，恐龙并没有与人类共存过），到《侏罗纪公园》的再生恐龙（其实它们在未来也不会与人类共存）。即使是《金刚》的电影人也不满足于一只巨大的可以爬上帝国大厦的大猩猩。我想，他们需要加入一个完全多余的有恐龙的场景。

为什么？因为恐龙太让人不可思议了。它们看起来和今天的动物非常相像，让人觉得似乎很熟悉，但是又非常不同，它们的奇异和古怪激发了我们的想象。恐龙有角和冠，还有骨质铠甲和刺。一些恐龙体型庞大、行动缓慢，而另一些却小巧而灵敏。一些恐龙生活在陆地上，有的用两条腿走路，有的用四条腿走路；而有一些则会在空中飞行。

然而，对许多人来说，当他们想到恐龙的时候，第一个划过脑海的事情是，那些体型庞大的动物永远也不能再在地球上行走了。虽然一些恐龙的确演化成了某些鸟类，并存活到了今天，但这些已经统治陆地长达几百万年的恐龙在大约6 600万年前灭绝了。有些人甚至带着淡淡的优越感来看待恐龙的灭绝，如此强大而机敏的生物怎么会愚蠢到让自己消失呢？而事实是，恐龙作为地球上霸主的时间，要远远超过人类或猿类可能生存的时间。恐龙的消失，并不是因为它们自身的错误。

是什么造成了陆生恐龙从这个星球上消失的？这在很长一段时间里，是困扰科学家和公众的一个巨大难题。为什么这个多样化的、强健的群体，而且似乎已经适应了当时环境的群体，突然在白垩纪末期消失了？这个话题好像离物理学，特别是与暗物质有关的物理学有些遥远。但这一章我会给出许多证据，这些证据已经表明，流星体的撞击几乎可以肯定是恐龙灭绝的罪魁祸首，而且这需要和太阳系的外来物体连接起来。而且，如果我和合作者所做的推测最终被证明是正确的，那么银河系平面上的一个暗物质盘便是触发流星体的致命轨迹的原因。

不管暗物质是什么角色，来自外太空的物体的撞击消灭了地球上至少一半的物种，这是肯定发生了的，这一事件把这次灭绝和太阳系的环境联系起来。地质学家、物理学家、化学家和古生物学家是如

何得出这一结论的呢？这个故事也许是现代科学领域最好的故事之一。

恐龙时代，一亿年的地球称霸

恐龙这个物种，除了它们的尺寸范围和冷血特性，还有一个引人注目的地方，就是它们存在时间之长——它们称霸地球超过一亿年。然而，尽管这个物种具备明显的强健特性，而且曾经一度十分繁荣，但是相当多的生命还是在6 600万年前突然结束了。为什么会发生这样的情况，它又是如何发生的？这些问题一直持续到20世纪末期。

在回答这些问题之前，让我们首先想想恐龙的年龄，以及当时的地球与现在有何不同。

恐龙生活在中生代，从25 200万年前到6 600万年前之间（见图11-1）。“中生代”（Mesozoic）这个名字来自希腊语，意为“中间生命”，这个时代的确位于显生宙三级地质时代的中间部分。中生代将古生代（Paleozoic，意为“古老的生命”）以及新生代（Cenozoic，意为“新的生命”）分开。这个时间段包括了我们所知道的最具摧毁性的物种大灭绝事件，即二叠纪-三叠纪灭绝事件。后者定义了第一个边界。中生代还包括白垩纪-第三纪灭绝（以前被称为K-T灭绝），定义了第二个边界，在这期间（非鸟类）恐龙和其他许多物种都消失了。

K-T里的K源于德语单词“Kreide”，意为“白垩”。“Retaceous”（即白垩纪）来自拉丁语词汇“creta”，字面意思是“克里特大地”（Cretan earth），意思也同样是白垩。K-T中的T来源于“Tertiary”，这是一个现在已停止使用的命名方案的一个遗俗，该方案将地球历史分为四个部分，Tertiary即为第三纪。^[27]即使如此，和许多人

一样，在提到此次灭绝的时候，我偶尔会使用口语化的名称K-T灭绝，不过从现在开始，我会经常使用更正确的术语——K-Pg灭绝。

纪元被分成时期，又被进一步划分成时代和阶段。中生代分为三个时期：

- 三叠纪时期，从25 200万年前至20 100万年前；
- 侏罗纪时期，从大约20 100万年前至14 500万年前；
- 白垩纪时期，从14 500万年前到6 600万年前。

可能“中生代公园”才是迈克尔·克莱顿和史蒂文·斯皮尔伯格的电影《侏罗纪公园》的准确名字。电影里介绍了两个侏罗纪时代的恐龙，但是还有一些直到白垩纪时期才出现的恐龙。尽管如此，我仍然愿意承认“侏罗纪公园”听起来更好，所以我不会质疑这个选择的明智性。

中生代时期，地球发生了很多改变。气候变暖和变冷以及显著的地壳构造活动改变了大气层和陆地的形状。被称为盘古（Pangaea）的超级大陆（泛古陆）在中生代分裂成我们今天看到的大陆，并随着时间的推移发生了大范围的陆地运动。

即使在白垩纪晚期，地壳的构造运动使地球更接近其现在的状态，但大陆和海洋尚未到达其当前的位置。印度尚未与亚洲相撞，而大西洋要窄得多。自那时起，大陆板块已经开始漂移，海洋在以每年几厘米的速度改变着大小。

单是这种效应就告诉我们：6 600万年前，大多数海岸离它们现在的位置远达几千公里的距离，因此，举例来说，美洲和欧洲当时非常接近。此外，当时的海平面很可能比今天的海平面高100米。气温，尤

其是在远离海洋的地区的气温，也比今天高。这些因素是破译在白垩纪-古近纪边界所揭示的一些线索的关键。虽然我们现在知道，意大利的沉积岩在形成的时候是水下几百米的大陆架的一部分，但研究人员最初并不知道这一事实。这个结论是由加州大学伯克利分校的地质学家沃尔特·阿尔瓦雷斯对意大利沉积岩中的黏土进行深入研究后得到的。

地球上的生命随着环境的改变而不断进化。被海洋分离的许多移动的大陆使新物种的大量出现成为可能。在三叠纪时期，节肢动物、海龟、鳄鱼、蜥蜴、硬骨鱼类、海胆、海洋爬行动物和第一代类似哺乳动物的爬行动物出现了。晚三叠纪也是许多独特的恐龙种类（包括陆生恐龙）首次出现的时期。它们后来成为侏罗纪时期主要的陆生脊椎动物。

在此期间鸟类也出现了，它们从兽脚亚目恐龙的一个分支演化而来。《侏罗纪公园》中的科学内容不一定都正确，但是这部电影让许多人知道了鸟类是从恐龙进化而来的。飞行的爬行动物、海洋爬行动物、两栖动物、蜥蜴、鳄鱼和恐龙持续生存到白垩纪时期，在此期间，蛇和早期鸟类首次出现，飞行爬行动物和银杏也出现了。此外，如苏铁、松柏类、杉树、柏树、紫杉等现代植物也出现了，我们今天还能看到这些类型的树木。哺乳动物也出现了，但它们那时还很小，通常在猫和鼠的大小之间。这一情形直到恐龙灭绝之后才有所改变，因为恐龙的灭绝留下了空间和资源，这使得哺乳动物能发展成更大体型的动物。

一个世纪的求索

在我写这本书的时候，我读过两本让人入迷的书，分别是沃尔特·阿尔瓦雷斯的《霸王龙和陨星坑》（*T. rex and the Crater of Doom*）和科幻作家查尔斯·弗兰克尔（Charles Frankel）的《恐龙灭绝》（*The End of the Dinosaurs*）。沃尔特是流星体假说的主要提出人，并且他的书非常有趣。我承认弗兰克尔的书对我显得非常特别的原因之一是，当我在亚马逊买它的时候，这本书已经绝版了，所以我收到的那本来自罗克波特公共图书馆（Rockport Public Library），它上面有一个大的印章标明“丢弃”。如果那本书没有被邮寄到我家（一个更为合适的栖息地），它显然也已经灭绝了。

这两本书讲述了非常精彩的故事：地质学家、化学家和物理学家如何证实了一个巨大的流星体（我也用“流星体”指代大的天体）是造成恐龙灭绝事件的最可能原因。与恐龙一起，很大一批其他物种也在那时消失了。诸多证据显示，这个流星体引发了白垩纪-古近纪过渡期化石记录中的戏剧性改变。在陨石坑边界的铱层附近，所有表征撞击坑的特点都被发现了，包括小颗粒、玻璃陨石和冲击石英。铱层将它下面丰富的生命遗迹和上方稀疏得多的化石记录分离了开来。

这两本书还涉及拍案惊奇的侦探故事，书中讲述了科学家们是如何发现与该流星体撞击相符合的陨石坑的，即便向领域专家的咨询后我知道书中一些内容是有点误导的。我将尽我所能地在这里把它正确地讲述出来。这是一个伟大的故事。

虽然流星体造成物种灭绝的想法直到20世纪后期才站住了脚，几百年来人们一直在猜测其潜在的可怕后果。当人们第一次注意到彗星时，就认为它会危及生命，但那多是基于迷信。

1694年，埃德蒙·哈雷大胆建议：彗星是《圣经》大洪水的来源。大约50年后，1742年，法国科学家、哲学家皮埃尔-路易·莫佩尔蒂（Pierre-Louis de Mau-pertuis）为来自彗星的潜在威胁找到了强

大的科学基础。他认识到：彗星撞击可能引发了海洋和大气的扰动，毁灭了许多生物。另一位法国人，伟大的科学家皮埃尔-西蒙·拉普拉斯（Pierre-Simon Laplace）也认为流星可能引发灭绝，他对太阳系形成的研究工作直到今天仍然是正确的。

不过他们的想法在很大程度上被忽略了，因为这些想法无法被证实，并且看起来似乎有点疯狂。另外一个被忽视的是美国古生物学家德劳本菲尔斯（M. W. de Laubenfels）的观点，他在1956年意识到：1908年击中西伯利亚并摧毁大片广袤森林的流星体的潜在重要性很高。德劳本菲尔斯指出，即使是彗星的一个碎片撞击地球也可能造成破坏，如大火和炎热。他还给出了惊人的具有先见之明的分析，他也认为这些环境的冲击会在不同程度上影响各类物种，而穴居哺乳动物可能存活。在白垩纪-古近纪事件之后，实际情况的确是这样。

1973年，地球化学家哈罗德·尤里（Harold Urey）建议：基于对熔岩的玻璃陨石的研究，流星体的撞击是造成白垩纪-第三纪灭绝事件的原因。大多数科学家依然选择视而不见。然而，尤里“热情不减”，他提议不仅是白垩纪-古近纪灭绝源于流星体的撞击，甚至所有其他物种的大灭绝都是由彗星撞击造成的。尤里对未来的研究作出了预言，并将早期的建议发展成真正的科学。尤里指出，详细的调查能识别出岩石的起源，它们的形状或构成只能由流星体击中时的热量和/或压力所解释。

不过，在活尔特提出他的建议之前，所有这些有先见之明的想法几乎都被人们忽略了。即使是在20世纪80年代，太空的撞击引起物种灭绝的想法仍然被认为是激进的、荒唐的。这让我想起一些理论，这是从参加我公开讲座的12岁的孩子们那听到的。当时他们试图炫耀所听说过的科学术语。这可能会引发非常有趣的场景。比如一个年轻人向我询问一个理论，声称他一直在琢磨，认为弯曲的额外维度中的黑

洞能解决宇宙的所有遗留问题。但当我告诉他，我认为他实际上并没有一直在思考这个问题的时候，他笑着承认了。

像那些最终被接受的激进理论一样，流星体的提议可以解释传统说法无法解释的观测。地球很多无法解释的现象却相继被发现了，这给了流星假说很多支持。使它得到了人们的信任，因为它作出了不少预言，而其中许多预言在后来又得到了验证。

一探6 600万年前的历史片段

沃尔特的科学侦探故事开始于意大利。在罗马以北几百公里，古比奥（Gubbio）附近的翁布里亚山丘上（Umbria），出现了从白垩纪晚期到第三纪（现古近纪）早期的海洋沉积岩。斯卡利亚罗萨（Scaplia Rossa）深水远洋灰岩，以其粉红色的颜色闻名，这种沉积岩由非常不寻常的深水石灰石，即方解石或碳酸钙组成，它们大多数由贝壳组成，也是骨补充剂常含有的成分。这种深水石灰石在海底形成，后来被向上推，逐渐露了出来。这意味着它们是物种灭绝的证据——也就是一层薄黏土将下面的白色岩石层和上面的红色层分离开了，可能会被一位细心的路人发现。在较低白色岩石中的化石大多是有孔虫，生活在深海的单细胞原虫的遗骸，并且对我们推导沉积岩的年龄非常有用，但只有最小的有孔虫在靠上的暗色层中被发现。有孔虫几乎和恐龙同时灭绝，这使得沉积岩上所展示的灭绝事件边界非常清晰。

一次，我在访问西班牙毕尔巴鄂的大学期间参观了复理层地质公园，那有一块白垩纪-古近纪边界的片段，表现为石灰石悬崖底部附近薄的暗线。像其他类似的黏土层一样，这一边界可以帮助我们追溯到灭绝的时间。我认为自己非常幸运，我的同事（物理学家）和他的地

质学家表弟帮助安排了对美丽的Itzurun海滩的一次游览。在那里我可以在退潮的时候跑进去，近距离地看一看这些边界。触摸这个6 600万年前的历史片段几乎让人激动不已（见图12-1）。虽然悬崖源自遥远的过去，但是它的信息宝库仍然作为我们世界的一部分与我们共同存在。

在白垩纪-古近纪边界

20世纪70年代，沃尔特研究了斯卡利亚罗萨的一个类似边界层。他把注意力集中将下面充满化石的浅色石灰石和上面没有化石的深色石灰石分离开的黏土层上。这种被沃尔特作为研究目标的黏土对于揭开发生在6 600万年以前的大灾难的原因至关重要。黏土的厚度取决于在较浅的和较暗的岩石沉积之间的时间间隔，并可能因此帮助他确定灭绝事件是快速发生的还是缓慢发生的。

20世纪70年代，沃尔特首先开始思考白垩纪-古近纪层时，渐进主义的观点依然是地质学的主流观点。这些观点被此前20年不断发展的板块构造理论所证实。整个大陆可能逐渐移动分开，山脉可以随着时间的推移形成，大峡谷可以通过渐进式影响出现。这些渐进式影响包括河流（如穿过地面的科罗拉多河）、水和冰的侵蚀、陆地板块的运动、岩浆喷发等，随着时间的推移，其中任何一种影响都能够彻底地改变地形。这些貌似戏剧性的变化并不需要灾难性事件来触发。



图12-1

和复理层地质公园的主任艾塞尔·希拉里奥（Asier Hilario）一起观看K-Pg分界线。照片摄于西班牙苏玛亚附近的Itzurun海滩。（感谢乔恩·尤勒斯提拉提供图片）

石灰石的形成似乎很神秘，因为石灰石的上部和下部的差别表明，当时出现了非常突然的转变，这和渐进的观点不一致。如果查尔斯·莱尔在那里，他会对白垩纪-古近纪的薄层作出简单而误导性的解释：尽管表现为很薄的一层，但是它的形成已经经过了许多年。达尔文可能认为这一薄层的形成只是由不充足的化石记录造成的一种错觉。

想要知道转变是不是突然发生的——当然也不仅仅是在几天时间里冲进来的黏土沉积而成的，只有一个办法，即测量两种被分开的、不同颜色的岩层黏土沉积所需要的时间。而这正是沃尔特给自己设定的任务。他一直对地质学事件的确定年龄有着浓厚的兴趣。沃尔特希望研究地磁倒转，以更多地了解关于白垩纪-古近纪边界的沉积时间，他知道这可能是触发这一事件的重要线索。（哈佛大学地球与行星科学教授安迪·诺尔[Andy Knoll]提到：沃尔特和他的妻子可能对中世纪的艺术和建筑更感兴趣。我想这两个兴趣或许都发挥了作用。）

但衡量黏土沉积所花费时间的更好方法其实是测量其铀含量。铀是一种稀有金属，并毗邻钍这一密度最大的元素。铀的耐腐蚀特性使其可以做成火花塞电极和钢笔笔尖等东西。它也被证明对科学很有用处。沃尔特和他的合作者们所发现的铀含量最终成为确立灭绝事件起源的关键。

虽然我知道铀含量的峰值已经有一阵子了，但最近还是了解到一个令我吃惊的事：沃尔特和他的物理学家父亲路易斯·阿尔瓦雷斯测量黏土中铀含量的初衷是和真正正确的原因背道而驰的，他们很快就意识到这一点。路易斯知道流星体的铀含量比地球表面高得多。虽然地球上铀

的含量应该和流星体的铀含量相同，但在地球形成早期，大部分原始的铀溶于铁熔岩并沉入地核中。因此，地球表面的铀应该起源于外来天体。

路易斯认为，陨石尘埃应该以一个相当稳定的速率沉积到地球上（他原本建议使用铍-10，但它的半衰期太短，并不实用）。如果不是因为这种天外来客的坠落带来的沉积，地球表面上的铀含量是非常低的。沃尔特和父亲路易斯有一个聪明的想法：通过研究地表铀的含量，他们可以访问这个宇宙的沙漏，确定白垩纪-古近纪边界的黏土用了多长时间才沉积下来。他们期待的是随着时间平滑的分布，稳定的、近乎常数的沉积率，因为这种沉积率能够被用来推断黏土层形成所花费的时间。

然而，当沃尔特和他的合作者检查实际的岩石的时候，他们的发现完全不同。这个惊喜让沃尔特相信这些奇怪的结果有点不对劲，因为来自黏土铀含量比预期高得多。1980年，在美国加州大学伯克利分校的科学家组成一个小组。小组成员包括路易斯和沃尔特的“父子兵”。此外，小组中还有核化学家弗兰克·阿萨罗（Frank Asaro）和海伦·米歇尔（Helen Michel）。他们能够测量极低丰度的铀含量，并发现黏土中铀含量明显高于周围的物质，比斯卡利亚罗萨周围的石灰石高30倍，这个数字后来被更正为90倍。

这种类型的地质结构不仅在意大利被发现（可悲的是，自沃尔特之后，太多人在斯卡利亚罗萨采样，白垩纪-古近纪边界的黏土现在很难找到），而且在全球各地都有，并且在这些地点铀的含量都存在明显的增峰。在丹麦斯泰温斯-克林特（Stevns Klint）陡崖有保存完好的白垩纪-古近纪证据，有类似的黏土层，铀的含量高达周围物质的160倍。其他实验室也确认了在其他地方的类似边界层中的铀含量同样有增加。

如果初始的假设（以及时进行测量的动机）是正确的，即陨星尘埃以恒定的速率落下，白垩纪-古近纪黏土将需要超过300万年才能形成。但是，对于白垩纪-古近纪边界的薄黏土层来说，这个时间太长了。另外，如果铱的含量在全球各地都有类似升高，那么大约有50万吨铱，这一被认为是地球上稀有的元素，就曾在白垩纪-古近纪灭绝的时候突然出现在地球上。对这种大量沉积的唯一解释只有起源于地外宇宙。地球表面本身的铱含量如此之低，如果没有外星现象，这样高的铱含量基本上是无法解释的。

伯克利小组还确定了其他稀有元素的相对丰度，使他们能够进一步缩小地外事件的可能范围。例如，也许外星来源是一颗超新星。在这种情况下，超新星应该会导致黏土里还有钚-244。初步的分析确实表明，这个元素也是存在的。但是，对于一个科学标准下可靠的研究，阿萨罗和米歇尔第二天重新进行了分析，并没有发现钚。最初的发现只不过是样品中的污染物。

对于铱的高含量，伯克利的科学家们绞尽脑汁想找到其他解释，最后基本只剩一个似乎合理的解释：大约6 500万年前，发生了来自外星物体的巨大撞击。1980年，由沃尔特和路易斯父子领导的小组提出了一个观点：一个大的流星体曾与地球相撞，并散落下稀有金属，包括铱。这样的流星体撞击由小行星或者是彗星导致，这是唯一可以同时解释铱的总量过高和黏土中的元素含量比与太阳系的特征元素含量比相符合的方法。

基于测量的铱含量和陨石中的平均铱含量，研究人员还可以进一步猜测撞击天体的大小。他们得出的结论是：撞击天体的直径有令人难以置信的10公里~15公里之大。

发现确证

鉴于一个巨大的流星体会带来多重灾难，而白垩纪-古近纪灭绝相关的地质事件证据又比较缺乏，相比更加传统的建议，如地质或者气候导致的物种灭绝，地外天体撞击造成生物灭绝的解释似乎是一个合理且明智的选择。然而，尽管这个假设有引人注目的性质，但任何科学家，无论多么大胆，在引进新想法时，也必须审慎行事。有时激进的理论是正确的，但更多的时候，一个更传统的解释会一直被忽视或没有被正确的评估。只有当现有的所有科学想法都失败，而更大胆的想法取得成功时，新的想法才能被牢固地确立下来。

出于这个原因，当考虑稀奇古怪的理论时，争论对于科学来说是一件好事情。虽然那些简单地想避开检验证据的人不会促进科学的进步，但盛行观点的追随者们提出的合理反对意见，则提升了将一个新想法引进到科学神殿的标准，从而迫使那些提出新假设的人，特别是那些怀有激进想法的人去面对他们的对手。这可以防止疯狂的或者完全错误的想法站住脚跟。阻力能鼓励提议者提升自己的游戏能力，以表明为何反对是不正确的，并尽可能多地为他们的想法找到支撑。沃尔特甚至写道：他很高兴过了好一阵流星体的想法才找到确凿证据的支持，因为这为他留出了充足的时间来发现更多辅助性的证据。

流星体的假设确实遇到过阻力。反对的人们认为这是一个不切实际的理论。其中许多人更倾向于渐进式改变的观点。令人困惑的是，虽然板块构造给了渐进式改变这一观点以支持，但是在大约同时进行的月球任务中，科学家们对很多陨石坑的近距离观察却给出了一个强力的论据来证明，冲击可能带来灾难性后果。这两种不同的论据使人们分成两派，地质学家倾向于渐进式的观点，而物理学家走向了灾难性的方向。

当然，月球上的环形山可能在月球形成的早期阶段就已经全部形成了。事实上，大多数环形山也的确如此。因此，它们的存在本身并

不是对后期流星体显著性撞击的有效论据。尽管如此，这些观点的流行应该让这样的假设不再那么令人惊讶，在太阳系中不仅有渐进过程的存在，还有灾难性过程的存在，而且后者也对生命的发展造成过影响。陨石坑是月球被撞击过的清晰而明显的证据。地球非常大，又非常接近月球，因此显然流星体也很可能曾经击中过这里。

在沃尔特提出撞击说建议的时代，很多古生物学家却依然青睐渐进的解释。有一些人认为，恐龙在白垩纪晚期消失只是简单地由于某种形式的不良环境条件，如气候变化或不良的日常饮食。其他一些人则认为火山活动是罪魁祸首。对于这个观点的支持来自印度的德干暗色岩，这些岩石是由大约在恐龙灭绝时代所发生的巨量的火山活动形成的。德干暗色岩覆盖的区域超过50万平方公里，相当于法国面积的大小；它的厚度约2 000米。这个熔岩总量是非常大的。使情况更加令人迷惑的是，暗色岩可追溯至的产生时间非常接近晚白垩纪-早第三纪界线。

的确，恐龙的一些族群，如蜥脚类恐龙^[28]，在这个时代末期的时候就已经灭绝了。但支持渐进式衰落理论的证据部分源自于调查初期不完整的化石记录。但是研究范围的扩大、化石数据的增多，让这些早期的化石记录变得不那么令人信服了。在蒙大拿州发现的化石显示：至少有10~15种恐龙物种存活到白垩纪末。最近在法国的挖掘发现证据表明：在白垩纪-古近纪边界的一米范围内存在恐龙化石。而那些在印度最新的挖掘给出了边界线下也有恐龙化石的证据。其他物种的化石记录，如菊石，确实在一开始显示出了物种多样性的下降。但是更准确和更广泛的研究再次发现：至少有1/3的恐龙物种存活到白垩纪-古近纪边界，虽然有些种类的恐龙确实在这之前就已经灭绝了。

除此之外，尽管人们最初认为暗色岩形成得非常快，但是后来的工作表明，它们的形成花了几百万年，而白垩纪-古近纪事件对应于暗

色岩中间的一层，奇怪的是：似乎是火山活动在这段时间里是受抑制的。也许能表明火山不是恐龙灭绝的全部责任的证据是：印度地质学家在正好构成白垩纪-古近纪边界的区域的沉积物中发现了恐龙的骨骼和卵的碎片。恐龙在这期间不仅活着，而且它们还曾居住在暗色岩存在的区域。

即便如此，再后来的研究进展将暗色岩的形成时间向灭绝的时间推得更近了，比人们以前认为的更接近。这给火山活动对灭绝的贡献给出了支持，尽管火山不用对毁灭负全责。有人推测，火山活动实际上是流星体冲击的结果。在这种情况下，不论火山作用如何发生，也都可以间接归因于流星体。无论它们的角色如何，火山不能解释许多其他同时存在的地质特性。这些特性为流星体的显著性给出了令人信服的有力证明。



事实上，一旦人们开始认可流星体假说，相关的证据就迅速积累起来。细节很重要，并可以帮助解决许多争议。1980年伯克利分校的建议被提出之后，人们对白垩纪-古近纪黏土层进行了精心的研究，足迹遍及意大利、丹麦、西班牙、突尼斯、新西兰和美洲各地。到1982年，全球各地的近40个地点已被仔细检查过。荷兰古生物学家扬·斯密特（Jan Smit）在西班牙发现了高含量的铱。其他古生物学家在斯泰温斯-克林特测量到类似的结果。斯密特还测量了其他稀有金属的丰度，如金和钼。他发现，铱和钼含量比地球上其他地方看到的高出1000倍。并且和以前一样，相对金属丰度与对流星体中的预期相符合。

赞成火山解释的一些科学家认为，大量的铱是伴随火山喷发从地球的地幔和地核中喷出的。我们知道，地幔和地核中的铱含量比较

高。但是，已知的所有火山所喷发出的铀也远远不足以解释世界范围内存在于白垩纪-古近纪边界的铀。据沃尔特和其他人的计算，包含其他的潜在聚集效应如海洋中的沉淀，白垩纪-古近纪边界上的铀含量达50万吨。无论如何，铀不是陨石中的唯一重元素，而其他元素的丰度也和火山排放出的量不符合。

白垩纪-古近纪层中和周围的其他观测为流星体假说提供了更多证据。例如，在不同地点发现了岩滴，如微晶（microkrystites），这是较小版本的玻璃陨石，这些有着圆润形状的玻璃质岩石产生于撞击时，是掉回地球之前在大气中飞驰并凝固的融化物。它们的发现为流星体假说提供了支持。

这些玻璃小球最初也只是混淆了视听，因为科学家被其误导了。它们的化学分类类似于海洋地壳，但是它原来很可能是来自撞击体，而不是被击中的目标。如果最初的误导性结论是正确的，并且流星体着陆于海洋而不是在陆地，这意味着：尽管对于撞击有着越来越多的证据，但是撞击地点很可能仍然被隐藏着。

当地质学家发现的证据表明流星体是在（可能到达的）大陆架着陆时，关于着陆地点的担忧就消除了。证据就是冲击石英的发现，这种石英起源于高压，而且只能产生于含石英的岩石之间的撞击。不融化的岩石被打碎，所以它们包含的矿物质可以移动并形成纵横交错的纹路（参见图10-3），这些纹路目前的已知来源只有流星体的撞击和核爆炸。想必核试验在6 600万年以前没有发生过（虽然一位研究人员告诉我，有个电台的采访者真的就这种可能性向他提问过），这就只剩下了流星体的撞击这个唯一可能的解释了。

1984年，冲击石英在蒙大拿州被发现，稍后在新墨西哥和俄罗斯被发现，新的发现为流星体的撞击提供了强有力的证明。这种石英的存在进一步证明了撞击坑（假设有一个的话）应该位于陆地上，因为石英在海洋中的岩石里非常罕见。

支持流星体假说进一步的证据持续地积累着。加拿大科学家在艾伯塔省（Alberta）的白垩纪-古近纪层发现了微小的钻石。这些钻石可能是由流星体从外太空携带过来的，也可能是在撞击过程中形成的。关于尺寸和碳同位素比率的详细研究结果更青睐于后一种解释。在加拿大以及丹麦，一些不应该存在于地球上的氨基酸在边界层被发现了，这些氨基酸在地球上任何其他地方都没有被发现过。这个证据的有趣特点在于，它支持彗星撞击说的解释，因为这些氨基酸在接壤的石灰石中也被发现了。如果彗星尘埃在形成边界层的时候存在于附近，那么就会出现这种情况。

另外一个为高压撞击提供证据的重要地质特性是被称为尖晶石（Spinel）的晶体。它们是金属氧化物，包含铁、镁、铝、钛、镍和铬，呈雪花状、八面体等各种奇异形状，这些结构反映了高温熔融之后的快速凝固。火山岩浆中也有尖晶石，但被发现的尖晶石含有镍和镁，这和火山爆发所形成的尖晶石不同，后者含有更多的铁、钛和铬。更好的证据是，氧的含量有助于确定形成尖晶石的地点。白垩纪-古近纪层中的氧化尖晶石表明：这些尖晶石产生于一个低海拔地区——低于20公里。该晶体只在一个薄层中被发现，这证实了发生在白垩纪-古近纪边界的灾难性事件是非常短暂的假设。

火山无法解释因冲击而创造出的材料。虽然火山确实会造成形变，但是现有的火山地区所产生的冲击石英不足以匹配来自灭绝时期的观察现象。产自火山爆发的冲击石英的结构错位是沿着单一平面的，而不是两个或更多个交叉的平面。后者是仅在高冲击压力下才会发生的现象。这些细节很重要，因为这些现象正是在标定白垩纪-古近纪灭绝边界的精确位置被找到的。

即便如此，虽然已经牢固地建立起了流星体破坏性影响的理论框架，我们也不应该完全忽视渐进式观点。很可能的是，在白垩纪-古近纪灭绝时间的前后，环境正在发生改变，增加了生态系统的脆弱性。因此，流星体的撞击带来了比环境没有发生改变条件下的更多损害。有证据表明，在剧烈的灭绝事件发生之前，一部分物种已经灭绝了。最近对德干暗色岩产生年代的更精确测量给出了进一步支持，结果表明，火山活动起了一定的作用，虽然不太可能对最终发生的灭绝事件负责，但是火山爆发以及其他现象很可能在灭绝事件前后都助长了灭绝所造成的破坏。

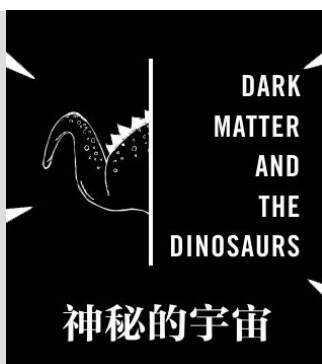
撞击并不需要任何帮助，就能造成巨大的破坏。

生命如何出局

要弄清流星体带来的灾难到底有多么巨大、多么有毁灭性，是很困难的。撞击的跨度大约是曼哈顿宽度的3倍。而且，这个流星体不只是大，它还在迅速地移动，以至少每秒20公里的速度移动，并且如果这个星体是个彗星，移动速度也许还要快3倍。这个物体的速度比在高速公路上以时速100公里前进的汽车快了至少700倍。这个冲击体就像一个快速移动的大城市，速度是在无限速高速公路上行驶车辆的500

倍。由于一个物体携带的能量正比于其质量和速度的平方，如此快速的、庞大的物体撞击地球将造成巨大的毁灭性影响。

从某种角度说，有着这种体积和速度的物体可以释放出的能量相当于百万亿吨的TNT炸药，是摧毁广岛和长崎的原子弹能量的10亿多倍。沃尔特的父亲路易斯为“曼哈顿计划”工作过，他提出过类似的评论。更广泛地说，“冷战”宣传核爆炸的影响激发了人们对撞击坑的兴趣。对二者的研究都得益于关于白垩纪-古近纪撞击体的长期环境影响认知的不断增加。



通古斯事件的物体和在亚利桑那州造成陨石坑的流星体所携带的能量只是上述天体所携带能量的一小部分，大概相当于10倍百万吨级的TNT炸药。在这两种情况下，冲击体的直径大概是50米，而不像形成白垩纪-古近纪事件的撞击体那样达到10公里~15公里。喀拉喀托火山的能量仅是这些较小流星体的几倍。这些小流星体的能量和至今所有生产的核武器相当（大约是现存核能量的50倍）。一个1 000米大小的流星体已经足够造成全球性的损害了。沃尔特建议的物体比这要至少大10倍，比珠穆朗玛峰的高度还高，超出海平面达9公里。

这个有着巨大速度的物体的撞击和影响都是毁灭性的。如第11章所述，在如此沉重的石头投掷到地球上之后，很多灾难将会随之降临。靠近爆炸的地方，大约1 000公里范围内，极端的风浪肆虐、巨大的海啸从爆炸地点向外辐射。这些潮汐波浪极端强大，但范围有限，因为事实证明，撞击地点的水深只有约100米。潮汐巨浪也会出现在地球的另一端，它是由地球有史以来经历的最强烈的地震引发的。极端的强风会从撞击位置向外吹，然后急速冲回。

这股强风会携带一大团的超热尘埃、灰烬和蒸汽云，它们是在流星体最初冲进地面的时候被抛起来的。这一阵风和水的能量释放可以用完大约1%的冲击能量。其余的能量会被用于熔化、蒸发，并在整个地球范围内传播地震波——相当于里氏10级地震的地震波。

上万亿吨物质被从撞击坑的地点喷出，并散布到四处。随后，当热的固体颗粒穿越大气层降落下来时，它们会被加热到白炽，并提升世界各地的气温。导致的结果是：到处大火肆虐，地球表面不断升温。1985年，化学家温迪·沃尔巴克（Wendy Wolbach）及其合作者发现，在白垩纪-古近纪层中存在大火的证据，这些证据以木炭和煤的形式呈现出来。他们发现的碳屑的丰度和形状证实了火灾曾经发生过，并摧毁了当时存在的动植物。研究人员得出结论，当时地球上超过一半数量的生物在撞击后的数月内被焚尽。

这还不是全部。水、空气和土壤中都携带了有毒物质。人类对彗星的恐惧并不仅仅是迷信。彗星的确包含了有毒的物质，如氰化物以及包括镍和铅的重金属。虽然有些化学物质在它们可能带来损害之前就会蒸发，但很可能重金属会从天上雨点般落下。

可能更多的破坏来自大气中产生的一氧化二氮。它可能通过酸雨降落到地面。硫也会被释放到大气中，形成硫酸，并可能停留在大气中，阻止阳光照射，造成全球降温，紧随着大灾难发生后立刻发生的全球加热，并可能持续数年。光合作用的缺失将对整个食物链产生严重影响。全球变暖和灰尘颗粒对地球的覆盖可能也会起到一定的作用，将异常的加热和冷却延展更多年。

的确，化石记录表明，在最初的撞击之后，遗留的破坏性仍然在持续。即使存活下来的物种的个体数目也严重缩水。海洋在几十万年内无法恢复，并且更可能的是，在之后至少50万~100万年中，破坏性影响仍然存在。化石记录表明，在包含很少或没有碳酸盐的深色石灰石处，缺乏浮游生物和其他生物的化石。取而代之的主要是碎屑颗粒的证据，也就是撞击后存留下来的被日晒雨淋侵蚀过的岩石小碎片。岩层中正常的颜色在至少几厘米内（有时达到几米）都没有重新恢复，这个厚度因地球的不同地点而不同。

如此多的灾难让动植物走向灭绝的可能性不断增加，几乎所有体重超过25千克（一只中型狗的重量）的生物都无法幸存。为了生存下来，必须有某种方式躲避灾难，通过冬眠或其他方法。由于特殊的繁殖方式（种子比其他繁殖方式更有机会幸存）和食物来源（以废弃物为食的物种更走运一些），部分物种确实生存下来，能逃到天空的动物们也有更好的机会。但是，大多数动植物都死亡了。一颗直径为10公里~15公里的流星具有巨大的破坏性——无论是对环境还是生命，都是如此。

重新发现陨石坑

尽管如此，研究人员当时就知道：即使有20世纪80年代发现的所有证据，以及对一个巨大流星体可能对地球生命所产生后果的日益增加的了解，但是找到一个可以触摸的、有着正确尺寸的、6600万年历史的陨石坑显然是撞击假说最强有力的证据。陨石坑不仅能证实这个

假说，还可以使更详细的研究成为可能，从而更好地估计其大小和撞击时间，并辅助确认撞击发生的其他特征。

陨石坑的大小以及它的年龄，是至关重要的预测。根据已测定的铱含量，沃尔特推断，流星体的直径应该至少有10公里。据此推断，陨石坑应该有大约200公里的口径，因为陨石坑通常是撞击体大小的约20倍。沃尔特不是唯一一个预测陨石坑有这种规模的人。另一位古生物学家独立预言出180公里的大小，他是基于这种假设：黏土包含了陨石物质的7%，其余物质来自被击中目标的粉碎的岩石。

一个有着正确尺寸、正确日期的陨石坑是沃尔特理论的确凿证据。然而，这个陨石坑的发现花费了超过10年的时间，成了现代科学最好的侦探故事之一。事实上，当人们最初开始寻找撞击地点时，成功的可能性似乎并不高。尽管在一些年中，有些大的陨石坑已经被发现，但更多陨石坑则无法被找到了。即使我们足够“幸运”，流星体击中陆地而不是海洋，但侵蚀、被沉积掩埋或地壳构造的破坏也可以消除陨石坑曾经形成的任何迹象。

对于造成白垩纪-古近纪灭绝事件的流星体来说，由于没有明显的撞击位置，因而更难发现其陨石坑。四处可测的铱以及其他地质证据在全球的分布是大致均匀的，这证实了流星体的全球影响，却没有指明任何特定的地区。要寻找特定的一次撞击所造成的陨石坑即使不是不可能的任务，但确定一个特定流星体在6 500万年前撞到了地球哪个位置，仍然十分困难。

然而，有利于陨石坑追踪者的是：已经发现的冲击石英表明它的起源是大陆或大陆架，因此在陆地上的搜索仍然有成功的机会去辨认出肇事者的遗迹。几个看似有希望的候选陨石坑出现了，但很快就被进一步的调查排除了，因为相关测量和更精确的细节相悖，包括对它们撞击时间的测量、大小的确定或矿物学研究。

一个很重要的独立观测在很长一段时间里都被忽视了。

早在20世纪50年代，工业地质学家们就辨认出了一个被掩埋的圆形结构，直径达到180公里，其中一半延伸到陆地上，位于尤卡坦石灰石平原之下；另一半在海里，被掩埋在墨西哥湾的海水和沉积物之下。来自墨西哥石油公司（Pemex）的地质学家们在这个结构内部钻井勘探。他们在大约1500米的深度找到了结晶岩，这使他们认为找到了火山的证据，而不是他们更希望的石油。

在20世纪60年代末期，地质学家罗伯特·巴尔托索（Robert Baltoser）为防止在第一轮调查时错过油矿的迹象，而参与了第二轮钻井勘探。随后他建议，这实际上可能是一个撞击坑。基于对这个地貌的引力势形状，即关于引力在这个圆形结构上如何变化的测量，他得出了这个结论。但它仍然不是油矿，而且墨西哥石油公司也不允许巴尔托索公开自己的观测结果。当时，这导致大多数了解这个结构的人都在为石油企业工作，而石油企业希望保护自己的成果。

但墨西哥国家石油公司坚持不懈地进行着对石油的探索，并在20世纪70年代进行了进一步的地质研究，其中包括对整个尤卡坦半岛的空中磁性测绘。一位美国顾问格伦·彭菲尔德（Glen Penfield）注意到了——一个强烈的磁异常，这个异常的区域大约50公里宽，被一个有着不寻常的低磁性的、直径约180公里的外环所包围。这正是对一个大型撞击坑所预期的模式：中央区域和冲击熔化物相关，而外围区域包含着变硬的被击中物的残骸。彭菲尔德把这种对应关系记录了下来。航空重力数据给出了对这种解释的进一步支持：升高和被降低的重力场与磁性信号的变化密切相关。

因此早在1978年，彭菲尔德已经意识到这个地质结构是撞击坑的强烈迹象。他意识到，发现一个以前未知的撞击事件的证据可能是一个意义相当重大的事件，因此他从墨西哥国家石油公司得到许可，发布了这些通常被认为是专有的数据。和墨西哥国家石油公司地质学家安东尼奥·卡玛戈（Antonio Camargo）一起，彭菲尔德于1981年在洛杉矶的勘探地球物理学家学会（SEG）会议上展示了他的结果。但是这一发现在当时并没有得到太多关注。大多数听众仍然不了解白垩纪-古近纪灭绝的撞击假设，所以当时没有人意识到其中的关系。

直到1990年，大多数对白垩纪-古近纪撞击坑感兴趣的人才开始研究这个特殊的陨石坑。但他们实现相关研究的过程也是一个难以置信的故事。那些为了证实沃尔特的理论而寻找一个特定的直径大约为200公里、6 600万年前的陨石坑的人，用和墨西哥国家石油公司的地质学家完全不同的角度对这个陨石坑进行了搜索。他们研究了白垩纪-古近纪层，搜索了撞击位置的线索。尽管在全球范围内铱沉积是均匀的，但他们知道一条线索，这个线索如果被发现，将给出更具体的位置信息。如果流星体击中海洋，但落在靠近海岸的地方，撞击就会制造出一个海啸，这个海啸强大到足以在大陆的平台留下它发生过的痕迹。根据已有的陆地撞击证据，这似乎是一厢情愿，但地质学家们对此保持着关注，并因为不懈的努力得到了奖励。

1985年，扬·斯密特和一位合作者研究了在得克萨斯州墨西哥湾附近的布拉索斯河（Brazos River）河床中露出地面的白垩纪-古近纪沉积物，结果表明这部分沉积物被某种力量打乱了，他们确信这一现象是由海啸造成的。华盛顿大学的地质学家乔安妮·布尔茹瓦（Joanne Bourgeois）认真地根据他们的工作做了进一步的调查，并出乎意料地发现了包含贝壳、树木化石、鱼的牙齿和黏土碎片的粗砂岩。这些碎片和当地的海洋底部相符合，并且这些物质的原始位置被确定在6 600万年前的海平面下100米。布尔茹瓦能够使用砂岩块的大小来估计水流的速度超过每秒一米，对应于至少100米的波浪高度，并进一步发现黏土具有一定的模式，这表明水流的移动既有冲向海岸的，也有从岸上离开的。通过假设最大可能的浪和整个海水深度的5 000米相同，布尔茹瓦推断出：撞击一定距离她的位置不到5 000公里，这意味着撞击地点在在墨西哥湾、加勒比海或大西洋西部。

撞击位置的另一线索来自地质学家布鲁斯·博尔（Bruce Bohor）和格伦·伊塞特（Glen Izett）。他们在1987年发现，冲击石英最大

和最丰富的储藏在北美西部内陆，这表明撞击在大陆附近。这与斯密特和布尔茹瓦的分析是一致的，他们都提出撞击靠近大陆的南端。

当海地地质学家弗洛伦廷·莫瑞斯（Florentin Maurrasse）在他祖国的白垩纪-古近纪边界层认证出一些有趣的碎片后，撞击地点的可能范围被进一步缩小。他对不寻常的沉积物的描述吸引了亚利桑那大学的研究生艾伦·希尔德布兰德（Alan Hildebrand）以及他的导师比尔·博因顿（Bill Boynton）和研究员大卫·克林（David Kring）的关注。虽然莫瑞斯认为碎屑的起源是火山活动，但是亚利桑那小组明白火山和撞击的碎屑有多么容易被混淆。他们一看到海地的样本，就辨认出了玻璃陨石，并决定亲自去海地访问。1990年，他们在那里发现一个半米厚的沉积凸起中似乎包含着玻璃陨石，并在其中还发现了冲击石英和铀黏土。这极有可能是与流星体撞击相关的区域。根据沉积层的厚度，他们得出结论：在撞击发生时间点，撞击坑离碎石片发现地应该不超过约1 000公里。

尽管希尔德布兰德最初青睐的一个加勒比海候选地区后来被他否定了，但是亚利桑那小组最终将目光聚集在尤卡坦这个10年之前就被认证的地区。但第一个找出陨石坑和恐龙灭绝事件之间关联的不是科学家，而是《休斯敦纪事报》（*Houston Chronicle*）的记者卡洛斯·拜厄斯（Carlos Byars）。听完希尔德布兰德在一次科学会议上展示的研究结果之后，拜厄斯告诉了希尔德布兰德关于彭菲尔德早期对一个撞击坑的发现，这帮助科学家们给丢失的陨石坑之谜找到了一个令人满意的答案。

墨西哥国家石油公司发现的陨石坑是正确的撞击位置，而且它也有合适的尺寸。这样的一致性支持将陨石坑联系到白垩纪-古近纪灭绝的一个重要论据。即便如此，当1990年希尔德布兰德向一份科学杂志提交了两份摘要，提出这一联系时，文章并没有得到发表。部分是

因为最初的证据没有足够的说服力。然而，当亚利桑那小组从陨石坑中辨认出冲击石英的时候，人们的观点开始发生改变。

由于陨石坑位于一个被淹没的大陆平台中，被泥沙覆盖着，这使陨石坑难以被发现被研究。但它的被掩埋在某些方面又是幸运的，因为它上面近公里的硬化泥土为陨石坑提供了防护，使它免受地表发生的侵蚀。为了调查这个被埋藏的，并且因此在最初也难以接近的陨石坑，亚利桑那的科学家们联系了彭菲尔德和卡玛戈来研究之前钻探出来的岩芯。他们获得了存储在新奥尔良的两份拇指大小的样品。果然，当亚利桑那小组研究这些古老的、来自墨西哥国家石油公司的岩芯时，他们找到了所期待的物质。他们鉴定出冲击石英和冲击熔化的岩石，这表明陨石坑形成于撞击而不是火山。1991年3月，克林在美国国家航空航天局约翰逊航天中心宣布了这一发现。

亚利桑那的科学家们将他们对岩芯的研究与墨西哥国家石油公司和彭菲尔德以及卡玛戈提供的地球物理数据相结合，与其他的合作者一起，为撞击-灭绝假说积累了有力的证据，并指出陨石坑是引发了白垩纪-古近纪大灭绝的撞击的遗迹。1991年，他们在《地质学》（*Geology*）杂志上发表了这一结果，以及他们对陨石坑尺寸为180公里的估计。在冲击石英和其他支持证据面前，许多科学家也开始关注这一理论。

亚利桑那小组用一个附近渔港的名字命名了陨石坑，这个渔港位于陨石坑中心的上方，不过它的名字很难发音：Chicxulub Puerto（希克苏鲁伯·波多黎各）。Chicxulub这个词发音为“CHICK-shuh-lube”，有时被翻译成“魔鬼的尾巴”——非常适合描述被沃尔特戏称为“末日大坑”（crater of doom）的特点。

亚利桑那小组的工作发表后不久，遥感专家意识到他们可以在卫星图像中检测陨石坑的周长。图像显示，陨石坑的周围被一个环形池塘围绕着，半径为80公里。这些也非常可能是由陨石坑的形成而导致

的。陨石坑的形成会造成地下水上涨并穿过地球的表面，因此这是与陨石坑关联的进一步证据。

更多后续的支持给出了一致的结论。亚利桑那的科学家们能够证明，在较年老的岩芯中的物质是冲击熔化物，它拥有类似周围海湾中的白垩纪-古近纪沉积物的微玻璃陨石的特征。克林和博因顿还观察到希克苏鲁伯陨石坑中的熔化岩石和海地的白垩纪-第三纪界线处的玻璃沉积小球之间的化学相似性。这是希克苏鲁伯陨石坑精确形成于毁灭生命的白垩纪-第三纪界线的确凿证据。这个时候，足够强有力的证据使这一发现成为头条新闻，并为公众所知。

地质学家接着寻找了尤卡坦陨石坑和白垩纪-古近纪灭绝之间的更多联系。在陨石坑附近，正好在边界区域，扬·斯密特和沃尔特精确认证出一种地理凸起，这是一处含有小球甚至玻璃的一堆角砾岩。玻璃的发现也是重要的。玻璃只能在一个快速的、像撞击一样的过程中形成，而不能在一个相对缓慢的过程（如火山喷发）中形成。在后者的情况下，其中的原子和分子有时间形成结晶。而且玻璃的条纹进一步表明它形成太快所以不均匀。

驻地墨西哥的地质学家们的勘探和讨论揭示出更多被撞击打乱的地区，这些位置都离撞击点较近。研究还表明，位于北美的边界喷溅区的厚度随着距撞击地点的距离增大而减小，这和用希克苏鲁伯作为来源所预测的一模一样。根据来自爆炸的连续抛射物的轨迹性质，地质学家苏珊·基弗（Susan Kieffer）帮助解释了铍、撞击熔融以及冲击石英的相对分布。

到1992年，根据所有已有证据，大多数地质学家相信，尤卡坦的特征地貌的确是一个撞击坑。但是他们仍然无法确定它和白垩纪-古近纪灭绝事件的关系。详细的年龄测定是明确确认这一关联的唯一途径，这需要研究陨石坑中优质岩芯的化学成分。

通过研究岩石中的氩同位素，科学家们设法成功确定了已经存在的岩芯的年龄，特别是来自三个保存完好的玻璃小珠当中的岩芯年龄。然后，他们给海地白垩纪-古近纪层中的球状岩石确定年龄，以检查撞击和灭绝的时间是否一致。第一个测量得到年龄为 $6\,498 \pm 5$ 万年，而第二个测得的年龄为 $6\,501 \pm 8$ 万年，这个结果表明这两个事件是同时发生的（在测量的不确定度之内）。这一出色的一致性让许多科学家相信：沃尔特和其合作者最先提出的恐龙-流星体灭绝理论是正确的。喷出物正好落在古生物边界上，证实撞击与灭绝发生在同一时期。

然而后来发现：最初对陨石坑和铱层的年龄估计（这对建立因果关系至关重要）与真实值偏差了大约100万年。相对日期并没有改变，但是对于指定年龄所必需的衰减常数最初有些许的不正确。这也是为什么我们现在认为白垩纪-古近纪灭绝发生在6 600万年前而不是6 500万年前。

对于流星体假说更强大的佐证是日期排列测量在近期的显著改善。在2013年2月，加州大学伯克利分校的研究人员保罗·莱纳（Paul Renne）及其同事们发现：希克苏鲁伯撞击与生物大灭绝发生的时间间隔少于32 000年。对于发生在如此久远的事件来说，这是一个令人难以置信的精确测量。伯克利团队使用的是氩-氩测年法，该技术依靠的是在之前章节提到的放射性同位素氩，结果表明，撞击和灭绝发生在这个非常小的时间间隔内。

他们的研究结果几乎肯定了：两个事件不仅仅是偶然的巧合，而是对撞击假说的非凡证明。尽管他们谨慎地指出，大灭绝可能是由火山喷发和天体变化而引起的，撞击事件也许只是压断骆驼背的最后一根稻草。但现在看来，则毫无疑问，创造了希克苏鲁伯陨石坑的大质量流星体的撞击才是至关重要的触发机制。

2010年3月，41位古生物学、地球化学、气候模型、地球物理学和沉积学专家相聚一堂，回顾了超过20年的、随着时间推移而积累起来的撞击物种大灭绝假说的证据。他们得出结论：确实是一颗6 600万年以前的流星体的撞击形成了陨石坑，并且引起了白垩纪-古近纪的物种大灭绝，其中最著名的受害者便是古老的恐龙。当年发表在《科学》上的一篇文章给出了对于流星体是灭绝原因的共识看法。几个月后，在同一个杂志上，持怀疑态度的古生物学家们发表了另一篇文章，指出：他们也同意，在那次物种大灭绝中，流星体至少是一个相当重要的贡献因素。

希克苏鲁伯陨石坑是地球上目前发现的最大的陨石坑之一。关于其认证的故事是一个令人难以置信的科学实践的例子，整个过程涉及巧妙归纳、测试和验证大胆假设；与此项研究相关的国家和地区也非常多，其中包括意大利、科罗拉多州、海地、得克萨斯州和尤卡坦。撞击尤卡坦的流星体对地球和地球生命造成了巨大影响。它的起源及其所产生的结果巧妙地展示了地球和宇宙间持久的联系。



正是这些天外来客与地球的碰撞，帮助我们建立了现在的生活方式。

我们绕了好大一个圈子来说明了我们的观点，即暗物质和陆栖恐龙的消失是如何可以联系到一起的。我们介绍了许多人类已经掌握的关于宇宙的知识，像宇宙中的物质，还有像星系这种结构在宇宙中是如何形成的。回到我们的家园地球上，我们回顾了五次生物大灭绝事件，包括被仔细研究又被改编成故事的白垩纪-古近纪灭绝。此外，通过对新发现的小行星和彗星观测，我们还研究了太阳系的组成。

科学的发展并不仅仅只包含已知的内容，更重要的是，它还包含了我们尚未知晓的东西。假说常常开始于尝试性的推测，用来解释那些微小的但有提示性的迹象，抑或在灵感显现时，用来构建出重要的新思想。科学方法的美妙之处在于，它让我们可以有很疯狂的想法——至少表面上看起来是这样，但是同时又要求我们着眼于那些细小的符合逻辑的结论来测试。我们可能会比较幸运，新的想法指出了前进的方向，但我们也可能失望地发现：开始那些很合理的假设使我们误入了歧途。

前进的道路一般都不是平坦的。有一个不经常滑雪但热衷此道的朋友用另一种有点夸张的方式表达了这个说法。当我在山坡上见到他时，他把他的进步说成“前进两步，后退两步”。但说真的，即使他觉得自己技术没有一点进步，但在雪山上不断地练习也能让他对雪山的地形更加熟悉，这也会提高他的滑雪技术。事实上，当我一年后在同样的雪道上又碰到他时，他的水平已经明显提高了。

我的朋友所表达的态度，其实是所有从事研究工作的人都可能遇到的。就算有个从不犯错误的人，他正确地解了所有的方程，恰当地解释了所有数据，但最后他可能发现自己的理论并不适合当前的宇宙，即使论证过程中他没有犯任何错误。尽管如此，和滑雪一样，不断地尝试至少能帮助我们更好地熟悉地形。而我们虚构的那些研究者，也能从他错误的理论中所学到的知识感到欣慰——至少他能确定这个想法确实是错的，即使一开始看起来不是这样。提出假设然后找到证实或证伪的方式毕竟是确认某一猜想有效性的唯一方法。在那些令人赞叹的例子中，一些猜想幸好是对的，并具有一定的启发性，与之相关的工作因此也就有所突破了。对一个科学家来说——对大多数人也是如此，当即将面临成功时，失败就渐渐消失了。

很快，本书就会列举一些关于暗物质的推测。但是本章会简单谈谈我们已知物质所产生的一个最有意思的结果：生命的起源和进化。我会解释一些对生命起源非常重要的因素：适合生命的环境条件，以及生命演化过程中那些外来天体可能的角色。尽管我讨论的许多观点是来自科学研究的，但这其中也包括一些猜想。它们通常是关于某种具体特征对地球上的生命是多么关键，或者在其他地方可能存在的新生命形式。

下面我们会专注于人们熟悉的物质形式，但这并不意味着没有关于暗物质的猜想，我只是暂时把它们搁在一边，并在本书的最后部分来讨论它们。尽管如此，我们完全不应该忽略暗物质对生命的作用。毫无争议地说，暗物质在太阳系周围的恒星环境的创生过程中扮演着

重要角色，从根本上讲，太阳系的恒星环境来自银河系的盘结构，这个盘结构又是由星系产生的，而星系则起源于基于暗物质原初聚集而汇集起来的普通物质。在这些初始的结构轮廓中，恒星和重原子核可以产生，而这在没有暗物质的贡献下是不可能发生的。暗物质还有一个功劳是，它帮助星系和星系团吸引了超新星爆发所产生的重元素，这些重元素是产生地球和生命的关键。

从暗物质到生命的产生是一段漫长的路。先要形成银河系的盘，然后是恒星和重元素，再之后是更复杂的结构。普通物质是这些微妙复杂过程的本质，至少对太阳系来说似乎是这样。虽然我无法告诉你们下面这些关于生命形成的猜测哪些是正确的，但我可以肯定地说：未来，科学还会进步。

生命的起源

生命的起源，是一个极具挑战的问题，特别是至今还没有人知道生命是什么。如果不研究地球上现有的极其复杂的生命，或者极端环境中本不该存在的生命，我甚至怀疑人类是否能猜出或弄明白人类生命的本质，或此类生命存在的必要条件。尽管人们意识到了许多深刻的、基本的问题尚待解决，人类还是经常高估了自己的已知范围。我发现，人择原理令人烦恼的一个原因是：对于任何形态的生命来说，还没有人知道其存在的可能的根本原因，也没人知道像星系这样的结构可以产生生命的根本原因。我不像有些人一样相信其他形式的生命和人类的生命是相似的。

在对想象中的抽象生命形式发问之前，我们可能首先想知道：

- 地球上的生命是如何开始的？

●从哪里开始？

●它是从地球上起源的还是从外太空飞来的？

一些人猜测，彗星或小行星带把已经形成的生命通过孢子带到了地球，这种情况被称为泛种论（panspermia）；有一些人认为，外来天体的影响帮助生命形成跨越了一些障碍；还有一些人更加保守，认为地球生命的发展没有直接受到地外作用的影响。最后一种假说有个优势，在太阳系内我们所有已知的地方，地球的条件似乎是最有助于生命出现的。尽管其他地方也可能存在类似的环境，但至今为止我们只发现地球上存在浅海的环境，例如潟湖或潮水坑、冻结的水溶液或者黏土的表面，在那里化学成分可以富集并反应。

组成生命的重元素肯定来自外太空。氢是在宇宙极早期时出现的，但是其他重要元素——碳、氮、氧、磷和硫，仅仅通过高温致密的恒星合成或超新星爆发产生，而且都是发生在太阳诞生之前。

我很高兴在一次采访中列举这一系列事件，那是在加那利群岛特纳里费岛（Tenerife）的望远镜，学生们组织的一次寻找近地小行星活动上。在常规的问题结束后，他们还为每一位受访者准备了一个古怪的问题：“你认为学生和年轻的恒星有什么共同特点？”当采访人表示满意我的回答时，我着实松了口气：“学生们学习新思想，并把它们传播给其他人，由此产生了循环——这很像恒星吸收星际物质产生了重元素，然后把这些重元素抛回宇宙空间再循环。当分子物质被排出散布到星际介质中，便聚集在稠密的云中，在那里一部分分子物质重新进入恒星形成区。这种循环模式很像思想的产生、传播和发展。”

不过在生命出现之前，重元素还需要进一步加工。在地球上，这些化学元素组成了越来越复杂的、稳定的有机化合物，最终产生了可以自我复制的RNA，然后是DNA，再后来是细胞，之后接着是（这期间花了很长时间）多细胞有机体。这些有机体一部分是由氨基酸组成，这一组成蛋白质的基本单位。随着对形成DNA、RNA以及细胞结构的必

要条件的深入了解，我们更加确定：那些极端条件对生命的起源必不可少。

关于生命的出现有许多有趣的问题，其中一个氨基酸是如何在星际介质和其他地方形成的。

20世纪50年代早期，芝加哥大学的斯坦利·米勒（Stanley Miller）和哈罗德·尤里曾做过一个著名的实验。他们给一个密封烧瓶中的水加热，这个容器中还包含甲烷、氨和氢。他们的目的就是模拟原始的地球大气下的海洋。在这个人造的“大气”中，对水蒸气放电的过程扮演了闪电的角色。米勒和尤里用这个简单的装置成功地制造出了氨基酸，证明了氨基酸的产生在太阳或太阳以外的环境中并不是什么出人意料的现象。

早期地球大气中很可能含有二氧化碳、氨和水，还有在试验中依然稳定的甲烷和氨。有意思的是，氨基酸在地球的分布和米勒-尤里试验所得到的结果出乎意料地相似。他们实验结果的关键结论是：在地球上形成有机物质是相对简单的，在太阳系或其他星系也是如此。需要记住的一点：化学中“有机”（organic）一词仅仅表示碳的存在，并不一定是生命的必须元素。据我们所了解的，某些（但不是全部）有机分子对生命是必须的，所以“有机”这个词的命名并不是巧合，尽管有些“不幸”。

事实上，包含碳元素的反应在宇宙中几乎无处不在。恒星的喷流、星际介质、稠密的分子云和原始恒星云，全都包含有机物。像太阳这样的恒星在周围制造大量的有机物，在稠密的分子云中也是一样。这使得有机合成相对来讲并不那么令人吃惊，但同样也使生命本质成分的起源更难下结论。这些成分中的其中一些可能产生于其他地方，但是一些科学家相信，许多有机物更有可能是土生土长的——或

者至少是从那些先从地幔中重塑的物质中产生的，然后才进入分子中去，并最终对生命有了贡献。

我们确实已经知道：至少一些有机物是被太阳系中的天体带到地球上来的。在小行星带内的有机物总量似乎显著地小于小行星带外侧，这是猜测地球上一些可观总量的有机物是来自外太空的原因之一。另外一个原因是，尽管地球早期的物质大都没有保留下来，但月球上数量巨大的陨石坑以及相比之下地球的尺寸比月球大得多的事实告诉我们：在早期时候，地球上也有许多陨石事件。这很有可能给地球带来了可观数量的有机物。

氨基酸还有嘌呤和嘧啶都确实在太空中被发现了，其实嘌呤和嘧啶也是产生DNA和RNA必需品。小行星和彗星上都发现有氨基酸，它们中的一些存在于生物体内，有一些则没有在地球上被发现过。至少有一种方法来鉴别非生物氨基酸，即旋向性（chirality）或者说分子的手性（handedness）（见图13-1）。在地球的生物体内只存在左旋性的氨基酸，而太阳系外圈发现的氨基酸包含左右两种旋性。手性与碳原子周围原子的排列有关，它有一个显著的方向性，就像你左右手的手指排列方向不同一样。然而至少有一例研究发现：在小行星的沉积物中，有一类氨基酸的左旋分子数量要多，多出的左旋分子和生命的联系更加难以判断。

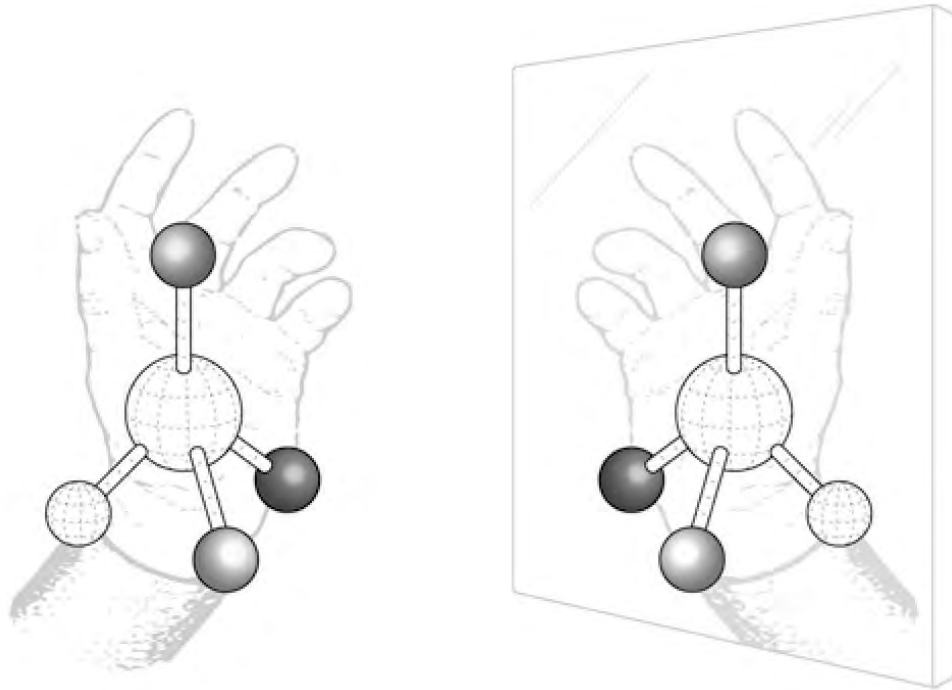


图13-1

给定旋向的分子的手性示意图。分子的旋性和镜子中分子的像的旋性是不同的。地球上生物的氨基酸是左手旋向的。

我们关于小行星上氨基酸的知识很多来自默奇森（Murchison）陨石，它是一块在1969年掉落于澳大利亚墨尔本附近的默奇森小镇的陨石。这块陨石是来自火星和木星之间的小行星。它被归类为“碳质球粒陨石”（carbonaceous chondrite），从这个名字你大概可以猜到它含有大量的有机分子，包括氨基酸。巧合的是，这家实验室还研究了阿波罗计划带回的月球物质样本。因此这些科学家们可以把默奇森陨石和其他类似的陨石相对比，例如在俄克拉荷马州发现的默里湖（Murray）陨石；还可以对比其他不同类的陨石，比如在法国发现的奥盖尔（Orgueil）陨石。

为了研究来自太空的氨基酸的命运，科学家们尝试在试验中重现宇宙的条件。研究显示，氨基酸能够在与彗星的碰撞中被保存下来，或者在地外物质撞击地面时产生。通过观察彗星的除气作用，科学家们发现：大部分小行星被星际材料多次加工过，一些彗星状的冰块包含着早期的原始星际物质。对陨石以及行星际尘埃的研究能帮助科学

家们了解彗星和小行星带到地球上的物质，也能帮助科学家们建立一些有关从太空飞来的分子的起源和数量的数据库。

水像碳一样，大概也是太阳系中产生生命的关键，哪里有水哪里就可能有生命。地球有一个特别值得注意的特征是，地球表面2/3被海洋覆盖——不是全部，也不是一点没有。海洋的部分覆盖使海岸线和潮汐区域成为可能，这些区域对生命的出现与发展非常重要。

正如我们所知道的，水对生命非常关键。岩石中的证据表明，在地球表面液态水已经稳定地存在了相当长的历史。岩石的纪年表明，水的存在可以回溯到38亿年前。更古老的锆石形成的一种结构似乎也需要早期地球（至少在43亿年前）有水的存在。

地球上的生命肯定受益于巨量的水，无论是什么东西将水带到了这个星球上。然而，正如一个最近迷上渡口骑行的朋友所说，我们周围这些巨量资源到底是从何而来的还是一个谜。海洋中的一些水可能来自地球表面下的岩石中，但想用这个假设来解释今天地球的存水量，还远远不够。

我们已经看到，陨石撞击可能带来的有机物质会促进生命的产生。在地球形成的早期，很可能是在晚期重轰炸期间，通过彗星和小行星所带来地外水当然也是可能的。这是一个棘手的问题，因为大部分通过陨石到达地球的水都是储存于矿物中的结晶水合物，所以需要一些过程把水分子从硅石上分离出来，尽管也有一些石缝间的冰被小行星带到地球上。

最初，彗星看起来更像是水的来源，因为它们的成分中很大一部分是冰。然而地球上的碳、氢和氧的同位素似乎和彗星上观测到的结果并不吻合，这表明彗星可能不是地球挥发物的主要来源。2014年，罗塞塔号卫星发回的结果证实了这一点，它测得的彗星上氢的同位素

的组成与地球上的不一致，这使得水主要来自彗星的假设变得不太可能。如果来自太空的天体确实扮演了重要角色，那么这些天体更可能是相对遥远的小行星，它们的同位素比例和地球更接近。这一点很重要。

另外一个关于水的问题是：地球形成早期，年轻的太阳输出的能量大概是今天的70%。在原始的太阳低光度的条件下，若没有其他条件的影响，即使水已经形成也不会是液态形式，这就是所谓的“暗淡太阳悖论”（The Faint Sun Paradox）。然而，年轻的地球可以产生热量，并通过很多方式释放出来：坍缩的引力能、火山喷发、穿过大气的陨石震荡，还有月球的潮汐热能（那时月球离我们比现在还近），此外还有地球内部不稳定同位素的放射性衰变。这些过程中的任何一种都可能使地球更温暖一些。大气中的温室气体像二氧化碳，能吸收一些太阳光，这些光开始是在可见光范围照射到地球上的，又以红外形式反射被大气层吸收。无论温室气体能否解释地球早期温度高于估计值，那时已经有液态海洋是非常清楚的事实了。所以，上述因素中的一个或几个因素肯定是起了作用。

宜居环境

在宇宙环境中，既有朋友也有敌人，太阳系内外都是如此。生命存在似乎依赖于本地的物理条件是否能够造就一套适合成长的生态系统。构建宜居生态系统的过程需要一些特别的条件，这些条件既能促使有利的方面发挥作用，又能改变或抑制那些不利的因素。理解生命的先决条件很可能被视为和理解生命起源一样令人畏惧。但是科学家们无论如何都希望找出“什么是宜居的环境”的答案——既包括基本的微生物，也包括更复杂的生命，而后者大概需要更特殊的条件。尽

管还没有人知道所有的答案，但所有使我们生存的环境如此特殊的东西都值得我们注意。

还有一点值得注意：太阳本身也有一些特殊。在大质量恒星中（比如恒星质量序列的前10%），以太阳的年龄来说，它的金属丰度比典型值要高，与星系盘中的恒星异乎寻常地接近。此外，与类似年龄的恒星相比，太阳的轨道似乎更圆了一些，还有它的位置使其轨道更像旋臂，因此很少与旋臂相交。我们不知道这些太阳的非典型特征有多重要，但是所有的非常规特点都可能是有趣的。

光合作用依赖于太阳的辐射，对地球生物至关重要。能量几乎对任何形态的生物都非常关键，因为它为制造并最终维持生命的过程提供能源。在地球上，太阳毫无疑问是主要的能量源。太阳光提供的能量是第二名（即地热能）的几千倍。其他更小的能量源包括闪电和宇宙射线，前者特供的能量是太阳提供能量的几百万分之一，而后者则是其几千分之一。

“液态水对所有生命形式都重要”的说法虽然只是一个推测，但对地球上存在的生物来说肯定是这样。我们想知道这些水是从哪里来的，还想知道水的液态形式在哪里会是稳定的。讨论这个问题不仅需要了解太阳到地球的距离，还要知道辐射的效力、其他可能的热源以及大气的压力。

仅仅根据地球的反射率、太阳的光度和到地球的距离，人们可以推算出：如果没有大气的温室效应，即使在今天，地表的水也是要冻冰的。尽管今天我们担心大气内的过多热量，但如果没有二氧化碳、甲烷、水蒸气和一氧化二氮的温室效应，地球将会变得非常寒冷。地球上液态水的存在仅仅是由于这些温室气体，它们吸收了地球的红外光，从而建立了平衡机制。

“宜居带”就是指一段区域内其条件满足生命能够存活。这是一个液态水能稳定存在的“金发女孩”区域^[29]。如果离热源（太阳）太远，水就会结冰；如果离得太近，水一开始就不会聚集在行星的表面。水也可能存在于地表之下，尽管这种情况不太可能像巨大的海洋一样，能够为种类众多的生物提供一个共同聚集的场所。

如果水的存在是宜居带的一个标准，那么宜居带的外沿界限的定义就是：大气层外二氧化碳开始凝结的位置。这给出了一个以太阳为中心、半径比日地距离还长1/3的区域。还有其他定义方式：在大气层内有足够多的二氧化碳和水，以确保水不结冰。这种定义给出了更大的范围，大约把宜居区域的半径较日地距离延长了2/3。具体来看，金星在两种定义中都属于宜居行星，火星只在第二种定义时是宜居行星，而其他的外圈行星都太远了。

尽管不清楚水是怎么出现的，但我们知道，几乎在行星出现时，水就已经存在了。然而太阳的光度发生了变化——自它形成以来有了显著的增加，大气层也发生了变化。因此有了一个对“持续宜居带”（the continuously habitable zone）更加严格的限制，在这个区域内，液态水可以在行星的整个寿命期间持续存在。根据当前的气候模型，连续宜居带的宽度被更严格限制在日地距离的15%之内。当然这是今天的定义。再过40亿年左右，太阳会变成一颗红巨星，然后再过几十亿年，它就会燃尽。根据现在的模型，地球上的生物没有哪个——无论简单生物还是复杂生物，能存活到那么远的未来。

在担心遥远而凄凉命运之前，我们还有更紧迫的事情。一个关键是地球温度的稳定性，我们知道这对生命意味着什么。在当前的社会，相对小的温度变化就会产生巨大的影响，例如海岸线、农业以及对于人类的适居性。但是对生命演化来说，人们倾向于使用了相对粗

糙的温度限制来衡量地球温度的稳定性。在地球上，碳是关键，大气中的碳必须有稳定的补充。

在其他行星上，甲烷和二氧化碳云也可能是相关的。在这个星球上，调整大气层中碳的反应至关重要。碳可以从大气中被除去，例如溶到雨水中或者被植物的光合作用吸收；还可以循环回到大气中，例如通过板块构造活动、岩石的风化。当在大洋中脊处诞生的海床随着地壳的运动消失在大陆板块间的潜没带时，此处的元素会发生反应从而产生二氧化碳，并通过火山喷发、温泉和其他地壳出口又回到大气中。山脉的产生和上升也会缓慢地释放碳。快速释放碳的方式是燃烧化石燃料。所有这些过程都影响着大气中碳的供给，这对调节地球温度至关重要。

长久的气候稳定可能是生命进化的另一个前兆。地球上，这种稳定不仅依赖海洋和内热源来驱动板块的结构演化，从而产生了温室层，还有恒星的演化、较少的小行星和彗星撞击，以及月球的存在稳定了地球的自转轴。在过去的5亿年中，这些条件很可能都是生命形成的关键因素，尤其对那些更大型的动植物来说，尽管在最初的30亿年里，气候的稳定可能对早期微生物也非常重要。

一个稳定的中心球对生命的出现也很重要。太多的宇宙射线射到地球上——很多小行星、彗星也一样，许多类型的生命没有机会诞生。任何成功出现的生命也可能很快就被消灭了。一个有生命的行星必须要离太阳足够远，避免过强的辐射；但又要足够近，可以被外圈的行星保护而不受小行星影响。不管这个条件是否是必须的，木星肯定扮演了地球大哥（或保镖）的角色，保护着它的“小兄弟”不受地外天体的攻击，使地球上生命的进化更容易一些。

保护地球的还有太阳风（在第8章讨论如何定义太阳系的边界时也提到过太阳风）。太阳风与星际物质发生反应产生了日光层。在这

个区域内，星系的宇宙射线数量相对低一些，这使得地球的气候稳定，也保护了所有出现的生物不直接受到来自宇宙线的毁灭性影响。

令人惊讶的是，我们当前生活的区域（大约有300光年的跨度）被称作本星系泡（Local Bubble），位于银河系猎户座旋臂上，是一个像真空一样的范围，其中的星际介质里的氢密度非常低。直到最近——大概几百万年前，我们才进入这个温暖的、低密度的、部分电离的区域，这是一种相对少见的星际环境。在此期间，这个区域被已经变得特别大的日光层边界包围着，太阳风主导着这一区域的星际介质。我们并不知道人类出现在地球的时间是不是一种巧合，人类出现的时刻恰好是在地球进入到本星系泡的空洞的时刻，或许正是这种不寻常的低气体密度和宇宙射线密度，对形成复杂的生命大有帮助。

外来天体和生命的进化

希克苏鲁伯陨石坑的陨星肯定扮演了重要的角色，在生命进化的后期，它造成了当时大部分物种的灭绝，并为其他物种的进化铺平了道路。尽管具体的数字不是很准确，但似乎大部分大型陨星落地的时间和物种大灭绝的时间是相吻合的。发生在物种灭绝时域边缘的铱层、微玻陨石和冲击石英增添了陨星撞击造成物种灭绝可能性的证据，这个课题值得进一步研究，因为陨石坑的出现时间似乎和陆地生物改变的时间确实存在一些巧合。

尽管如此，下面的许多想法还只是推测。有人提出，虽然彗星引发的后阿尔瓦雷斯波（post-Alvarez wave）非常强大，但小行星和彗星的撞击肯定不足以解释行星上生物的灭绝或者起源。白垩纪-古近纪灭绝事件是仅有的、可以相信的彗星撞击造成的灭绝。气候变化和大量的火山爆发在几个时期的生物灭绝事件中也扮演了重要角色，这些

证据可能比彗星撞地球更令人信服，包括早寒武纪、二叠纪、三叠纪和中新世中期的物种灭绝都更倾向于这种解释。所以请读者对于我将要列举的猜测不要过于激动。确实有证据指出，撞击与灭绝是一个互相联系的系统。因为一些大型撞击发生的时间大致也相当于地球年龄、生命的起源、文明的开始。研究我们能找到的两者之间任何可能的联系都是值得的，即使证据不是那么强有力。

在五次大灭绝中，如果从地外影响的证据角度看，发生在泥盆纪结束期的灭绝（大约发生在3.6亿~4亿年前）是仅次于白垩纪-古近纪灭绝的一次物种大灭绝，历史上排在第二位。在这一时期内，大概发生了多起地外天体撞击事件，且很可能是由一个碎裂的小行星引起的，或者是多起彗星撞击引起的扰动（我们很快会说到）。但是精确的时间测量并不支持彗星是这次灭绝的主导，因为在这个时间范围内，物种的减少似乎是由于物种形成的有限而不是实际上的灭绝。有意思的是，1970年，也就是沃尔特提出“地外天体撞击导致的白垩纪灭绝”这一理论的很多年前，古生物学家迪格比·麦克拉伦（Digby McLaren）已经提出了地外天体的撞击，可能是这个更早的物种灭绝的原因。

许多联系了撞击和灭绝的事例解释了一些小事件，例如7400万年前在北美地区的局部灭绝。许多鳄鱼的种类、一些水生爬行动物、哺乳动物，还有部分恐龙在这期间灭绝了，似乎和艾奥瓦州的曼森（Manson）陨石事件符合。始新世晚期事件大概是在3500万年前，发生了多起海洋生物、爬行动物、两栖动物和陆地哺乳动物的灭绝，大约也和一些流星陨石撞击事件相符。证据包括在俄国的波皮盖陨石坑，在华盛顿特区附近的切萨皮克湾近期发现的90公里宽的陨石坑，还有在新泽西大西洋城外的一个小一些的陨石坑。华盛顿的那个陨石坑是在一个大型石砾场的沉积物下发现的，这个石场是陨石撞击的海啸所产生的。接着，科学家们还应用地震测型和钻探岩心等方法对其进行了

更深入的研究。高于常规的铷等级和超高含量的行星际间尘埃显示：彗星雨是此次多起撞击事件的“责任人”。

始新世晚期的事件也有地外干涉的证据，不过是用不同的方法（地球化学的证据）得到的，这一方法可能是最终弥补那些少的可怜的撞击纪录的有效途径。加州理工学院的肯·法利（Ken Farley）及其同事找出了一种研究天体撞击事件的新方法：通过测量氦的同位素含量来追踪行星际间的尘埃，因为氦的同位素的含量在彗星雨期间会升高。他们的结果很有意思，3 600万年前产生的波皮盖陨石和切萨皮克陨石前100万年到之后的150万年间，氦-3的含量是有所提升的。这些尘埃是彗星雨的强有力证据，它们可能是由奥尔特云中的某次扰动引发的（我们会在后面几章回到这个话题）。

现在我们把这个撞击猜测的列表补全。在中新世晚期，大约1 000万年前发生过一次小型灭绝事件，它和铷异常和玻璃陨石似乎相吻合。有意思的是，法利也发现了此次事件前后的氦-3含量的增多。在这个例子中，合适的时间加上尘埃的演化和小行星撞击的猜测非常吻合——特别是已知的碰撞产生了威尔塔斯族（Veritas）小行星。

因碰撞而产生生命的证据比毁灭生命的证据还要少，但有些人还是乐在其中。我将会提到一个想象的可能性，这是在《圣经》和神话中提及过的戏剧性事件，以及尚无解释的史前遗迹，如巨石阵，这些看起来很神秘或者有人为了某些目的而创造出来的神秘事件，也可能起源于地外流星的撞击。科学地看，研究者们认为，早期撞击事件可能炸飞了部分大气层甚至海洋，从而减慢或限制了地球上生命的进程。但这些也可能制造出了适合生命的环境，例如产生了水热对流系统，促进了生命起源前的化学反应。

科幻作家查尔斯·弗兰克尔在《恐龙灭绝》中指出：20亿年前的前寒武纪的复杂性和两次已知的那个时候的巨大陨石坑之间可能存在

联系。尽管不是很确定——因为氧的角色在那次事件中可能更关键，但是他指出，时间的吻合性很吸引人。另一个同样遥远的可能相关性是一次5.5亿年前的碰撞及其后发生的寒武纪生命大爆发（这里，大爆发是指逐步扩大的生物多样性）。科学家们相信，也许这次撞击消灭了已经存在的物种，从而给新物种腾出了生存的空间。尽管没有一种已知的机制和生命联系起来，但是在澳大利亚和其他地方也可以找到撞击的证据。澳大利亚的阿克拉曼陨石坑（Acraman Crater）是个直径超过100公里的湖，被一层喷发物所包围，这些喷发物中包含铍和冲击石英，并向东延伸了300公里，一直到伊迪卡拉纪的化石群，而寒武纪生物大爆发紧接着这个化石群的形成就发生了。更多证据在中国三峡地区也有。引人注意的是，三叶虫的化石出现在紧挨着撞击层之上的一层。这表明，不管怎样，某些事件种下了来自异星的种子（带来了地外的元素），复杂的生命立刻就开始在海洋中出现了。

另外一些猜测基于已经成为化石的陨石的成分、受冲击的物质组成以及陨石坑的观测数据，这些数据都明确地指出奥陶纪存在一系列的撞击事件，这组撞击事件在4.72亿年前，即奥陶纪的中期达到峰值，相关数据特别精确地吻合了海洋生物物种的一段蓬勃发展期。陨石化石的想法给人留下了深刻的印象，所以这里我会介绍一下这个发现，尽管它和生物发展的巧合是明显的猜想，不足以当真。最早的线索来自一个孤立的大圆石，它于1952年在瑞典的一处沉积岩中被发现，但显然它不属于那里。研究人员花了25年才确认它是一块陨石的化石，这块陨石的成分除了对风化的抵抗力比较强的铬铁矿（一种岩石的形式）之外都被侵蚀掉了。后来，科学家们又在周围发现了将近100块陨石，合起来就是5亿年前的一块100公里~150公里宽的物体，它造成了大量的陨石和微小陨石尘埃飞落到地球上，这种情况持续了几百万年。这些碎片甚至可能形成了一个行星带，现在这些小行星还在缓慢而持续地落向地球表面。

上面提及的关于地外星体对生命的产生和灭亡作用的想法中，其中一些是有问题的。但是我会以一个可靠的地外星体作为这部分的总结，那就是：坠落于地球的地外星体是我们这个星球上的重要资源。有意思的是，陨石所带来的物质对社会很重要，特别是在铁器时代之前，早期人类会使用陨星铁来制造工具、武器和文化用品。

这些矿物质现在也仍然非常重要。许多金、钨、镍和贵重元素能够从地壳中获得，就是因为这些天外来客砸开了地球表面。尽管行星和小行星的组成是相同的，但地球的引力会把重元素吸到它的地核中，这些重金属中的大部分不会回到地球表面上来。地表的重金属物质主要靠来自地外并坠落在地球上的陨石加以补充。大概有25%的流星撞击造成了潜在的有益沉积，这些沉积中至少50%已经被开发利用了。因此，即使流星撞击地球不是产生生命的必要因素，它们在生命的形成及进化过程中的重要性也是毋庸置疑的。正是这些天外来客与地球的碰撞，帮助人类建立了现在的生活方式。



外太空物体会以规律的周期撞击地球，间隔大约在3000万~3500万年之间。

20世纪早期，物理学家欧内斯特·卢瑟福（Lord Rutherford）以他划时代地发现原子核而闻名于世。他曾有一句名言：“科学要么是物理学，要么是集邮。”尽管这一观点有些傲慢和令人不快，但这句话的确包含了一些真理。科学不仅仅只是列举一些现象，无论这些现象是多么漂亮或卓越，而是去尝试理解这些现象。科学家们可以通过精巧和先进的方法收集事实，例如今天的生物学家们使用DNA序列和其他技术促进了数据的快速积累。但是仅仅当数据被彻底理解后，这些信息才会变成真正的科学，即最好能建立一个完备的理论来解释现有的数据，并根据这个理论给出可检验的假说和预测。

我们已经研究了在太阳系中的天体，并知道这些天体中哪些与地球发生过撞击，以及从化石记录里得知了生物灭绝的相关信息。一个好的科学研究会提炼、理解并诠释所有的数据。但是一些重大问题还

是没有解决，例如“哪些现象是有联系的”“如果有，那又是如何联系的”。



一个最吸引人但非常具有猜测性的天体物理与生物灭绝的联系是：外太空物体会以规律的周期撞击地球，间隔大约在3 000万~3 500万年之间。如果这是真的，这个周期将会是一个非常重要的线索，科学家们可以根据它来研究那些安全轨道上的物体偏离轨迹的原因，并知道这些偏离了安全轨道的天体已经变成了一个潜在的危险炸弹，且有可能飞向地球。有许多“某种扰动导致了这种周期性”的猜想，但是没有哪个猜想能给出一个与已经发现的陨石坑记录吻合得很好的碰撞周期。

米兰科维奇循环

以地球物理学家、天文学家米卢廷·米兰科维奇的名字命名。是一个地球气候变动的集合影响，以10万年为主要周期，与地球绕日运动轨道的三个变化有关，即：地球公转轨道离心率的变化、地球自转轴倾斜角度的变化、地球自转轴进动的变化。

我会尝试简短地以卢瑟福的观点来评判陨石撞击事件是否展现了足够强的周期性，而这个周期性是否拥有科学的解释。我首先会指出一个不相关但已经建立得很好的关联理论：关于地球在太阳系内的周期运动和地球气候的周期性变化之间的联系。这个温度变化周期，被称为米兰科维奇循环（Milankovitch cycles），是发生在一个短得多的时标上，当然这是和我后面将要讨论的时标相比。这个是以塞尔维亚的地球物理学家、天文学家米卢廷·米兰科维奇（Milutin Milanko

vic) 的名字命名的，他在第一次世界大战期间建立了这个理论，而当时他还是一名囚犯。

米兰科维奇研究了地球公转轨道的离心率、地球自转轴倾斜角度和地球自转轴进动的变化对气候的影响。基于这些考虑，他和后来的科学家们建立了一个以大概2万年和10万年为周期的温度变化模式，全球冰河时代的存在印证了他们的发现。在我访问西班牙巴斯克地区的苏玛亚时，当地的向导为我指出了岩石结构中清晰的层面。这些层面同样是温度变化的结果，因为温度的变化导致了沉积率随时间周期性发生了变化。

虽然有米兰科维奇循环被发现这一成功案例，但对陨石坑周期的探索（发生在一个更大的时间尺度上）必然是一项需要极大勇气的事业，而我不想过分吹嘘它。地球上几百万年前发生的事，保存到今天的证据已经十分稀少了，而且有很大不确定性，例如它们发生的确切时间。只有在极少数的情况下，这些远古的事件才会保存下些许的信息，而那些留下足够多信息、让人们能从细节上理解的的事件就更少了。然而，只要假设与发现的数据不冲突，而且能够给我们带来对世界的进一步认识，科学家们对它们的探索就是有意义的。好奇的人们不但想知道发生了什么，而且想了解背后可能的原因。

我们现在将要考虑的课题是大型的、规律的、来自空间的撞击，它们发生在几百万年的时间尺度上，并希望把这些同我们的运动轨道联系起来，当然这里的轨道指的是太阳系在银河系中的轨道，而不是地球在太阳系内的轨道。我们通过研究陨石坑来尝试着解释所观测到的数据，并期望基于这些研究更好地理解太阳系和银河系的动力学，以及它们之间的内在联系。最有趣的假说应该能给出能被检验的预言，这些假说和怀疑论者所认为的假说是两码事。虽然许多关于周期性的想法还处于猜想阶段，但本章的目的是仔细地解释什么是我们认可的，以及为了进一步研究我们还需要什么。

建立周期

马修·里斯（Matthew Reece）和我没有立刻着手研究暗物质能够解释太阳系内某些周期性现象的可能性。在介绍我们的观点之前，我们想先确认一下这些周期性现象的证据是否足够充分，并且是否值得做进一步研究。另外一个重要的考量是：我们的工作是否对指导将来的观测和分析有所帮助。

开始时，我们在我凌乱的办公室里讨论了仍然有些混乱的想法，目的是搞明白什么目前我们已知什么、下一步研究的最佳方案是什么。我们工作的第一步就是调研周期性的证据然后看看它是否可靠，或者周期性是否仅仅只是科学家们随口抛出的一个有趣说法。

我们参考了许多前人的研究。在大量文献中搜寻观点、区分出主张和真相，比想象的更具挑战性。通常一个结果之后会出现另一个结果，例如有些研究者在一些论文中找到了周期性的证据，而另外一些人在前人的结果中却发现了错误或者疏漏的地方。争议会继续，并没有实质性的解决。在我们完成了最近的论文后，那些对周期性证据持怀疑的人确实让他们的观点为人所知。幸运的是，我们没有什么私心，而只是对这个问题好奇，这使得我们更客观一些。

这个工作中的数据处理所需要的统计分析确实不简单。地质记录非常少，不可避免地有一些大的空白期。由于数据的不完整，研究者分析记录时定义精度的方式会影响其结果。脚踏实地把数据视为根本是吸引人的，但是许多解释注重如何展示和评估测量，而这种测量的统计性质却较差。

例如，数据的分组会产生差异。当科学家观察数据随时间的变化时，他们作出一些关键的选择，而这些选择能影响后面的结论。比如，应该用多少数据点，在哪个具体的时间间隔上应该放哪个数据。科学家们还需要估计这些事件持续的长短，还要理解增强的信号强度的影响，因为正是他们选择了在某段时间内的信号强度。

回应周期性证明的论文还强调了许多统计错误，它可能影响研究的结果。在德国海德堡，马克斯·普朗克天文研究所（MPIA）的科兰·贝勒-琼斯（Coryn Bailer-Jones）是这个领域的翘楚之一。他提出了很多异议，包括上面提到的那些。他还很关心“确认偏差”（confirmation bias）问题，即人们倾向于注意和报道他们认同的结果。贝勒-琼斯认为，也许有些作者过多地尝试拟合了这个结果，因为这个周期和太阳系运动周期很接近。尽管他的很多异议是正确的，但是这个周期的接近并不一定是个坏事。数值的吻合可能仅仅是巧合；或者可能暗示了某些内在的科学联系，并将会被进一步的理解。

贝勒-琼斯等人还指出了一个常见的错误：你不能拿一个假设和一对简单的矛盾模型进行比较，然后认为这是简单的二选一，而忽略所有其他选项。例如，人们总是问，数据和哪个模型拟合得更好一点？彗星周期性地撞击地球的假说或者撞击的概率是否随着时间长度的增加而变为一个常数？即使周期模型比完全随机的假说更好，数据也可能更好地遵从一个不同的模型，例如越古老的天体撞击导致的陨石坑被找到的概率就越小。换句话说，在简单的二选一情况中，更好的模型并不意味着它就是对的。幸运的是，研究者们可以扩大可对比模型的列表来避免这个错误。如果没有出现明显不同的概率，尝试多种模型并测试周期性模型是否最好，是可行的。

确认周期性的信号至今还有一些障碍。1988年，地质学家理查德·格里夫（Richard Grieve）和他的同事们指出，不精确的日期能够抹去所有的周期性信号——无论这个信号是否是真实的。1989年，普林斯顿大学的一名本科生朱莉娅·海斯勒（Julia Heisler）和当时在多伦多加拿大理论天体物理研究院工作的斯科特·特里梅因教授（Scott Tremaine，现在是普林斯顿高级研究院的主任）进一步量化了这种效应：在多大的不确定前提下依然能可靠地确定周期性现象。在1989年发表的一篇论文中，海斯勒和特里梅因认为：数据中13%的不确定性

就不可能得到一个好于90%的置信度的周期性。如果不确定性达到23%，那么测得周期信号的可能性就降到了55%。这样的不确定性不会使建立可靠的周期效应变为不可能，但却会使问题变得更有挑战性。

灭绝事件的周期性

这些有警示作用的论文的焦点集中在天体物理学中的周期效应，这也是我工作的焦点。但最初引起人们研究陨石坑随时间变化的起因是一个表面上看起来不一样的题目：灭绝事件明显的周期性。普林斯顿的地质学家阿尔弗雷德·费舍尔（Alfred Fischer）和迈克尔·亚瑟（Michael Arthur）最先发现，生命似乎有规律地呈现出繁荣和衰落的迹象。1977年他们总结出，化石记录似乎符合一个3 200万年的周期。芝加哥大学的大卫·劳普和杰克·塞科斯基在1984年发表了一篇更有影响力的论文，他们展示了其对周期性灭绝的研究。一开始劳普和塞科斯基发现了一个可能的周期范围：介于2 700万~3 500万年之间。之后他们又重新分析，修正了这一估值，得到了2 600万年的周期，之后大多数科学家们也都得到了这样的结果。

任何激动人心的观点都会被不断检验，而后来的研究确实发现了支持的证据，不过时标上发生了一些变化。2005年，加州大学伯克利分校的两位物理学家罗伯特·罗德（Robert Rohde）和理查德·穆勒（Richard Muller）使用同样的化石记录，但重新校对了时标，得到了不同的周期结果，即6 200万年。有趣的是，后来的结果变来变去，但2 700万年和6 200万年却都被保留下来。近期一次细致的分析，是由堪萨斯大学天文学教授艾德里安·梅洛特（Adrian Melott）和华盛顿特区的史密森尼国家自然历史博物馆的古生物学家理查德·班巴奇（Richard Bambach）完成的。他们发现，大部分灭绝发生在2 700万年周期模版的300万年内，此外在6 200万年的框架上总是发生物种多

样性的减少。这表明，这两个时间框架可能实际上都与周期撞击事件相关。所有对周期性的警告还是适用的，但支持周期性的微弱证据也的确存在。

即使化石记录的规律被证明是真的，仍然不能改变一个事实：没有一个支持周期性灭绝的学者能够解释为什么灭绝是周期性的。导致物种灭亡的因素有许多，气候变迁、火山爆发、天体撞击和板块活动等都扮演过重要的角色。流星体可能导致大范围的物种灭绝，而且白垩纪-第三纪灭绝事件肯定是流星体造成的。但是灭绝事件中任何一个被提到的周期性，都不是单一原因的结果。如果想了解这背后明确的物理机制，科学家们最好期望所得到的周期性是不同周期性现象叠加的结果；然而若是记录不够完备，叠加后的结果看起来会很像随机结果。

任何把可能的周期性灭绝和引起它们的物理过程联系起来的尝试，都比根据具体物理现象来理解周期性更具不确定性，例如只把外来天体的碰撞与周期性灭绝联系起来。研究彗星撞击已经足够有挑战性了。把它们和灭绝事件的不确定联系起来肯定是像兔子窝一样^[30]。

由于这些不确定因素——除了唯一确定的彗星/白垩纪-第三纪联系，后文将会避开关于灭绝的猜测，尽管这个题目非常吸引人。取而代之，我将会把重点放在如何联系宇宙中的周期现象和那些足以留下陨石坑的大型陨石上。这些研究的好处是，陨石坑的记录是和天体物理直接联系的——不像造成灭绝的潜在因素，不受期间天气、环境和生物学等因素的影响。

天体坠落提供了一个绝佳的机会来研究地球上的现象和太阳系中发生的事件之间的联系，并可以把它们看作一个整体，这为更好地研究宇宙提供了一个独特的途径。不定期的流星体并没有什么特别的解释，而周期性的流星体则很有可能暗示了什么。如果流星体出现的

时间真的是规则的，那么时域相关性则很可能反映着背后的宇宙学起因。

在第21章我将会谈到我和同事推定的一个理论，它预言的周期性已经得到了现有数据相对充分的支持，在将来，会得到更可靠的数据对其进行验证。这里我会讲述一些过去文献中的代表性结论，并不涉及精确的统计方法或者数据选择。

我们将看到，历史文献中展示出的一些支持周期性的证据，但是它们不足以给我们以确凿的结果。在有了更好的数据和更仔细的分析后，那些模糊不清的结论也许就不会存在了，又或者会变得更清楚。现在，我们需要把这些结果看成是科学家们在陨石坑的数据中寻找周期性事件的阶段性收获（也许包括一些乐观的结论），而不是什么全面的调研或结论。

陨石坑记录的周期性

无论如何，在寻找陨石坑的周期性时对数据还是有一些限制的。人们应该集中分析较大和较近期的陨石坑。那些过于久远的陨石留下的印记肯定不如近期的可靠。另外，尽管小陨石坑的数量远比大的多，但是对周期性的研究应该只专注于那些大的陨石坑。小天体随时都可能撞到地球上——除了小行星带的瀑布式碰撞，这些撞击大多数都是随机的。那些造成小陨石坑的天体的行为是无序的。我们下一章会看到，真正具有周期性的只是彗星，而且只是那些从遥远的奥尔特云飞来的彗星。

因此，在数量（陨石坑越小数量越多）和可靠性（陨石坑越大对周期现象越可靠）上要有一个取舍。最佳的选择是未知的。很多文献都使用了不同的标准，因此大家在看早期的研究结果时一定要记住

这一点。在马修和我的工作中，我们最终决定选择在过去的2.5亿年落到地球上的，直径大于20公里陨石坑。我们选择的时间2.5亿年看起来足够长，统计上比较合理，但还不会太久远以至不可信；20公里也是一个比较好的取舍点，它要求陨石尺寸要达到公里量级，又不会太大而丢掉相关的统计性数据。

即使有了这些限制，从陨石坑记录中辨认可靠的周期性也是难以完成的任务。陨石坑的印记在地球历史进程中并不能被完整地保留下来，它们中只有一小部分今天还清晰可见。此外，陨石坑的时间并不总是能足够精确地推测出事件的时间特征。更麻烦的事情还有，研究者们需要使用不同的数据组。即使同样的数据，分析人员有时会使用不同的时间间隔或者以不同方式分组。如上文所述，这使得情况变得更加混乱，即使有些陨石是周期性出现的，还有一些却是随机的。这意味着最好的情况也只是在一个随机组分上叠加一个周期组分，此类叠加进一步“连累”了本来就可怜统计记录。

尽管如此，受沃尔特·阿尔瓦雷斯提议（彗星撞击导致白垩纪-古近纪大灭绝）的启发，以及周期性灭绝的证据，科学家们继续寻找周期性天体撞击的证据。1984年，沃尔特和加州大学伯克利分校的同事、物理学家理查德·穆勒发起的这项工作，当时他们用2.5亿年前的、半径大于5公里的陨石坑得出了2.84亿年的周期。他们研究的样本只有11个陨石坑，而且没有严格考虑数据的不确定性，但是很多更全面的分析接踵而至。

1984年早些时候，纽约大学的生物学家迈克尔·兰皮诺（Michael Rampino）与美国国家航空航天局戈达德太空研究所的理查德·斯图斯尔（Richard Stothers）合作，研究了另外一个包含41个陨石坑的样本，这些陨石坑产生于2.5亿~100万年前，他们得到了一个3.1亿年的外来天体撞击周期。1996年，日本的科学家们得到了类似的结果。

——过去3亿年间的陨石坑有一个3 000万年的周期。2004年，京都大学的一名应用数学家藪下·新（Shin Yabushita），也是这项研究的参与者之一，进行了一项更细致的分析，使用了过去4亿年的陨石坑，并根据其大小给它们以不同的权重。由此他从一组91个陨石坑中得到了3 750万年的周期。这些分析都从陨石记录中找到了规律性的证据。但是周期的结果都不一样，不足以支持这些结果。

2005年，英国白金汉天体生物中心的教授威廉·纳皮尔（William Napier），进行了一项有趣的研究。他认为天体撞击是成组出现的，每期持续大约一两百万年，间隔2 500万~3 000万年。他的样本是2.5亿年里直径大于3 000米的40个陨石坑。他发现，最大的一组撞击发生在相对较短的间隔，而且白垩纪-第三纪大灭绝也发生在其中。这个周期性的证据比较薄弱，但他依然推断出了一个周期的长度范围——这取决于如何诠释数据，这个范围内似乎2 500万年和3 500万年占主导。

纳皮尔意识到他的证据不足以成为很强的论据。他指出：尽管已经比沃尔特开始的数据多了很多，但还是无法期望得到更强的信号或者信号会完全消失。他认为这种模糊不清的结果的一种可能解释是：他的结果是稳定的随机事件和周期性事件的叠加，所以即使增加了3倍的数据量，信号也不会清晰地出现。

纳皮尔还提出了一些有趣的建议，是关于可能的天体是彗星还是小行星的。尽管他认为在分析中排除的那些较小流星体可能源自小行星带，但他怀疑那些彗星——而不是小行星，是那些大陨石坑的主要制造者。他给出的原因是，大型小行星不足以解释轰炸时期所需要的密度，相反，太多的大型小行星与我们发现的短时标周期不符。纳皮尔还指出，不充足的陨石记录实际上支持了他的论据。如果大部分陨石坑都没有保留下来，那么撞击的数量应该比从地球上鉴别的陨石坑

数量多很多。如果我们知道一些大型陨石坑集中在一个轰炸期，那么我们也应该知道发生过很多撞击，但证据已经不存在了。

纳皮尔进一步解释道，由于扰动而进入地球轨道的小行星只有不到4%撞到地球上。大部分都被甩出太阳系或掉到太阳里了。考虑这两种效应，纳皮尔从他的数据中总结出：一个至少20公里~30公里大的元行星碎裂而产生的几百个小行星到达了近地轨道。这种碎裂肯定是由于碰撞。但是这么大的小行星碰撞出这个数量的小行星的事件太不常见了。因为一两百万年的影响时间太短，而且这个时间范围对小行星的解释也不合理，所以他认为彗星更像是这个周期爆发的根源。尽管他的结论并没有得到证明，而且我们也知道一些小行星有一个一两百万年的“快速通道”，但是他们的确提出：对一些显著的撞击事件来说，彗星可能比小行星更重要，这也许是最终区分此类撞击事件的方法。

“旁视效应”

以上都是令人好奇的观测。然而，上面这些结果都没有给出统计显著性，这是最后建立周期性所需要的。当分析统计显著性时，另一个棘手的问题出现了，这个问题还解释了为什么之前的结果会互相矛盾。不过这个问题是可以解决的。

你也许认为，如果假设数据是周期性的，就可以简单地把数据匹配一个周期性函数，然后评估周期性函数的拟合情况，并用来解释观测结果。然而，这会给出一个过于乐观的估计。当你不是测试单一的假设而是许多猜测时——在这种情况下，函数会得出不同的周期。如果给出足够多的可能性，周期性函数的拟合结果几乎肯定会被证明比随机情况更好，但这并不一定正确。

这个微妙但是有些明显的（至少事后看来是这样）问题，在粒子物理学界被称作“旁视效应”（look-elsewhere effect）。当希格斯玻色子在大型强子对撞机（LHC）中被发现时，这个现象成为很多讨论的主题。^[31]大型强子对撞机就是在日内瓦附近的欧洲核子研究中心（CERN）的巨型粒子加速器，科学家们试图用高能质子互相碰撞来制造新粒子，并由此可以洞察还没被发现的物理理论。这尽管不是本书的主题，但寻找希格斯玻色子的结果，阐明了一个科学家在寻找周期性时需要面对的问题。

实验人员寻找希格斯玻色子的方法是在数据中试图寻找希格斯玻色子衰变成的粒子，然后测量它们被发现的频率。因为粒子碰撞的大多数时间里是不产生希格斯玻色子的，而希格斯玻色子存在的迹象出现在数据突出的信号里，那些光滑的背景曲线则来自希格斯玻色子没有出现的那些实验事件。如果可以正确地做图，这个突出的地方应该发生在正确的希格斯玻色子质量上。所以当实验人员展示数据时，他们集中在了数据“鼓包”的区域，在那里，某种物质（希望是希格斯玻色子）贡献了高于背景很多的信号。

旁视效应

科学实验中，尤其是粒子物理学实验中统计分析的一种现象。由于参数空间的改变，会导致统计显著性偶然地明显增加。

需要提醒的一点是，统计的侥幸（技术上所谓的涨落）总是使数据上下变化。有时会有很大的涨落。尽管特定在某点的涨落是不太可能的，但是如果你分析的质量范围足够大，一个不太可能的涨落会恰好出现在某个地方。这个不太可能的事件的“信号”看上去可能和希格斯玻色子的信号一样。但这个“信号”只是背景事件中的“不可能涨落”在某一质量处的积累。

当实验人员一开始寻找时，他们并不知道希格斯玻色子的质量^[32]。因为衰变产物中的能量和质量可以用来计算希格斯玻色子的质量，所以当找到了适当的证据时，他们就能够测量这个质量了。但是研究人员只能在他们看到一个“鼓包”后才能得到质量，而不是相反的过程。

当实验人员展示他们的数据，并讨论他们辨识出的每一个“鼓包”有多可能或不可能出现希格斯玻色子时，他们需要考虑其质量值的不确定。因为统计涨落可能出现在任何地方，它们中的每一个都可能被认为是一个希格斯玻色子的衰变，每个“鼓包”的统计显著性被其他什么地方更大的统计涨落削弱了。实验人员知道这一点，所以他们展示的是已经考虑到了旁视效应的显著性结果。这个旁视效应会告诉你，如果你事先已经知道了希格斯玻色子的质量，结果会更加显著。如果你不知道，那么一个“鼓包”更可能是一个涨落，因为你可能把数据中一个非常规正信号的可能性乘以了它们的发生次数。仅仅当实验产生了足以探测的希格斯玻色子并显示出统计显著的结果时——旁视效应也要考虑进去，物理学家们才能最终声明他们的发现。

当寻找陨石记录的周期性时，如果事先不知道你要找的周期，类似的考虑也适用，而天体物理学家喜欢把这个效应称作“实验因子”（the trials factor）。如果你使用了足够多的不同周期，它们中总有一个同数据的拟合结果比没有周期的结果要好，即使数据是完全随机的。原来，结合了周期性天体撞击的模型与数据符合得更好——至少比假设完全随机的撞击事件要好。但是因为没有人知道预期的周期，研究人员基于简单拟合能推测的任何统计显著性都低于他们天真的结论。如果有足够多的可能性，而每一个可能性又都具有各自的统计不确定性，最终，一些周期函数必然会看起来与数据合理吻合。

我们费了好大力气才解释了科兰·贝勒-琼斯的结果与其同事的分歧：他没有发现周期性的统计证据，而他的同事却发现了。他们各自都做了正确的分析，但是贝勒-琼斯并没有像我们一样，提前把周期因素考虑进去。没有额外输入，这个信号需要非常强才能压倒这种中和效应。但一开始这个信号并不够强。

好消息是，我们确实有额外输入，并且能够把它考虑进来。在某种程度上，我们知道星系的组成，因为天文学家已经测量了其组分和引力场。如果周期效应是由太阳系的运动造成的，我们可以把星系的性质和太阳的位置结合起来，并把预测的太阳系的运动和已知数据进行比较。在下一章，我会介绍一个周期性陨石雨的触发机制及其应用，这正是马修和我决定研究的题目。

A large, stylized number '15' is centered at the top. The number is composed of a dark grey '1' and a lighter grey '5'. The '5' has a trail of small black dots extending upwards and to the right, resembling a comet's tail or a starburst effect.

从奥尔特云中逃逸的彗星

DARK MATTER AND THE DINOSAUR

一个足够强的扰动就可以将一个奥尔特云中的天体慢慢地移出它应该在的位置，就像那些步伐不够精确的外圈舞者一样。

你也许曾经在纽约的电波城音乐厅（Radio City Music Hall）观看过洛克茨（Rockettes）舞蹈团表演的集体舞，或者在一些电视节目中看过一些舞蹈团的类似表演。在这种舞蹈里，很多衣着华丽的舞者会围成一个圆环并同步作出一些优美的动作。这些舞者有时会组成从中心往外延伸的辐条式图形，有时会组成大圈套小圈式的多重同心圆。这些圆环上舞者的动作流畅、互不碰撞，这让观众很容易忘记，维持这种运动需要精确地维持舞者之间的相对位置，而这是极其困难的。特别是对那些处于外围的舞者，她们需要动得更快并且需要从最初的位置跑得更远。你可能偶尔会看到一个处在最外圈的、动作难度更大的舞者跳砸了，并因此无法与其他人的动作同步，不过只要她没有摔倒，并不会造成太大问题。这种错误虽然会破坏所有舞者动作的同步性，进而破坏整个表演的完美性，但并不会造成灾难性的失败。

奥尔特云距离太阳的距离是地球和太阳距离的几万倍。奥尔特云中那些多冰天体面临着和外圈舞者类似的挑战。这些天体距离太阳如此之远，以至于太阳的引力只能使其处于不太稳定的平衡状态。一个

足够强的扰动就可以将一个奥尔特云中的天体慢慢地移出它应该在的位置，就像那些步伐不够精确的外圈舞者一样。如果一个奥尔特云中的天体靠太阳系内圈足够近的话，那么一个大力的推或者拉，将会完全改变它的轨道。当这种情况发生时，这个天体与原轨道的偏离将远甚于那些跳砸了的舞者，它将面临冲向太阳系内圈的危险，甚至有可能冲向地球。

近地小行星以及一些偏离正常轨道的短周期彗星，也会受到行星或者附近天体的推挤，偶然也会撞上地球。但是这些作用力基本都是随机的，只有来自奥尔特云的彗星才被认为会受到周期性的扰动。奥尔特云既是进入太阳系的长周期彗星的唯一来源，也是大部分能靠近太阳的彗星的来源，而且也被认为是等距轰炸地球的彗星的唯一发源地。前文提到的生物灭绝的周期性以及陨石坑的记录，让我们有了强烈的兴趣去研究，到底是什么触发了这些能将奥尔特云中的冰冷天体送进内太阳系的扰动。

在这一章，我将首先简单地阐述彗星或者小行星是否会造成重大的冲击。随后我会总结一些最初的关于如何从奥尔特云中移出天体，并形成可以冲击地球的彗星的想法。虽然这些老的想法不能解释那些定时的冲击，但不可否认，它们鼓励了人们思考星系相互作用的新方向。这些旧想法也为我们之后提出的基于暗物质的更可靠新想法，铺平了道路。

小行星还是彗星

如果造成希克苏鲁伯陨石坑的是一颗小行星的话，那么暗物质是与它没什么关系的；但如果造成这些破坏的是一颗彗星的话，那么暗物质很有可能是这一事件背后的“主谋”。地质学家沃尔特·阿尔瓦

雷斯在他的《霸王龙和陨星坑》一书中默认用了“彗星”一词，来描述造成白垩纪-第三纪大灭绝的撞击者。虽然他认为没有人能够明确搞清楚到底是彗星还是小行星造成了这些撞击。区分这些陨石坑到底是彗星还是小行星造成的非常困难，尤其是对那些在几百万年前撞击而成的陨石坑。如果没有观测到撞击天体的轨迹的话，我们基本上没有办法确定撞上来的是彗星还是小行星。因此，我们也无法“起诉”这个造成恐龙灭绝的凶手彗星。

但我们确实知道彗星及其碎片并不怎么会撞到地球上。据估算，彗星撞地球的概率大约是小行星撞地球的2%~25%。这个相对概率比较小，因为在地球附近的彗星数目本来就比较少。在一万多个所知的近地天体中，大约只有100个是彗星，剩下的都是小行星或是更小的流星。

较大的撞击并不一定完全来自接近地球的天体。遥远的彗星也会时不时地偏离轨道，偶尔撞上地球。杰出天文学家吉恩·苏梅克（Gene Shoemaker）通过一项有趣的研究声称，虽然较小的撞击基本都是小行星造成的，但是较大的撞击很可能是由彗星造成的。苏梅克画了一张撞击数-撞击大小的关系图。他发现，这张图似乎表明存在着两种不同的撞击。那些比较小的撞击都落在一条简单的曲线上，而有很多较大撞击却不能被这个简单的曲线所描述。既然已知是小行星造成了这些较小的撞击，苏梅克认为应该有另外一种东西造成了那些更大的撞击——也就是说，他所画的那张图应该是由两种不同的撞击源的曲线组合而成的。他猜想，那种造成更大撞击的天体就是彗星。

彗星有一个更好的特性，那就是，它们携带着相比小行星多得多的能量，因为它们运动的速度比小行星快得多——每秒可达到70公里，甚至更快，而小行星的速度每秒只有10~30公里。一般来讲，弹道导弹的速度基本在11公里/秒以内，小行星的速度大约是20公里/

秒，短周期彗星的速度大约为35公里/秒，长周期彗星的速度大约为55公里/秒，当然也会有更快的速度（见图15-1）。动能不仅与质量成正比，而且与速度的平方成正比。彗星具有较高的速度，这意味着，虽然它们撞击地球的概率比较低，而且具有较小的质量，但是它们却能比速度较慢的小行星造成更大的破坏。

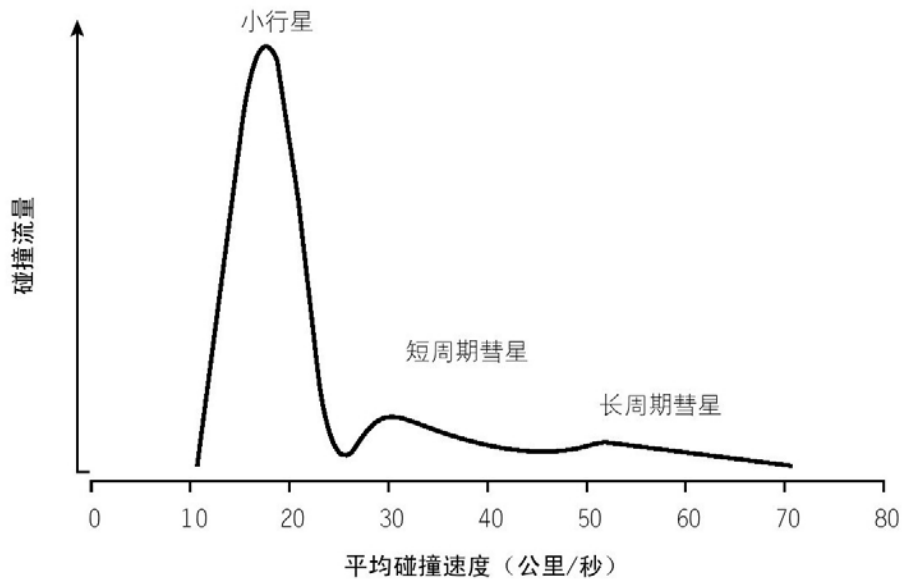


图15-1

小行星、短周期彗星以及长周期彗星撞击地球的平均速度。图中的曲线也描绘出理论上对这三种天体流量的预期值。

苏梅克又进一步做了化学分析，以佐证他的彗星撞击说。当然为了公平起见，我们应该提一下，做这种化学分析的科学家们既有赞同也有反对这种假说的。一个支持小行星撞击说的证据是，同位素含量的比例以及存留下来的流星碎片和那些球粒状陨石小行星相吻合。这些球粒状陨石小行星含有45.6亿年前在太阳系形成的时候，由星云暴风雨中融化而成的毫米大小的球体。但这些证据并不是决定性的。我们并不知道彗星中的同位素比例，而且彗星的同位素比例其实也差不

多。另外，最近的一些研究声称，铍元素和钪元素的含量比以前所认为的要低，这样，反而使证据更倾向于彗星撞击说。

1990年，天体物理学家凯文·扎恩勒（Kevin Zahnle）和大卫·格林斯普恩（David Grinspoon）用不同的方式，论证了希克苏鲁伯陨石坑是由彗星撞击而成的。他们提出，彗星的尘埃在白垩纪-第三纪大灭绝之前和之后分别进入了地球，并以此来解释包围着白垩纪-第三纪岩层周围的沉积物中的氨基酸。既然尘埃颗粒可以悬浮在大气中并慢慢沉降到陆地上，而不会受到破坏，那么这些尘埃原则上是可能来自彗星的。这个彗星在漫长的时间内被分解，并将这些分解开的物质像雨一样洒落到地球上。

有一个因素可以导致彗星的撞击比预期的更加频繁，即当木星扫过附近彗星时，木星有时会将它们撕裂成碎块。如果发生这种状况，彗星撞击地球的概率就会增加，因为可能有好几块这种碎块会冲到地球的轨道上来。有些天文学家推测，这种情况最近一次发生在几千年前，因为在内太阳系里，有大量的彗星尘埃。

前几年刚刚撞到木星上的苏梅克-列维彗星就是这种彗星碎片的一个鲜活例子。卡罗琳·苏梅克（Carolyn Shoemaker）第一个在1993年发现了这颗当时处于木星附近的彗星，随后她和她的丈夫吉恩·苏梅克以及另外一位同事大卫·列维共同追踪了这颗彗星的踪迹。他们注意到，这颗彗星看起来不太寻常，因为它看起来并不像是划过天空的一条线，而是一个由亮点组成的圆弧。很快，在一次更加精确的观测中，天文学家珍妮·刘和大卫·朱维特在其中发现了17个分立的小点，这些散列在圆弧上的点就像是一串珍珠。

当时在美国中央天文电报局（CBAT）的天文学家布莱恩·马尔斯登（Brian Marsden）根据这颗彗星的轨迹推断，它的独特结构是因为一次太过接近木星的飞行而造成的，在那次飞行中，木星的引力将彗星撕成了较小的碎块。他提出了这颗彗星以后还会再次近距离接近木星，甚至撞上木星的可能性。其他天文学家跟进了他的提议，并且通过计算发

现，木星的引力确实会困住这些碎块并使它们可能在1994年7月16日到7月22日期间撞上木星。

果不其然，第一个碎块如期以高于60公里/秒的速度飞入了木星大气。木星上看起来受到影响的区域至少有一个地球那么大。木星大气被先于碎块闯入的尘埃所照亮，造成了绚烂的巨闪。这些撞击造成的破坏与希克苏鲁伯陨石坑周围的差不多，只不过，这次受到破坏的地点是在木星上。这个碎块直径只有不到300米大，而最初形成这些碎块的彗星最多也只有几公里大。因此，这次撞击所释放的能量远小于当初造成希克苏鲁伯陨石坑的那颗天体的能量。虽说如此，这次慧木大撞击依然令人震撼。

木星各个卫星上的陨石坑们显示：这种彗星被撕裂并撞击到木星或者其卫星上的现象，在这一区域并不是第一次发生了。而且，如果周期性流星的想法是正确的话，那么它将是彗星在太阳系演化过程中一直扮演重要角色的又一重要证据。这种天体物理现象与行星表面的关联提醒我们，即便是抽象的理论研究，也最终可能会帮助我们解释人类的存在。

神秘的触发

虽然这种来自奥尔特云的彗星造成了当年的大撞击的学说尚未被证实，我在本书后文会假设它是成立的。这是我们现今所知的唯一一个能够解释周期性撞击的理论。虽然这种外太阳系的冷冰冰的天体由于受扰动而冲到地球轨道上的说法，听起来像是个科幻小说——但它并不是，虽然我们经常在科幻小说里看到类似情节，这一系列事件也有科学依据。

太阳系最外围就是奥尔特云，我们可以将其想象成一个由很多较小天体组成的，可能延伸至超过5万倍日地距离处的球状天体。这个巨型的彗星发源地存在的间接证据——因为它太远了而无法直接观测，就是我们可以看到的进入内太阳系的彗星。

与我们之前提到的外圈舞者的情况不同，使奥尔特云中的冰冷天体保持在它们轨道上的，是太阳的引力，而不是奥尔特云中各个天体的相互作用。但是太阳的引力只能很弱地束缚住这些天体，因为它们离太阳实在是太远了。引力的强度随距离的平方递减，所以在一万倍原距离的情况下，引力的影响将是原来强度的一亿分之一。太阳对奥尔特云中彗星的引力影响和对地球的引力影响比起来，就是这个比例。在如此弱的束缚下，即便是相对较小的扰动，都可能改变奥尔特云中天体的轨迹，并最终将其踢出轨道，或者甩出太阳系，或者将其送入一个飞奔向太阳的轨道。

天体物理学家恩斯特·奥匹克早在1932年就提出，对太阳系外边缘彗星的扰动可以将这些冰冷的天体送入太阳系的内圈。奥尔特后来为奥匹克的想法找到了更可靠的基础，因此奥尔特云有时也被称作奥匹克-奥尔特云。奥匹克的理论基本上是正确的，他推断，有些冰冷的天体最终会变得不稳定，容易受到扰动的影响，外界影响有时会将它们推出轨道，并进入一条通往地球的道路。他甚至认为，这会对地球上的生命造成影响，虽然他并没有预想到这种破坏会导致白垩纪-第三纪生物大灭绝。

奥匹克的理论虽然让人印象深刻，但也留下了问题：为什么轨道会变得不稳定，是什么触发了它们的逃离？人们一直无法解答这些问题，直到很多年后，沃尔特的提议（以及“冷战”带来的破坏）进入了公众视野，并重新激发了人们的兴趣。

天文学家提出了很多种可能引起扰动的天体，比如从附近经过的恒星或者巨型分子云——质量在1 000~1 000万倍太阳质量的巨型分子气体聚集体。虽然前者会推挤轨道，而后者也会对轨道有一些影响，但两者都不是将彗星送入内太阳系轨道的主要因素。推挤对于彗星轨道的影响取决于推挤的强度和推挤发生的频率，以及被推挤彗星的密度和质量。而恒星和分子云推挤的力度和频率都不足以解释我们看到的这些彗星。

1989年，朱莉亚·海斯勒和斯科特·特里梅因研究了一种更为巨大的影响力——银河系的潮汐力。月亮对地球上离月亮距离不同的地方的引力是不同的，这种畸变的引力会造成海洋的上升和下降，并形成我们所熟悉的海洋潮汐。同样，银河系上的银河系潮汐也在以相似的方式影响着外太阳系的天体。银河系对位于不同位置的天体的引力是不同的，从而将本来是球形的奥尔特云往朝向太阳的方向拉长，并在另外两个方向压缩。

经过一段时间，银河系的引力会将小天体的轨道慢慢修改成非常细长的轨道，或者说是偏心的轨道。一旦轨道足够偏心，它的近日点——距离太阳最近的距离，将会变得很小，以至于这些天体会很容易射入内太阳系。这样，潮汐力就足以将那些冰冷的天体从奥尔特云送出，并增加彗星流入太阳系内部的流量。这将会造就一个缓慢而稳定的朝向地球的彗星流。

更有意思的是，将这些冰冷的天体抽出来并变成彗星送到内太阳系的主导因素，并不完全是潮汐，而是恒星和潮汐扰动的联合作用。虽然恒星扰动并不是造成彗星流的最终因素——因为恒星扰动的作用时标比潮汐的作用时标要长得多，但恒星仍然很关键，因为它们是它们将奥尔特云预先送到一个潮汐力可以起很大作用的点上。这就像是环法自行车赛中的一个自行车队，车队中其余的车手会帮助领头的赛

车手占领好的位置，从而使他能够完成最后一击并赢得冠军。我们一般只知道最先跨越了终点线的胜利者的名字，而忽略掉队伍中其他人。即便如此，其他骑手也扮演了很重要的角色。与此类似，虽然最终将彗星抽出来的是潮汐力，但是它之所以能够对彗星施加足够的推力，是因为恒星扰动已经预先将彗星的轨道送到了不稳定的位置，以至于只需要相对较小的推拉，就可以将彗星送入内太阳系。恒星扰动很重要，但最终引发彗星变轨的，也即人们知晓的因素，主要是潮汐力。

银河系潮汐力超过太阳引力而起主导作用的区域大约是距离太阳10万~20万倍天文单位（日地距离）的地方。在奥尔特云的外边缘，太阳引力将不足以维持轨道的稳定。我们已经看到了潮汐力如何影响处于稳定边缘的轨道，并不时将一个小的太阳系天体送入内太阳系。在更靠近里面一些的地方，也就是我们能观测到的区域，潮汐力就不如太阳引力强了。所以只在奥尔特云中，潮汐作用才能很大程度地推挤处于弱束缚态的彗星。在来自奥尔特云的彗星中，可能有90%是因为这些潮汐影响而产生的。

现在，物理学家和天文学家已经理解了银河系的这种通过引力作用对彗星作出扰动，并将其送入内太阳系的方式。这些机制虽然很重要而且很有意思，却并不足以解释所有的彗星雨或者彗星撞击的周期性。如果没有其他因素的话，上面描述的潮汐力只能造成缓慢却稳定的彗星流。

天文学家们为了解释彗星流的周期性增强，做了各种努力和尝试，来解释为什么这些彗星的触发并不是完全的随机，而是可能每过几千万年就发生一次。我首先要说明，下面我将要描述的一些理论并没有成功解释上述现象。但是，通过理解这些理论以及它们为什么会

失败，将会帮助我们找到新的理论。其中的一个理论就是后文我将要描述的暗物质盘理论的前身。

内梅西斯提议

第一个，也是最出彩的一个用来解释周期性撞击的理论，是太阳有一个名为内梅西斯（Nemesis）^[33]的伴星，而内梅西斯和太阳组成了一个大的双星系统。天文学家为这颗想象中的太阳伴星规划了一条非常椭圆的轨道，它和太阳每过260万年将会在3万天文单位的距离内擦身而过。这个1984年提出的理论是为了用内梅西斯在每260万年离太阳最近时增强的引力，来解释大卫·劳普和杰克·塞克斯基的生物灭绝周期性理论。这个理论提出，在内梅西斯离太阳最近的时候，它的引力将会从奥尔特云中抽出一些小天体，并使之成为太阳系的成员，这些小天体之后将变成彗星并对地球进行轰炸。

将近300万年的周期性相遇（以及因此带来的彗星撞击的增多）需要一个非常大的系统。这个大系统的半长轴（椭圆长轴的一半）大约是一到两个光年的量级。这个理论的一个问题是，恒星或者星际气体云会使这个巨大的双星系统不稳定，这将会摧毁之前预设的相遇的周期性，并造成在过去的2.5亿年内，次双星系统的周期是变化的。然而，科学家们没有发现这种变化。

最终让这个想法成为板上钉钉的，是一个很好的红外全天巡天星表。如果内梅西斯真的存在的话，那它也应该在这个星表里。虽然1984年的测量并不足以完全排除内梅西斯的存在，现在的观测手段已经有了巨大的提高。美国国家航空航天局的大视场红外巡天探测仪（WISE）于2009年发射升空，之后一直收集相关的数据直到2011年2月。如果这颗红矮星真的存在，那么这颗卫星应该已经看到它了——然而并

没有。同时，由于人们也没有发现木星大小的巨型气体行星，另一个相似的理论也被排除了——这个理论假设，导致彗星流周期性变化的是一颗新的行星（被称作X行星）。

源于银河系运动的触发

在参考了这些失败的想法之后，一些非常不同的基于太阳系、穿行于银河系内已知的不同部分的理论，看起来是很可靠的选择。这些理论并没有引入新的或者奇怪的东西，而是认为，太阳系通过星系旋臂或者穿过星系盘时所经受的密度变化，会引起对奥尔特云扰动的变化。重复通过高密度区域，原则上可以解释彗星的周期性。

我们知道，银河系是一个盘星系，也就是说大部分恒星和气体是在一个直径大约为13万光年、厚2 000光年的薄盘上。太阳位于距银河系中心约2.7万光年的地方，而此时它在垂直方向上非常靠近星系盘的中心平面——只有不到100光年的距离。它也在一个旋臂的边缘。

银河系的旋臂从银河系的中心开始向外弯曲着延展（见图15-2）。这些旋臂比旋臂之间的区域包含更多气体和尘埃，因此也是年轻恒星更容易形成的地方。这里也是高密度的巨型分子云的所在地。当太阳穿过这些密度较高的区域时，这些分子云会施加更强的引力，原则上会导致更强的扰动，而造成撞击的周期性增强。

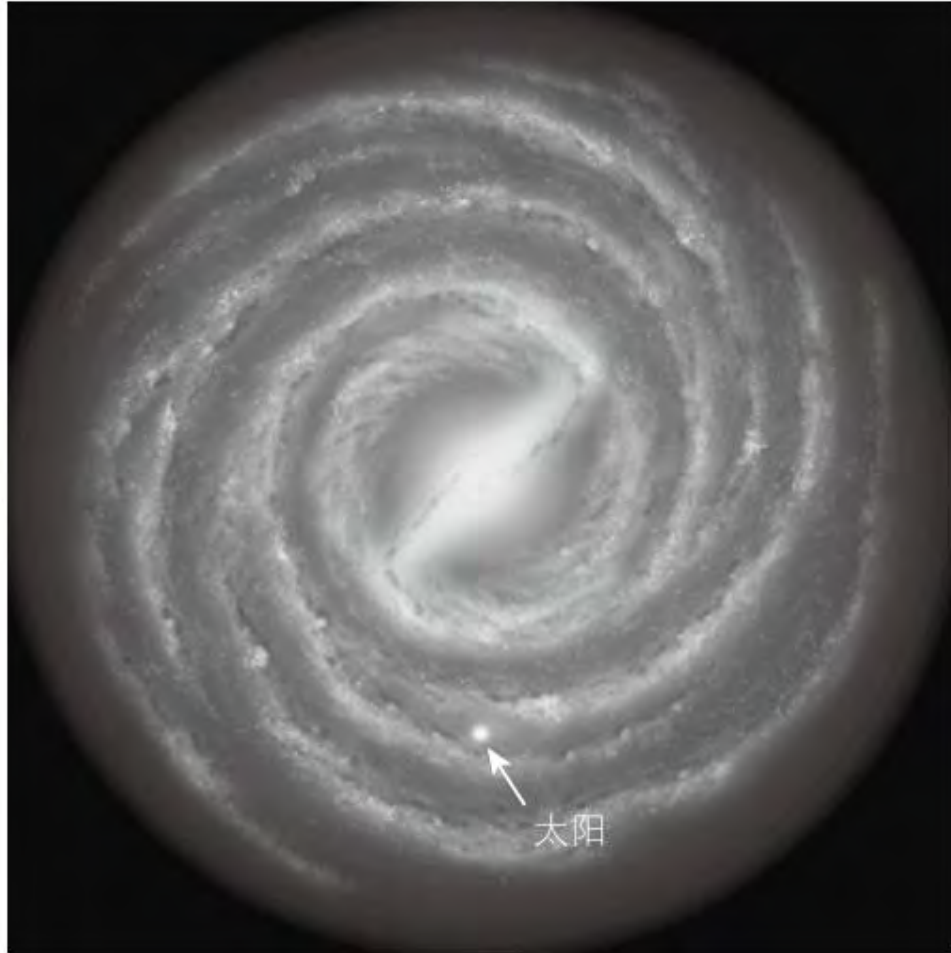


图15-2

银河系的旋臂以及太阳（大小并不是按比例绘制的）的位置。

这个理论的一个潜在问题是：这些旋臂并不是完全对称的，甚至没有一个相对太阳的固定旋转速率，即太阳很可能并不会以精确的周期穿过它们。然而，现在人们对旋臂的结构、动力学以及演化并不十分了解，仅仅通过这些就排除这个旋臂理论，也很武断。除非周期性非常完美，否则这种缺乏完美周期的理论很难被数据排除，因为这些数据的周期性也可能只是近似的。

不过，另外两个因素令旋臂理论无法很好地解释观测到的撞击率的提高。第一个因素是，旋臂中的平均气体密度不够高，并不能解释

周期性的撞击率提高。如果密度的改变不够大的话，地球在穿越旋臂过程中所受到撞击率的提高将小到无法察觉。

另一个因素是，太阳系穿过星系旋臂的频率并没有那么高。银河系只有四个大旋臂，可能还有两个小旋臂，而围绕银河系转一圈的时间是很长的，因此这就意味着在过去2 500万年中，只有不到4次穿过大旋臂的机会。而实际上，这些旋臂运动的方向和太阳系运动的方向相同（虽然速度不同），这样的话，穿过大旋臂的周期大约是800万~1 500万年，远远不够解释生物灭绝或者陨石坑的出现。

虽然穿越旋臂无法解释撞击的周期性，但这并不代表可以排除银盘垂直方向上的密度变化作为触发机制的可能性。而这种理论很可能被证明是一种很有希望的理论。太阳系在做圆周运动的同时，还在银盘的垂直方向做上下的振荡。如图15-3所示，这种振荡的尺度是非常小的（相比于太阳距离银心的距离2.6万光年）。太阳系围绕着银心做着近似圆周的运动，它围绕银心一圈大约需要2.4亿万年，这个时标就是所谓的一个星系年。在做圆周运动的同时，太阳还会上浮和下沉。太阳在垂直方向振动的振幅很小，而且与星系盘上的物质分布有关。但是一个合理的估算给出的值大约是200光年，不过我们现在处在一个更加接近中心盘的位置，离中心盘大约有65光年。

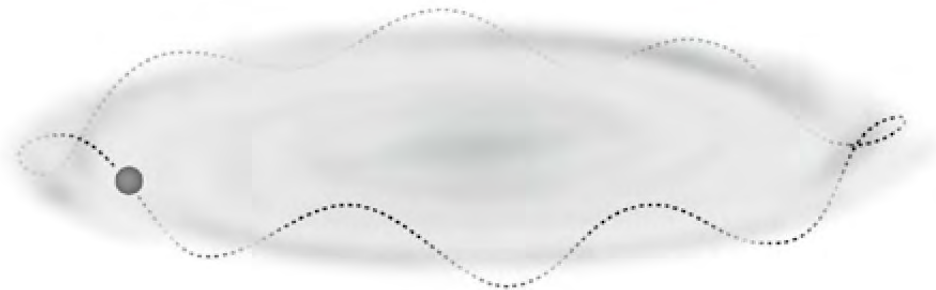


图15-3

太阳在围绕银河系旋转的同时，还会穿过银盘上下振动。在穿过银盘的时候，它会受到更强的潮汐引力。注意，为了更清楚地描画这种振动，图中的振动周期被刻意

缩短了，而太阳实际上在绕银河系一周的过程中，只会做3~4次垂直振动。

太阳系的这种垂直方向的振动也许可以说明潮汐效应随时间的变化，并且可以用一个合适的时标来解释任何周期性。因为，随着太阳系穿入或者穿出某些银盘上较致密的区域，太阳系周围的恒星和气体含量是变化的，因此太阳系在振动穿过银盘时，会经历不同的环境。如果太阳系在穿过银盘时，周围的密度急剧升高的话，那么它受到的扰动也会急剧增大，结果是，彗星撞地球的概率在这种情况下也会提高。星系潮汐是奥尔特云受到的主要扰动，所以银盘上垂直方向的密度变化原则上可以造成足够的影响。这项理论是由纽约大学的迈克尔·兰皮诺教授和布鲁斯·哈格蒂（Bruce Haggerty）教授提出的。他们也给这个理论起了一个多彩的名字——“湿婆假设”（Shiva Hypothesis，湿婆是印度的毁灭和重生之神）。



为了符合观测结果，“湿婆假设”需要银河系中的物质分布满足两个条件。第一，中心盘的密度必须能够提供一个可以造就正确的垂直振动周期的引力场。这个条件与任何精确的扰动机制无关。如果太阳系不能以正确的频率穿过银河系的中心盘，那么穿过中心盘时的任何引力增强，都不会使理论符合观测数据。

第二个条件是，形成一个能解释周期性彗星雨变化率的必要条件，也就是需要一个足够明显的密度变化。这样太阳系在穿过银盘的时候使奥尔特云受到的潮汐力会随时间变化。这两个条件与所有基于银河系中心平面的密度增强，来解释周期性彗星流的理论是相关的。根据这两个条件，我们排除了现在正在讨论的这个理论，而且它们还可以解释为什么比普通物质盘更致密、更薄的暗物质盘可能会是一个合适的选择，下面我会讲到。

兰皮诺和理查德·斯图斯尔在1984年试图利用一个更加标准的银河系组分，来解释彗星流的周期性，并计算了这种组分（即巨型分子云）的密度变化，这些巨型分子云在靠近银河系中心平面的地方最为致密。他们的逻辑类似于之前提到的旋臂穿越理论——当太阳系穿过这些分子云的时候，物质的分布更加集中。这个理论在第二年就被排除了，因为天文学家们发现，分子云的尺度太大了，它几乎可以延伸到和太阳垂直振动的振幅差不多远的地方。这样的话，在太阳垂直轨迹上的密度变化将会太小，而不能起到所需的作用。如果没有其他物质的话，仅仅靠与分子云的相撞频率来解释大约3 000万年的周期，还远远不够。

朱莉亚·海斯勒和斯科特·特里梅因与天体物理学家查尔斯·阿尔科克（Charles Alcock）共同研究了另一个可能的理论。他们在证明了银河系潮汐引力的巨大影响之后指出，虽然潮汐的影响会造成一个非常稳定的彗星流，但是一个近邻恒星的临门一脚，也可以造成彗星雨。接下来的问题就是，这种和近邻恒星的遭遇，会以多大的频率以及多大的力度发生呢？后续造成的彗星撞击地球的概率变化会有多大呢？

这支研究团队通过回答以下问题估算了遭遇恒星的频率：一个太阳质量的恒星（在速度为40公里/秒的情况下能造成所需影响的最低质量）进入距离奥尔特云2.5万个天文单位（能够扰动奥尔特云的最小距离，因为这个距离和奥尔特云与太阳的距离差不多）区域的概率是多少？他们发现，这种遭遇发生的频率大约是7 000万年一次。这个频率并不足以解释前面提到的周期，但是从理论上讲，这种机制确实可以在过去2.5亿万年中造成一些类似事件。

为了作出更好的预言，海斯勒及其合作者后来做了一个更详尽的数值模拟：他们还考虑了潮汐能施加的额外推力。他们发现，理论成立所需要的恒星离太阳的距离要比之前认为的更近一些。因此，根据此理论而给出的真实的彗星雨概率比之前算的还要小，大约是1亿年甚至1.5亿年一次，比所有观测到的周期都要长得多。在后续的更详尽的数值模拟中发现，恒星遭遇对引发彗星撞击地球的影响，要比他们这次计算结果大一些，但是仍然不足以解释观测数据。

这些研究的结论就是，如果没有新的成分加入，太阳系的引力不可能在短期内发生足够巨大的改变，并造成彗星撞击率每隔一段固定的时间，就发生一次远超平均水平的可观爆发。虽然太阳系会周期性地穿过银河系中心平面，但是由于普通物质的分布造成的彗星雨，并没有在穿越的时候增加。

总体来讲，这种情况很像之前提到的旋臂触发理论。这些理论给出的周期太小，密度的变化也不够大，以至于不能产生理论提出者所希望的那种周期性陨石撞击。虽然最初所测得的密度分布有可能使理论符合观测，但是后续关于星系的更精确的观测数据显示，之前的那些理论都无法产生符合陨石坑记录的正确频率或者正确的周期性增强。除非银盘上有某种新的未曾被发现的物质，否则所有关于星系平面的理论都会因为所预言的周期过长而被排除。

将所有现有的测量数据，比如周期性随着时间变化了很多的证据，综合在一起之后，我和马修·里斯终于得出了这样的结论：如果盘中没有一个现在未曾探测到的物质成分的话，那么这种上下振动的频率将会因为太长，而无法解释观测到的周期。这不仅仅是因为，普通物质的分布太平滑，而不能产生那种突然爆发的彗星撞击率，而且如果银盘只由普通物质组成的话，将会因为太过稀薄而无法产生正确的周期。

虽然上述这些理论并不能解释各种周期性，但它们仍然给了我和马修继续前行的基础。我们学到了，潮汐效应会在靠近以及穿过银盘时，造成足够大的扰动来驱动彗星进入内太阳系。但我们也知道了，已知的天体物理成分是无法造成我们希望看到的周期性的。它们都不能产生一个足够突然的潮汐效应，来增强撞击地球的彗星数目。

这给了我们两种可能。也许最可能的就是，我们观测到的周期性并不是真实的。首先，周期性的证据并不是那么强，而且很多偶然事件都可能将其伪装成周期性的样子。另一个更富想象力但也更有意思的可能是，银河系的结构并不是我们认为的样子，这样，潮汐影响有可能更大，并且具有比之前期望的还要大的变化率。我们打算按第二条路继续探索下去，而且，我们成功了。

在第三部分我会讲述，当我和马修·里斯考虑了银盘上所谓的普通物质，以及测量到的太阳位置和速度时，我们发现，陨石坑的记录与我们提出的暗物质模型符合得更好。在银盘上加上一个具有合适密度和厚度的暗物质盘，会调整银盘潮汐力的强度和um间相关性，这样一来，撞击的周期和强度都可以与观测数据匹配得很好。

沿着这个思路，我们还有一个意外收获，即第14章提到的旁视效应并没有之前想的那么差。我们不需要再去考虑所有可能的周期，只需要考虑已经测到的星系中的普通物质的密度。利用这些并不精确的太阳系测量数据，以及一个合适的暗物质盘模型，我们可以将可能的振动周期限制在与已经测到的银盘密度分布相符合的范围。我和马修发现，在考虑了现有的数据之后，周期性的假设大约是随机碰撞可能性的三倍。虽然我们提出的暗物质盘的存在性并没有很强的统计学证据，但这样的结果已经很好了，足够鼓励我们进行进一步的研究。

这种方法最好的地方在于，我们对银河系引力场的认识会进一步提高。我们的方法考虑了所有关于银河系的已知数据，它将随着对银

河系和太阳运动更精确的测量数据，而变得越来越可靠。科学家们现在正在测量银河系中的物质分布。现在，太空中的卫星正在记录恒星的位置和速度，以帮助我们推测它们所感受到的银河系的引力场——也就是将它们束缚在银河系内的场。这些都会为我们进一步揭示银河系平面的结构。

在这些注定会激动人心的结果中，理论和观测将会把太阳系的运动和地球上的数据结合起来。未来观测到的更多数据将会带来更可靠的预测和更可靠的结果。

第三部分将会回到暗物质模型，并以可以解释陨石坑数据所反映的周期性模型作为结束。关于周期性和地球历史的研究，可以帮助我们思考那些充满宇宙人类却看不见的暗物质的本质，这也是人们探索地球周围的可见世界和飘渺的暗物质世界的一个非常好的理由。

DARK

MATTER

AND

第三部分
破解暗物质的
秘密

THE

DINOSAURS

THE ASTOUNDING INTERCONNECTEDNESS OF THE UNIVERSE



16

隐形世界中的神秘物质

DARK MATTER AND THE DINOSAURS

任何一个与已有实验和观测结果不符的事实都会判一个新模型“死刑”。

过去的一个世纪里，天文学、物理学和宇宙学等领域中，研究数据和理论的长足进步，教给了我们很多不可思议的知识。但宇宙中的很大一部分我们还没有看到，甚至可能永远都看不到。这很大程度上要归罪于我们有限的视力。比如，有些天体无法被观测到的原因就是太远。由于距离遥远，这些天体即使会辐射或反射出光，这些光也会随着距离的增大而急剧消散，并迅速变暗。

首先，尘埃或者天体会遮挡或阻碍我们的视线。尽管空间探测对宇宙中天体的直接探测，在某种程度上的确帮我们避免了这些障碍，但目前还没有一个探测器能到达离我们最近的恒星，更别说最近的星系了。鉴于空间探测器的有限到达范围和低下的分辨率，它们所能提供的帮助是非常有限的。

还有很多因素限制了我们可以看到的東西。即使在我们周围，有些事情也因为太小而被忽略。人眼的成像过程限制了我们在不使用干涉技术的情况下，所能看到的東西。人眼能看到的光的波长在可见光

范围内，所以任何小于可见光波长的物体，都超出了人类裸眼的可视能力。在大型强子对撞机（它代表了人类目前最先进的技术）的帮助下，人们可观测到前所未有的小尺度物理学过程。尽管如此，这个令人瞠目结舌的机器也只能展示一百亿亿分之一米尺度上的物质。没有技术的进一步发展，我们依然无法研究与更小尺度上的距离和力学的相关问题。

对暗物质来说，我们之所以无法观测到它们，有着更无懈可击的借口。暗物质本身就不发射光或吸收光，而光是人类视觉有效的根本，这是必须面对的事实。暗物质只能通过引力与其他物质相互作用，除此之外，目前我们还没有发现其他相互作用的途径。第2章已经解释过为什么我们知道暗物质的存在，而且我们知道暗物质的一些大致特性，但我们还不能准确地指出暗物质到底是什么。因此，探寻暗物质的本质成为一个紧迫的研究项目。

鉴于本书的终极目标是把暗物质和彗星联系起来，本章我们会从研究太阳系转换到研究暗物质，并介绍几种最有可能解释暗物质本质的候选理论。

一个简洁可预测的模型

尽管相信暗物质是真实存在的，对暗物质的本质我们却一无所知。通过研究宇宙微波背景辐射（CMB），我们知道了暗物质在宇宙中的平均能量密度；通过研究星系中恒星的旋转曲线，我们可以估算出银河系附近暗物质的密度；暗物质是“冷”的，也就是说，暗物质粒子的运行速度要比光速小很多，这个结论是从宇宙中小尺度结构的存在而得出的；根据对子弹头星系团形状的观测可以推测，暗物质之间

的自相互作用非常弱，而暗物质直接探测信号的缺乏又说明，暗物质和普通物质的相互作用也非常弱；暗物质本身不携带电荷。

目前为止，我们知道的只有这些。即使暗物质由基本粒子组成，我们依然不知道暗物质粒子的质量，不知道暗物质粒子和外界是否存在非引力作用，更不知道它是如何在宇宙极早期形成的。我们知道暗物质在宇宙中的平均密度，但是具体是一立方厘米内有一个暗物质粒子，还是一立方厘米中有一个质量大1 000万亿倍的粒子，就不得而知了。大量的微小粒子和弥散的巨大粒子都能给出相同的暗物质平均密度，而目前天文学家们只能观测到暗物质的平均密度。

大多数物理学家认为，暗物质由一种新的基本粒子组成，这种新的基本粒子不具备标准模型中基本粒子的相互作用方式。想了解这种新的基本粒子，意味着我们必须了解它的质量和它与其他物质的作用方式，而且还要弄清楚这些性质是未知粒子性质的全部，还是只是其众多性质中的一部分。很多物理学家都有自己喜欢的候选未知粒子，但是在没有足够的观测证据之前，我不会轻易地拒绝任何建议。

幸运的是，朝着揭开暗物质神秘面纱这个最终目标前进的过程中，人类的视觉限制只是个不太重要的因素。就算人类目前的技术有能力观测到暗物质，由于极度专注而产生的盲点或者粗心都可能让我们错过它。

人们经常会忽略他们不期待的事物。例如，当观看热播美剧《生活大爆炸》中的某一餐厅场景时，很少有人会注意到我的出现，其实我自己也没有注意到。就算我离主角非常近，也仍然会被忽略（见图16-1）。



图16-1

一个关于专注和盲点的例子：《生活大爆炸》中被忽略的我。（感谢吉姆·帕森斯 [Jim Parsons] 提供图片）

不过，由于专注而产生的盲点是可以被修复的。与魔术师利用人类的这种弱点不同，科学家需要克服这种弱点。我们的目的是确定在不经意间错过了什么。和我一样，建立模型的理论物理学家们尝试着想象：实验物理学家们还有哪些方面没有查看，或者其实信号已经在已有数据中，而他们还没注意到。在模型里，我们猜测着那些被隐藏的信息，而这些信息可以帮助我们解释已知现象。根据指定模型的预言，实验物理学家们可以专注于这个方向的实验，并做有针对性的数据处理，这样就可以确认或否定这个模型。即便是非常难发现的物质也可以找到。

经常有人问我：当我建立粒子物理学模型的时候，是否遵循某种标准。当然，任何好的模型都应该基于合理的物理理念，即这个模型应该是拓展或利用了已有的关于物质、力或者空间的数学理论。但除了这个基本准则外，还有什么其他的指导原则吗？

我和我的同事都比较喜欢一个附加指导原则：**模型越简洁、越有预言性就越好**。如果一个模型有太多的自由参数，那它什么也解释不了；如果一个模型包含了太多的可能结果，也是不科学的。一个能激发人们兴趣的模型需要给出足够明确的预言，这个预言必须是可验证的，并且要明确地区别于其他模型。

如果一个特性能够和现有的模型有所联系，就会很吸引人，尽管这个特性不是必要的。暗物质模型的建立就是一个具有这种特性的例子，我们希望暗物质的候选者来自普通物质标准模型。尽管这种搭配不一定是对的，但建立此种联系的原因在于，它可以避免引入关于全新的粒子和力的附加猜测。

最重要的是，**新模型必须和已有的实验和观测结果保持一致**，任何一个与已有实验和观测结果不符的事实都会判一个新模型“死刑”。这些标准适用于所有模型，包括我们即将讨论的最流行的几种暗物质模型。

候选模型1：最有潜力的弱相互作用大质量粒子

弱相互作用大质量粒子（weakly interacting massive particle, WIMPs）在过去几十年里，一直是物理学家和天体物理学家心中最有可能用来描述暗物质粒子的候选者。这里的“弱相互作用”指的并不是“弱相互作用力”。绝大多数弱相互作用大质量粒子的相互作用，甚至要比标准模型中弱相互作用的中微子还要弱。当暗物质粒子在宇宙中穿行时，它们不会被散射很多（如果存在散射截面的话），因为暗物质粒子的相互作用实在是太小了。

此外，弱相互作用大质量粒子的质量也在“弱尺度”（weak scale）范畴。粗略来说，和新发现的希格斯玻色子质量相当，这个质能级

别也正是大型强子对撞机现在能探索的范围。需要特别说明的是，希格斯玻色子十分不稳定，而且存在自相互作用，它显然不是暗物质粒子的基本组成。但是，具有类似质量的其他粒子就有可能是暗物质了。如果这个假设是正确的，暗物质简直就在我们的鼻子底下，它的身份将很快被揭晓，至少对于为大型强子对撞机工作的实验物理学家们来说，是这样的。

弱相互作用大质量粒子

假想的大质量粒子，是暗物质数种假想成分之一，这类粒子的相互作用非常弱。

支持弱相互作用大质量粒子假设的证据，是一个引人注目的观测结果，而这个观测结果有可能是偶然的，也有可能是研究暗物质本质的有利线索。如果存在一种质量与最近发现的希格斯玻色子相当，而且自身又非常稳定的粒子，那么存在于宇宙中的这种粒子所携带的能量，也许正是现在宇宙中存在的暗物质粒子所携带的能量。

关于暗物质粒子质量的演示性计算是基于以下观测的：随着宇宙的演化，宇宙的温度会降低，存在于宇宙极早期的大量大质量粒子会变得越来越少。因为如果在宇宙温度降低的过程中，大质量粒子与其反物质（正反物质是一对等质量的粒子，并且它们会发生湮灭）会因湮灭机制而消失，但由于温度和能量的降低，其逆过程发生的概率却要低得多。结果就是，宇宙冷却后，大质量粒子的数密度会急剧减少。

如果粒子保持了原有的热分布，即特定温度的粒子具有特定的数目，温度的降低会导致大质量粒子最终全部互相湮灭掉。然而，由于大质量粒子丰度会随温度降低而降低，上面的推测就显得过于简单

了。如果正反粒子要湮灭，首先它们要找到彼此。但数密度的降低导致它们分布得非常弥散，正反粒子彼此碰面的机会就会小很多。这会导致正反粒子的湮灭效率随着宇宙年龄的增长和冷却而越来越低。

所导致的结果就是，今天宇宙中粒子的数密度要远大于简单的热力学理论所给出的预言。某种程度上，粒子和反粒子会由于空间数密度的过低，而无法碰面并消除彼此。残存粒子的数量取决于暗物质粒子的自身质量，以及推定的暗物质粒子的相互作用截面。通过合理估算，我们可以得到一个神奇又醒目的结论：和希格斯玻色子质量相当的稳定粒子在宇宙中的丰度，恰好和当今暗物质在宇宙中的丰度相当。

我们还不知道这些计算结果是否绝对正确，我们还需要知道更多的粒子性质来确定以上结论的正确性。尽管计算过程有些粗糙，但是表面上看起来毫不相干的两个物理量却神奇地在这里联系到了一起。也许，这一现象反映了一个事实：弱尺度物理学可以解释宇宙中暗物质的存在，并预言其性质。

这个结论也导致很多物理学家怀疑暗物质到底是不是由弱相互作用大质量粒子组成的，因为其性质是已知的。弱相互作用大质量粒子模型的优点是其与标准模型的联系。相比于其他暗物质候选模型，弱相互作用大质量粒子更容易被证实或者证伪。弱相互作用大质量粒子暗物质与普通物质除了引力作用，还存在其他相互作用方式。它和标准模型中的其他粒子有非引力的微小相互作用。就算这种相互作用十分微小，人们依然可以用极高敏感度的实验来记录它们的涨落，这种探测行为被称作暗物质的直接探测实验，我会在下一章中描述相关的细节。

到目前为止，弱相互作用大质量粒子的相关探测结果却依然是零。当然不能说完全是零，有时候确实会出现一些诱人的蛛丝马迹，

然而没人相信这些蛛丝马迹就是发现暗物质的确凿证据。另一方面，探测设备自身的问题和天体物理背景的理解不当，都会产生类似于我们本来期望观测到的暗物质信号。暗物质存在的证据还不那么一目了然。

尽管缺少观测数据，许多物理学家还是非常喜欢这个想法，依然思考着粒子物理和暗物质的近似于巧合的吻合。而且，他们相信这个吻合如此地真实，完全不应该只是巧合。以上这些还不是最乐观的物理学家，最乐观的物理学家们已经开始研究特定的弱相互作用大质量粒子模型了，比如超对称模型。超对称模型指出：任何一个已知粒子都存在一个还未被发现的超对称的伙伴粒子，这个伙伴粒子具有与前者一样的质量和电荷。但是目前为止，由于无论是超对称粒子，还是弱相互作用大质量粒子都没有被发现，一些相信这些模型的死忠科学家也开始动摇了。

至于我自己，我倾向于尝试评估各种模型的可能性。在最近参加的一个婚礼上，我认识了一个神父，他对粒子物理学非常好奇，并一直问我个人认为暗物质应该是什么。我回答：“让大自然来决定暗物质到底是什么吧。”显然这个答案让他很失望。作为一个模型的建立者，我对应用超对称模型解释希格斯玻色子的质量，持怀疑态度。甚至在大型强子对撞机的科学家们公布最新结果之前我依然如此，因为我知道，在这种情况下超对称模型要完全自圆其说有多难。我没有，也不会判超对称模型“死刑”，因为这是实验物理学家要做的事情；但我不会说超对称模型绝对正确，甚至也不会说它“好像是对的”。

同样，我愿意讨论任何关于暗物质的候选模型。正如我对神父所说的一样：“我并没有偏好的模型。”我在尝试建立可检验的模型，因为只有这样，我们才能知道答案。和超对称模型一样，由于没有观测结果的支撑，以前弱相互作用大质量粒子阵营的支持者们开始怀疑自己相信的东西到底是不是对的。当然在没有观测结果和实验支持的时候，选择相信看上去最有希望的选项，无可厚非。我不知道哪些是

最有希望的选项，但也许其他表面上是巧合的自然行为，会给我们提供更好的线索。

候选模型2：有趣的非对称暗物质模型

除了弱相互作用大质量粒子之外，还有很多暗物质模型，非对称暗物质（asymmetric dark matter）模型是其中最有趣的一个。这种模型下的暗物质和已有模型存在另一个令人惊讶的吻合。这种吻合也许只是个巧合，但也可能帮助我们洞悉暗物质的本质：宇宙中暗物质的总量和普通物质的总量是可比的。这太让人惊讶了！

我想，当你第一次听说暗物质能量是普通物质能量的5倍时，你也许会得出这样的结论：宇宙中暗物质所携带的能量要远远大于普通物质所携带的能量。但结论恰恰相反，它们的能量密度已经出乎意料地相近了。暗物质的总量曾经可能是普通物质总量的700万亿倍，也有可能是一古戈尔^[34]分之一。当然，在这些情况下，宇宙的演化过程也会完全不同，但这些比值都是可能存在的。

宇宙中暗物质总量和普通物质总量大致相同。换一种说法：在描述暗物质、暗能量和普通物质的扇形图中，不存在一个扇形图中全是一种物质，或者一种物质所占比例极其微小的情况。暗能量、暗物质和普通物质都是宇宙学“饼”的组成部分（如果我们把宇宙看作一个馅饼的话），只是它们分别占有的权重不同。如果不存在任何未知原因的话，这显然是个巧合。

公平地说，在我们所观测的尺度上，暗物质能量密度和普通物质能量密度是相当的，太小了就观测不到了。有意思的是，今天的观测中，不同的组成部分都具有足够高的能量密度，并且在观测总量中的贡献是可比的。从理论上讲，如果一个物质组分的能量密度远远高于

其他物质组分的能量密度，那么其他具有微小能量密度的组分将无法被观测到。但实际情况完全不是这个样子，暗物质和普通物质具有非常相似的能量密度。

非对称暗物质模型

该模型认为，宇宙中暗物质能量5倍于普通物质能量，但它们的质量密度出乎意料地很相近，且具有非常相似的能量密度。

根据非对称暗物质模型，暗物质的能量密度和普通物质的能量密度的相似性并非偶然，而是一个理论预言。非对称暗物质模型可以给出的预言和弱相互作用大质量粒子模型的预言是不同的，这个预言更多的是与暗物质部分湮灭之后所残存的暗物质的能量密度。我们不知道，这些所谓的巧合哪个能真正帮助人们加深对暗物质的理解，但这些模型所给出的预言都足够强，以至于我们必须去注意它们。而且它们其中的一个很有可能是对的。

20世纪90年代初期，一些物理学家比较关注非对称暗物质模型，其中包括位于西雅图华盛顿大学的核理论研究所的现任所长大卫·卡普兰（David B. Kaplan）。这个模型的想法来自21世纪第一个10年后期的宇宙学观测，领导这个宇宙学观测的人是另一个同名大卫·卡普兰，他曾经在华盛顿大学学习，师从物理学家马库斯·鲁蒂（Markus Luty）和凯瑟琳·祖瑞克（Kathryn Zurek）。还有许多物理学家，包括我，也曾研究过这一类型的模型。

那么这个模型的概念是什么呢？为了理解该模型的设想和动机，我们先看看普通物质的一些性质。像第3章提到的那样，无法确定本质的暗物质并不是神秘物质的唯一形式。我们熟悉的普通物质也有一些奇怪的性质，特别是我们今天观测到的宇宙中普通物质的总量。普通

物质的能量大部分以质子和中子为载体，也就是我们说的重子（baryon），这些重子是由一种叫作夸克的基本粒子组成。如果大部分由重子物质组成的普通物质的分布都遵循宇宙极早期的最简单假设，即随着宇宙的冷却，普通物质会互相湮灭掉，那么今天宇宙中存在的普通物质的密度要远远低于现有观测。

宇宙有一个关键的特点，我们人类自己也有类似的特点，那就是，和标准热力学理论的预言相反，普通物质会存留下来，并能够存留出足够的数量来形成动物、城市和恒星。这种可能性也许只是因为物质的总量超过反物质的总量，即正反物质在宇宙中的非对称分布。如果正反物质的总量一直相等，那么正反物质粒子会找到彼此然后互相湮灭，最后消失。

显然，在宇宙演化的过程中，正物质的总量要超过反物质的总量。如果没有这种正物质的总量溢出，现存正物质的很大一部分都会消失。但我们并不知道为什么正物质的总量会超过反物质，正反物质总量非对称性的主因，只会是宇宙极早期的一些特殊相互作用和特殊环境。其中的一些物理过程一定已经超出了热力学平衡的范畴（例如某些物理过程非常缓慢，以至于无法跟上宇宙的膨胀），或者正反物质粒子的总量在创生期就是不相等的。进一步讲，也许看起来很自然的对称性在多余正物质被创造之后，是无法实现的。

我们既不知道是什么产生了对称性的破坏，也不知道是什么导致物理行为偏离了热力学平衡。尽管在大统一理论、轻子模型（轻子是指像电子和中微子一样不参与强相互作用的粒子），以及超对称模型中，都给出了一些建议。没人知道哪一个模型是正确的（如果有一个是正确的话），除非有人给出确凿证据。然而，目前这些领域并不存在直接的可观测证据。

尽管如此，我们依然可以相信，一种被称作重子生成（baryogenesis）的过程确实发生过，在此过程中产生了比反物质更多的正物质（正反物质不对称）。如果没有重子生成过程，我们也不可能在这里讨论这个故事了。

非对称暗物质模型指出，因为暗物质的能量密度和普通物质的能量密度如此相似，也许暗物质和反暗物质的产生过程中也存在正反暗物质的不对称。马修·巴克利（Matthew Buckley）和我一起研究过这一课题，那时他还是加州理工学院的博士后。我发明了“X合成”（Xogenesis）一词来描述这一过程，X表示未知的暗物质总量。真正有意思的事情是，这些模型允许暗物质可以像普通物质一样产生，这也是普通物质和暗物质之间有联系的最有意思的一个例子。如果暗物质和普通物质存在相互作用——即使它们之间的作用非常弱小或者曾经比较强，那么暗物质能量密度和普通物质能量密度之间的可比性，就不仅仅是巧合。这也是这些模型值得被相信的最好理由。

候选模型3：轴子模型

弱相互作用大质量粒子和非对称暗物质模型是比较常规的模型。弱相互作用大质量粒子类模型引入了弱尺度稳定的粒子，非对称暗物质模型指出了暗物质粒子和非暗物质粒子的非对称性。这两个不同的想法启发我们，要考虑模型建立的多样性，也许存在全新的粒子和相互作用。

轴子模型（Axion model）处理了一个更严格的情况。一个轴子只会出现在与一个特殊情况有关的模型中，这种特殊情况在粒子物理学中被称作强电荷宇称性问题（strong CP problem），其中，C代表电荷（charge），P代表宇称（parity）。电荷守恒定律告诉我们，正

负电荷的粒子是密切相关的。宇称守恒告诉我们，没有物理定律可以分辨出左右，例如右自旋的粒子和左自旋的粒子具有完全相同的相互作用。然而，在自然界中，不仅这两点对称性各自破坏了，连它们的组合也破坏了，也就是单独破坏C和P，并不会互相补偿。

轴子模型

轴子是物理学家及天文学宇宙模型中假想的暗物质构成粒子之一。在宇宙中，大量轴子常处于凝聚状态，轴子间通过极微小的力相互作用。

由于某种未知原因，电荷-宇称对称性破坏——因为C和P组合的对称性是可行的，只会发生在一些特定情况下。为什么CP守恒只限制在一些相互作用中，在标准模型里还无法解释，我们称之为强电荷宇称性问题。轴子被创造出来的目的，就是解决这个问题。

解释这些是为了方便读者理解。如果没有粒子物理学的相关知识储备，或者没有读过一本关于相关理论的书籍，我担心读者理解这些概念会非常困难。幸运的是，如果只为了理解轴子在宇宙尺度的预言及其成为暗物质粒子候选者的可能性，你不需要掌握粒子物理学的专业知识。轴子在宇宙尺度上的预言，只依赖于轴子是否足够轻，以及是否只具有极其微弱的相互作用。

你也许会认为，这些特性会使轴子变得不具破坏性，事实上，大部分物理学家一开始也是这么想的。但在一篇著名的文章里，理论物理学家约翰·普瑞斯基尔（John Preskill）、弗兰克·维尔切克（Frank Wilczek）和马克·怀斯（Mark Wise）解释了为什么即使超级轻、超弱相互作用的轴子也不一定是毫无破坏性的。他们指出，由于轴子太轻，而且相互作用又非常弱，所以轴子的存在不会影响早期宇

宙的能量构成。没有任何物理过程明确指出轴子粒子的存在总量。只有当宇宙演化到足够冷后，它们才会起作用。

因为轴子的密度在宇宙早期无关紧要，当最终轴子可以影响宇宙演化的时候，它的存在总量将不遵从宇宙中最普遍的方式，例如最低能量。宇宙中会因此出现大量的轴子粒子处于凝聚状态，所以即使轴子粒子非常轻，轴子凝聚态下的能量也会非常可观。一个令人惊喜的转折是，轴子的相互作用不会过弱或者说宇宙中的能量要比之前理论预言的多。

基于上述考虑，我们可以限制出轴子的相互作用范围。把这一假设推演到观测中：如果轴子的相互作用比较弱又不是非常弱，轴子会携带较高的能量密度，但又不用高到违反观测。事实上，如果轴子的相互作用恰好满足我们的假设，暗物质是有可能由轴子组成的，并精确地等于暗物质的能量密度。

轴子的质量完全不同于前面所描述的其他暗物质候选者。在那些理论中，暗物质粒子的质量应该在弱相互作用尺度上或者是其百分之一，而轴子理论预言的是极轻粒子，质量大约是弱相互作用尺度的十亿分之一。

轴子的相互作用也不同于其他暗物质候选者。从宇宙学和天体物理学的观测，轴子模型可以被限制到一个非常狭窄的质量窗口与相互作用强度的窗口。相互作用不能太弱，否则轴子携带的能量密度会过高；相互作用也不能太强，否则我们就可以在粒子物理学实验中或者恒星的内部观测到相关现象。因为在恒星内部，如果轴子的相互作用足够强的话，恒星会被冷却掉。关于超新星冷却率的观测表明，并不存在非标准模型没有预言的冷却贡献，这限制了轴子相互作用的强度。

理论上讲，考虑到轴子的相互作用窗口，我发现轴子模型有一个比较奇怪的地方：通过实验测定的相互作用窗口看上去像是随机的，它与其他物理机制并不存在明显的相关性。我有点怀疑通过实验寻找轴子是否会有正面的结果，但我的很多同事却都比较乐观。在目前正在实施的轴子探测试验中，人们假设轴子与光之间有着非常微小的作用。轴子探测器被放置在一个巨大的磁场里，人们通过探测轴子和磁场相互作用后所产生的辐射来搜寻轴子。对于这类实验，只有时间能告诉我们，自然界中是否存在轴子，以及它们是否确实组成了暗物质（如果轴子存在的话）。

候选模型4：中微子，被否决的暗物质候选者

目前为止我所介绍的模型有一个共同点：它们都包含着一些暗物质与普通物质的联系。例如，弱相互作用大质量粒子模型中的相似质量巧合，非对称暗物质模型中的近似能量密度，以及用于解决强电荷宇称性问题的轴子。轴子模型本是为解决粒子物理学问题而提出的假设，但也许对暗物质有用。弱相互作用大质量粒子属于粒子物理学范畴，它基于超对称理论。非对称暗物质模型也许同样存在于理论中，尽管暗物质和普通物质的相互作用假设是与已有理论无关的一个附加假设。

暗物质可能是纯引力相互作用的，或者只有某些暗物质是这样的。也许暗物质还存在自相互作用力，这些力无法被普通物质所感知到。

在提出“暗物质可能以一种独立于普通物质的形式存在”的假设之前，物理学家们最先考虑的是，是否存在一种普通物质，这种物质相互之间存在一种自相互作用，从而使之看起来像是暗物质。这样，

问题就变成了：在标准模型的理论框架下，是否存在一种由标准模型粒子组成的物质，它可以成为暗物质的候选者，而不需要引入额外的粒子？

这种假设最早想到的一种基本粒子叫作中微子（neutrino）。在一个被称作 β 衰变的辐射过程中，中子会衰变成光子、电子以及中微子（严格来说应该是它们的反粒子：反中微子）。与电子及它的对应重粒子（被称作 μ 子和 τ 子）一样，中微子不参与原子核的强相互作用，并且中微子自身不携带电荷，也就是中微子自身不直接参与电相互作用。中微子的一个有趣特点是，除了引力之外（当然所有粒子都参与引力相互作用，尽管作用程度非常微小），它们只直接参与弱相互作用。另外一个个性质是，中微子非常轻，质量至多是电子的百万分之一。

中微子

轻子的一种，是组成自然界的最基本粒子之一，它自身不带电，不参与原子核的强相互作用，也不直接参与电相互作用，可以自由地穿过地球，被称为宇宙“隐身人”。

由于它们的相互作用非常弱，所以中微子开始看起来可能是暗物质的候选者之一。但现在这种猜想已被否决，原因如下：中微子在标准模型中通过弱相互作用力实现相互作用，但是在下一章提到的直接探测实验中却没有发现相关的暗物质信号。最主要的是，我们所了解的常规中微子不可能是暗物质，因为它们的能量密度实在是太低了。如果宇宙中的中微子想达到已知暗物质的能量密度，那么中微子的质量要比已知中微子质量重得多。

事实上，轻中微子可以形成热暗物质（hot dark matter），这些暗物质粒子的运行速度接近光速。热暗物质会抹除小于超星系团尺度的所有结构，然而，我们可以观测到星系和星系团的结构，所以热暗物质在解释小尺度结构形成的时候，会出现问题。因此，标准物理模型范畴内的中微子并不能成为暗物质的候选者。随后物理学家们尝试着修改标准模型，结果也失败了。像中微子一样相互作用的粒子，就算是修改标准模型之后，也无法顺利解释第5章所介绍的结构形成。

从理论上讲，如果小尺度结构不是直接形成的，而是从大尺度结构碎裂而成的，那么热暗物质依然有存在的合理性，而且这个理论已经通过数值模拟给出了一些预言。但是，这些预言都无法与观测吻合。所以尽管存在一些新的轻中微子和关于中微子暗物质的头条报道，但中微子真的不是暗物质。中微子至多可能是目前存在的暗物质密度的一小部分。这就是为什么物理学家更愿意关注冷暗物质（cold dark matter）模型。在这种模型下，暗物质粒子的运动速度会很慢，质量通常更高。热暗物质模型（即像中微子一样质量很轻且运动极快的粒子模型）已经被排除了。

候选模型5：晕族大质量致密天体

最后，让我们了解一下暗物质。这种物质不需要新的基本粒子，由无法燃烧的粒子组成（即无光的辐射）、无反射的宏观结构的可能性。我们无法观测到这些天体，就像我们在黑暗房间里什么也看不见一样。从某种程度上讲，并不是这些物质不与光发生相互作用，而是因为你周围没有足够的光来发现它们。在接受暗物质存在之前，大多数人都想知道（无论具有科学的态度与否）：为什么看起来如此显而易见的可能性是不对的？

具有上述特性的暗天体被统一称作晕族大质量致密天体（Massive Compact Halo Objects, MACHOs），这个名字的命名和弱相互作用大质量粒子异曲同工。因为晕族大质量致密天体仅有一点甚至没有可观测的光的辐射，所以这类天体尽管由普通物质组成，依然藏身于黑暗的宇宙中无法被直接观测到。晕族大质量致密天体的候选者包括黑洞、中子星和褐矮星。

晕族大质量致密天体

表面上和暗物质行为很相似的一种普通物质，由无法燃烧的粒子组成，无反射的宏观结构。这类天体尽管由普通物质组成，依然藏身于黑暗的宇宙中无法被直接观测到。

我们前面介绍过，黑洞是物质超级紧密的一种引力束缚态，不发光也不反射光。中子星，可能由超新星爆炸后坍缩而成，是质量较大的恒星的超新星爆发后的遗迹，这些恒星的质量不足以形成黑洞，但其超新星爆发之后会形成致密的由中子组成核心。褐矮星是大于木星却小于正常恒星的一类天体的统称，它们由于质量太小而无法触发核聚变，只能通过引力收缩来加热自身星体。

上面提到的天体看起来有很大的可能性成为暗物质的候选者。但就算在以前的观测中，这种可能性也已经得到了严格的限制，晕族大质量致密天体不太可能是暗物质的候选者。第4章中提到了一个关于标准宇宙大爆炸理论的早期检测：原子核来自早期宇宙的核合成过程，这个过程被称作原初核合成。而这一过程只会发生在普通物质的某一特定能量密度范围内。大部分的晕族大质量致密天体模型都需要过多的普通物质，才能给出正确的核丰度预言。最重要的是，就算普通物质形成了这些密度很高的天体，但是搞清楚它们为什么通常分布在星系晕中而不是星系盘中，会是另一个重要的挑战。

即使如此，天体物理学家依然对各种模型保持一种开放的态度。暗物质是个非凡的研究课题，任何能证明常规解释无效的努力都是值得的。20世纪90年代，物理学家通过一种叫作微引力透镜（microlensing）的方法寻找晕族大质量致密天体。根据这个细致而美妙的想法，晕族大质量致密天体会偶然地在一个恒星的前面通过。因为光线会在晕族大质量致密天体（或者其他大质量天体）周围产生弯曲，所以这类天体会扮演着一个透镜的角色：它周围的引力扰动会临时放大背景恒星的亮度，使背景恒星先变亮，然后又恢复到正常亮度。当然，要观测到这一现象，这一光变过程的时间尺度需要足够小，光变尺度也需要足够大。通过这个方法，天文学家们给出了这样的结果：质量在1/3月球质量到100个太阳质量的范围内的晕族大质量致密天体，不可能是暗物质，因此其很多候选者都被排除了。

尽管晕族大质量致密天体的相关观测已经排除了中子星和白矮星是暗物质的可能性，但是一个狭小质量窗口内的黑洞依然可能是暗物质。不考虑没有足够的理论原因让我们相信在任意给定质量区间的黑洞的总量，恰好满足解释暗物质的要求，黑洞给出的引力扰动以及黑洞存在的时间，会对这个假设给出进一步的限制。太小的黑洞会在很短的时间内通过辐射光子而衰变掉，这个黑洞辐射过程被称作霍金辐射，以第一个指出这一过程的物理学家史蒂芬·霍金的名字命名。而大质量黑洞的预言现象还没有被观测到。这些预言包括：双星系统的引力扰动，可以加热和拓宽银盘结构的散射，黑洞对其他物质的吸积和辐射，以及通过精确测量脉冲星得到的黑洞引力波信号。把这些限制都放在一起，黑洞质量可能会被严格限制在一百万分之一个月球质量到一个月球质量之间，而在这个质量范围外的黑洞将被排除成为暗物质的可能性。对于中子星性质的细致观测，也许会把仅存的这个质量范围或多或少地再排除掉一些。

就算这个非常小的质量窗口依然存在，但弄清楚为什么只有黑洞会在这个质量范围内产生并保留下来，仍然非常困难。当然，考虑这种可能性无可厚非，可是根据核合成理论以及创建模型应该遵循的一些限制，黑洞（尤其是只有普通物质创造出来的黑洞）是暗物质的可能性是极小的。

是时候重新开始了

上面的模型包含了最常见的几种暗物质候选模型，这些候选模型也是物理学家们认为“具有合理存在的可能性”。但我们几乎可以断定，它们并不是仅有的选择。尽管其中一些想法看起来很有前途，但在这些特定模型被实验证实之前，我们有足够的理由怀疑其正确性。

另一方面，我们十分确信暗物质是存在的，尽管还不知道它到底是什么。现在，是时候让理论物理学家和实验物理学家重新考虑一个完备度更高的暗物质模型候选范围了。这些新理论大部分都具有不同的探寻策略。不同的模型会有助于制定探寻策略。

不过，在开始尝试新的想法之前，我会首先回顾一些已经存在的暗物质探寻技术，以便读者们具有足够的背景知识。我们会看到：虽然已经拥有了丰富的天体物理学数据，但既有模型观测的证据却仍然十分匮乏，这是督促实验物理学家和观测物理学家们抛弃过去的观测技术和策略的好理由，去开发出一些更加先进的寻找暗物质的观测方法或实验方法。



寻找暗物质是一段充满荆棘的旅程。

布朗大学的理查德·盖茨夏尔教授（Richard Gaitshell）是暗物质探测实验LUX（the Large Underground Xenon Detectro，大型地下氙探测器）项目的理论研究员和联合发言人。他于2013年12月在哈佛大学做过一次报告，台下众多全神贯注的听众来自哈佛大学物理系。他高兴地谈及，他及合作者们还没有发现暗物质。其实验的成功之处在于，他们已经排除掉很多暗物质候选者，其中包括一大类模型，甚至还得到了一些目前来看不可信的实验结果。尽管现在让人有些失望的是，还没有暗物质被实验发现的物理学新闻，但盖茨夏尔教授的高兴却并不是毫无理由的。他们设计并进行了极具挑战性的实验，实验按照他们当初的期望正常运行着。没有发现暗物质并不是盖茨夏尔教授和他同事的错，而是大自然不合作：大自然没有提供一个粒子质量足够大并相互作用足够强的暗物质候选者，以便让他们的仪器观测到。

这只不过是LUX实验公布的第一组实验结果。科学家们会继续通过LUX实验来收集更多数据，而这些新数据会超过之前结果的涉猎范围，而且一些设计更加巧妙的新实验设备也将投入使用。盖茨夏尔教授和

他的合作者们创立了一个非常干净的环境，这个环境保证了实验的最初结果就足够令人信服，并取代了以前其他仪器的结果。在暗物质探测试验中，如果一个粗心的实验物理学家不小心留下了指纹，这个指纹所造成的放射性信号的强度会是暗物质粒子辐射信号的10亿倍，而盖茨夏尔教授的实验在避免污染方面做得非常出色。这些仪器得到的纯净又可靠的数据，反映了他们所设计出的仪器精确地按照设计初衷工作着，既以极高的敏感度探寻暗物质信号，又能可信地排除任何一种误导信号。

许多数据都是通过当今最先进的技术收集到的，而这些在当今人们的消费领域中是看不到的。今天数据的积累会给粒子物理学、天体物理学、宇宙学以及其他科学领域带来长足的进步。尽管目前为止没有试验确定地发现了暗物质，但很多实验都得出了令人兴奋的结果。有时候，类似于盖茨夏尔教授等人的实验能排除的许多模型的可能性，这是以前的设备所无法做到的。盖茨夏尔教授的实验和其他先进的实验会继续搜索暗物质，并期望在不久的将来寻找到足够可信的暗物质信号。

寻找暗物质是一段充满荆棘的旅程。由于引力是一种非常弱的作用力，寻找组成暗物质的粒子需要准确地发现暗物质固有的相互作用，而对此，我们却一无所知。

如果暗物质只通过引力相互作用，或者普通物质无法感受到的新相互作用，传统的暗物质探测方法是永远也不会找到它的。即使标准模型的力也作用于暗物质粒子，我们依然无法确定这种作用是否强到可以被当今的仪器探测到。

如今的暗物质探测实验都依赖一个冒险的假设：尽管暗物质几乎是不可见的，但与其相关的相互作用可以被由普通物质制造的探测器捕捉到。这个假设算是一种一厢情愿，这种乐观也来自我们前文所

讨论的弱相互作用大质量粒子模型的一些暗示。大部分弱相互作用大质量粒子暗物质候选者应该与标准模型中的粒子存在微小的相互作用，虽然小，至少存在被当今的精确实验器材观测到的可能。现在的探测已经到达了一个临界点，超过这个临界点，要么绝大部分弱相互作用大质量粒子模型被排除，要么这些模型的粒子被观测到。我们在等待着这些实验的最终结果。

在第16章中，我们研究不同的暗物质模型，按理说这里我应该展示这些模型彼此不同的观测特性。但在本章，我只专注弱相互作用大质量粒子暗物质模型，以及现在比较流行的三种观测方法（见图17-1）。暗物质很难被捕获，但实验物理学家们利用他们的实验器材，无所畏惧地寻找着这些微小的可观测信号。

暗物质的直接探测实验

第一类搜寻弱相互作用大质量粒子的实验被归类到暗物质的直接探测实验中。直接探测暗物质的实验主要是应用地球上巨大但极其敏感的人造仪器。之所以要建设巨大的实验设备，是为了补偿暗物质极小的相互作用强度。这类实验的基本想法是，暗物质在穿过实验器材时，会与其原子核产生微弱的相互作用。这种相互作用会产生碰撞的热量或能量，理论上可以被观测到。当然，观测仪器的温度要非常低，或者建设观测仪器的材料要非常敏感，这样暗物质与原子核相互作用产生的能量才会被吸收或者记录下来。如果暗物质粒子穿过直接探测的实验器材，并被探测器的原子核轻轻弹开，那么实验器材就会记录下这个微小的能量变化。这个微小的能量变化可能是暗物质通过的唯一可观测证据。尽管这种相互作用发生的概率很低，但加大器材尺度和提高敏感度可以提高这一概率，这也是为什么这类器材都建得非常巨大的原因。

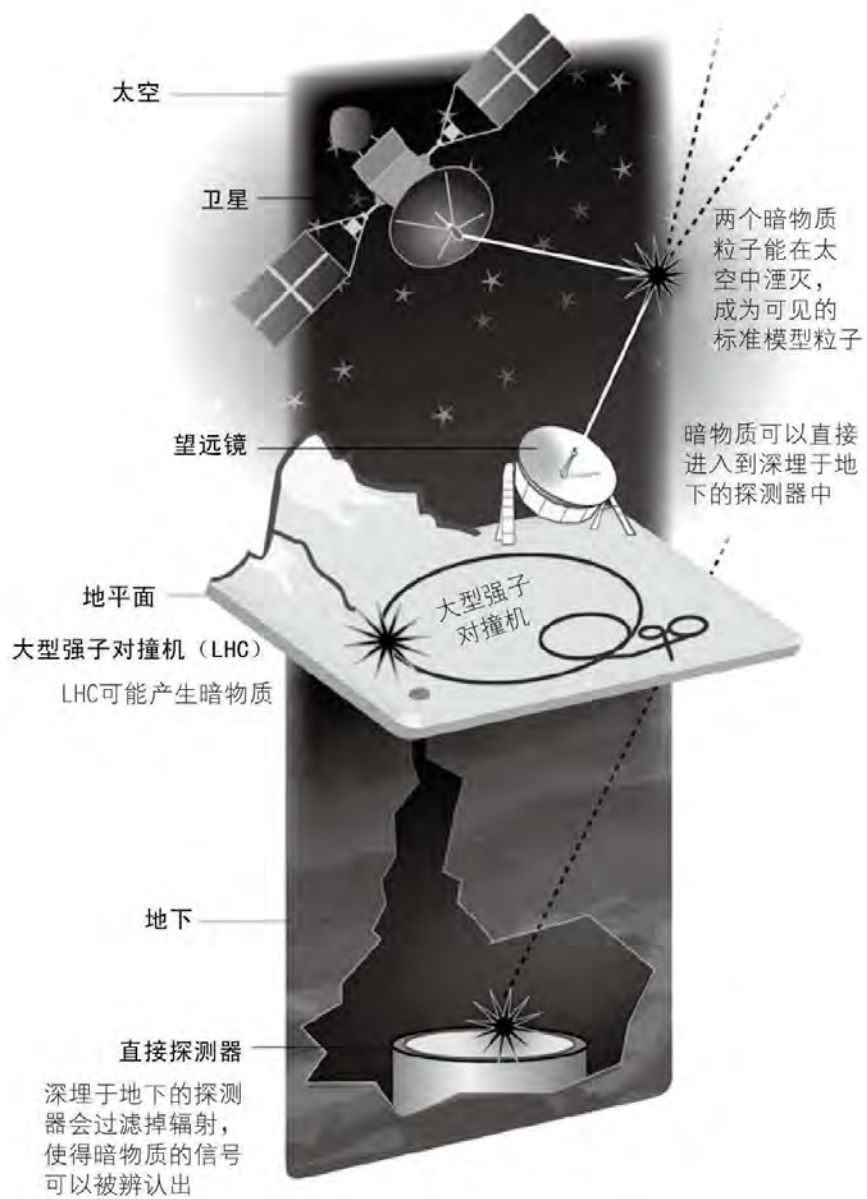


图17-1

寻找弱相互作用大质量粒子的三种主要方法。地下探测器主要寻找直接撞击靶核的暗物质。在大型强子对撞机中进行的实验主要寻找由大型强子对撞机产生的暗物质。卫星和望远镜则通过观测暗物质湮灭所产生的可见物质，来间接寻找暗物质粒子。

低温探测器（Cryogenic detectors）是由温度非常低的结晶吸收体（例如锗元素）构建而成。它们会通过仪器内部的超导量子干涉仪（superconducting quantum interference devices, SQUIDs）来对极少的热量作出反应。这种接收器的工作原理是，只要暗物质与原子核碰撞产生的极小热量接触到探测器，探测器中的极低温超导体就会失去超导性，从而把这一可能的信号记录下来。这类实验器材包括低温暗物质搜寻（Cryogenic Dark Matter Search, CDMS）、低温罕见事件探测器（Cryogenic Rare Event Search with Superconducting Thermometers, CRESST）、地下弱相互作用大质量粒子探测站（法语：Expérience pour Détecter Les Wimps en Site Souterrain，即 Experiment to Detect WIMPS in an Underground Site, EDELWEISS）。

低温探测器并不是暗物质直接探测的唯一方式。另一种方式要用到一种非常昂贵的液体，这种方式现在变得越来越重要。尽管暗物质与光不直接作用，但暗物质与氙或者氩相互作用所产生的附加能量，有可能会产生一种特性明显的闪烁。这类实验包括氙基暗物质探测实验XENON100和大型地下氙探测器LUX。这里所提到的实验同样存在氙基版本的探测器，分别被叫作ZEPLIN、DEAP、WARP、DARkSide和ArDM。

XENON和LUX在接下来的几年里都会有一个尺度和精度上的升级，它们被称作XENON1T和LUX-ZEPLIN合作项目。为了表明升级带来的进步，前面提到的“100”以千克为单位，而升级后的1T则表示1吨。LUX-ZEPLIN的升级会更加明显，届时探测器中的液体体积将达到5吨。

低温探测器和惰性气体探测器都是为了记录暗物质与其原子核相互作用而可能产生的微小能量。尽管这种信号的存在令人印象深刻，但是探测到如此小的信号并不足以说明暗物质粒子确实在此通过。实验物理学家们还需要确定他们所记录的信号不是来自背景辐射。因

为，这些背景辐射同样可以产生类似于暗物质与试验物质相互作用而产生的微小能量。而且，背景辐射与实验物质的作用要比暗物质还要强。

这种情况就比较棘手了。扰乱高灵敏度暗物质探测器的辐射到处都是。宇宙线中的 μ 子是电子较重的“兄弟”，可以打到岩石上并产生其他粒子，其中包括一些中微子^[35]可以产生类似于暗物质的行为。就算对暗物质粒子的质量与相互作用强度作出乐观假设，背景电磁事件依然主导着信号，并且这一信号的强度是暗物质信号的1 000倍都不止。而且，这一估计还没算上所有的原始效应，大气中存在的人造辐射污染，环境以及探测器本身等。

设计这些实验器材的科学家对这些背景噪音非常了解。天体物理学家和暗物质实验物理学家管这个游戏叫作“遮挡和鉴别”（shielding and discrimination）。为了防止仪器受到危险辐射的影响并分辨出潜在的暗物质信号，实验物理学家们把探测器建在很深的矿井中或者大山之下。这样，宇宙射线就会打到周围的石头而不是探测器本身。这样，大部分背景辐射都会被屏蔽掉，而暗物质与其他物质的极弱相互作用就会无阻挡地作用到探测器上。

幸运的是，大部分为商业目的而建造的矿井和隧道都能够容纳这些实验设备。矿井存在的部分原因（如前面提到的一样）是寻找向地心沉积的重元素，但某些重元素偶尔也会上升而储存到地下的矿石中。DAMA实验是XENON10的延伸项目，并且比XENON100还要大，甚至达到CRESST的尺度。DAMA实验主要使用钨，隶属于格兰萨索国家实验室（Gran Sasso National Laboratory），被建造在意大利1 400米深的地下隧道内。

LUX实验位于一个1 500米深的洞穴中，位于南达科他州的霍姆斯特克矿井（Homestake），这个矿井最初是为了挖掘黄金而建造的。霍

姆斯特克矿井在物理圈非常著名，因为另外一个著名的探测器也建设于此：该探测器是用于探测太阳中微子的，正是这个实验为人类揭示了中微子的非零质量。低温暗物质搜寻实验则建在位于地下750米地下的苏丹（Soudan）矿井中。还有位于加拿大安大略的萨德伯里矿井，它最初用于挖掘20亿年前撞击地球的一个小行星中的矿物质，这个矿井里面也有几个暗物质直接探测器。

并不是所有矿井和隧道上的石头都能够保证对探测器的零辐射影响。实验物理学家们会尝试不同的方法来保护探测器不受外界辐射影响。我所知道的最有趣的防护屏障是一种古老的铅，这些铅来自一艘已经沉没于海底的法国大帆船。铅是一种致密的吸收性材料，年代久远的铅则已经没有了自身的辐射，所以它能有效吸收外来辐射，而不产生自己的新辐射。

另外，随着更先进的屏蔽措施不断出现，例如，聚乙烯会因为有相互作用而发光，这些相互作用因为过强而不可能来自暗物质。在惰性液体探测器中，例如氙基探测器，容器本身就是一个屏障。这些探测器的吸收区域非常之大，以至于实验物理学家们可以忽略来自探测器外部的信号。其实，这些外部的氙就是用来屏蔽背景辐射的。

鉴别实验结果也很重要，粒子物理学家称其为粒子识别码（particle ID）。与屏蔽过程相反的鉴别过程，主要是区分信号中属于暗物质事件的部分和背景噪音的部分。通过测量电离和固有闪烁，实验物理学家们可以确定哪些信号是来自背景的。

闪烁暗物质探测实验DAMA的负责团组已经声称发现了一些暗物质事件的可能信号。但这个实验缺少背景信号和事件信号的鉴别过程，只有“信号”出现的时间信息，加上其他人的实验并不能重复DAMA实验的结果，大部分物理学家对其结果的真实性持怀疑态度。

其他实验也记录了一些潜在的信号，但次数极少，并且都是在低能标区域（人们当然也有足够的理由怀疑这些结果）。这些探测器主要是测量反弹时的能标变化。如果能标过低甚至低于探测设备的敏感度极限，探测器将无法记录这些事件。这些事件的最低能标已经非常接近低能标探测的极限。所以对低能标信号的怀疑是合理的，只有更多的数据或其他实验对某一结果作出双重确认，人们才会相信这些信号来自暗物质。

暗物质的间接探测

直接寻找穿过地球的暗物质实验可能会成功，并发现暗物质粒子。其他一些可行的暗物质探测方法也开始兴起，比如基于暗物质粒子湮灭的间接探测（相同种类的粒子可以发生湮灭）。这些湮灭可以将暗物质粒子的能量转化成其他形式的能量或物质，我们希望这些能量或物质是可见的。暗物质湮灭也许不会发生得非常频繁，因为宇宙中暗物质的分布非常稀薄。这并不意味着暗物质完全不会湮灭，湮灭概率取决于暗物质粒子的自身性质。

当暗物质湮灭发生时，地球上或者太空中的实验器材也许会发现湮灭过程中所产生的粒子，这一过程被称作间接探测。这种探测寻找的是暗物质粒子湮灭消失过程中所创造出来的粒子。如果我们足够幸运的话，这些新创生的粒子中可能包含标准模型框架下的粒子和反粒子，例如电子和它的反粒子，即正电子。或者，湮灭的过程可以产生光子对，而这些光子有可能被地球上或太空中的探测器所捕捉到。反粒子和光子信号是暗物质间接探测中最有希望的搜寻目标，因为反粒子在宇宙中非常稀少，所以暗物质湮灭所创生出的反物质会很容易被观测到。湮灭过程产生的光子信号也非常有用，因为这些源自暗物质

湮灭的光子，会和天体物理学中的其他过程产生的光子具有不同的能量和空间分布。

大部分用于寻找暗物质湮灭所产生标准模型粒子的探测器，一开始并不是为暗物质探测而专门设计的。地球上或太空中用于暗物质间接探测的望远镜和探测器，最早主要用于记录太空中的天体所发射的光信号或者粒子信号。目标天体都是我们目前比较熟悉的恒星、脉冲星或其他天体，而这些天体所发射出的信号对于搜寻暗物质的实验物理学家们来说，都是背景或噪音，这些信号甚至还有可能给出假的暗物质间接探测信号。

从另一个角度来说，由于天体物理学的背景源和假定的暗物质湮灭过程，都能发出相似的粒子辐射，所以现有的望远镜同样有可能告诉我们有关暗物质的信息。如果天体物理学家能够很好地理解传统源的粒子辐射，他们就可以通过扣除已知辐射，来计算暗物质湮灭所产生的多出来的粒子辐射信号。尽管还有一些模糊不清的解释，但如果传统源已经完全搞清楚并能确保它们可以完全被排除出去，那么暗物质间接探测就算成功了。

有一个暗物质间接探测装置位于国际空间站上。来自麻省理工学院的诺贝尔奖获得者丁肇中有一个聪明的点子：人们可以在国际空间站上安装一个寻找正电子和反质子的探测器。这个被称作阿尔法磁谱仪（Alpha Magnetic Spectrometer, AMS）的粒子探测器被最终发射到太空中。它的科学目的是继续一个原来由意大利负责的PAMELA卫星的科学任务。PAMELA卫星于2013年发布了第一份科学报告，而现在它的使命已经完成了。

尽管一些实验数据在一开始看起来非常有趣，但源自暗物质的可能似乎不太大，因为PAMELA和AMS的信号要求宇宙极早期有大量的暗物质，而这些暗物质对宇宙微波背景辐射的扰动应该能够被普朗克卫星探测到。目前来看，一开始令人惊讶的结果应该是由于天体物理学家

对类似脉冲星这类天体的理解不透彻。如果传统天体可以给出这些信号的合理解释，那么这些信号是不足以证明暗物质的存在的。

暗物质也有可能湮灭成夸克和反夸克，或者直接变成胶子，而胶子是可以直接参与强相互作用力的。事实上，大部分弱相互作用大质量粒子类的模型预言，这一湮灭过程所释放出的信号是标准模型框架下最有可能观测到的信号。天体物理学领域里最明显的目标应该是反质子，而那些低能量的反氦原子却十分稀少。因为，反质子和反中子会非常微弱地束缚在一起。实验有可能通过暗物质湮灭到这些低能态所放出的能量，寻找暗物质。一个建在气球上的暗物质探测设备GAPS将于2019年在南极大陆升空，用来寻找此类数据。

不带电的中微子通过弱相互作用力也可以帮助人们实现暗物质的间接探测。假设暗物质也许会被束缚在太阳或者地球的中心，这些暗物质的密度要比宇宙中暗物质的典型密度高出很多，从而大大提高了湮灭概率。在这一湮灭过程中，唯一可以逃离并可能被探测到的粒子便是中微子，因为中微子和其他粒子最大的不同就是，中微子与其他粒子的极弱相互作用无法阻止其逃逸。一些建设在地下的大型探测器，例如AMANDA、IceCube和ANTARES，就是用来寻找这些高能标中微子的。

还有一些探测器，人们用它们寻找高能光子、电子以及正电子。高能全息系统（the High Energy Stereoscopic System, HESS）位于纳米比亚，高能辐射成像望远镜阵列系统（the Very Energetic Radiation Imaging Telescope Array System, VERITAS）位于亚利桑那，它们都是建在地球上为寻找星系中心的高能光子而建造的。下一代高能伽马射线观测站，契伦科夫望远镜阵列（Cherenkov Telescope Array），灵敏度会更高。

在过去几十年里，最重要的暗物质间接探测仪器可能要数费米伽马射线太空望远镜，常被简称为费米，其命名是为了纪念已故意大利物理学家恩利克·费米，费米子也是以他的名字命名的。费米天文台放置在一个卫星上，它于2008年初发射，轨道高度是550公里，绕地球一周为95分钟。地球上光子探测器的最大优点是，人们可以把设备做得很大。但是费米卫星上的精确设备具有更好的能量分辨率和定位信息，对低能光子更加敏感，并具有更大的视场。

费米卫星一直被认为是诸多暗物质有趣推断的源头。自从费米卫星开始工作以来，有几次看上去比较真实的暗物质间接探测的信号出现过，但没有一个是决定性的证据。然而，所有这些数据都可能帮助我们了解暗物质到底是什么。最强的信号报告来自费米国家加速器实验室（FNAL）的物理学家丹·胡珀（Dan Hooper），这个实验室位于伊利诺伊州的巴达维亚，离芝加哥很近。丹·胡珀和他的合作者通过仔细的观测和数据处理发现，银河系中心的弥散光子辐射要超出天体物理学家给出的预言。

和先前令人惊讶的正电子结果一样，观测数据确实表明与理论预言的不符。问题再次变成：这些未被观测到的物质到底是一些被忽略的天体物理源，还是真正的暗物质？天文学家们还在为寻找答案而努力着。目前为止，并没有一种完全直接或令人信服的解释。

另一个信号是一个几电子伏的X射线发射线，大约只有电子携带能量的1%，然而目前看来，并没有传统的天体物理源可以产生这类信号。观测表明这个X射线信号是一个发射线，也就是说，那些高出理论预言的光子来自一个特定的能量区间，并且区间很窄。另外提醒一下读者，在原子和分子级别的辐射过程中，不同能级的跃迁辐射可以给出类似的发射线，但信号通常都不会非常强，所以我们还不知道这不是一个关于暗物质的重大发现。由于可信证据的匮乏，基于轴子或衰变暗物质源的研究也未停止。在进一步的数据和理论研究表明这个

信号确实是一个扰动，或是背景，或是真正发现了一些新东西之前，我们什么也不知道。

我要提的最后一个推测信号是一个具有130GeV能量的光子信号，这是一个费米望远镜所能观测到的信号量级。这个信号着实非常有趣，它和希格斯玻色子的质量相当（希格斯玻色子的质量大约是125GeV）。由于缺少合理的天体物理辐射源的解释，一些天文学家认为，这个信号可能来自暗物质湮灭。

我过去认为这些先前的证据都经不起时间的考验（或者说新数据的考验），现在看来都是错的。在尝试解释这些信号可能来源的过程中，我的同事马特·里斯、范吉吉（JiJi Fan）、安德雷·卡茨（Andrey Katz）刚刚完成了一类有趣模型的研究，这类模型为我们展示了以前从未被发现的其他方面。经过很多科学性的发展，这类模型变成了一个超越其初始动机的模型，我会很快对这一模型加以解释。

大型强子对撞机中的暗物质

尽管目前看来弱相互作用大质量粒子是暗物质的可能性越来越小了，但它们依然有可能出现在大型强子对撞机中。大型强子对撞机是位于法国和瑞士交界处的日内瓦附近的环状巨型粒子加速器。两束质子流会以相反方向环绕周长约为27公里的环形粒子加速器，并最终以极高的能量相互碰撞。大型强子对撞机跨越了很大的能量范围，其中包括希格斯玻色子创生与发现的能量段。也许大型强子对撞机还可以产生其他假设的粒子，例如稳定的弱相互作用大质量粒子。如果这是可能的，那么弱相互作用大质量粒子和标准模型中粒子的相互作用，也许会在大型强子对撞机中得以发现。

就算大型强子对撞机发现了新的粒子，人们依然需要其他实验设备的辅助，来确定这些新创造出来的粒子就是暗物质。例如，我们需要其他地下或者太空中的暗物质探测器来对新粒子的性质进行反复确认。不管怎样，在大型强子对撞机实验中发现弱相互作用大质量粒子都将是一项重大发现。我们也许还会发现暗物质粒子的其他性质，而对这些性质的研究，是其他探测方法很难实现的。

暗物质粒子与质子的相互作用就算在大型强子对撞机的高能碰撞条件下，依然很弱，因为其碰撞截面实在是太小了。尽管如此，其他粒子却可能通过衰变转化成暗物质粒子。这样，问题就变成了，如何确定大型强子对撞机已经创造出了暗物质。因为暗物质不会与探测器相互作用，因此也无法产生相应的可观测证据。

带电粒子的衰变是一个值得关注的点。带电粒子在衰变时不会简单地变成中性的暗物质粒子，因为这样的衰变过程是电量不守恒的。通过搜寻衰变后的带电粒子，人们可能发现，最终被发现的粒子所携带的能量与开始时的不一样了。这是由于暗物质可能带走了一部分能量和动量，一类微弱相互作用粒子可能由此被发现。



暗物质产生的征兆是，除了和预言的事件发生率以及数据中的信号一致之外，能量无法准确地等同于实验装置探测到的能量。除非这一过程中存在的物理定律与已知物理定律有着根本不同，否则丢失的能量以及角动量的唯一解释是，这一过程创造出了没有探测到的未知粒子，而这种粒子有可能就是我们说的暗物质。

尽管弱相互作用大质量粒子与普通物质的相互作用极弱，它们还是可以直接被成对地创造出来。两个碰撞的质子很有可能制造出两个弱相互作用大质量粒子。这个过程是两个弱相互作用大质量粒子湮灭产生普通物质的逆过程，而关于弱相互作用大质量粒子湮灭的计算，则能给出宇宙已知残余物质丰度的结果。质子碰撞生成弱相互作用大质量粒子的发生概率，在不同模型下会非常不同，毕竟弱相互作用大质量粒子不是必须湮灭成质子，所以这个逆过程也无法保证质子碰撞就产生这类粒子。但对于很多模型来说，这一逆过程是寻找暗物质粒子的一个有效途径。

重申一遍，实验物理学家们需要解决的，不是探测暗物质粒子本身，只有在生成暗物质过程中所产生的伴随粒子可以被观测到。但实验物理学家们可以观察产生暗物质粒子过程中伴随生成的光子或胶子（一种联系夸克间强相互作用力的粒子）的行为。而且，理论物理学家们也指出，这种观测策略可能得到足够强的观测信号。

目前为止，大型强子对撞机的研究结果中并没有发现有关暗物质产生的信号。物理学家们不确定，是不是因为机器的能标过低，或者因为理论上值得关注的伴随粒子会在这一能标区间被发现的预言，是一种误导。但在大型强子对撞机的能标区间内寻找因碰撞而产生的伴随粒子，是非常可能的。而这些粒子也许有一种正是暗物质。

如何用引力寻找暗物质

与“绝地武士”欧比旺·克诺比（Obi-Wan Kenobi）不同，弱相互作用大质量粒子不是我们唯一的希望，尽管目前所有的探测方法都喜欢寻找这类粒子，它们却只是我们众多选择中最好的一个。暗物质直接探测可行的基础，是暗物质与标准模型粒子之间存在着相互作

用，而弱相互作用大质量粒子模型确保了这种可能。另外，热产生保证了暗物质粒子与反暗物质粒子（或者说暗物质粒子就是它自己的反粒子）的总量相等，这样暗物质湮灭理论是完全可能的。但我们如何寻找其他模型预言的暗物质呢？

然而，其他还没被排除的暗物质候选者好像比弱相互作用大质量粒子更加难以探测。针对不同的模型，人们要选择特定的探测策略。有些探测策略的可行性甚至超出了人类目前的技术极限。也许我们比较幸运，暗物质不是完全透明的，它只是非常透明，基于目前标准模型相互作用的乐观假设还无法发现它。但根据既有的不确定性，我的观点是：是时候更多地关注一下如何用引力寻找暗物质了。暗物质与自身以及其他不可见物质的相互作用也许不会直接呈现在我们面前，但我们可以通过宇宙中物质的分布来寻找这些物质的相互作用，而这些相互作用会让暗物质自己呈现在我们面前。

A large, stylized number '18' in a light gray font. The number is surrounded by several small, dark gray dots and lines, suggesting a starry or particle-like background. The '1' is a simple vertical bar, while the '8' has a more complex, rounded shape with a small dot in the center of its upper loop.

暗物质的社交天性

DARK MATTER AND THE DINOSAURS

暗物质并不是我们想象的样子。就像当我们过于关注黑与白的时候，往往会忽略灰的存在，更不用说斑点和条纹了。

城市化已经成为现代社会进步的必要因素。把足够多的人连接起来可以让思想闪耀，让经济繁荣，让其他大量有益的事情得以发生。城市的成长源于城市的扩张，城市会不断地吸引越来越多的人到来，因为这里有更多的工作机会、有更好的工作和生活的条件。但是，一旦城市变得非常拥挤，房价不断升高，犯罪率居高不下，或者出现其他城市综合征，就会致使人们搬到人口没有那么密集的城市周边，甚至更远的地方，完全远离市中心。城市的其他部分也许会按照规划正常地发展，但是城市开发者那过于乐观的态度，很可能被市中心高楼林立却依然不足以提供住处的现状所打击，这一现象讽刺地反映了城市的过快发展。另一方面，没有稳定的市中心，郊区的社区也无法繁荣起来，这样，购物中心的开发者们也会非常失望。

这些现象看起来和宇宙结构的形成与演化颇为相似。我已经解释了人们今天对暗物质的理解。现在的很多观测以及预言都表明，暗物

质与普通物质的相互作用非常微弱。在基于暗物质的数值模拟中，假设暗物质与普通物质只存在引力相互作用，由此给出了关于星系和星系团的尺度、密度、集中度分布以及形状的相关预言。与对大尺度城市的成长预言一样，宇宙大尺度结构的预言与观测符合得非常好。

精确的宇宙学数值模拟在小尺度上与观测符合并不怎么好。星系或者星系团中心部分的质量分布，以及银河系周围的小质量矮星系的数目，与理论预言相差很多。就像人口密度没有那么高的城市中心部分和欠发达郊区一样，数值模拟对星系中的密度以及卫星星系数目的预言都太高了。仙女座星系以及其他星系中矮星系的数目，与理论预言的空间分布也有很大偏差。

这也许说明了，数值模拟存在问题或者观测数据还非常不完善。但在小尺度结构上理论与观测的不符也给了我们一些暗示：暗物质并不是我们想象的样子，也许暗物质自身的相互作用一点都不弱。

尽管暗物质与普通物质的相互作用很小，一个暗物质粒子与另一个暗物质粒子的相互作用却可能很大。这种自相互作用暗物质由于探测设备限制，并没有受到很强的观测限制，因为目前的暗物质直接探测实验，都是研究暗物质与普通物质的相互作用。这些自相互作用也许会大到值得人们去关注。

我们现在可以确定宇宙中结构的形成与演化的基本概念，但这些可能的理论与观测的不自洽表明，科学还没有发展到给暗物质本质下定论的程度。对于研究者来说，这非常好。我们必须学习很多新东西，不管结果怎样。这一章我会阐述宇宙小尺度结构（Small-scale structure）的相关问题，并说明为什么自相互作用暗物质可以解决这些问题。

小尺度的不吻合问题

第5章提到，暗物质粒子的引力作用如何决定了宇宙的结构形成。暗物质通过早期宇宙的密度涨落，演化成拥有较深引力势阱的暗物质晕，星系则从这个暗物质晕中诞生并成长，这里大部分利用的是引力的吸引作用。星系一旦形成，便开始向位于大尺度薄片结构和网状结构中的星系团中合并，这样就为其他结构的形成提供了基本的平台。尽管单个星系或者星系团的演化细节取决于其所在位置的原初条件，但天文学家们可以预言星系和星系团统计上的性质分布，而且这些预言大部分与观测符合得非常好。

对于小尺度结构（即矮星系尺度的结构）的预言却不那么可靠。关于星系中心密度分布的计算给出的预言太高了，关于银河系周围的卫星星系数目的计算，给出的预言也高出观测很多。观测天体物理学家无论是在大质量的晕中，还是独立的、小质量的晕中，都没有发现足够数目的小尺度结构。这和等级式成团理论的预言完全不符，并且这个问题一直持续到今天。

此类问题最著名的应该是核-尖峰问题（core-cusp problem）。天文学家和宇宙学家不但预言宇宙中存在天体的种类，还能预言出这些天体中物质是如何分布的。关于星系中心质量分布的预言是：星系或星系团的密度轮廓（density profiles，即密度随半径变化的方程），在中心处应该存在一个尖峰结构（cuspy）。这就意味着，暗物质的密度在星系或者星系团的中心会升高得异常快，并在中心区域形成一个极端高密度的区域。

然而，观测天体物理学家们所测量的密度，在某种程度上无法和这个预言相一致。事实上，根据现有的观测数据，大多数星系不存在核-尖峰结构，而是存在一个平缓的核结构（cored profiles，见图18

-1)。这让人迷惑不解，并不只是因为三星智能手机的“Galaxy Core”。对于大多数人而言，核结构应该是一个很高密度的结构，比如地球的熔岩核。这里的核结构却恰好是相反的情况。核结构意思是说，本来的中心高密度区域被磨平了，就像你从一个苹果中切掉了苹果核一样。当然，没人把星系的整个内核挪走。但观测表明，星系中心的物质密度并不像预言所描述的核-尖峰结构，而是密度分布相对平缓的核结构。星系团的情况也十分类似。

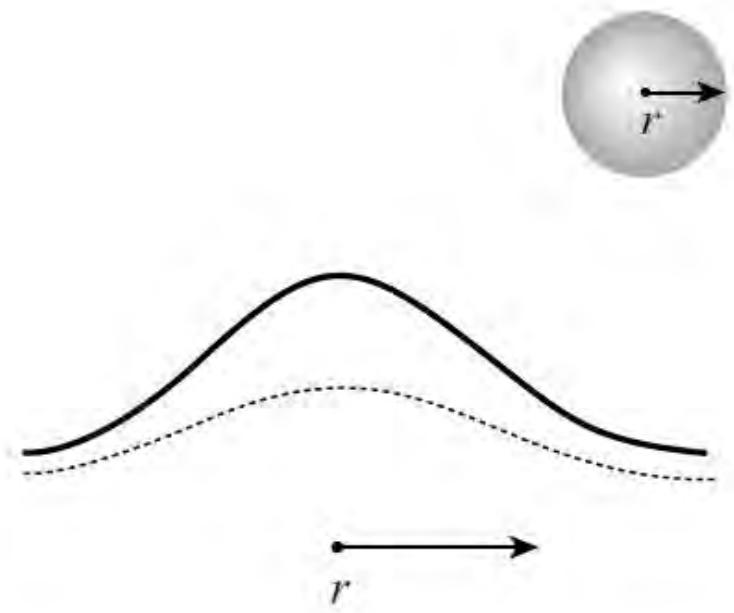


图18-1

数值模拟预言：在星系中，暗物质的质量分布会在星系中心呈核-尖峰结构，也就是说，很大一部分物质会存在于星系中心。但是，观测给出的却是一个平滑低密度的核结构。图中给出了这两种密度分布的示意图，但是核-尖峰结构在中心的密度分布比图中展示的更为陡峭。

解释为什么星系或者星系团中心的质量分布是平缓的核结构，而不是暗物质数值模拟所预言的核-尖峰结构，已经成为最简单暗物质模型的最重要挑战。这个问题和丢失的卫星星系问题（missing satelli

te problem, 中心星系的卫星星系数目远小于理论预言), 以及大质量星系理论与观测不符问题 (too big to fail problem, 即最致密、最大质量星系的相关观测与理论预言不符), 也许指出了标准暗物质模型的不足。

另外, 最近还有更多值得注意的暗物质预言问题, 例如暗物质盘模型, 这个我很快会谈到的。还有, 大星系周围的卫星星系空间分布并不是天文学家所预料的那样。以前, 天文学家们认为, 卫星星系应该或多或少呈球对称分布, 但是仙女星系中的大约30个卫星星系中, 几乎有一半都在一个近似平面内运行, 而且这些卫星星系的轨道方向也大体相同。银河系的卫星星系有些也遵循这种奇怪分布。

矮星系的这种近似在同一平面且相同旋转方向的性质也许说明, 它们起源于合并星系的盘结构。就算合并能解释矮卫星星系的分布, 却无法解释矮卫星星系中包含了太多的暗物质。这时也许需要一种非标准的暗物质模型, 来解释暗物质主导的矮星系为什么会沿着一个平面分布。

对于现在已知的差别, 数值模拟的结果和观测结果都还很初级。如果我们假设观测结果或者数值模拟有一个是错的, 其中一些问题就能被解决。更高精度的数值模拟也许会证明原来结果的不准确, 或者我们对普通物质的行为了解不足, 比如超新星对结构形成的影响。在这种情况下, 保守的暗物质模型也许可以解释现在在宇宙中观测到的结构, 甚至不用对暗物质的性质做任何修改。但如果问题依然存在, 这种差别会让人怀疑最简暗物质模型的可靠性, 并进一步说明我们需要更加复杂的暗物质模型。

仔细观察现有的结果, 来自20世纪90年代的一些记忆也许会让我们稍微振作一点。早期的数值模拟结果与观测数据非常不符, 因为当时并没有引入暗能量。许多科学家都有类似的想法: 这些早期的数值

模拟和观测结果都不可信，人们需要不断改进观测和数值模拟的结果，来使理论预言和观测数据符合得更好。当人们把暗能量（一个全新的发现）加入到数值模拟中时，理论和观测的不符消失了。为此，我们应该向这些早期的数值模拟致敬，正是这些数值模拟给出的精确结果表明：想解决这些问题，人们需要考虑新事物。也许当今关于小尺度结构与观测不符的窘境和当时的情况类似，当我们发现宇宙中物质和能量的新的不为人知的物理学特性时，这个问题就会得以解决了。观测科学和计算科学在接下来10年的发展将带来新的结果。

可能的暗示

尽管还没有确定的结论，很多天体物理学家和宇宙学家已经开始认真地研究这种小尺度的不吻合问题，而且已经开始研究暗物质粒子除了参与引力作用之外的其他相互作用的可能性。有些科学家甚至走的更远，例如，认为爱因斯坦的引力方程并非完全正确。尽管有一少部分物理学家专心于修正对引力的认识，但是我发现这个方向也许要比暗物质的相互作用更加让人难以置信。因为我前面已经讨论过，常规引力与暗物质作用的证据还是非常令人信服的。

解释类似于子弹头星系团的观测，是对修正引力最有难度的挑战。对于这种或者类似于这种正在合并的天体，在合并过程中那些相互作用较强的热气体会被留在天体的中间，而相互作用较弱的暗物质会相互穿过而留在天体的外部，这类观测的最佳解释就是，微弱作用的暗物质只参与了引力的相互作用。任何情况下，在考虑没有理论基础的激进变革之前，我们应该先考虑其他一些比较“乏味”的尝试。例如，有些理论预言也许给了我们误导的结果，比如普通物质在结构形成中的作用也许比我们想象的要大，或者暗物质的性质要比保守理论的期望复杂得多。

我最近参加了两个会议，会议的主题是小尺度结构问题和可能的解决方案。第一个是一个小规模研讨会，是由在哈佛大学物理系粒子物理学领域的同侪组织的，主题是关于自相互作用暗物质。来自哈佛大学天体物理中心的天体物理学家们组织了第二个会议，会议的主题是“关于暗物质的辩论”，这个会议在2014年春天举行。幸运的是，这次辩论是关于本质而非看法。因为，有时过于强调看法也许会使讨论偏离科学性。

我认为这两次会议值得参加的原因是，它们给与会者提供了很多与来自哈佛大学的物理学家和天文学家交流的机会。哈佛大学天体物理中心是天文学家工作的场所，创建于1847年，位于坎布里奇城的最高点，主要用于存放一个直径大约为130厘米的望远镜，这在当时是最大的望远镜。尽管这个望远镜现在不再用做科学用途，但天文学家们依然在此工作。不过这使得天文学家和物理学家相隔2 000米左右，所以两帮人从未在饮水机旁或者咖啡机旁偶遇过。这个会议使得我们以及很多来自世界各地的物理学家和天文学家得以共处一处。

会议的主要成果是，它所展示的结果都是原创的，而且非常新颖。会议报告的题目涉及一些关于小尺度结构问题令人兴奋的证据，以及可能的解决方案。与会者对此也讨论了到底是对普通物质作用的错误估计导致了小尺度问题，还是像暗物质自相互作用这类新颖的理论可以解决小尺度问题。

会议报告讨论了，为什么普通物质可以在小尺度上影响结构的形成，并且在数值模拟中引入了普通物质，这算是小尺度问题最无新意的解决方案。但要解决观测和理论的差别，这一尝试还有很长的路要走。原来的数值模拟假设：暗物质主导动力学和宇宙的结构形成与演化，相反，普通物质只是简单地落入暗物质晕的引力势阱中。尽管普通物质会在恒星形成后照亮大质量、高密度区域，但普通物质除了可以点亮暗物质高密度区域外，它们对结构形成的影响可以忽略不计。

物理学家们则不这么认为，普通物质不会明显地影响结构的形成和演化，只是为了计算简单，而不能作为一种可靠的假设。就算今天，当天文学家们尝试着引入普通物质的一些作用时，也会同时引入大量的不确定性。根据现有的存储能力和计算能力，没人能够把所有物理细节都模拟到，所以天文学家们在做数值模拟时，需要做一些近似和假设。尽管目前数值模拟存在诸多限制，但引入普通物质的行为，确实缓解了一些小尺度结构不吻合问题。

几个效应改进了数值模拟和观测的不符合。标准模型下的普通物质除引力之外还存在其他相互作用，所以尽管它们原初的引力扰动相对小一些，但它们对结构的影响，尤其是小尺度结构的影响也许并不会那么小。例如，丢失的卫星星系问题的一个潜在解释是，这些卫星星系太暗了，我们还无法观测到。星系间气体可以被来自恒星的紫外辐射加热，这个辐射与普通物质相关。一旦这一过程发生，小质量的暗物质晕会很难再有效地吸积气体。如果这些暗物质晕没有足够的气体，它们就无法形成恒星，从而使得它们由于太暗而无法被现有的望远镜观测到。

超新星爆发

它是某些恒星在演化接近末期时经历的一种剧烈爆炸。它所产生的亮度足以照亮其所在的整体星系。超新星爆发所释放出的能量会电离和加热星系外围的气体。

卫星星系不足以及星系中心核结构问题的另一个解释是，超新星爆发会把宿主星系内部的一部分质量推出去，这样宿主星系核结构的密度就会被大大地降低。这使得星系中心的暗物质分布和城市中心的人口分布具有一定的可比性。例如，城市中心若发生骚乱，暴力事件发生过后，会使附近的人口停止增长，且留下一个荒芜的核心。星系

中心会发生非常多的超新星爆发，而这些超新星爆发不会使中心密度增长，这个效应的结果要比人口稀少的城市中心更加明显。

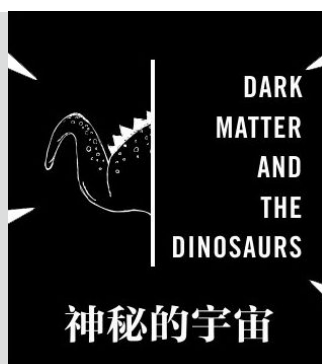
进一步说，超新星爆发所释放出的能量会电离和加热星系外围的气体。这个效应会吹走本应在暗物质子结构中形成恒星的气体。这些被吹走的气体会围绕一些更大的暗物质晕运动，或者阻止其他气体落入暗物质晕的引力势阱并形成恒星。这些外围矮星系会相应地具有比较少的普通物质而且也许会更暗，从而使它们更难被观测到。

支持与反对普通物质在小尺度结构形成中更大效应的证据，会随着计算能力和研究方法的改进而不断演化。这两次会议上还出现了几个很有意思的讨论，尽管这些讨论是由粒子物理学家提出的，天文学家们却非常开心与欣慰。因为这里每个人都希望寻找到正确答案，并不只是为维护自己的领域，而图一时口舌之快。尽管这些强调重子物质重要性的人指出，重子物质也许不是解决所有差异的万能钥匙。例如，小尺度结构问题存在于独立矮星系的时候，超新星爆发对结构形成的反馈非常之小。如果这一论点是正确的，那么根据目前的观测，一些超越普通暗物质的模型仍然有必要讨论。尽管与会的每个人都同意，普通物质的引入是解决数值模拟和观测数据差异的正确方向，但物理学家和天文学家都意识到，想要彻底解决丢失的卫星星系问题，对于标准无相互作用暗物质模型的大幅度修正，也许是必要的。

暗物质间的“黑暗之力”

当数值模拟和观测数据相碰撞时，会发生有意思的问题，所以修改暗物质模型来解决这些问题也变得有意思起来。最让人感兴趣的可能性便是，以前无相互作用的暗物质模型是错误的，而暗物质的自相互作用会影响宇宙中的结构形成。这种可能性会启发物理学家研究暗

物质粒子的自相互作用、可能存在的作用力。不管结果如何，已有的观测以及不断改进的数值模拟，都可以告诉我们越来越多有关暗物质的性质。就算前面讨论的问题不存在，我们依然可以更好地理解暗物质的性质，以及暗物质和普通物质是如何影响宇宙结构的形成与演化的。但如果这些问题依然存在，也可以证明暗物质自相互作用存在的合理性。



自相互作用暗物质算是一项比较可靠的建议，部分原因是，我们对暗物质的性质知之甚少。就像普通物质会参与像电磁相互作用一样的引力之外的相互作用一样，暗物质也有可能参加除引力之外的相互作用。尽管常见的假设是，暗物质会参与引力作用，并有可能与重子物质存在非常弱的相互作用，但现有的暗物质直接探测装置，并不会告诉我们有关暗物质自相互作用的任何信息。自相互作用的暗物质粒子会吸引还是排斥其他暗物质粒子，都和我们熟悉的物质完全不同。暗物质可能参与的是目前为止还没有被观测到的黑暗之力，只会影响暗物质粒子，而不会影响普通物质。就像电磁力只能作用于普通物质一样，黑暗之力只能作用于暗物质，这样，暗物质粒子和普通物质粒子最终彼此依然是毫不相干的。

自相互作用暗物质与普通物质一样，具有一定的社交性，同时又具有一定的排外性，即它们只和同类相互作用。暗物质粒子可能只会散射其他暗物质粒子，但是普通物质对于暗物质来说是隐形的，就像暗物质对于普通物质是隐形的一样。因为暗物质直接探测只寻找暗物质与普通物质的相互作用，所以自相互作用并没有被排除掉，并且这个假设还可能支持结构形成的相关研究。

如果暗物质确实是自相互作用的，那我们对其性质知之甚少。但我们可以推断：存在于暗物质粒子之间的力以及暗物质粒子的自相互作用力，不可能非常强。还记得那个存在于由普通星系团合并，而产生的子弹头星系团中非常著名的暗物质信号么？这个例子很好地限制了暗物质与其他物质的碰撞强度。引力透镜的观测告诉我们，在一个星系团的暗物质穿过另一个星系团的暗物质晕时，整个过程几乎是没有任何阻挡的。因此会产生两个圆鼓鼓的结构，而气体却被留在中心区域。

如果所有暗物质都存在非常强的自相互作用，甚至和重子物质一样强，那么暗物质会像气体一样存在于中心区域。但双球状的外部结构告诉我们，暗物质并没有存在于中心区域，而是会穿过彼此。这并没有告诉我们，暗物质完全不自相互作用。但这确实限制了相互作用的强度和距离尺度。暗物质自相互作用的强弱也受星系形状的限制，因为星系的形状受暗物质自相互作用的影响也非常大。

这些并没有排除自相互作用暗物质存在的可能性。它们只是为自相互作用的强弱和形式设置了一些限制。就算遵循这些限制，暗物质的自相互作用在理论上讲，依然可以解决小尺度结构问题。参加这两个会议的报告人讲述了，为什么暗物质的自相互作用可以帮助我们部分地解决一些潜在的结构问题。例如，它们可以降低那些大质量卫星星系的中心密度，从而使理论和观测符合得更好。

例如，自相互作用暗物质可以解决星系中心密度被预言得过高的问题。在没有非引力相互作用时，暗物质运动得非常慢，容易落入已有结构的中心势阱，这使得暗物质晕中心的密度变得非常高。而互相排斥的自相互作用，会让暗物质粒子彼此分开，从而防止它们离得非常近。这就像在一个拥挤的火车站，每个人都被自己的行李围

绕，那么人与人之间的距离就会保持在一个手臂的长度。暗物质粒子间的自相互作用会通过类似的过程，为各个粒子之间建立一个壁垒，这个壁垒可以防止暗物质的密度过高。

引入自相互作用的暗物质数值模拟确认了这一预言，并且确实产生了一个低密度的核结构。这些核结构具有平缓密度分布的内部区域，而不是产生类似于核-尖峰的高密度结构。物质的密度分布在向星系或者星系团的中心变化时，在暗物质粒子饱和之前，密度有一个快速提升，而饱和之后，密度则不会有明显的变化。现存的所有有关小尺度结构的问题，都可能通过引入暗物质的自相互作用得以解决。

自相互作用暗物质模型关于星系和星系团的预言，一定能在将来的观测和数值模拟中告诉我们关于暗物质以及星系的更多信息。并且，不同自相互作用模型的预言也会不同，这样通过比较数值模拟与观测的结果，我们甚至可以确定哪种自相互作用模型更加合理。

通过考虑不同的可能性，我们还可以根据大量有关宇宙中结构形状的数据，进一步了解其意义。也许暗物质存在自相互作用，从而影响了结构的形成和演化，所以数值模拟和观测会符合得更好。或者，也许普通物质在结构形成与演化的过程中，扮演了比理论预言更重要的角色，甚至比自相互作用暗物质的角色还要重要。我们可以在数值模拟和观测数据都足够可靠的时候，排除上述某种或全部可能性，并给出确定的答案。无论结果怎样，我们都可以从获得一些超越保守的弱相互作用大质量粒子模型探测实验的信息。

自相互作用暗物质本身虽然非常有趣，但它并不是我研究的重点，下一章我会着重介绍一下。毕竟，自相互作用暗物质和非自相互作用暗物质都不是唯一的可能性。就像当我们过于关注黑与白的时候，往往会忽略灰的存在，更不用说斑点和条纹了。如果我们相信暗物质要么不相互作用，要么完全相互作用，就会使我们忽略世界的多

样性。下一章我会介绍一个有趣的新想法，即暗物质和普通物质一样复杂。也许，暗物质既有不相互作用的组分，又有自相互作用的组分，而这些组分会共同作用于宇宙结构的形成与演化。



黑暗的速度

DARK MATTER AND THE DINOSAURS

尽管科学家更喜欢简单的想法，但简单的想法往往不能囊括全部。

不论是外行还是科学家自身，在评价科学方案时都经常使用“奥卡姆剃刀”作为指导。这个经常被提及的原则是指：在解释一种现象时，最简单的理论极有可能是最好的。它用听起来很合理的逻辑指出，在更精简的方案存在的情况下，再构建一个复杂的结构并不是什么好想法。

不过，有两个因素会削弱“奥卡姆剃刀”的权威，至少，它们提醒我们在把“奥卡姆剃刀”作为“拐杖”使用时要谨慎一些。我曾艰难地学会了小心使用这些“拐杖”——在心理上和生理上都是，为了复原受伤的脚踝，我曾经使用过拐杖，由于使用姿势不正确，导致我胳膊上的神经受损。同样，那些服从“奥卡姆剃刀”原理的理论，会在解决一个显著难题的同时，在其他地方产生问题——通常是这个理论的其他方面。

最好的科学原则总是应该解释尽可能广泛的观测，或者至少与之保持一致。真正的问题是，什么才能够最有效地解决整个体系中无法

解释的现象。一个起初看起来简单的解释，在面临更多问题的时候，可能会演变成鲁布·戈德堡（Rube Goldberg）^[36]式的漫画。另一方面，一个解释可能在应用到最初的问题时显得过分烦琐，但通过科学的眼光来看，也许会显露出它隐藏的优雅。

我对“奥卡姆剃刀”的第二个顾虑与一个真相有关。这个世界比我们任何人所能够设想的都更为复杂。一些粒子及其特征似乎于任何重要物理过程而言，都无关紧要——至少迄今为止我们发现的是这样，但它们依然存在。有时候，最简单的理论并不一定是正确的那个。

第18章提到的“关于暗物质的辩论”的会议中，多次对这个话题进行了讨论。在关于可观测到但却无关紧要的粒子的报告中，粒子物理学家娜塔莉亚·多罗（Natalia Toro）辩称，一个比“奥卡姆剃刀”更恰当的理论指导是“威尔逊的手术刀”（Wilson's Scalpel）。她用物理学家肯·威尔逊（Ken Wilson）来命名。威尔逊开发了一个大体框架，来理解如何通过只用可测量元素来做科学实验。娜塔莉亚提出，以威尔逊名字命名的手术刀可以用来调整（而不是剔除）一个理论。无论我们能否将其归属于某个基本的意义，都要完好地留下所有的可测试元素。当我接着她发言时，我开玩笑地建议“玛莎的桌子”（Matha's Table）原则才是一个更好的概念。至少你不会在餐桌上只摆刀子。你会摆放上能让你优雅地吃上一顿饭所需要的所有餐具。若拥有玛莎·斯图尔特（Martha Stewart）^[37]的才华，你可以让一切保持井然有序，不论你要布置多少盘子和银器。

科学同样也需要合理的摆桌——用一个理论来解释许多我们观测到的现象。尽管科学家更喜欢简单的想法，但简单的想法往往不能囊括全部。

以上讨论是对我接下来要介绍的理论的一个序曲。这种被我与合作者称为“部分相互作用暗物质”（partially interacting dark matter）的概念，引出了“双盘暗物质”（double-disk dark matter,

DDDM) 这一类模型。两种体系的模型都认为，暗物质的结构也许并不简单。就像普通物质的粒子一样，暗物质粒子也许不止一种。具有不同类型相互作用的新型暗物质也许是存在的，未来也许能被观测到前所未料的结果。即使具有相互作用的只是很少一部分暗物质，它们也可能对太阳系和银河系产生重要影响，也许对恐龙也产生过影响。

普通物质沙文主义者

尽管我们知道普通物质仅占宇宙总能量的5%，约占全部物质的17%（剩下的部分由暗能量组成），我们仍然认为普通物质是重要的组成成分。尽管普通物质在总能量组分上微不足道，但除了宇宙学家外，人们的关注点都在普通物质上。

当然，我们更关注普通物质是因为我们由它组成——我们生活在有形的世界中。但我们关注它也是因为它具有丰富的相互作用。普通物质的相互作用有电磁力、弱相互作用力和强相互作用力，帮助周围世界中的可见物质形成复杂致密的系统。不只是恒星，也包括岩石、海洋、植物以及动物，它们的存在都是因为普通物质在通过自然界这种非引力相互作用。就像是在啤酒中只有很少量的酒精，喝多了照样能让人大喊大叫一样，携带了很少能量密度的普通物质，相对于仅仅会路过的天体，会更明显地影响它自身以及它周围的环境。



我们熟悉的可见物质可以被认为是物质中有特权的一部分，约有15%。正如在政治和经济中，有影响的1%人口主导了决策和政策的制定，而剩余99%的人口则提供基础设施支持和其他各种支持，例如维修建筑、保证城市运转、为人类的餐桌提供食物等。同样，普通物质几乎主导了我们所有的观测；而暗物质，虽然丰

富且无处不在，帮助创造了大星系团和星系，促进了恒星的形成，却对此刻我们周围的环境影响有限。

对于近邻结构，普通物质在起主导作用。它支配着我们躯干的运动，社会经济的能量来源，你读这本书时面对的电脑屏幕或者纸张，以及任何你能想到的和关心的东西。如果某些事物都有可被测量的相互作用，它就值得被关注，因为它会更快地影响周围的事物。

在通常的理论中，暗物质缺乏这种有意思的作用和结构。一个通常的假设是，暗物质是连接星系和星系团的“胶水”，但只存在于其周围非固定形态的云中。但有没有可能，这个假设不对，只是我们的偏见和无知造成的，从而让我们走上了这条可能是错误的路？如果一部分暗物质就像普通物质那样也存在相互作用，又会怎样呢？

标准模型包含6种不同的夸克，包括电子在内的3种带电轻子、3种类型的中微子。这些粒子都会承载力，包括最新发现的希格斯玻色子。如果暗物质的世界——即使没那么普遍，粒子类型也相当丰富呢？在这种情况下，大部分暗物质的相互作用可以被忽略，但少数会像普通物质那样受力相互作用。标准模型中粒子和力的丰富而复杂的结构，创造了许多有意思的现象。如果暗物质中也有一种会相互作用的成分，这个成分也许会产生一定的影响。

如果我们是由暗物质组成的生物，并假设普通物质的粒子都为同一种类型，那么我们将大错特错，也许我们这些由普通物质构成的人类在犯着同样的错误。标准粒子模型描述了大部分我们所知物质的基本成分，鉴于其复杂程度，假设所有的暗物质只由一种类型的粒子构成，是非常奇怪的。为什么不设想，一定比例的暗物质在受它自己独有的力呢？

在这种情形下，就像普通物质包含了不同类型的粒子，且这些基础组件通过不同的荷组合相互作用一样，暗物质也许有不同的基础组建，并且在这些完全不同的新粒子类型中，至少一种可以发生非引力相互作用。标准模型中的中微子不会有强相互作用力和电力，但6种夸克可以。相似地，也许一种类型的暗物质粒子会有微弱的或者没有引力之外的相外作用，但是它的一部分（也许5%）会有。基于我们在普通物质领域所了解到的，也许这种理论比通常假设的具有单一弱相互作用或无相互作用的暗物质粒子更可行。

从事外交关系的人常会犯的一个错误是，一概而论地评价其他国家的文化，假设他们不像我们自己一样拥有多样化的社会。一个好的谈判专家在试图平等对待两种不同的文化时，不会假设一个社会团体优先于另一个。同样，一个无偏见的科学家不应该假设暗物质不像普通物质那样有趣，并假设它一定缺少物质的多样性。

科学作家科里·鲍威尔（Corey Powell）在《发现》杂志上报道我们的研究时，以宣称他自己是“轻度物质沙文主义者”来开头，并且指出实际上每个人都是。他用这个词来表示，我们常认为自己熟悉的物质是目前最重要，也是最复杂有趣的。这种想法实际上已经被哥白尼革命所颠覆^[38]。然而大部分人仍坚持臆想，他们关于自身重要性的观点和信念，是与外部世界一致的。

普通物质的许多组分具有不同的相互作用，对世界有不同形式的贡献。所以暗物质也许同样有不同的粒子，有不同的行为，并可能会对宇宙的结构具有可测量的影响。

部分相互作用暗物质

我与合作者把这种小部分暗物质会发生非引力相互作用的理论称为“部分相互作用暗物质”。我们第一次调研了最简单的模型，它仅包含两种成分：主导成分仅通过引力相互作用，是常规的暗物质，存在于星系和星系团周围的球形晕中；第二种成分也通过引力相互作用，但此外还有另一种类似电磁作用的力。

这两类暗物质理论也许听起来有点怪异，但是请记住，同样的描述方法也适用于普通物质。夸克受强相互作用力，但像电子一样的粒子并不受这种力影响。这就是为什么夸克被束缚在质子和中子中，而电子没有。同样，电子受电磁力影响，但中微子对之浑然不觉。所以，如果推翻我们惯用的沙文主义，允许暗物质的世界有类似的多样性，就不难想象，暗物质的一部分也通过类似但不同于构成我们的物质的力相互作用。

不过请牢记，部分相互作用暗物质与标准模型物质有一点不同。比如，尽管电子不直接受强相互作用力，但它们可以和夸克相互作用，所以会受到间接影响。新提出的暗物质粒子在它自己的相互作用中是完全孤立的，大部分暗物质甚至不会受这种新引进的力的间接影响。我们还不知道暗物质的成分之间是否会相互作用，甚至不知道暗物质是否有多种成分。最初最简单的假设是，只有新类型的电磁力和新的带荷粒子受力外，没有其他新的相互作用。在这种理论中，大部分暗物质粒子完全不会受这种新型力的影响。

出于好玩，我把这种相互作用的暗物质成分所受的力称为“暗光”（dark light），或者称其为“暗电磁力”（dark electromagnetism）。选择这个名字是为了提醒其他人，这种新的暗物质受到一种类似电磁力但又与我们世界的普通物质无关的力。普通物质带电荷，所以它才能辐射和吸收光子；这种新型的暗物质成分只能辐射和吸收这种新型的光，普通物质则根本感觉不到它的存在。

这种暗电磁力会类似于平常的电磁力。但这是作用在一种全新粒子荷上的完全不同的作用，通过一种全新的粒子来传递——你可以称其为暗光子（dark photon）。尽管这种新型的暗物质成分不会和普通物质相互作用，它会有自相互作用，从而使得它表现得像我们熟悉的普通物质，也不会和其他暗物质相互作用。

普通物质和暗物质都带电并受力，但这些电和力并不相同。这些受新暗力影响的带电粒子会彼此吸引或排斥，其形式就像普通带电粒子的行为一样。但是暗物质层的相互作用会对普通物质透明，因为暗物质用它自己独特形式的光作用，而不是通过我们熟悉的那种光。仅有暗物质粒子会受这种新型力的影响。

尽管遵守着相似的物理学定律，也许在空间上也相互接近，暗物质和普通物质依旧拥有着各自的世界。普通物质和暗物质甚至可以空间上重叠却又无任何相互影响。它们会通过各自独特的力相互作用，除了它们极其微弱的引力外，带电普通物质和带电暗物质对对方的存在都浑然不觉。

在同一位置的两种类型的带电粒子而不与对方相互作用，并非匪夷所思。这有点像普通物质通过Facebook互动，而带电的部分相互作用暗物质模型通过Google+互动一样。它们有相似的作用方式，但仅在各自的社交网络上接触。互动发生在一个网络或者另一个上，但不会两个都用。

我们再打一个比方。这就像是左派和右派各自的电视节目，他们遵循着差不多相似的程序规则，都能在同一个电视上播放，但是他们实体完全不同，并各自在加强自己确信的偏见。尽管他们形式类似，有采访的主持人，有专家和嘉宾，用图表阐述他们的观点，并且在屏幕下方弹过随机、无关的新闻快讯，但这两类节目的真正内容和结论以及其中的广告非常不同。很少有相同的嘉宾或议题会同时出现在两类节目上，他们的产品及拥护的参选人也会不同。

就像罕有个人用户既看福克斯新闻，也听NPR（美国国家公共电视台）一样，大部分（也许是全部）粒子会通过这种或那种力相互作用。这个模型就像是媒体一样，提倡坚持一种观点。即使理论上应该有一种中间态的粒子会参与两种相互作用，但大部分粒子要么带有一种类型的电荷，要么是另一种，因此相互之间不会传递作用。

公平地讲，物理学家排斥接受暗物质受到一种新型电磁力影响的观点，并不仅仅是因为偏见。相互作用通常有可被探测的结果。物理学家对暗力和自相互作用暗物质这种想法退避三舍，是因为他们认为这种理论很受限，甚至可以被排除。然而，如第18章中介绍的那样，即使所有暗物质都受这些力，这些限制条件也不会太严苛。但是，相互作用必须在观测所得到的限制范围内才可以。

不过，如果只有少量暗物质有自相互作用的话，限制就会弱一些。回想一下关于自相互作用两种类型的限制条件。第一种与暗晕自身的结构相关：暗晕必须是球形的——仅有被称为三轴结构（triaxial structure）这一点非均匀性。第二种与暗晕的并合有关，举个最著名的例子，子弹星系团，就是星系团并合的结果。气体很清楚地位于中心区域，而从引力透镜观测到的暗物质，无障碍地相互穿过并创造出两个额外的半球结构。这有点像米老鼠的两个耳朵。

如果所有暗物质都相互作用，那么这两种限制都将非常显著。但没有哪个能清楚地告诉我们，是否只有一小部分暗物质有相互作用。如果只有一小部分的成分相互作用，那么大多数的暗物质晕会是球形的。相互作用也不会抹消三轴结构，除非有相互作用的是主导成分，或者弥散比预期的大得多。

类似地，子弹星系团中的暗物质和气体比例也没有好到能标记出很小一部分暗物质成分，因为它们毕竟只是星系团的很小一部分。这种成分有可能会相互作用，并与气体一样留在中心区域——没有更好

的地方了。也许，将来这类关于子弹星系团的测量会精确到可以限制我所描述的部分相互作用理论。目前肯定的是，部分相互作用暗物质依然是一个可行的、有希望的可能性。

灵感来源

我与新进入哈佛大学物理系的马修·里斯以及两个博士后范吉吉和安德雷·卡茨一起提出的这个想法，一开始并不明确。就像其他一些后来变得极为有趣的研究项目一样，我们最终关注的并不是一开始想要研究的东西。而我们本来是在尝试理解来自费米卫星的一些有趣数据。费米卫星隶属于美国国家航空航天局的空间天文台，用来扫描全天的伽玛射线，这是一种比可见光和X射线能量更高的光。

大多数天体物理过程都会产生辐射，这些辐射在一个宽广的频率范围上平滑分布。这说明，光子的数目不会在某个特殊波段显著变化。所以当阿姆斯特丹大学的克里斯多夫·韦尼格（Christoph Weniger）及其合作者发现费米数据中的过量辐射全集中在单一频率上时，这引起了我们以及许多其他物理学和天文学同行们的兴趣。

韦尼格及其同事们辨认出的这个辐射密度（这里的辐射只代表光子和光）上的峰值显示，它在星系中心出现，那里暗物质高度聚集，但却不可能由常规的天体物理源产生的信号。在缺乏更常规的解释或者一个错误的解释的情况下，光子数的这个峰值只能代表某些新的东西。

最有趣的解释是，这个信号可能是暗物质湮灭产生光子（第17章介绍的一种间接信号）的结果。暗物质粒子之间也许会相互碰撞，然后通过公式 $E=mc^2$ 转变成为光子，之后被费米卫星探测到。对这个理论更进一步的支持是，被观测到的超额光子的能量，刚好位于预期的暗

物质能标范围内。它也刚好接近希格斯玻色子质量^[39]，或许这暗示了更深层的关系。这个测量第三个有趣的方面是，产生辐射的相互作用率与保证暗物质残留密度所需要的相互作用率一致。如果暗物质以被观测到的速率湮灭过，那么刚刚好遗留下今天的暗物质数量。

除了这些乐观迹象，如果信号真的来源于暗物质，也有一些现象不好解释。暗物质不会直接产生光子，因为它不会和光相互作用。也许暗物质会和某些我们还没有观测到的带电重粒子相互作用，而这些粒子反过来又和光相互作用。如果是这种情况的话，我们会期望，当暗物质湮灭产生能量时，这个能量也会产生带电粒子。然而费米卫星并没有观测到这一现象。

另一个问题是，尽管暗物质的总体数量取决于它湮灭了多少，但信号的强度只取决于有多少暗物质粒子湮灭为光子。给定宇宙中的暗物质密度时，我们发现暗物质湮灭为光子的速率，除了在人为精调模型中以外，在其他模型中也都太低了。这说明，关于这个信号，这种特定的暗物质解释，可能仅在很窄的一个参数范围内才与观测符合。这要求一个足够大的湮灭速率来产生光子，而不会产生可测量的带电粒子。目前没有可靠的理论可以使得这种情况发生。

我和范吉吉、安德雷、马修认为，这是一个研究可行暗物质模型的有趣机会。我们想知道，是否有什么合理的方案可以使所有这些速率都符合其观测值。我们从关注费米的结果开始并且问自己：我们能否想到一个方法，以让大自然可以比其他物理学家提出过的模型做得更好？我们非常明白，可能结果会是，这些数据一直在误导我们。费米的数据很诱人，却没有强到可以断定一个新信号是来自暗物质，还是其他什么源的结论。这个观测可能只是因为统计上的巧合，或者对仪器信号的不当处理，而非一个真正的物理过程的信号——我可不想读者读到这里时产生任何过高期望，可能这个信号只是个错误。

但这个结果也非常有趣，尤其是在早期，它使得我们发问：是否存在合理的物理过程可以产生这个信号？毕竟寻找奇异的新型物质很困难。我们想知道所有可能找到它们的方法。不论这种信号会不会被证实，我们都能学到一些未来可能有用的知识。

我们四人曾试了若干个想法，试图在保障所需信号的同时，又能巧妙地避免各种问题，但是没有有一个提议能好到值得继续下去。那些可以成功满足所有限制条件的想法都与“奥卡姆剃刀”的精神不符。更糟糕的是，它们不容于任何现有的概念。

尽管如此，我们曾否定的一个模型引发了一系列思考，而这最终比我们的其他尝试都更有趣。我们最初的努力全都是为了找出一个特定的模型，来塞入现有的框架内。但退一步想：假如局部暗物质比我们原本设想的更致密一些呢？也就是说，我们其实曲解了这些暗示。假如因为暗物质密度更大，可以比预期的湮灭更多呢？

如果密度更高，暗物质粒子就可以更有效地接触彼此并相互作用。这反过来会创造更强的信号，也因此更容易被观测到，就像你在人潮流动的火车站比在晚上11点的郊区更容易撞到人一样。高密度环境中的暗物质粒子比在通常暗晕中那种稀疏环境下，更容易与另一个暗物质粒子相互作用。如果一些暗物质比暗晕中的物质更加紧致，那么其他限制条件就很容易满足了。

接下来的问题是更基础的原因。为什么暗物质或至少某些暗物质，会比我们认为的更致密呢？这就是我们会想到“部分相互作用暗物质”这个主意的原因了——当然，还有“暗物质盘”这个概念。事实上，即使我们现在发现费米信号是假的，这个崭新的想法也有非常多未被深究的推论，我们会意识到它本身也非常值得探索。其中一个结果就是，比通常假设的要致密得多的暗物质盘。

暗物质盘，嵌在银河系中的薄盘

有次我打扫房子的时候（好吧，实际上是我的吸尘机器人在工作），我清理了吸尘器的灰尘盒，发现了我很久以前保留的一张幸运饼干里的纸条。这张纸条问了一个神秘的问题：“黑暗的速度是什么？”那时我不知道这些话其实是一个运签，不过这句话或多或少地预言了我接下来要开始的研究。

暗物质盘

兰道尔及合作者提出的“部分相互作用暗物质”理论所假设的一种盘。相互作用的暗物质会耗散能量，降低速度，并冷却形成一个盘。其温度和普通物质盘相当。

第5章解释了，普通物质之所以会聚集在一个致密的薄盘中，是因为它会辐射光子，而光子能有效地带走能量。这种能量耗散的结果是，那些热的、高能的、高速粒子逐渐变成慢而冷的粒子，不会再有大范围的活动。物质会坍缩，因为较低的能量使得它速度变低，不能四处扩散。普通物质通过消耗能量来降低速度，坍缩成一个盘——就像银盘一样，你可以在晴朗夜空看到它。

在我与合作者想出“部分相互作用暗物质”这个概念以后，我们探索了它对银河系和更多现象的可能后果。我们假设相互作用的暗物质是存在的，并假设它的行为和普通物质相似。我们已经知道，普通物质在星系中心冷却、速度降低，于是形成了一个盘。

在我们的理论中，仅仅有一小部分暗物质相互作用，大部分暗物质仍旧会形成球形的暗晕，这与天文学家已经观测到的相一致。然而，新的相互作用暗物质也能耗散能量，所以它也能像普通物质一

样，冷却并形成一个盘。相互作用的暗物质成分会通过暗光子的相互作用，辐射出能量来降低速度。在这一方面，它会像普通物质那样活动。并且由于角动量的累积，阻止了除纵向之外的其他方向上的坍塌，因此相互作用的暗物质会形成一个盘。

再者，就像普通原子由带相反电荷的质子和电子组成一样，这种暗物质成分也许包含带有相对电荷的粒子。这种带电粒子会持续辐射能量，直到它们冷却到可以被束缚在暗原子中。之后这种冷却会变慢，暗物质原子就像普通物质的原子那样，存在于盘中，而盘的厚度与束缚的暗原子的温度相关。根据合理假设，普通物质和暗物质冷却停止后的温度会相当。因此，会留给我们一个温度相当的暗物质盘和普通物质盘。

尽管如此，暗物质盘不会有和银盘完全一样的结构。事实上，它会更有趣。暗物质盘卓越的特征是，如果一个暗物质粒子比一个质子更重却有相同的温度，那么暗物质盘和银盘相比会更薄。

一个粒子带多少能量与它的温度相关，但动能也与质量和速度相关。同样温度的更重粒子必须有更低的速度，这样它们的能量才相当。所以，大质量粒子会导致更薄的盘。对于一个比质子质量重100倍（这是我们通常假设的暗物质质量）的暗物质粒子来说，暗物质盘可能比银河系的窄盘薄100倍（这是一个值得注意的可能性），正如我们在接下来的两章里将看到的那样，这会带来许多有趣的观测结果（见图19-1）。

同样重要的是，两个盘尽管各不相同，却应该对齐——暗物质盘镶嵌在银河系平面中广阔的盘里。这是因为，普通物质盘和暗物质盘会通过引力相互作用，不是完全相互独立的。之前关于福克斯新闻和美国国家公共电视台的比喻有个缺陷，暗物质盘和普通物质盘都会受到引力的拉力，实际上会使这两个实体指向同一方向。左翼和右翼的

电视台也不是完全相互独立的，因为它们会通过固定且经常重复的广播的集体效应相互影响对方，且大部分的反应是负面的，这让它们的互动会相互排斥。不同的是，暗物质和普通物质的盘通过引力发生相互作用并会对齐。

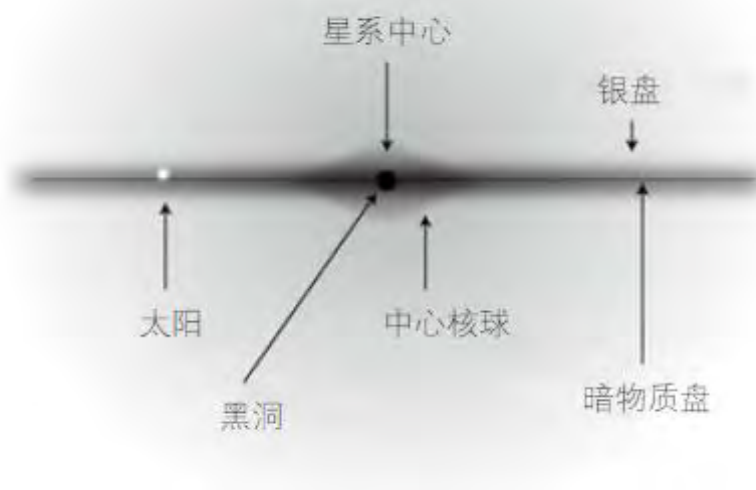


图19-1

小部分相互作用暗物质成分可以在银河系的中平面形成一个非常薄的暗物质盘，由图中的黑色实线表示。

我们研究的重要且惊喜的结果是，普通物质盘会伴随着一个暗物质薄盘，并且这个暗物质盘会镶嵌在银盘中。我与合作者对这个结果感到非常兴奋，并热切地想与其他物理学家分享。我的同事、哈佛大学的霍华德·乔吉（Howard Georgi）也非常喜欢这个主意，但明智地想到，这个理论值得拥有一个比我们所建议的更吸引人的名字。他还帮忙提议了“双盘暗物质”这个名字，这如此贴合我们的目的，我们后来一直使用它。这个名字非常合适，因为根据我们的研究，星系的确拥有两种类型的盘，一个镶嵌在另一个中。

恒星观测显示，我们离开银河系平面不会超过几百万年，这只是宇宙尺度上的一瞬间。这告诉我们，如果双盘暗物质存在，太阳系也会在大约同样的时间尺度上于暗物质盘上震荡，所以我们离得并不太远（以天体物理的角度来看）。实际上，如果盘更厚一点，我们也许正在其中，并可能会有些可观测的结果。而且，这个盘也许会影响太阳系的运动——也许会产生一些戏剧性的影响，尽管是在超长的时间尺度上。我们所提出的部分相互作用暗物质也能在其他星系中制造盘，这也许恰巧解释了它们的某些特征。

当然，重要的问题是，相互作用的暗物质成分和暗物质盘是否真的存在。通过测量它的结果来发现一个暗物质盘，会帮助我们了解以上任何一个建议的重要意义。幸运的是，与普通物质一样，即使它只占宇宙中暗物质总量的很少一部分，相互作用的成分产生的高密度，也许会让它比普通离散的暗晕中的暗物质更容易被发现。我们下章会提到，由这种加强的暗物质密度所产生的许多粒子物理和天文信号，应该能告诉我们暗物质盘是不是可行的或者更好的选择。

如果我足够幸运，也许一个或更多此类观测会最终揭示暗物质盘的存在。



寻找暗物质盘

DARK MATTER AND THE DINOSAURS

你不知道黑暗生物有多美，而且你也许永远都不会知道。

最近，我参加了一个关于言论自由的开放讨论，参加讨论的人有律师、学者、作家以及人权工作者。大家都认同这个问题的重要性，但却无法对人权的真正内涵达成最终一致。或者说，大家没有对如何平衡言论自由以及人类其他权利达成最终一致。在什么时候，言论自由的潜在危害会超过它能带来的好处？到底是否应该花钱促使通过某个特定的法律，或限制候选人的数量？一位律师解释了美国最高法院是如何基于言论自由权去给某些公民案件进行宣判的，以及在这个过程中是如何花钱的。在这个过程中，公司对政治的贡献是不加限制的，但参与讨论的其他人则比较关心，如果不对公司的付出加以限制，那么个人的声音将被掩盖。况且，言论自由这一概念是针对个人的，而不是针对企业。毕竟，无论是钱还是企业都不能自由地发声，而是企业里的人在说话。

但美国最高法院的规定就是如此，其结果是，大量金钱被花费到了政治上。让我们看看个人和企业通过花钱来改变公众看法的不同方

式。

金融机构可能更专注于在城市或城镇的附近做广告。在这些地方，他们可以轻易改变人们对某些事物的看法，并影响最后的投票结果。或者捐赠者会把广告的范围做得更广泛，自己掏钱把他们的主张覆盖在更大范围，从而产生一些相对宽泛的观念，但其效果不是很清楚。两种不同的策略放在一起产生的效果，比单独做任何上述单一广告行为的效果更强。在目标区域内，受影响的人数应该更多，因为在较小、人口稠密的地区，广告密度更大。

物理学也有类似的情况。一个薄而致密的物质盘对恒星的运动影响，要比厚的、弥散的物质盘所造成的影响更加剧烈。就像局域广告投放得越多，广告效果越明显一样，一个薄而致密的盘结构，对银河系平面上飞进飞出的恒星的速度及其位置的影响更明显。

银河系是由重子物质和暗物质组成的盘结构，处在星系盘平面内或者外部的恒星的运动，会同时受到这两种盘结构的影响。当你从星系盘的致密区域移动到稍微低密度的区域时，这种影响的变化形式也从开始的迅速变化，转化成后来的平缓变化，这种转化形式就像结合局部广告和全局广告对选民的影响。如果一个薄的暗物质盘镶嵌到厚的普通物质盘当中，暗物质的吸引力就会和比较弥散的普通物质盘的吸引力混合到一起，从而产生一个特性鲜明，并可测量的作用力作用到恒星上。这个作用力随着恒星到银河系中盘的距离的不同而不同。

我们生活在一个数据充裕的时代，并且我们确实也不想忽略任何可能的观测对象，尤其是在寻找那些匪夷所思又很难找到的对象时，例如暗物质盘。在这一章，我会解释，如何通过研究银河系星系盘引力对恒星运动的影响，来确定或者排除暗物质盘的存在。但在此之前，我们先讲讲其他关于暗物质盘的可能性，并讨论一下应用目前的

暗物质探测方法找到暗物质盘的可能性。在此之后，我还会介绍几个暗物质盘在天文学上的有趣预言。

多样的暗物质

在开始研究部分相互作用暗物质时，我很惊奇地发现，竟然没有人考虑过“只有普通物质是有多种粒子和相互作用的”这一假设的不合理性。尽管有些物理学家已经尝试过一些模型，例如镜像暗物质（mirror dark matter，指暗物质在模仿普通物质的一切），但类似这种模型太过特殊和奇异，因为这种模型给出的预言很难和我们已知的一切相吻合。

有一小部分物理学家曾研究过更加普遍的相互作用暗物质模型。但就连他们也假设，所有暗物质都是相同的粒子，因此参与唯一的作用力。没有人提出过下面这个非常简单的假设，即尽管大部分暗物质粒子不发生相互作用，但一小部分暗物质粒子却可以发生相互作用。

这里有一个比较明显的潜在原因。大多数人希望出现一个新的暗物质形式能与大部分观测现象相关，就算这种新的暗物质只占暗物质总量的一小部分。但在没有观测到占主要组分的暗物质之前，强调小部分暗物质的组分好像有点太早了。

这个逻辑的错误之处在于：普通物质只占暗物质的20%，但却几乎是我们所关注的一切。通过更强的非引力相互作用的物质，要比那些具有更大组分但只参与微弱相互作用的物质更有意思，并且更有影响力。

这一事实对普通物质来说是对的。普通物质的影响要远超其在宇宙中的组分，因为它们



形成的物质盘中，形成了恒星、行星，甚至形成了地球和生命。一个带电的暗物质组分（不用带太多的电）可以形成与银河系的可见星系盘类似的暗物质盘。这个盘甚至可以碎裂成像恒星一样的天体。这个新的盘状结构理论上是可观测的，甚至可能会比占主导的传统暗物质组分所形成的巨大球状暗物质晕，更容易被证明。

一旦你开始沿着这些线索思考，可能性会迅速变得多起来。毕竟，电磁场是标准模型粒子参与的几种非引力相互作用之一。除了这个能把电子束缚在原子核附近的力之外，其他两种相互作用是弱相互作用和强相互作用。也许可能还有其他相互作用存在于普通物质的世界中，但它们在人类所能达到的能量范围内极端微弱，至今还没有任何人发现一点迹象。但即使根据现有的三种非引力相互作用力，我们也可以猜想，相互作用的暗物质也许不止参与暗电磁相互作用。

也许除了暗电磁力外，暗强、弱相互作用也会作用在暗物质粒子上。在这个更复杂的领域里，暗物质恒星也许会形成，暗物质恒星里的核燃烧会创造出的结构和我以前描述的暗物质模型所产生的结构相比，更类似于普通物质。在这种情况下，暗物质盘由暗物质恒星组成，而暗物质恒星周围又围绕着暗物质行星，这些行星都是由暗物质原子组成的。这类能形成盘的暗物质模型可能具有和普通物质一样的复杂性。

部分相互作用暗物质模型一定可以为新的猜想提供肥沃的土壤，从而鼓励我们去思考我们不曾思考过的可能性。科幻作家和电影工作者也可以寻找到一个新的创作领域，新的相互作用以及其产生的黑暗领域还是很吸引人的。他们甚至可以绘制一个和我们一样的黑暗生

命。这个构想下，人类与外星人不再是简单地相互攻击或者合作，因为暗物质生物的军队已经穿过了屏幕，并已经统治了一切。

这个故事不会因太过天马行空而无人观看。问题是摄影师如何为黑暗生命成像：显然我们看不到黑暗生命，同时黑暗生命也看不到我们。就算黑暗生命就在那里（也许他们早已在那里了），我们也无从得知。你不知道黑暗生物有多美，而且你也许永远都不会知道。

尽管上面这些对黑暗生物的猜测充满娱乐性，但想观测到这种黑暗生物，或者说通过更直接的方法确定黑暗生命的存在，是十分困难的。寻找与人类一样同样由重子物质组成的智慧生命，已经颇具挑战性了，尽管太阳系外行星的搜寻工作正在艰难地进行着。如果黑暗生命存在的话，他们存在的证据要比遥远外星上存在普通生命的证据，更难以找到。

我们还没有直接探测到来自单个星体发射出来的引力波。就算是天文学家们通过其他方法寻找到了黑洞和中子星，却仍然没有引力波的相关观测^[40]。我们探测到，黑暗生命存在的可能几乎为零，不管他们离我们有多近。

理想一点儿说，我们希望通过某种途径联系这片新区域，或者让他们通过某种特定的方法回应我们。但如果这些生命不能感知人类的那些作用力，那么交流将不会发生。就算我们都参与引力相互作用，小尺度物体或者生命所产生的引力肯定是低得无法探测的。只有非常大的黑暗天体，比如贯穿于银河系平面的暗物质盘结构，才能产生可观测的结果。

黑暗天体和黑暗生命可能离我们很近，但如果它们的净质量不是非常大，我们就无法确定它们的存在。即使用最先进的技术，或者那些我们还在想象的技术，也只能在某个非常特定的可能性范围内加以

验证。“影子生物”（Shadow life）不一定能产生我们可以注意到的可观测结果，它们给了我们一个诱人的可能性，却无法被我们观测到。

坦白说，黑暗生命是一个高级现象。科幻作家可以毫无问题地创造出来，但是宇宙还有很多障碍需要逾越。穷尽所有化学的可能性，到底有多少种组合能够维持生命，我们还不清楚，即使对已知的生命，我们仍然不知道哪种环境是必要条件。尽管黑暗生命的假设很吸引人，而检验黑暗生命并不是唯一的难点。宇宙要如何产生黑暗生命才是真的困难。因此我把这个可能性放在一边（至少目前得这样做），而只专注于对银河系中大质量致密盘的寻找。我个人认为，后者比前者更有希望。

来自暗物质盘的信号

为了系统工作和最少的假设，我同范吉吉、安德雷·卡茨、马修·里斯四人从我们能想到的最简单的双盘暗物质模型开始研究。与普通的极弱相互作用的暗物质模型相比，我们模型中的暗物质粒子还带有黑暗电荷，并且相互之间会有类似于电磁力的相互作用，这一黑暗电磁力是通过暗物质粒子所带的黑暗电核而产生的。我们的模型还包括一种相对较重的暗物质粒子，与质子相似，并带有正电荷，另一种粒子则带有负电荷，与电子相似。

研究一个还未列入标准物理准则世界的新想法和上坡作战一样困难。对于一些物理学家和天文学家来说，双盘暗物质模型有点夸张。甚至对于粒子物理学家来说，尽管他们的工作就是用大胆的想法来揭开物理学中最基本的问题，例如物质的种类，但是我的大部分同事（一般都是科学家）仍然比较保守。这一现象也并非完全不合理：如

如果一个观测有个保守的解释，那么这个解释差不多总是对的。革命性的新尝试只有在它们能够解释旧理论无法解释的现象时，才会被别人接受。只有很少的例子需要新理论来解释。

即使科学界认同，现在的物理学需要新的想法和模型，但如果新想法超出几种公认的猜想准则，那么这种猜想一般会遇到阻挠。例如，超对称理论和弱相互作用大质量粒子模型经常被粒子物理学家认为是能马上建成的理论，尽管证明它们存在的实验证据还不存在。面对越来越多的数据限制，这个领域的人开始怀疑旧理论，并开始考虑新的可能性：那些超出已经确定的研究准则的可能性。

一旦新的概念得到认可，人们会竭尽所能地研究这个概念的所有细节，检验参数空间的所有可能，甚至对那些还没有被证实为正确的假说也不会放过。在一个想法到达这个层次之前，会有很多批评声音，甚至是带有偏见的批评。但有一些粒子物理学家在面对很多不确定性时，更倾向于保持一种开放的态度，我和我的合作者也是如此。我们也许会因为某个理论更优雅简洁而倾向于它，但在数据完全确定之前，我们不会断言什么就是正确的。

我们四人很快意识到，相互作用的暗物质的行为和无相互作用暗物质的行为非常不同，人们应该可以通过观测区分两种模型。但考虑到双盘暗物质的最初提出动机，我还是倾向于利用保守的探测方法来探测这一新型暗物质。例如，那个第一次刺激我们研究的暗物质间接探测信号，也许这正是传统暗物质无法解决，而双盘暗物质模型可以解决的领域。所以我决定从暗物质间接探测信号出发开始这项研究，例如来自费米望远镜的光子信号。

一个薄的暗物质盘是致密的，也就是说，暗物质粒子在此区域的汇聚度很高。与传统冷暗物质晕的粒子分布相比，在这个致密的盘里，更多暗物质粒子会相遇，因此更多的湮灭也将产生。这并不说明双盘暗物质模型都是被如此观测到的。我们研究的起因便来自双盘暗

物质模型可以产生间接的光子信号，而这一过程需要我前面所提到的带电暗物质。由于类似于费米卫星观测到的信号需要暗物质能变成光子——光子是普通物质的一种形式，且只有当一个粒子既有常规电磁力又有黑暗电磁力时，这种可观测的相互作用才会出现。这就像一个人既看福克斯新闻同时又听全国广播电台，或者同时登录在Facebook和谷歌上一样。如果一个粒子同时具有常规电荷和黑暗电荷，那么暗物质湮灭时所产生的间接粒子就会释放出光子，而这种间接粒子同时连接着普通领域与黑暗领域。这样，费米卫星观测到的光子信号就可能是双盘暗物质模型的一个预言。当然这并不是一个普遍性的预言。

这个致密的暗物质盘确实意味着，如果可观测的相互作用存在，它们发生的频率要比我们想象的快。一个更好的消息是，如果双盘暗物质模型能够产生某种可观测的间接信号，例如光子、正电子或者反质子，那么这个结果会与其他暗物质模型区分开来。在常规暗物质模型给出的间接可探测信号的预言中，银河系中心的信号是最强的，因为那里是暗物质密度最高的地方。双盘暗物质模型预言的信号强度也是，越靠近银河系中心越强，但来自整个暗物质盘的信号应该和来自银河系中心的信号一样强，因为整个平面都是高密度区域。这种贯穿整个星系平面的可见湮灭观测，将是双盘暗物质模型的最大特点。

双盘暗物质模型直接探测的可能性也非常有意思，毕竟直接探测暗物质是很多暗物质探测器的终极目标。还记得暗物质直接探测依赖于暗物质和普通物质的微小相互作用么？探测器是通过测量微小的反冲能量来寻找暗物质的。和间接探测一样，任何双盘暗物质模型的直接探测信号也依赖于对其粒子的乐观而不寻常的假设：暗物质粒子与普通物质存在相互作用，这个相互作用既可以小到与现有的观测相吻合，又可以高到产生可观测信号。

直接观测信号也依赖暗物质的局部密度，因为毕竟暗物质越多越好。暗物质盘可能在或不在普通物质的周围，这取决于暗物质盘的厚度，但如果存在暗物质盘，那么它的密度会远远高于暗物质晕中的暗物质密度。

另一个众所周知的事情是，暗物质的探测概率与暗物质粒子的质量相关，决定了暗物质粒子的反冲能是否大到可以被仪器记录下来，如果能达到，这个能量就能被记录下来。信号的可探测性和另一个常被忽略的暗物质性质息息相关，即粒子的速度，它与动能直接相关，所以与反冲能也直接相关。速度快的暗物质与比较慢的暗物质相比，更容易被观测到，因为碰撞时产生的能量更高。

双盘暗物质在星系盘内外的速度要比普通物质低很多，因为它们已经冷却了。进一步讲，暗物质在银河系中的轨道与太阳系的轨道风格类似，所以和地球的相对速度也比较小。这种相对的低速度意味着，即使有相互作用，双盘暗物质也只会直接探测仪器中留下极小的能量，这个能量几乎可以肯定会低于探测器的探测下限，因此也就不会被探测到。没有更敏感的仪器或者模型中新的成分出现，传统双盘暗物质的相互作用不会被现有的仪器观测到。

更低探测极限的试验已经开始建造，并且对模型的修改也允许现有的仪器得到有效的观测。有意思的是，如果这个信号被观测到，那它将会非常鲜明，以可以确认一个双盘暗物质的起源。因为这种低速暗物质会产生一个比其他之前提过的暗物质候选者都要集中的能量信号。

关于双盘暗物质模型（或其他带电的可形成原子的暗物质）更有意思的检测，是来自宇宙微波背景辐射的细致研究。一些天文学家和物理学家在宇宙微波背景辐射和星系分布数据中寻找黑暗原子以及双盘暗物质粒子的证据，这是个有趣的新方法。

需要记住，来自普通物质的辐射会抹掉带电物质中的密度变化，就像沙滩上的风会吹平海浪留下的痕迹一样，而暗物质则主导着之后的进一步演化。这种特别的效应会在宇宙微波背景辐射上留下印记，这个印记可以用来区分暗物质和普通物质。当带电物质组成中性物质的时候，普通物质也可以在宇宙微波背景辐射上留下印记，这和沙滩上海浪拍打过后，在海浪的最远端会留下一个非常特别的印记一样。

如果暗物质或者部分暗物质，与黑暗辐射相作用，那就像普通物质改变宇宙微波背景辐射一样。因为双盘暗物质模型中包含了一类重的暗物质粒子和一类轻的暗物质粒子，而且它们带有相反的电荷。这很像质子和电子，这两类暗物质粒子还有可能组成黑暗原子，这个过程也与普通物质非常相像。

对宇宙微波背景辐射的细致研究还可以给出我们所提出的有相互作用的暗物质的比例。如果暗物质和普通物质的温度还算相似，那说明暗物质与普通物质在早期已经进行了充分的相互作用，那么相互作用暗物质占暗物质总的百分比会低于5%，这大约是普通物质的25%。幸运的是，这个值也很有意思，用后面介绍的方法我们可以观测到它。而且这个值依然在能解释周期性彗星的范围内，我将在下一章给出详细的论述。

测量星系的形状

上文描述的研究的有趣之处是，它不仅介绍了宇宙微波背景辐射的能量，还说明了如今宇宙学中大数据的重要性，而且天文学家们已经开始着手处理这些数据了。我们现在有了假设的模型，加上今天的技术发展以及数值计算能力，寻找非传统暗物质的影响变得比以前都

更有可能性——尽管这些影响在观测到的结构分布中反映得非常少。我和我的合作者们意识到，最有趣、最有说服力的信号也许并不是来自现有的普通暗物质探测器所专注的领域，暗物质盘的可观测证据更有可能来自盘本身的引力。在大数据时代，似乎只要在现有数据中寻找可辨别的暗物质特性就足够了。

双盘暗物质模型最明显也最具决定性的理论预言是，在银河系的中央盘内还存在一个薄的暗物质盘。如果暗物质粒子比质子重，那么这个盘会比恒星和气体的星系盘更窄，从而使得银河系的引力势与传统暗物质模型预言的结果完全不同。就像有目标的广告，暗物质盘会在星系的普通物质组分附近加上自身的引力势，这个影响会在星系盘附近达到最大。因为在星系盘附近，暗物质的密度非常高。由于暗物质分布导致的引力分布的改变会影响恒星的运动，当足够多的恒星位置和速度被非常精确地测量后，其结果可以证实或证伪暗物质盘，至少对极高密度的模型作出分辨。

一个最不可思议的进展发生在2013年的夏天。当时，我和范吉吉、安德雷、马修刚刚开始思考暗物质盘模型。与此同时，对银河系内恒星的相关观测也正在筹备当中。在2013年的秋天（在南美洲的法属圭亚那发射中心，此时也正是春天，这点差异是我一个澳洲同事指出来的）计划发射一颗卫星，它可以用于测量这一辨识度很高的引力影响。

这个卫星的名字叫作盖亚（GAIA），主要用于高效率地测量星系的形状。5年之内，我们就会知道结果。在我们准备第一篇论文的时候，这颗卫星的数据准备工作已经顺利开展起来，但如果他们在准备的过程中询问我们的建议，我们也许会提出让盖亚观测暗物质盘等有关现象。

事实上，尽管没有准确的模型和基本的方法，天文学家们已经讨论过重构银河系的质量分布是盖亚的主要任务之一，不管是什么形式的物质，也不管它们位于银河系的何处。尽管卫星的发射时间比原计划晚了几个月，最终还是在那年12月发射成功，这个时间也只比我们完成论文晚了几个月。这一切看起来都是那么偶然。

粒子物理学家不会经常遇到类似的惊喜。当某实验可行时，我们会设法调节它，使它能用来测试我们的新理论。工作于欧洲核子研究中心大型强子对撞机的实验物理学家们就研究了类似的科学申请，例如我和拉曼·桑卓姆（Raman Sundrum）以及其他曾设计实验解释希格斯玻色子的质量。尽管大型强子对撞机最初是为了探测其他模型而设计的，但当拉曼和我在建立一个额外卷曲维度空间时，完全了解它的潜力。

另一方面，当一个想法特别令人感兴趣且可以被检验时，实验物理学家们就会对此作出反应，设计一个相对小尺度的实验来排除或者证实这个猜想。比如，物理学家们曾经设计过一个实验用来更精确地测量引力，从而研究空间大尺度的高维性质。

不过，一个实验仪器一开始的目的就是用于研究一个完全不相关的想法的情况却很少见。然而它却确实发生了。盖亚卫星拥有一个空间观测站，用于观测银河系中数十亿颗恒星的位置和速度，目的就是精确地重构星系的三维结构。这个观测将直接应用于重构星系的三维引力势，从而告诉我们银河系的三维质量分布。如果这个质量分布能够证明一个暗物质盘的存在，那么盘的厚度以及密度分布可以告诉我们新型暗物质粒子的质量，以及可相互作用暗物质占暗物质总量的百分比。

研究上述课题的方法是基于天文学家简·奥尔特的一个想法，他因为发现奥尔特云而被人们熟知。奥尔特意识到：恒星进出星系盘的速度依赖于星系盘的形状以及密度分布。因为恒星的运动是盘引力势的结果。测量恒星的速度以及恒星的位置震荡，可以直接限制盘中物质的空间分布。

这正是我们想要知道的信息，它可以检验或者证实我们关于暗物质盘的假设。暗物质盘的引力会影响恒星的运动，因为恒星的运动

被星系的引力势所主导。准确了解如此多恒星的位置和速度，对重构星系引力势并指出星系中是否有盘结构，有着重要的意义。应用盘的引力势的细节信息，以及盘中物质的空间分布，我们希望确定更多盘的性质以及形成它的可相互作用暗物质的性质。

我们没有必要等待盖亚的数据来检验这个方法，因为我们已经掌握了一些有用数据。这些数据来自依巴谷（Hipparcos）卫星，它于1989年由欧洲航天局发射，并且一直服役到1993年。依巴谷是第一个用于测量恒星位置和信息的太空望远镜，但它的测量精度和恒星数量比盖亚要低。尽管这些结果不如盖亚的最终结果完备，但是已经可以用于研究暗物质盘结构的相关课题了。

这个方法对于粒子物理学家来说可能比较新颖，但天体物理学家们已经对其非常熟悉了。事实上，应用这个方法，有些人甚至已经得到暗物质盘不存在的结论。这些有关暗物质盘的否定迷惑了很多人，其中就包括我们论文的一个审稿人。第一反应会告诉你，至少在当前的数据情况下，这个结果是不可能的。不管测量有多么准确，密度都可以足够低，以至于逃离所有现有观测的限制。但其实天体物理学家真正想说的是，不需要这样一个暗物质盘。如果密度场的不确定性完全来自已知的气体和恒星，那么已经测到的引力势用已知的物质就可以解释。

但有时正确的问题是，还存在其他可能，以及相应的其他可行理论。打破这种简并性的唯一途径就是，通过证实或证伪特定的假设给出的特定结果，来排除或确认这些模型的可能性。我和合作者问了一些天文学家另外一个问题，我们没有问如何证明盘的存在，而是问到底多么显著的盘结构会和现有的所有观测保持自洽。我们还问了另一个问题：引入暗物质盘结构是否会使模型和观测结果吻合得更好。

这种不同的思考方式反映了粒子物理学家（特别是建模者）和许多天文学家在科研思维方式上的不同。客观地讲，天体物理学家确实教会了我们很多。我们了解到他们如何逼近问题，以及现在都有哪些数据，他们的方法也非常有用。但从不同的角度研究问题，会给人们带来新的启发，并呈现出新的可能性。不管暗物质盘是否存在，我们只能通过假设它们的行为，和观测进行对比，然后看看这个假设是否合理。其实到最后，我们是共赢的。

我们想知道，观测数据是否允许暗物质盘的存在，或者是否更倾向于支持一个暗物质盘的存在，而对是否可以不用加入暗物质盘就与已有数据拟合得很好，并不感兴趣。每一种被用于计算银盘引力势的物质都被研究得非常清楚了，观测中不确定性的存在，为新物质提供了存在的空间。我为我的学生埃里克·克莱默（Eric Kramer）设计了一个课题。他当时正在研究依巴谷卫星的数据，以及星系盘中气体的密度分布。我们发现，天体物理学家在做数据处理时的很多假设都需要重新考虑。尽管对依巴谷数据的粗处理得到的结论并不支持暗物质盘，但最新一次更仔细的分析表明，这些数据不足以宣判暗物质盘的死刑。

依巴谷卫星的数据本身就有一些不确定性，对银河系内可见物质比较差的测量是不确定性的重要来源。不确定性越大的地方，暗物质盘存在的可能性就越大。最重要的是，银河系内所有物质都会感受其他物质的引力，只有在一开始的时候把所有物质（包括暗物质盘）都计算在内，人们才可能得出真正的限制。这是一个模型存在的价值。它为人们评估观测结果提供了明确的目标和确定的计算策略。

经过认真分析，我们发现了暗物质盘存在的可能性，而且可能性还很大。但在更有说服力的数据释放之前，我们无法简单地证明双盘暗物质模型就是对的，还是说简单的标准暗物质模型就可以了。

这让我想到一个问题：我们到底希望暗物质盘的密度为多少？也就是说，多强的限制会比较有趣？从多方面考虑，任何数值都是值得期盼的。无论暗物质盘的密度有多低，找到一个暗物质盘都会根本地改变我们对宇宙的认识。另一方面，我们还要考虑暗物质盘对周期性彗星撞击的诱导作用。简单地讲，我们所需要的引发彗星撞击的数值与当前数据是不冲突的。

此外，部分相互作用暗物质也可以帮助我们解决一些传统暗物质领域里的著名问题，虽然这并不是我们的最初目的。例如，卡内基·梅隆大学的天体物理学家马修·沃克（Matthew Walker）教授曾指出，双盘暗物质模型有可能帮助我们解决仙女星系内卫星星系数目的问题（这一问题我在第18章提过了）。如果宇宙中只有普通物质或者只有传统暗物质，这个问题是无法被解释的。哈佛大学的博士后雅各布·肖尔茨（Jakub Scholtz）和我的研究表明，自相互作用的暗物质模型也许是暗物质主导的矮星系倾向于分布在星系盘平面这一现象的唯一可行解释。我和雅各布·肖尔茨、马修·里斯也研究了双盘暗物质模型对原初黑洞的影响，这些原初黑洞的质量要比标准模型的预言高出很多。

费米卫星的伽马射线信号促使我们开始了对非标准暗物质模型的研究。现在看起来，这像是一个幌子，因为这个信号随着时间慢慢变暗了。不过，暗物质盘模型却在这一过程中被创造并完善起来。它现在已经有广泛的预言，可以有其他方法观测双盘暗物质模型。这个模型也许能在星系形成和动力学中扮演非常有趣的角色，我们已经开始了这方面的研究。

在漫长的宇宙和太阳系探索之旅后，让我们把旅程中的众多想法汇总起来。接下来要讨论的是，暗物质将如何影响我们的家园，如何影响地球周围的恒星运动，甚至如何影响太阳系外围天体的稳定性。

书籍免费分享微信jnztxy



暗物质！彗星！恐龙！

“Boffins”这个词对大多数美国人来说，并不熟悉。因此，当科技作家西蒙·沙伍德（Simon Sharwood）给我和我的合作者马修·里斯在英国期刊《纪事》（*Register*）中授予此项称号时，我一开始并不清楚是怎么回事。沙伍德是在批评我们和我们愚蠢的工作，或者这个词就像“pulchritude”（特指外表漂亮）一样，听起来挺糟糕的？又或者，这可能实际上是一个恭维的词？

当我知道“boffins”这个词是指科技专家时，还是感到很安慰——尽管担心有点多余。我开始担心这个词的意思是“狂妄的人”，抑或是沙伍德在暗指我们关于暗物质和彗星的工作完全无事实基础时，简单用了这样一个称号。我们认为，暗物质能有效地拽出奥尔特云中的彗星，它们由此会周期性撞到地球上，甚至可能引发了大型的物种灭绝。

即使像马修和我这样的粒子物理学家，总是尽量开放地包容各种理论，对于把彗星撞击这种混乱的现象和太阳系以及暗物质动力学联

系到一起，看起来也是不太可能的想法。更何况是暗物质！彗星！恐龙！人们对太阳系感到好奇并探索它，更不用说我们是科学家，本来就喜欢研究这些变化的片段，并把它们联系在一起。我们还必须要测量暗物质盘的存在。在接下来的5年里，太阳系内的一个太空望远镜要测量10亿颗恒星，并由此解释这个问题，以及我们的观点是否正确。

即使这个方案有多种观点，或者接下来的卫星测量不足以执行这项调查，在我问马修是否愿意考虑这项研究的那天，车里雅宾斯克小行星撞上了地球。尽管大多撞到地球或大气层的小行星都太小了，我们一般都不会注意它们，但这颗在2013年2月15日爆炸的小行星，有15~20米大——它大到可以产生耀眼的光芒，并爆发相当于50万吨TNT炸药的能量。在这颗小行星爆炸前3天，一位来自亚利桑那大学的听众的问题让我开始思考关于周期性小行星碰撞。在我向马修建议深入挖掘这个问题的那天，这颗小行星“撞到”我们了，这非常有趣。就在我们迟疑是否应该研究外来天体撞击地球的时候，它偏偏发生了。我们怎么能不继续前进呢？

接下来我将会讲述我们的研究，这把本书介绍的许多理论联系起来，并解释了暗物质如何在大约3 000万~3 500万年的时间尺度上影响我们的星球。如果我们是正确的，那么不仅有一个15公里大的小行星在6 600万年前撞到地球上，而且引发这次碰撞的是银河系中暗物质盘的引力影响。

一场脱轨引发的恐龙灭绝

我们现在对银河系的理解是一个明亮盘，里面是气体和恒星，也许还有一个更密的盘是由有相互作用的暗物质构成。银河系形成于130亿年前，当时暗物质和普通物质坍缩到一起，形成了一个被引力束缚起来的结构。大约在星系晕形成10亿年后，普通物质开始辐射能量，形成了我们今天看到的亮盘。如果一些暗物质也有相互作用并能足够

快地辐射黑暗光子，它也会坍缩到一个薄平面，也就是盘。这也许会花上一些时间，但这个薄盘应该在很早前就形成了。

同时，大约45亿年前，太阳和太阳系形成。随后，围绕太阳的物质盘中，行星也出现了。行星形成后，木星向内迁，其他巨行星则向外移动，由此盘中的物体被打散开来。它们中的一些飞到了很远的奥尔特云区域，而那些很小的冰块则被太阳很弱的吸引力束缚着。

太阳系以2.4亿年的周期绕着银河系旋转。除了主要的公转运动，太阳系还以一个大约3 200万年的周期在银河系平面上下振动。盘的引力自始至终都作用在太阳上，每当太阳系在垂直方向上离开盘面最远的时候，这个引力就成了复位力，并拉它回来。因为银河系几乎没有摩擦阻力，所以太阳系一直持续这种垂直方向的周期性运动，盘的复位力也一直作用在太阳系上。

此外，当太阳系在盘内或离盘非常近时，盘引力的潮汐扭曲效应就会变强。在这些特别激烈的过程中，暗物质薄密盘的潮汐作用可能扰乱了奥尔特云中一些轨道不稳定的物体的平静，否则它们会一直平稳地运行在它们遥远的轨道上。一旦进入暗物质盘，奥尔特云中的冰块似乎会由于轨道崎岖，而不会待在原本的位置了。

当所有这些无生命体在“各自忙碌”时，地球上的生命在35亿年前出现了，大约5.4亿年前，复杂的生命形式开始繁荣。从那时起生命也开始了起起落落，生命多元化与物种灭绝此消彼长。五次生物大灭绝从时间上把所谓的显生宙分割开来。最后一次发生在6 600万年前，当时一个小行星对地球造成了重击。

直到撞击之前，恐龙们都没有注意到遥远太阳系的混乱。一些冰块的轨道穿过奥尔特云，来自遥远银盘的拉力偶尔有很小的变化，



并根据太阳到中平面的距离而变化。它们中的很少一部分，在引力作用下脱离了初始轨道，轨道变形进入了太阳系的内部。这些冰块中至少有一个可能变成彗星，并进入与地球碰撞的轨道。

从奥尔特云方面来看，这只是非常小的一次骚乱。一个或者最多只有一小部分冰块被带走了。但是对地球上75%的生命，包括可爱的恐龙们来说，这颗要撞到地球的流星体可是世界末日。即使恐龙是有感觉、有意识的物种，当彗星第一次出现时，它们也不会注意到任何反常。尽管彗星核非常亮，白天也能看见，且它拖出的长尾巴整夜可见，它也不会泄露半点它即将带来的毁灭的消息。一旦彗星落地，感觉就完全不一样了，大火和碎片照亮了天空。无论这些生物们可能看到或想象到了什么，当引力扰动改变了彗星轨迹的那一瞬，命运的车轮将无法避免地碾压到这些动物身上。

很快，彗星就猛冲进尤卡坦州，摧毁它的目标，然后以引发全球性大灭绝的方式结束了这次旅程。当彗星撞到希克苏鲁伯陨石坑时，撞击气化了彗星和撞击周围的土地，它激起了尘埃的卷流并很快散布到了全球。大火烧焦了地表，海啸吞没了周围的海域，甚至传播到了地球的另一端，而随着雨水降落的有毒物质则更危险。食物供应骤，那些从更早期灾难中存活下来的陆生生物大概在其后的几个星期或几个月中，逐渐被饿死了。大部分生物根本没有机会面对这样突然且剧烈的变化，从全球气候到各种栖息地。当各种条件最终改善，生命的延续进入到了另一个充满不确定性的时代时，只有陆地穴居的哺乳动物和空中的鸟类存活了下来。

这是一幅戏剧性的场面，但基本的事实是：彗星撞击肯定发生过。地质学家和古生物学家们已经发现了许多证据，6 600万年前一个大家伙撞上地球，造成地球上至少75%的生物死亡。很快我将会介绍，暗物质盘如何引起了彗星的变轨然后导致了这场灾难。但首先我来解释一下这个想法的初衷。

无心插柳触发的灵感

通过书本和讲座把物理知识分享给公众，有许多额外好处。但因为花在这些活动上的时间会耽误正常工作，所以我常常得从中选择。在一些幸运的场合中，我的研究却因此受益，我曾担心参加公众活动会使我分心，最终却常会得到一些通常没有机会遇到或想到的、我自己不会有的新颖想法。

2013年2月，我从天体物理学家保罗·戴维斯（Paul Davies）的邀请中得到这样一个机会。他在亚利桑那州立大学的超越中心（BEYOND Center）举行年度讲座。尽管我担心旅途太长，但那所大学有一个非常好的宇宙学研究组，所以我不仅很高兴地接受了一个公开讲座，还在之后的几天参加了一个专业研讨会。这个研讨会聚焦于我近期的工作，就是双盘暗物质模型。

出席研究会的物理学家问了许多关于模型的精彩问题。例如，它的可探测性，它对宇宙微波背景辐射的牵连。但当保罗问我暗物质盘是否造成了恐龙的消失时，我大吃一惊。我要承认在那之前，我还没有想那么多，在我的科研中没有恐龙！我一直都集中在基本粒子和宇宙的基础上。但保罗提示我，有潜在的证据显示存在周期性的陨石碰撞，但缺少合理的解释。他问我，暗物质盘能否满足要求，然后告诉我关于彗星引起了恐龙灭绝的猜想。

保罗的问题太精彩了，我怎么也忘不了。答案并不简单，我还需要研究很多东西才能给出明确的回答。暗物质和恐龙之间貌似存在的联系，以及参会的很多科学家，教给了我很多东西。我问马修是否感兴趣研究一下由我们提出的暗物质盘所引起的彗星碰撞。对一个物理学家来说，这种联系要比仅仅关注恐龙，要更吸引人。

显然，马修是个不错的合作伙伴人选。他在双盘暗物质模型的最初研究中是关键人员；他有着冷静的技术头脑，而且在科学上对新想法非常包容——这与他果断保守的行为完全不同。马修不会在任何假设上犯一般性的错误，总是能正确猜出所有事。

最重要的是，马修是位优秀的物理人，有很高的科学修养。当他做事时，每一步都走得很稳。当然，我并不肯定他是否会被这种显然的疯狂建议打动。当马修认为这个想法很有趣并认为它有很大科学意义时，我感到很高兴。保罗·戴维斯也很感兴趣，但是他已经有太多的科研工作要做了，于是选择保持联系但不直接参与。

当听到车里雅宾斯克小行星的消息后，我们当天就开始讨论这个问题。我和马修加紧调研相关内容。我们的目的是把这个疯狂的想法变成可证实的科学。作为建模和物理学家，我和马修尝试接受新的观点和解释。但我们也时刻注意保持无偏见和仔细认真，这些品质在我将要讲述的研究中非常重要。

暗物质盘和太阳系

如在14章所述，为实现我们的目标，我和马修决定开始先简化我们的调研结果。尽管我们对恐龙很感兴趣，但开始时我们先把这些关于灭绝问题的挑战放到一边，集中研究彗星和太阳系动力学以及陨石坑记录的周期性。当把灭绝问题放到次要地位后，我们就能集中研究

暗物质盘对彗星可能的影响，以及它是否造成了周期性彗星的问题。之后我们可以看看我们对具体某次彗星的预测如何，包括引起白垩纪-第三纪灭绝的那次。

我们发现，之前的理论没有能解释如何周期性地把奥尔特云中的天体拽出，并留下周期性信号。如果一个常规机制能够奏效，那将没有人，包括我们自己，愿意给陨石记录设计一个更奇异的方案——无论这项工作有多酷、多诱人。

然而，传统起因此时行不通了。仅靠标准银盘，星系的潮汐效应太温和，恒星的扰动不够频繁。传统的潮汐效应，内梅西斯星、未知行星以及银河系旋臂，都不足以引发如此频率和如此数量的彗星雨。这些早先的理论既没有给出穿过银盘的正确时间，也不能给出充分的撞击吻合陨石坑的记录。例如，如果只考虑盘中普通物质对运动的影响，太阳在垂直方向的谐振周期应该在5 000万~6 000万年左右，这和现有数据相比太长了。

这留下了两个可能的结论：一是，这个周期不是真的；二是，更有趣也合理的解释是，这一起因是非常规的。仅靠普通物质并不能解释所需的周期和变化率，在排除了这些理论后，我们可以更合理地回头看看我们的暗物质盘理论能否实现。事实上，暗物质盘的确有普通物质盘不具备的必须特征。一个薄而密的暗物质盘，它的潮汐力能成功地解释周期和奥尔特云扰动对时间的依赖。

从它们出现起，奥尔特云中的天体就受普通物质盘的潮汐力影响，此外还受偶尔路过的恒星的影响，这其实也很重要。这些效应都会引发奥尔特云内天体的扰动，从而被拽到太阳附近。银盘的潮汐力能最后推它们一下，使这些冰块进入不稳定的、离心率非常大的轨道。它只有日地距离的10倍左右，而那些大行星的引力很可能把这些冰块拽出奥尔特云的轨道。这些彗星既不会飞出太阳系，也不会进入太阳系内圈轨道。这些扰动说明长周期轨道彗星的产生，每年都会有

一些新进入太阳系的。偶尔，被扰动的天体可能一起被甩出轨道，这时这些偏轨的彗星可能发生碰撞。

仅靠这些扰动并不足以解释周期性的陨石碰撞。要引发周期性的碰撞，对奥尔特云扰动的快速变化必须发生在一个规律的间隔上。此外，为了符合现有证据，这个周期需要在3 000万~3 500万年左右。即使其中某个标准没有达到，它都不能解释周期性彗星碰撞。而且，常规解释没有符合这些标准的任何一个。

在考虑一个密且薄的暗物质盘后，这些问题被很好地解决了。一旦承认了周期性彗星的事实，暗物质盘就成为一个非常可行的理论。暗物质盘施加的影响比普通物质盘的强度更大，变化得也更快——这是产生彗星数量峰值的两个根本要求。

如果银河系平面上包含了暗物质盘，那太阳的垂直振动周期将会变短，因为增加了暗物质盘的质量后，引力变强了。除此之外，根据现在的物质密度测定结果，太阳系在银河系平面上下的运动仅有70秒差距，这个范围比只有普通物质盘出现的情况要小得多。这个窄暗物质盘，覆盖了很大一部分太阳系的轨迹，随着太阳系在银盘面上下摆动，能够对太阳系的运动产生很大影响。

薄暗物质盘的另一个功效是，太阳系很快地从盘中穿过，由此产生了一个彗星率的峰值，每次持续100万年左右。因为它的效应具有强烈的时间依赖性，太阳系每次穿过银盘面，暗物质盘都会引发额外的扰动，由此会规律发生彗星雨——就在太阳系每次穿过银盘面时——否则，彗星只能偶尔由接近的恒星引发。太阳系穿过暗物质盘狭窄的区域时，增强了的潮汐效应就会发生。在穿过盘面时，以及其后的一两百年间，彗星的撞击会增多。

当太阳系以这个时标穿过盘面，并受增强了的潮汐力影响——如果发生足够快就会有一个力的峰值，此时奥尔特云中的冰块可能会被拽出，其中一些可能会以每秒50公里的速度抛向地球。一旦开始进入这种轨迹，进程就非常快，大概只有几千年。但最初使它走向这一步的扰动却要慢得多，一般要花费几次轨道周期的时间。这意味着，在10万~100万年的时间周期里，这些彗星是否会飞到太阳附近的命运就会被决定，而其中一些则导致了地球上看到的彗星雨。

我和马修对轨道做了预测，尽管有些数据不确定，方案还挺成功的。但还有最后一步检查我们没有完成，这是当我们向顶级物理学期刊《物理评论快报》投稿时，审稿人给我们指出的一个问题：在暗物质盘的存在下，除了要确定太阳系的运动，还要计算当太阳系穿过暗物质盘时其环境密度的涨落。我们需要知道密度，因为我们假设对奥尔特云的扰动是与这个物质聚集的程度成正比的。质量越大，潮汐影响越大，扰动也就越强，我们因此假设密度是彗星撞击率的代表，事实上也的确是这样。

对暗物质盘作用在奥尔特云上的潮汐扭曲是否足以使云中的冰块形成正确数量的彗星雨，我们还没有十足把握。幸运的是，斯科特·特里梅因和朱莉娅·海斯勒在10年前已经完成了许多困难的工作，所以我们可以借鉴他们的成果。我们的假设确实是正确的：增大了的密度在恰当的时候对彗星产生了所需的拉拽力。

事实上，我很钦佩《物理评论快报》审稿人提出的建议。审稿报告，是论文在被批准发表前由同行专家审阅后的评论，但如今很多都是人云亦云，不假思索给予批准或者成了要求作者引用其文章的手段。这次审稿人的建议事实上教给了我们一些物理想法。虽然口气有点轻视，但我们确实从中学到了一些东西。我们同时还要解释一些误

会的批评，因为我们之前已经仔细检查并请同行审阅过，我们可以很容易地找出这些评论中的错误。

最后，我和马修计算了与陨石记录符合得更好的暗物质盘密度和厚度，发现它们与我们之前已有的双盘暗物质模型一致，那时我们已经知道它与银河系的测量一致。我和马修发现，新的方法不仅在暗物质盘的条件下可行，而且如果你真的把盘当作引发彗星出现的原因的话，你会发现我们的模型真的更好。

暗物质盘的密度约是普通物质盘的17%，这很有趣，而且不会推翻当前对任何现象的理解。这是相当大的一块暗物质，不是百万分之一，而是百分之十几。万一这个暗物质成分存在，它将可以产生可测量的效应，因此也就值得人们注意。此外，暗物质盘的厚度也许只是普通物质盘的1/10，不到几百光年，而普通物质盘的厚度大约为2 000光年。正是暗物质盘如此薄，才解释了为什么它能够令人信服地引发了周期性的戏剧效果。

我们发现，暗物质盘的恰当密度值倾向于大3倍。这个新结论的一个关键原因是，它有更好的统计支持，也即我们在前文提到的旁视效应。给定了一个可以引发周期性彗星的明确模型，我们不仅可以更好地预言周期，而且这个预言更可靠。实际上，我们论文的目的不仅是证明暗物质盘能比普通物质盘能更好地解释周期性彗星雨现象。还有一点——这需要一些统计方法，即如何评估相应结果的重要性。

正如14章所说，大部分方法都是把太阳系上下运动的数据拟合到一个周期函数上，比如正弦波。这也挺有意思，但不能抓住全部内容。我们不需要猜测太阳系的运动。如果我们已经知道关于星系和太阳的初始位置、速度、加速度等所有信息，就可以从牛顿定律中算出太阳的运动，并预测出周期。毕竟，太阳系的运动不是杂乱无章的，

它符合动力学原理。即使密度分布和太阳的数据不是很完善，可能的运动范围和由此得到的可能周期，也是限制在一定范围内的。

我和马修一起找出了已知的银盘密度，给出了当前测量支持的全部范围，并将暗物质盘考虑在内。我们的目的是，如果我们把盘所有已知的成分都考虑进来（恒星、气体等），再加上暗物质盘成分，看看太阳系的运动是否与周期性的陨石记录相符合。

测量普通物质的作用限制了太阳系可能的运动轨迹，这是因为盘中物质（包括普通物质和暗物质）的引力作用在太阳上，影响了它的运动，因此降低了旁视效应的影响。我和马修使用测量的密度预言了太阳系的周期运动，比较了穿过银盘的时间和产生陨石的时间，并对比记录看看它们是否吻合。尽管没有具体模型，预言不能很好地区分具体情况，我们仍发现，根据现有的测量数据，统计上倾向于周期为3 500万年的流星体撞击。近期的数据表明，这个周期可能要短一些——大概为3 200万年。

暗物质盘是使这个方案运转的关键，而且产生了预想的撞击率。回到我们的主题，陨石坑更好地吻合了太阳系的运动，说明暗物质盘实际上是更好的因素。将来的数据分析应该考虑这种模型，以便得到更好的统计显著性。结果将会进一步支持我们的结论——或者把它否定。

关于恐龙灭绝

当我和马修整理完所有资料时，我们的论文也被《物理评论快报》接受了，马修把结果放到了互联网^[41]上。我们保守地给论文命名为《暗物质引发了周期性的彗星撞击》（*Dark Matter as a Trigger for Periodic Comet Impacts*）。不过，马修在上传结果时在注释栏

中标注了“4张图片，没有恐龙”，我觉得这非常有趣，因为我们一直在论文中有意地避免提到恐龙，只集中在了陨石记录及其直接的物理联系上。我们当然一直知道这个联系，并开玩笑地把这项工作称为“恐龙论文”。我们的研究结果得到了很高的在线关注度，许多博客和新闻网站，包括“boffins”那个，在转发论文时几乎都会配上有趣的图片。

这使我回到了恐龙的话题。在初次建立起数据和模型的联系、预言了碰撞周期，并知道这肯定不是最终结果，还会随着将来的测量发生变化，我们想试试我们的模型和希克苏鲁伯陨石坑的时间是否相符。计算显示，在改进了对银盘普通物质的测量后，流星体撞击会发生在3 000万~3 500万年的间隔上。因为我们用200万年穿过星系盘面，一个从奥尔特云中被拽出的彗星完成了一次振荡（两次穿过银盘），并可能在6 600万年前（即白垩纪-第三纪灭绝）被快速甩到了地球上，造成了那次浩劫。需要补充的一点是，如果地球是在不到100万年前穿过了银盘，地球可能还在彗星流的末端，我们今天还有可能看到剧烈的撞击。但更可能的情况是，除了那些无规则、极小概率事件，地球已经穿过危险区，在下一个3 000万年将不会再看到另一个希克苏鲁伯陨石。

由于太阳位置的不确定，并缺少精确的周期，我们只能大概估计穿过盘面的时间。如果地球是在200万年前穿过星系中平面，那么3 200万年的振动周期将会是引起6 600万年前那次事件的最佳值。我们粗糙的初始分析得到的3 500万年的周期，对比希克苏鲁伯陨石的时间稍微长了一点——模型中的不确定性和剧烈彗星碰撞的时间允许这一合理的可能性。修改后的银盘模型，考虑了最近对星系成分的观测，给出了较小的周期，它与白垩纪-第三纪灭绝的时间符合得更好。但即使使用最初的简陋模型，暗物质盘的预言也有可能给出希克苏鲁伯陨石事件的合理时间。

我们的结果不够精确的一个主要原因是，从我们着手分析以来，对银河系物质的测量结果改变了。而且我们还没有考虑星系环境随时间发生的变化，例如旋臂，至今我们对其知之甚少。以上这些效应引起的密度变化不足以引起彗星撞击，但在几百万年的时间里，它们足以改变模型对于撞击发生时间的精确预言。

对于猛烈彗星雨确切时间的预言，还有其他不确定因素。首先，太阳系穿过银盘大概要用100万年的时间——如果盘更厚，花得时间会更长。其次，从事件开始触发到彗星真正撞到地球上还要有大概几百万年的时间。第三，对陨石坑时间的测定精度很低。找出更多陨石坑，或者更精准地测定它们的时间，将对我们很有帮助——尽管很少能发现新陨石坑。不仅仅陨石坑，岩石中积累的尘埃，也能帮助更精确地测定陨石撞击的时间。

关于太阳在银河系平面上下的振动周期（3 000万~3 500万年），还有意想不到的证据。我和马修完成了论文后，一位粒子物理学同事知道我对天文、地理和气候都很着迷，但当时他还不知道“恐龙论文”，他偶然一次告诉了我有关尼尔·沙韦夫（Nir Shaviv）及其同事的工作。沙韦夫是耶路撒冷希伯来大学的教授，他主要研究整个5.4亿年间显生宙的气候变化。出乎意料，他们发现了一个3 200万年的气候变化周期，这和我们得到的周期非常相似。如果他们的结果靠得住，而这个气候的周期又确实是由太阳在星系中的运动引起的，那么这个3 200万年的周期将是暗物质盘存在的证据。因为，普通物质不足以产生这么短的周期。

当然，我们不需要深究过去来看暗物质的影响。如果暗物质真的有相互作用成分，并改变了宇宙中物质分布的结构，我们将会很快知道——快于所有寻找暗物质的研究。只有有限的一部分暗物质盘密度

能解释陨石的数据。进一步的测量几乎肯定会缩小预言的范围，证实或者否定我们的理论。

我与我的学生埃里克的研究证实，暗物质盘所需的密度和厚度是在当前观测允许范围内的。来自盖亚卫星的数据将会进一步确定盘的位置、密度、厚度。当这个卫星完成绘制银河系近邻区域的恒星三维图像后，暗物质盘是否存在将会更加清楚。通过这种间接途径，我们不仅可以了解星系和暗物质，还能知道一些太阳系的过去。如果盖亚卫星的数据能够证实暗物质盘的存在及其厚度和密度，那它也是陨石理论的有力证据。

一个点睛之笔当然是，我们精确地算出了恐龙灭绝的时间，这是一个复杂的主题，十分具有挑战性。尽管如此，过去50年的科学发展从不缺少惊奇。暗物质在很多方面都比地球、太阳系和宇宙中的许多可见元素更难以捉摸，但通过研究，物理学家们正在通过新方法去寻找它。无论发现了什么，我们都非常肯定星系、宇宙和其中的物质，都隐藏着迷人的“礼物”。

结语

一个奇妙的宇宙正等着我们去探寻

我是一个幸运的人，常被邀请参加各种不同领域的高层会议，包括商业、法律、对外政策、艺术、媒体，当然还有科学。当我和其他参会者持不同的观点时，我们的讨论更能激发出新鲜的想法，而这些想法都是各个领域内的重要课题。但最好的问题，尤其是关于我的研究，往往不是来自会议的参与者。例如，某次关于物理学的交流特别让人开心，那次会议之后，一个名叫杰克的年轻普通人对物理学的深刻思考让我很震惊。杰克是个司机，当时正载着我去蒙大拿机场。

一般情况下，当人们知道我是个物理学家的时候，常会把他们关于事物的理解强行地灌输给我，这些理解包括对科学的爱与恨、喜欢与迷惑。我发现有时这还挺有意思，毕竟我们中的大多数都不会向律师聊太多自己关于法学的想法。但这些对物理学的另类“忏悔”有时还是会有所回报的。杰克向我提到，几年前上高中时，他曾接触过一些大学级别的物理学课程，现在他依然渴望继续学习。尽管现在他已经不再上课了，但他希望聆听到更多有关物理学的前沿知识，并进一步了解我们对宇宙的理解。

杰克并没有局限于询问物理学的最新发展，他想知道从他所了解的物理学开始，物理学都有了哪些新发展。所以我向他解释：20世纪物理学的发展告诉我们，牛顿力学尽管在我们熟悉的环境中非常准确，但在小尺度、高速度或者高密度环境中不再适用了，人们要用量子力学、狭义相对论以及广义相对论来替换牛顿力学。

在仔细思考了一番之后，杰克提出了一个看似随意，但又十分深奥的问题。他问我，如果我能回到过去，我会用我现有的物理知识做些什么；我是否会告诉过去的人们一些只有今天才知道的物理学新发现。

杰克认识到了这个问题的两个方面。第一个方面是，到底会不会有人相信我，或者他们只是把我看成一个疯子。毕竟，没有试验证据的支持（尤其是这些实验都是今天最先进的技术实现的），这些20世纪物理学家发现并推演出的令人印象深刻的现象和联系，在“过去”看起来确实有些疯狂。这些发现已经颠覆了人们对这个世界的常识性理解。

他问题的另一个方面更令人压抑。杰克想，就算人们相信这些新观点，他们也会被吓到而忽略这些观点，或者走向另一个极端，即急迫地想应用这些新的物理概念。他的第一直觉是，应该在想象的时间旅行里保守这些秘密，因为我们最好还是让世界以它本应该有的形式发展：获得科学知识不应该有捷径。

公众通常不喜欢长远思考，所以杰克担心，突然的信息爆发也许并不安全。他并不认为改变是坏的。但当他看到自己的弟弟妹妹们沉迷于电子游戏或者智能手机中，而忘记了锻炼和户外活动以及他在年轻时喜欢的探险活动时，杰克感到非常忧虑。他也非常担心发生在他家乡的事情。在那里，一旦有新的技术引进，工业化就会大肆抢夺资源，完全不顾这些行为对当地人或者全球带来什么影响。自杰克出生以来，他看到自己的家园和家人的生活都发生了不可挽回的变化，然后他得出了自己的结论：若未经过认真了解和制定长期战略，公众也许很难有足够的时间去适应最新的科学发现和技术进步。

本书也介绍了地球上发生的几次无法控制的主要变动对这个星球造成的巨大影响。其中一次发生在6 600万年前，一颗来自太空的高速

彗星——可能是由暗物质触发，导致了一次大型生物灭绝。也许3 000万年以后，另一颗彗星又会撞到地球上。解开这些神秘事件令人非常着迷，这就是为什么我会在书中介绍它们，并且还在继续研究它们。

理解这些碰撞对地球以及地球生命的影响，也许更为重要。这可以帮助我们预测，今天的我们对环境所做的一切会带来可能后果。在与文明以及物种多样化有关的时间尺度上，“一颗脱轨的彗星”，也许只是杞人忧天。但正像人口爆炸导致地球资源的过快消耗等类似问题一样，我们无法忽略上述问题。其影响甚至相当于一颗慢速的彗星，这一切的始作俑者是人类自己。但和遥远彗星和地球的碰撞相比，我们也许可以尝试着去控制人类制造的这次“冲撞”。

对暗物质的研究看起来和人口爆炸、资源匮乏毫无联系。《暗物质与恐龙》描述的是我们周围的大尺度环境（即地球所在的宇宙环境），科技进步带来的显著效应，以及未来它能给我们带来什么。通过思考暗物质的本质，可以让我们思考银河系，进而思考太阳系，以及我们关心的彗星。这让我们更好地了解了恐龙的灭绝，并让我们由衷地为现在地球上存在的所有生命的平衡及其茁壮成长感到赞叹，为大自然的精致感到赞叹。如果扰乱了这一平衡，我们也许会存活下来，这个星球也会存活下来。但是，我们却无法确定和我们一起生活的物种，以及我们赖以生存的物种，会不会在打乱平衡而导致的重大变更中存活下来。

宇宙大约有138亿年的历史，地球围绕太阳也转了将近45亿圈。人类降临到地球上也就仅仅200万年，而文明的历史更是只有区区不到两万年。然而在我生活的年代，全球人口已经翻了一番还多，增长40多亿。当人类过快地消耗地球资源时，这已经影响到地球以及地球上的其他生命了，人类正快速毁灭着宇宙百万年甚至百亿年所积累的一切。在一个人的人生跨度上，这个威胁也许没那么紧迫。然而，对这

一威胁保持警惕，却可以帮助我们设计出最优化的方式，去更合理地应用将来会出现的新科学技术。

人类喜欢把自己想象成为适应能力很强的物种，但事实是，当今的世界要比我们想象的更加不稳定。不断变化和被破坏的栖息地以及大气环境都影响着物种的多样性，甚至促成了第六次灭绝。尽管人类并不会很快消失，但人类生活方式中的某些重要方面可能会消失。人类现在的改变已经威胁到我们的环境，更不用说对社会稳定性以及经济稳定性的影响了——虽然有些是我们自认为正确的解决现有问题的方法。而且，尽管有些结果从长远看也许会有益于全球的环境，但这些改变对于现在生活在地球上的物种来说，也许并不是那样的。

我们可以尝试改造自然环境的一些方面，但世界是一个极其复杂的系统，它有很多看起来如奇迹一般的特性，而关于这些特性我们现在知之甚少。即使现有方法能解决一些问题，但很难保证它们能跟得上这不断发展的变化。没有对平衡的不断调整，结果将不可避免会不可控地增长，并为这一过程付出代价。在一个支持技术发展的政治、社会以及经济环境下，如果我们期待获得积极的回应，把技术集成到更复杂的策略中是非常必要的。我们面对着巨大的挑战，但它不应该是阻碍我们向这个值得一试的目标前进的理由。

指数型增长在一开始会相对缓慢，但之后会急剧增长。为了维持新平衡状态，需要的资源会超过我们过去所消耗的所有资源。我们这个精致的生态平衡系统，加上其复杂性以及脆弱的基础，就算是非常微小的扰动，也会导致非常大的影响。因此，三思而后行变得尤为重要。就连教皇弗朗西斯（Pope Francis）在他2015年的教皇同谕中都警告世人，人类活动过多与过于频繁是个不容忽视的问题，他称之为“过快的改变”。尽管有些改变是有益的，但潜在的有害结果需要提前估计。如果单纯地从一个角度看问题，我们会只拘泥于眼前。

大家不要误解我，我相信一切都会有所改善，知识是美妙的。但我也相信，理智地应用先进的科学技术正意味着长远的考虑。一个智慧物种不应该把自身的存在建立在与其他物种的恶性竞争和破坏稀缺的资源上，而这些资源要经过几十亿年或至少几百万年才能形成。尽管技术可有益、可有害，而且有些危害是不经意的，但知识的不断增长让我们有能力创造出人类需要的机器，作出更好的预测，为潜在问题找出解决方案，并评估人类理解的极限。是否能合理地利用知识，完全取决于我们自己。

科学正在悄悄改变我们的世界

我们应该记住，从全局来看，所有科学发现的应用在一开始都是不明朗的。然而，科学的进步却悄悄地改变着这个世界，以及人们的世界观。如果应用得合理，科学技术会为人类带来不可估量的好处。即使很多基于抽象理论的观点，一开始没人相信它们能有任何实际应用，但它们确实极大地改变了我们的世界。

基因科学最早是对DNA的研究，是为了解决纯理论的问题，但如今被用来治疗癌症。一些医疗仪器，比如核磁共振成像（MRIs）最早用于研究原子核。核能的应用也许争议很大，而关于核能的知识都源于对原子结构的研究。电子行业的革命源于原子晶体管的发展，而这一研究属于量子物理学的范畴。互联网是计算机科学家蒂姆·伯纳斯-李在欧洲核子研究中心（现在用于放置大型强子对撞机）工作的副产品，它的主要目的是更好地联系来自不同国家的科学家，从而更合理地协调工作。今天普遍存在的GPS系统则融入了爱因斯坦的相对论理论。在电被发现的时候，甚至没有人认识到电会如此重要，而今天我们的生活已完全离不开电。

当地质学家沃尔特·阿尔瓦雷斯刚开始研究工作的时候，他认为地质学比物理学乏味。地质学家们绘制着看似简单的河流和陆地的图谱，而20世纪的物理学却根本地改变了人类对这个世界的认知。但随着板块构造学、地层学以及地质演化学的发展进步，储藏在地下的石油和矿产不断被人发现、开发并利用，这开启了人类研制机器来寻找石油和矿产的时代。这个转变开始于18世纪，但地质学变得更重要是在20世纪。地质学为我的世界创造了可观的收益，它实实在在地为现代工业的复杂机器提供了燃料，这使得我们的经济和生活方式都有所改进。但同时，这也带来了今天的环境问题。

沃尔特以及其他地质学家对地理学发展的贡献，并不只是为了工业发展，也同样为了理论研究，他们的工作使得我们对地质学有了新的认识。把流星体和太阳系联系到了更广阔的领域——银河系的结构，这种过程看上去正是智能生物正确的扩张冒险步骤：把我们的世界和周围的宇宙联系起来。

正如沃尔特的工作引发了地质学、化学以及物理学的交融，我希望本书的研究也是对其他学科的一种延续。暗物质可能是让人类更好地理解这种联系的一块拼图。我们不仅可以应用地质学去理解宇宙学事件，而且对暗物质本质的详细理解，也许会帮助我们搞清楚把彗星带到地球轨道附近的动力学系统。

大部分人更关心流星体（天文学家和小行星矿产投资人除外），因为这些星体可能和生死相关，但这些天体产生高危事件的概率非常低。小行星和彗星基本上会在各自的轨道上稳定运行，而且它们与地球相撞的概率很小。只有极少数确实巨大的天体会过于偏离轨道，而离开太阳系或者攻击地球。基于我已经介绍的，希望你现在对什么样的地外天体会在将来撞击地球、会带来多大的危险，有更好认识。

这本书解释了一个可能的这种撞击所产生的几个原因，这次撞击造成了6 600万年前的白垩纪-第三纪大灭绝。从广义上来讲，我们都是希克苏鲁伯的后代。这是人类历史的一部分，我们应该研究并了解这段历史。如果这是真的，那么根据本书其他的相关分析，暗物质不仅不断在改变着人类的世界，而且对于人类的出现扮演着至关重要的角色。在这个框架下，从恐龙的视角来看，暗物质就是魔鬼，科学家命名其为“暗物质”也是很明智的。但从人类的角度出发，这种可能的新暗物质类型是这些重要意外的触发者，而这些意外会改变地球的发展进程，甚至决定了你是否有机会坐下读这本书。

在这本书中，我尝试着展示科学研究的本质，即人们如何根据已有的知识去探索未知的领域。另一方面，宇宙和地球的历史都把我们送上了一次通往过去的、刺激却充满挑战的旅程。如果你认为家族历史已经很难追根溯源了，尽管长辈会告诉你很多相应的线索。那么请想一想，通过研究那些只存在于不会说话的石头上的信息来研究过去的困难程度，而且很多信息已经随时间的流逝而被风侵蚀掉了，甚至，有些随着地壳的运动而到达地幔中；又或者理解暗物质结构的复杂性，而对于这些暗物质，我们根本看不到。

科学的进步已经揭示了基础物理学和我们眼前可见世界之间极其复杂的联系。暗物质的坍缩产生了星系，恒星中心所产生的重元素则为生命的形成做好了准备，地幔深处原子核衰变所释放出来的能量又驱动着地壳的运动从而形成了高山。我发现，人类这种能够理解宇宙中深层次联系的能力确实让人振奋不已。每当科学家探索已知世界的边界时，总会有意想不到的结果发生。

这个世界是丰富多彩的。它如此丰富多彩，以至于粒子物理学家会问的两个最重要的问题是：“世界为什么如此多样？”“我们所见的所有东西是如何联系到一起的？”在我的研究中，我知道这些研究

也许会或不会最终直接连接起人类对现有周围世界的了解。但不管怎样，我希望研究结果会对将来的研究作出一点贡献。我专注于现有理论框架的研究，我明白，任何不符合标准理论的图像和计算的结果，可能是源自对传统模型的不够了解，或者是在提示我们作出一些新的尝试。

暗物质及其在宇宙演化中的作用是当今最令人兴奋的科学主题之一。我们能够真实地理解所有形式的物质（正如不同文化的复杂社会），但这需要我们了解它们的性质，并能够定量地计算不同物质的组分。而且，按照这种组合方式组合在一起的物质，可以重现我们周围环境的物质丰度。改进我们理解物质的最好方法就是，构建最优美可靠的，并且符合观测的物质成分表。我在前面所展示的暗物质研究也许只是一次思想实验，但它将会在未来的测量中被证实或证伪。数据和理论的自洽，是结果正确与否的最好标准。

本书书名所暗示的影响彗星轨道的暗物质，并不是我们所指出的新型暗物质的唯一可能。暗物质盘结构可以影响恒星的运动、矮星系的形状，以及实验室中的实验结果或者太空中的观测结果。尽管与探索地球和太阳系比起来，研究暗物质从许多方面来看都更难以捉摸，但科学家们正在寻找新的方法来追踪它。相应的结果会帮助我们揭开银河系以及宇宙的神秘面纱。

地球上的生命也许不是宇宙中唯一的生命形式，但人类的存在需要宇宙或者行星具有一些显著的特征。根据现有的了解，人类的问题最终变为了：人类是如何来到地球的？通过研究星系以及人类的起源，人类可以判断，人类到底是因为一次偶然的故事来到地球的，还是经过了平缓的演化，即人类出现在地球上是可预言的？我们已经了解了非常可观的信息，并希望可以揭示更多的联系。在过去的50年中，从不缺乏惊人的进步。

尽管我们经常会看到新闻报道的坏消息和全球不断发生的令人失望的事件，但传播科学知识可以帮助人类改善自己的生活，并指导我们朝着对将来最有利的方向发展。科学研究不断呈现着关于人类生命与环境之间的联系、过去与未来的联系，等等。所以，人们应该感激这个世界是如此丰富多彩，它为人类应用已有的智慧和技術提供了场地和对象。

我经常觉得，奇妙的宇宙学环境是如此令人振奋。来自外界的争吵和忧虑都不会打扰我们从大尺度上洞悉科学的意义：科学能教会我们什么？我所说的，也许听起来并不像是最实用的建议。但是仰望星空，就在你的周围，一个奇妙的宇宙等着我们去追求、拥抱和探索。

致谢

这本书的灵感源自我在物理学中的研究。许多天文学、地质学和生物学研究，特别是我书中提到的那些研究工作，都给了我启发。这些研究是许多科学家的贡献，感谢所有的物理学家和天文学家，感谢他们在研究中共享了他们的知识。《暗物质与恐龙》反映了我对这个世界的着迷和激情，以及我对其发展方向的关心。这些想法最终能成型，得感谢我的很多朋友，多年来和他们的许多交流使我深受启发。我想感谢在这一路上每一个帮助过我的人。

特别感谢我的许多同事，他们与我有着同样的科学兴趣；尤其感谢在暗物质盘研究方面的诸多合作者：JiJi Fan、Andrey Katz、Eric Kramer、Matthew McCullough、Matthew Reece以及Jakub Scholtz。还要感谢Paul Davies的建议，把本书中导引的观点联系起来，并感谢Matthew和我一起着手做这项工作。非常感谢Matthew和Lubos Motl审阅这本书的初稿，感谢他们的意见和鼓励（尽管Lubos在某些争议问题上的观点还有待商榷）。

我还要感谢物理学和天文学的同事帮我校正书中不同章节：Laura Baudis、James Bullock、Bogdan Dobrescu、Doug Finkbeiner、Richard Gaitskell、Jakub Scholtz以及Tim Tait。Adam Brown对即将定稿的稿件做了检查，这项工作是非常重要的。科学内容上的意见来自Jo Bovi、Matthew Buckley、Sean Carroll、Chris Flynn、Lars Bergstrum、Ken Farley、Lars Hernquist、Johan Holmberg、Avi Loeb、Jonathan McDowell、Scott Tremaine、Matt Walker，他们的工作对本

书内容的进一步完善非常有帮助。同样提出深刻见解的不少天文学家也对本书贡献颇多，特别是，Francesca DeMeo、Dmitar Sasselov、Maria zuber。特别感谢Martin Elvis和Chris Flynn非常慷慨地付出了很多时间和精力。Jerry Coyne、Nathan Mhyrvhold，特别是Walter Alvarez、Andy Knoll以及David Kring提供了关于白垩纪-第三纪灭绝的颇有价值的看法，他们的建议以及所指出的错误，对我来说非常珍贵。同样非常感谢Jose Juan Blanco、Asier Hilario、Miren Mendea以及Jon Urrestilla帮助我安排参观位于西班牙的白垩纪-第三纪界线。

真正懂得科学是一回事，写书又是另外一回事。除了从同事那里获得的深刻见解和不同想法外，我非常幸运地得到了许多朋友的支持。感谢Andi Machl的付出、支持，以及他慷慨的评估。Cormac McCarthy的认真推敲带来的高标准（还有他表达反对的沉默）让本书内容更准确。朋友们的知识、建议和鼓励对我修正本书的观点和措辞有很大帮助，他们是：Judith Donough、Maya Jasanoff和Jen Sacks。David Lewis对语言的把握，以及编辑Anna Christina Buchmann的建议，极大地帮助我改正书中的不当之处。Jim Brooks、Richard Engel、Timothy Ferris、Milo Goodell、Tom Levenson、Howard Lutnick、Dana Randall和Michael Snediker都对本书有所贡献，感谢他们。

特别感谢我的编辑Hilary Redmon，感谢她在出版过程中的建议、鼓励和耐心；感谢她的助理Emma Janaskie帮助把那些零碎的意见整合起来。英国兰登书屋的Stuart Williams也提出了很多有价值的建议。我还要感谢Dan Halpern和Ecco出版社的工作人员，感谢Alison Saltzberg与我构思封面。天才Rose Lincoln帮助制作了肖像照，Gary Piko vsky提供了新的插画，Elisabeth Cheries、Robin Green、Emma Janaskie、Eric Kaplan、David Kring、Emily Lakdawalla、Tommy McCall以及Bill Prady帮我处理了一些插图。感谢Kathleen Rocheleau对参

考文献的处理，感谢Elisabeth Cheries对其进行了校对。感谢Yaddo以及住在那里的人们为我提供了欢乐、高效的环境，感谢Marty和Sarah Flug在重要时刻对我的热情招待，感谢哈佛大学提供高效的物理学研究环境。非常感谢Andrew Wylie帮助启动这个项目，感谢Andrew Wylie和Sarah Chalfant对这本书初稿的鼓励。还要感谢Wylie代理团队的其他人，包括James Pullen、Kristina Moore等人所做的诸多事情。

我还欠Jeff Goodell一个特别的感谢。他多次和我分享他作为一位高超作家的广阔视野，告诉我如何讲一个好故事、如何更好地表达它们。我还非常高兴地与他分享了我的好奇心。感谢Jeff Goodell的家人、我的家人，感谢我的朋友，感谢他们珍贵的好奇心以及他们对我著作的兴趣和鼓励。

最后，我要感谢我的父母。虽然很遗憾他们看不到这本书，但他们对我的影响贯穿了这本书。感谢他们教会我相信，我的目标是是可以实现的，包括我雄心勃勃的事业——这本书就是其中一项。

译者后记

相信不少读者童年时最感兴趣的话题都与宇宙和恐龙有关。“暗物质”与“恐龙”这两个词放到一起，即使对一些专业的科研人员来说，大概也会觉得十分奇怪。正如作者兰道尔教授所说，人们一般会想到的可能是即将上映的某部科幻大片。当然，这也是《星球大战》系列电影和《侏罗纪公园》系列电影在世界各国都非常流行的原因。

相比于大家没怎么听说过的暗物质，恐龙大概是大家非常熟知，并颇受小朋友喜爱的一种动物了。然而，这个曾经一度统治地球的“大爬虫”为什么在6 600万年前突然就灭绝了？这个困扰了科学家多年的问题似乎能被位于由恒星构成的银盘中的暗物质盘所解释。

其实，暗物质作为宇宙的主要构成成分之一，一直以来都是物理学家最感兴趣的研究对象之一。可惜迄今为止，我们对于暗物质性质的了解还非常少。天文观测表明，太阳系在围绕银河系旋转的同时，还在银盘中上上下下地振动。科学家发现，其振动周期和地球上大规模生物灭绝的周期有着某种关联。作者兰道尔教授及其合作者由此推测，太阳系每次经过银盘时，都会引起位于太阳系边缘处的彗星的轨道发生变化，其中一些脱轨的彗星会飞到地球附近，并因地球的强大引力而改变轨道，最终可能冲向地球。你可不要认为这些天体只是晴朗夜空中浪漫的流星，它们撞到地球上的冲击力相当于几十颗原子弹，其后引发的地震、海啸和酸雨等灾害，更是具有大规模的杀伤性——曾经的“地球霸主”恐龙就是在这一过程中灭绝的。

《暗物质与恐龙》一书不但展示了相关学科最前沿的进展，而且细致描述了这些学科中与暗物质和恐龙相关的课题的发展历史，以及一些不为人知的趣事。在开始翻译这本书之前，我们曾尝试着猜测暗物质与恐龙背后的联系：暗物质作为宇宙的主要构成成分之一，它对恐龙灭绝的最大贡献可能是其自身引力势对彗星轨道的改变，因此诱发了6 600万年前那次导致了恐龙灭绝的大撞击。随着翻译工作的深入，我们发现，事情比当初想象的要复杂得多。作者兰道尔教授分别从宇宙学、天体物理学、地质学、古生物学以及粒子物理学等多个学科，对此问题作出了详细且严谨的论述和分析，涉猎之广、之深都是我们翻译前所没有想到的，所以对于知识有限的我们还是有一定挑战的。

翻译过程虽然不是一帆风顺的，但是整体来说是被惊叹与快乐之情主导。惊叹的是兰道尔教授知识面之广，她从西班牙海滩的化石聊到北美洲矿井下的暗物质直接探测器；从奥尔特云附近的周期性彗星现象聊到当代人类活动对物种多样化的影响。我们在每一个章节几乎都能学到新的东西。快乐的是兰道尔教授用诙谐的笔法把那些看上去晦涩难懂的专业名词描述得直观且活泼，甚至有时，我们会为她睿智的类比而拍案叫绝。我们在翻译时，力争和原书的风格保持一致，即趣味性和科学性相结合。希望读者在享受阅读的同时，也能学习新知识来充实自己。

此外，我们想强调的是，兰道尔教授在书中提出的暗物质模型也只是一种理论猜想，其正确性还有待实验观测来验证。然而，自然界中看似毫不相关的事物之间却有着千丝万缕的联系，这让我们在看待这个世界时，永远充满好奇之心与惊讶之情。更重要的一点是，正如兰道尔教授所指出的，科技进步虽然给人类社会带来了巨大的发展、为人类提供了巨大的便利，但是其后续影响是否真能有利于人类的发展，还难以事先准确估计。所以，我们在对待新的科学技术时，应该

持有十分谨慎的态度和真正科学的观点。兰道尔教授提到的第六次生物大灭绝已经在很多地方被科学家们指出，这次大灭绝完全是由于人类活动造成的。不同于彗星那一次相当于几十颗原子弹的能量，人类日常生活的种种小事同样拥有如同原子弹一样的能量，并能逐渐破坏地球环境。为了我们唯一的生存家园，我们也希望大家能更多地关注科学的发展，并为保护我们的地球环境尽一份力。

对于我们来说，翻译此书不仅是文字、语言上的简单转换，更是学习、思考与理解的一次自我升华之旅。真心希望我们的这个翻译作品能帮助读者领略宇宙的奇妙，感受科学的魅力。由于知识、人力以及时间等限制，书中的错误和疏漏之处在所难免，还望读者朋友们不吝指正。

最后，感谢湛庐文化，将如此优秀的科普图书引入了中国；感谢湛庐文化的编辑和其他几位老师，在翻译过程中，她们给了许多帮助和建议。

以此来纪念我们共同翻译的时间。是为后记。

译者简介

尔欣中：

男，德国波恩大学天体物理博士，曾为中国科学院国家天文台副研究员（2011—2014），现为意大利罗马天文台博士后。研究方向：宇宙中的各种引力透镜效应。

李楠：

男，国家天文台天体物理学博士，现为美国芝加哥大学Kavli天体物理研究所博士后。研究方向：引力透镜数值模拟，暗物质空间分布。

王岚：

女，北京大学天体物理学博士，现为中国科学院国家天文台副研究员。研究方向：星系的形成和演化，以及温暗物质宇宙学模拟。

郑征：

男，美国约翰·霍普金斯大学博士，现为中国科学院国家天文台博士后。研究方向：近邻星系的结构，星系形成和演化。

谢利智：

女，国家天文台天体物理学博士，现为意大利的里雅斯特天文台博士后。研究方向：暗物质晕及星系的形成与演化。

苟利军：

男，美国宾夕法尼亚州立大学天体物理博士，美国哈佛大学哈佛-史密森天体物理中心博士后及研究人员，国家青年千人计划入选者。现为中国科学院国家天文台研究员，中国科学院大学教授，国家天文台恒星级黑洞研究创新小组负责人。研究方向：黑洞基本性质及爆发现象。

参考文献

第01章 ~02章

Bergström, Lars. “Non-Baryonic Dark Matter: Observational Evidence and Detection Methods.” *Reports on Progress in Physics* 63.5 (2000) : 793-841.

Bertone, Gianfranco, Dan Hooper, and Joseph Silk. “Particle Dark Matter: Evidence, Candidates and Constraints.” *Physics Reports* 405. 5-6 (2005) : 279-390.

Copi, C J, D N Schramm, and M S Turner. “Big-Bang Nucleosynthesis and the Baryon Density of the Universe.” *Science* 267. 5195 (1995) : 192-9.

Freese, Katherine. *The Cosmic Cocktail: Three Parts Dark Matter*. Princeton University Press, 2014.

Garrett, Katherine, and Gintaras Duda. “Dark Matter: A Primer.” *Advances in Astronomy* 2011 (2011) : 1-22.

Gelmini, Graciela B. *TASI 2014 Lectures: The Hunt for Dark Matter*. (2015) . <http://arxiv.org/abs/1502.01320>.

Lundmark, Knut. *Lund Medd.* 1 No125=VJS 65, p.275 (1930) .

Olive, Keith A. "TASI lectures on dark matter." *arXiv preprint astro-ph/0301505* (2003) .

Panek, Richard. *The 4 Percent Universe: Dark Matter, Dark Energy, and the Race to Discover the Rest of Reality*: Mariner Books, 2011.

Peter, Annika HG. "Dark Matter: A Brief Review." *Frank N. Bash Symposium 2011: New Horizons in Astronomy*. Ed. Sarah Salviander, Joel Green, and Andreas Pawlik. University of Texas at Austin, 2012.

Profumo, Stefano. "TASI 2012 Lectures on Astrophysical Probes of Dark Matter." (2013) : 41.

Rubin, V. C., N. Thonnard, and Jr. Ford, W. K. "Rotational Properties of 21SC Galaxies with a Large Range of Luminosities and Radii, from NGC 4605/R=4kpc/to UGC 2885/R=122 Kpc/." *The Astrophysical Journal*238 (1980) : 471-487.

Rubin, Vera C., and Jr. Ford, W. Kent. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions." *The Astrophysical Journal*159 (1970) : 379-403.

Sahni, Varun. "Dark Matter and Dark Energy." *Physics of the Early Universe*. Springer Berlin Heidelberg, 2005. 141-179.

Strigari, Louis E. "Galactic Searches for Dark Matter." *Physics Reports*531. 1 (2013) : 1-88.

Trimble, V. “Existence and Nature of Dark Matter in the Universe.” *Annual Review of Astronomy and Astrophysics* 25 (1987) : 425–472.

Zwicky, F. “Die Rotverschiebung von Extragalaktischen Nebeln.” *Helvetica Physica Acta* 6 (1933) : 110–127.

Zwicky, F. “On the Masses of Nebulae and of Clusters of Nebulae.” *The Astrophysical Journal* 86 (1937) : 217.

第03章

Humboldt, Alexander von. *Kosmos: A General Survey of Physical Phenomena of the Universe*, Volume 1. H. Baillière, 1845.

第04章

Baumann, Daniel. “TASI Lectures on Inflation.” (2009). <http://arxiv.org/abs/0907.5424>.

Boggs, N. W. et al. “The COBE Mission—Its Design and Performance Two Years after Launch.” *The Astrophysical Journal* 397 (1992) : 420–429.

Freeman, Ken, and Geoff McNamara. *In Search of Dark Matter*. Springer, 2006.

Guth, Alan H. *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*. Perseus Books, 1997.

Hinshaw, G. et al. "Five-Year Wilkinson Microwave Anisotropy Probe Observations: Data Processing, Sky Maps, and Basic Results." *The Astrophysical Journal Supplement Series* 180.2 (2009) : 225–245.

Kamionkowski, Marc, Arthur Kosowsky, and Albert Stebbins. "A Probe of Primordial Gravity Waves and Vorticity." *Physical Review Letters* 78 (1997) : 2058–2061.

Komatsu, E. et al. "Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation." *The Astrophysical Journal Supplement Series* 180.2 (2009) : 330–376.

Kowalski, M. et al. "Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets." *The Astrophysical Journal* 686.2 (2008) : 749–778.

Leitch, E. M. et al. "Degree Angular Scale Interferometer 3 Year Cosmic Microwave Background Polarization Results." *The Astrophysical Journal* 624.1 (2005) : 10–20.

Penzias, A. A., and R. W. Wilson. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." *The Astrophysical Journal* 142 (1965) : 419–421.

Seljak, Uros̃, and Matias Zaldarriaga. “Signature of Gravity Waves in the Polarization of the Microwave Background.” *Physical Review Letters* 78.11 (1997) : 2054–2057.

Weinberg, Steven. *The First Three Minutes: A Modern View of the Origin of the Universe*. Basic Books, 1993.

第05章

Binney, J., and S. Tremaine. *Galactic Dynamics*. Princeton University Press, 2008.

Davis, M. et al. “The Evolution of Large-Scale Structure in a Universe Dominated by Cold Dark Matter.” *The Astrophysical Journal* 292 (1985) : 371–394.

“Hubble Maps the Cosmic Web of ‘Clumpy’ Dark Matter in 3-D.” (7 January 2007) . <http://hubblesite.org/newscenter/archive/releases/2007/01/image/a/grav/>.

Kaehler, R., O. Hahn, and T. Abel. “A Novel Approach to Visualizing Dark Matter Simulations.” *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012) : 2078–2087.

Loeb, Abraham. *How Did the First Stars and Galaxies Form?* Princeton University Press, 2010.

Loeb, Abraham, and Steven R. Furlanetto. *The First Galaxies in the Universe*. Princeton University Press, 2013.

Massey, Richard et al. “Dark Matter Maps Reveal Cosmic Scaffolding.” *Nature* 445. 7125 (2007) : 286–90.

Mo, Houjun, Frank van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010.

Papastergis, Emmanouil et al. “A Direct Measurement of the Baryonic Mass Function of Galaxies&Implications for the Galactic Baryon Fraction.” *Astrophysical Journal* 259. 2 (2012) : 138.

Springel, Volker et al. “Simulations of the Formation, Evolution and Clustering of Galaxies and Quasars.” *Nature* 435. 7042 (2005) : 629–36.

第06章

Blitzer, Jonathan. “The Age of Asteroids.” *New Yorker*. (2014) . <http://www.newyorker.com/tech/elements/age-asteroids>.

DeMeo, F E, and B Carry. “Solar System Evolution from Compositional Mapping of the Asteroid Belt.” *Nature* 505 (2014) : 629–34.

Kleine, Thorsten et al. “Hf–W Chronology of the Accretion and Early Evolution of Asteroids and Terrestrial Planets.” *Geochimica et Cosmochimica Acta* 73. 17 (2009) : 5150–5188.

Lissauer, Jack J., and Imke de Pater. *Fundamental Planetary Science: Physics, Chemistry and Habitability*. Cambridge University Press, 2013.

Rubin, Alan E., and Jeffrey N. Grossman. “Meteorite and Meteoroid: New Comprehensive Definitions.” *Meteoritics and Planetary Science* (2010) : 114–122.

第07章

Bailey, M. E., and C.R. Stagg. “Cratering Constraints on the Inner Oort Cloud: Steady-State Models.” *Monthly Notices of the Royal Astronomical Society* 235.1 (1988) : 1–32.

“Europe’s Comet Chaser.” *European Space Agency*. (2014) .
http://www.esa.int/Our_Activities/Space_Science/Rosetta/Europe_s_comet_chaser.

Gladman, B. “The Kuiper Belt and the Solar System’s Comet Disk.” *Science* 307. 5706 (2005) : 71–75.

Gladman, B., B. G. Marsden, and C. Vanlaerhoven. “Nomenclature in the Outer Solar System.” *The Solar System Beyond Neptune*

ne. Ed. M. A. Barucci et al. University of Arizona Press, 2008. 43–57.

Gomes, Rodney. “Planetary Science: Conveyed to the Kuiper Belt.” *Nature* 426. 6965 (2003) : 393–5.

Iorio, L. “Perspectives on Effectively Constraining the Location of a Massive Trans-Plutonian Object with the New Horizons Spacecraft: A Sensitivity Analysis.” *Celestial Mechanics and Dynamical Astronomy* 116. 4 (2013) : 357–366.

Morbidelli, A, and H F Levison. “Planetary Science: Kuiper-Belt Interlopers.” *Nature* 422. 6927 (2003) : 30–1.

Olson, RJM. “Much Ado about Giotto’s Comet.” *Quarterly Journal of the Royal Astronomical Society* 35. 1 (1994) : 145.

Robinson, Howard. *The Great Comet of 1680: A Study in the History of Rationalism*. Press of the Northfeld News, 1916.

Walsh, Colleen. “The Building Blocks of Planets.” *Harvard Gazette* 12 Sept. 2013.

第08章

Francis, Matthew. “The Solar System Boundary and the Week in Review (September 8–14) .” *Bowler Hat Science*. (2013) . ht

[tp: //bowler hatscience.org/2013/09/14/the-solar-system-boundary-and-the-week-in-review-september-8-14/](http://bowlerhatscience.org/2013/09/14/the-solar-system-boundary-and-the-week-in-review-september-8-14/).

McComas, David. “What Is the Edge of the Solar System Like?—NOVA Next|PBS.” (2013) . [http: //www.pbs.org/wgbh/nova/next/space/voyager-ibex-and-the-edge-of-the-solar-system/](http://www.pbs.org/wgbh/nova/next/space/voyager-ibex-and-the-edge-of-the-solar-system/).

第09章

Gehrels, T., ed. *Hazards Due to Comets and Asteroids*. University of Arizona Press, 1995.

The Earth Institute, “The Growing Urbanization of the World, ” Columbia University, New York, 2005.

“IAU Minor Planet Center.” (2015) . [http: //www.minorplanetcenter.net/](http://www.minorplanetcenter.net/).

Kring, David A., and Mark Boslough. “Chelyabinsk: Portrait of an Asteroid Airburst.” *Physics Today*67. 9 (2014) : 32-37.

Levison, Harold F et al. “The Mass Disruption of Oort Cloud Comets.” *Science*296. 5576 (2002) : 2212-5.

Marvin, U. B. “Siena, 1794: History’s Most Consequential Meteorite Fall.” *Meteoritics*30.5 (1995) : 540.

Marvin, Ursula B. “Ernst Florens Friedrich Chladni (1756–1827) and the Origins of Modern Meteorite Research.” *Meteoritics&Planetary Science*31. 5 (1996) : 545–588.

Marvin, Ursula B. “Meteorites in History: An Overview from the Renaissance to the 20th Century.” *Geological Society, London, Special Publications*256. 1 (2006) : 15–71.

Marvin, Ursula B. , and Mario L. Cosmo. “Domenico Troili (1766) : ‘The True Cause of the Fall of a Stone in Albereto Is a Subterranean Explosion That Hurlled the Stone Skyward.’” *Meteoritics&Planetary Science*37.12 (2002) : 1857–1864.

“Meteorites, Impacts, &Mass Extinction.” (2014) . [http: //www.tulane.edu/~sanelson/Natural_Disasters/impacts.htm](http://www.tulane.edu/~sanelson/Natural_Disasters/impacts.htm).

National Research Council. *Defending Planet Earth: Near-Earth Object Surveys and Hazard Mitigation Strategies*.National Academies Press, 2010.

Nield, Ted. *Incoming! Or, Why We Should Stop Worrying and Learn to Love the Meteorite*.Granta Books, 2012.

Shapiro, Irwin I. et al. , with National Research Council. “Defending Planet Earth: Near-Earth Object Surveys and Hazard Mitigation Strategies.” 2010: 149.

Tagliaferri, E. et al. “Analysis of the Marshall Islands Fireball of February 1, 1994.” *Earth, Moon, and Planets*68.1–3 (1995) : 563–572.

第10章

“Astronomy: Collision History Written in Rock.” *Nature* 512. 7515 (2014) : 350.

Barringer, D. M. “Coon Mountain and Its Crater.” *Proceedings of the Academy of Natural Sciences of Philadelphia*, Vol. 57. 1905. 861–886.

“Earth Impact Database.” [http://www. passc.net/EarthImpactDatabase/](http://www.passc.net/EarthImpactDatabase/).

Grieve, R A F. “Terrestrial Impact Structures.” *Annual Review of Earth and Planetary Sciences* 15 (1987) : 245–270.

Kring, David A. *Guidebook to the Geology of Barringer Meteorite Crater, Arizona*. Lunar and Planetary Institute, 2007.

Tilghman, B. C. “Coon Butte, Arizona.” *Proceedings of the Academy of Natural Sciences of Philadelphia*, v. 57 (1905) .887–914.

第11章

Bambach, Richard K. “Phanerozoic Biodiversity Mass Extinctions.” *Annual Review of Earth and Planetary Sciences* 34. 1 (2006) : 127–155.

Bambach, Richard K. , Andrew H. Knoll, and Steve C. Wang. “Origination, Extinction, and Mass Depletions of Marine Diversity.” *Paleobiology*30.4 (2004) : 522–542.

Barnosky, Anthony D. *Dodging Extinction: Power, Food, Money, and the Future of Life on Earth.* University of California Press, 2014.

Barnosky, Anthony D et al. “Has the Earth’s Sixth Mass Extinction Already Arrived?” *Nature*471 (2011) : 51–57.

Carpenter, Kenneth. *Eggs, Nests, and Baby Dinosaurs: A Look at Dinosaur Reproduction.* Indiana University Press, 1999.

Eldredge, Niles. *Reinventing Darwin: The Great Debate at the High Table of Evolutionary Theory.* Wiley, 1995.

Jablonski, David. “Mass Extinctions and Macroevolution.” *Paleobiology*31. sp5 (2005) : 192–210.

Kelley, S. “The Geochronology of Large Igneous Provinces, Terrestrial Impact Craters, and Their Relationship to Mass Extinctions on Earth.” *Journal of ; the Geological Society*164. 5 (2007) : 923–936.

Kidwell, Susan M. “Shell Composition Has No Net Impact on Large-Scale Evolutionary Patterns in Mollusks.” *Science*307 (2005) : 914–917.

Kolbert, Elizabeth. *The Sixth Extinction: An Unnatural History.* Henry Holt&Co. , 2014.

Kurtén, Björn. *Age Groups in Fossil Mammals*. Helsinki: Societas scientiarum Fennica, 1953.

Lawton, John H., and Robert May, eds. *Extinction Rates*. Oxford University Press, 1995.

MacLeod, Norman. *The Great Extinctions: What Causes Them and How They Shape Life*. Firefy Books, 2013.

“Modern Extinction Estimates.” (2015). http://rainforests.mongabay.com/09x_table.htm.

Newell, Norman D. “Revolutions in the History of Life.” *Geological Society of America Special Papers* 89. Geological Society of America, 1967. 63–92.

Pimm, S. L. et al. “The Biodiversity of Species and Their Rates of Extinction, Distribution, and Protection.” *Science* 344 (2014) : 1246752.

Rothman, Daniel H et al. “Methanogenic Burst in the End-Permian Carbon Cycle.” *Proceedings of the National Academy of Sciences of the United States of America* 111. 15 (2014) : 5462–7.

Sanders, Robert. “Has the Earth’s Sixth Mass Extinction Already Arrived?” *UC Berkeley NewsCenter* 2 Mar. 2011.

Schindel, David E. “Microstratigraphic Sampling and the Limits of Paleontologic Resolution.” *Paleobiology* 6. 4 (1980) : 408–426.

Sepkoski, J. J. “Phanerozoic Overview of Mass Extinction.” *Patterns and Processes in the History of Life*. Ed. D. M. Raup and D. Jablonski. Springer Berlin Heidelberg, 1986. 277–295.

Valentine, James W. “How Good Was the Fossil Record? Clues from the California Pleistocene.” *Paleobiology* 15. 2 (1989) : 83–94.

Wilson, Edward O. *The Future of Life*. 1st ed. Vintage Books, 2003.

第12章

Alvarez, Walter. *T. Rex and the Crater of Doom*. Princeton University Press, 2008.

Alvarez, L. W. et al. “Extraterrestrial Cause for the Cretaceous–Tertiary Extinction.” *Science* 208. 4448 (1980) : 1095–108.

Caldwell, Brady. “The K–T Event: A Terrestrial or Extraterrestrial Cause for Dinosaur Extinction?” *Essay in Palaeontology* 5p (2007) .

Choi, Charles Q. “Asteroid Impact That Killed the Dinosaurs: New Evidence.” (2013) . <http://www.livescience.com/26933-chicxulub-cosmic-impact-dinosaurs.html>.

Frankel, Charles. *The End of the Dinosaurs: Chicxulub Crater and Mass Extinctions*. Cambridge University Press, 1999.

Kring, David A. et al. "Impact Lithologies and Their Emplacement in the Chicxulub Impact Crater: Initial Results from the Chicxulub Scientific Drilling Project, Yaxcopoil, Mexico." *Meteoritics & Planetary Science* 39.6 (2004) : 879-897.

Kring, David A. et al. "The Chicxulub Impact Event and Its Environmental Consequences at the Cretaceous-Tertiary Boundary." *Palaeogeography, Palaeoclimatology, Palaeoecology* 255.1-2 (2007) : 4-21.

Moore, J. R., and M. Sharma. "The K-Pg Impactor Was Likely a High Velocity Comet." *44th Lunar and Planetary Conference; Paper#2431*. 2013.

Ravizza, G, and B Peucker-Ehrenbrink. "Chemostratigraphic Evidence of Deccan Volcanism from the Marine Osmium Isotope Record." *Science* 302. 5649 (2003) : 1392-5.

Sanders, Robert. "New Evidence Comet or Asteroid Impact Was Last Straw for Dinosaurs." UC Berkeley News Center 7 Feb. 2013.

Schulte, Peter et al. "The Chicxulub Asteroid Impact and Mass Extinction at the Cretaceous-Paleogene Boundary." *Science* 327. 5970 (2010) : 1214-8.

"What Killed the Dinosaurs? The Great Mystery-Background." <http://www.ucmp.berkeley.edu/diapsids/extinction.html>.

“What Killed the Dinosaurs?The Great Mystery–Invalid Hypotheses [http://www. ucmp.berkeley.edu/diapsids/extincthypo.html](http://www.ucmp.berkeley.edu/diapsids/extincthypo.html).

Zahnle, K, and D Grinspoon. “Comet Dust as a Source of Amino Acids at the Cretaceous/Tertiary Boundary.” *Nature*348. 6297 (1990) : 157–60.

第13章

American Chemical Society. “New Evidence That Comets Deposited Building Blocks of Life on Primordial Earth.” *Science Daily* (2012) : 27 March. www.sciencedaily.com/releases/2012/03/120327215607.htm.

“Astronomy: Comets Forge Organic Molecules.” *Nature*512. 7514 (2014) : 234–235.

Durand-Manterola, Hector Javier, and Guadalupe Cordero-Tercero. “Assessments of the Energy, Mass and Size of the Chicxulub Impactor.” (2014) . arXiv: 1403: 6391.

Elvis, Martin. “Astronomy: Cosmic Triangles and Black-Hole Masses.” *Nature*515. 7528 (2014) : 498–499.

Elvis, Martin. “How Many Ore-Bearing Asteroids?” *Planetary and Space Science*91 (2014) : 20–26.

Elvis, Martin, and Thomas Esty. “How Many Assay Probes to Find One Ore-Bearing Asteroid?” *Acta Astronautica*96 (2014) : 227–231.

Knoll, Andrew H. *Life on a Young Planet: The First Three Billion Years of Evolution on Earth*. Princeton University Press, 2003.

Livio, Mario, Neill Reid, and William Sparks, eds. *Astrophysics of Life: Proceedings of the Space Telescope Science Institute Symposium, Held in Baltimore, Maryland May 6–9, 2002*. Cambridge University Press, 2005.

Melott, Adrian L., and Brian C. Thomas. “Astrophysical Ionizing Radiation and Earth: A Brief Review and Census of Intermittent Intense Sources.” *Astrobiology*11.4 (2011) : 343–361.

Poladian, Charles. “Comets Or Meteorites Crashing Into A Planet Could Produce Amino Acids, ‘Building Blocks Of Life.’” *International Business Times*15 Sept. 2013.

Rothery, David A., Mark A. Sephton, and Iain Gilmour, Eds. *An Introduction to Astrobiology*. 2nd ed. Cambridge University Press, 2011.

Steigerwald, Bill. “Amino Acids in Meteorites Provide a Clue to How Life Turned Left.” 2012. <http://scitechdaily.com/amino-acids-in-meteorites-provide-a-clue-to-how-life-turned-left/>.

第14章

Alvarez, Walter, and Richard A. Muller. “Evidence from Crater Ages for Periodic Impacts on the Earth.” *Nature* 308 (1984) : 718–720.

Bailer-Jones, C. A.L. “Bayesian Time Series Analysis of Terrestrial Impact Cratering.” *Monthly Notices of the Royal Astronomical Society* 416.2 (2011) : 1163–1180.

Bailer-Jones, C. A.L. “Evidence for a Variation—but No Periodicity—in the Terrestrial Impact Cratering Rate.” *EPSC–DPS Joint Meeting 2011* (2011) : 153.

Bailer-Jones, C. A.L. “The Evidence for and against Astronomical Impacts on Climate Change and Mass Extinctions: A Review.” *International Journal of Astrobiology* 8.3 (2009) : 213.

Chang, Heon-Young, and Hong-Kyu Moon. “Time-Series Analysis of Terrestrial Impact Crater Records.” *Publications of the Astronomical Society of Japan* 57. 3 (2005) : 487–495.

Connor, E. F. “Time Series Analysis of the Fossil Record.” *Patterns and Processes in the History of Life*. Springer Berlin Heidelberg, 1986. 119–147.

Feulner, Georg. “Limits to Biodiversity Cycles from a Unified Model of Mass-Extinction Events.” *International Journal of Astrobiology* 10. 02 (2011) : 123–129.

Fox, William T. "Harmonic Analysis of Periodic Extinctions." *Paleobiology* 13. 3 (1987) : 257-271.

Grieve, R. A.F. et al. "Detecting a Periodic Signal in the Terrestrial Cratering Record." *Lunar and Planetary Science Conference* (1988) : 375-382.

Grieve, R. A.F., and D.A.Kring. "Geologic Record of Destructive Impact Events on Earth." *Comet/Asteroid Impacts and Human Society: An Interdisciplinary Approach*. Ed. Peter T. Bobrowsky and Hans Rickman. Springer Berlin Heidelberg, 2007. 3-24.

Grieve, Richard A. F. "Terrestrial Impact: The Record in the Rocks*." *Meteoritics* 26.3 (1991) : 175-194.

Grieve, Richard A. F., and Eugene M. Shoemaker. "The Record of Past Impacts on Earth." *Hazards Due to Comets and Asteroids*. Ed. T. Gehrels. University of Arizona Press, 1994. 417-462.

Heisler, Julia, and Scott Tremaine. "How Dating Uncertainties Affect the Detection of Periodicity in Extinctions and Craters." *Icarus* 77. 1 (1989) : 213-219.

Heisler, Julia, Scott Tremaine, and Charles Alcock. "The Frequency and Intensity of Comet Showers from the Oort Cloud." *Icarus* 70. 2 (1987) : 269-288.

Jetsu, L., and J. Pelt. "Spurious Periods in the Terrestrial Impact Crater Record." *Astronomy and Astrophysics* 353 (2000) : 409-418.

Lieberman, Bruce S. “Whilst This Planet Has Gone Cycling On: What Role for Periodic Astronomical Phenomena in Large-Scale Patterns in the History of Life?” *Earth and Life: Global Biodiversity, Extinction Intervals and Biogeographic Perturbations Through Time*. Springer Netherlands, 2012. 37–50.

Lyytinen, J. et al. “Detection of Real Periodicity in the Terrestrial Impact Crater Record: Quantity and Quality Requirements.” *Astronomy and Astrophysics* 499.2 (2009) : 601–613.

Melott, Adrian L. et al. “A ~60 Myr Periodicity Is Common to Marine-⁸⁷Sr/⁸⁶Sr, Fossil Biodiversity, and Large-Scale Sedimentation: What Does the Periodicity Reflect?” *Journal of Geology* 120 (2012) : 217–226.

Melott, Adrian L., and Richard K. Bambach. “A Ubiquitous ~62-Myr Periodic Fluctuation Superimposed on General Trends in Fossil Biodiversity. ; I. Documentation.” *Paleobiology* 37.1 (2011) : 92–112.

Melott, Adrian L., and Richard K. Bambach. “Analysis of Periodicity of Extinction Using the 2012 Geological Timescale.” *Paleobiology* 40.2 (2014) : 177–196.

Melott, Adrian L., and Richard K. Bambach. “Do Periodicities in Extinction—with Possible Astronomical Connections—Survive a Revision of the Geological Timescale?” *The Astrophysical Journal* 773.1 (2013) : 1–5.

Noma, Elliot, and Arnold L. Glass. "Mass Extinction Pattern: Result of Chance." *Geological Magazine* 124.4 (1987) : 319-322.

Quinn, James F. "On the Statistical Detection of Cycles in Extinctions in the Marine Fossil Record." *Paleobiology* 13. 4 (1987) : 465-478.

Raup, D. M., and J J Sepkoski. "Mass Extinctions in the Marine Fossil Record." *Science* 215.4539 (1982) : 1501-3.

Raup, D. M., and J.J. Sepkoski. "Periodicity of Extinctions in the Geologic Past." *Proceedings of the National Academy of Sciences* 81.3 (1984) : 801-805.

Raup, D., and J. Sepkoski. "Periodic Extinction of Families and Genera." *Science* 231.4740 (1986) : 833-836.

Stigler, S M, and M J Wagner. "A Substantial Bias in Nonparametric Tests for Periodicity in Geophysical Data." *Science* 238. 4829 (1987) : 940-5.

Stothers, Richard B. "Structure and Dating Errors in the Geologic Time Scale and Periodicity in Mass Extinctions." *Geophysical Research Letters* 16. 2 (1989) : 119-122.

Stothers, Richard B. "The Period Dichotomy in Terrestrial Impact Crater Ages." *Monthly Notices of the Royal Astronomical Society* 365. 1 (2006) : 178-180.

Trefil, J. S., and D.M. Raup. "Numerical Simulations and the Problem of Periodicity in the Cratering Record." *Earth and Planetary Science Letters* 82.1-2 (1987) : 159-164.

Yabushita, S. "A Statistical Test of Correlations and Periodicities in the Geological Records." *Celestial Mechanics and Dynamical Astronomy* 69. 1-231-48.

Yabushita, S. "Are Cratering and Probably Related Geological Records Periodic?" *Earth, Moon and Planets* 72. 1-3 (1996) : 343-356.

Yabushita, S. "On the Periodicity Hypothesis of the Ages of Large Impact Craters." *Monthly Notices of the Royal Astronomical Society* 334. 2 (2002) : 369-373.

Yabushita, S. "Periodicity and Decay of Craters over the Past 600 Myr." *Earth, Moon and Planets* 58. 1 (1992) : 57-63.

Yabushita, S. "Statistical Tests of a Periodicity Hypothesis for Crater Formation Rate-II." *Monthly Notices of the Royal Astronomical Society* 279. 3 (1996) : 727-732.

第15章

Davis, Marc, Piet Hut, and Richard A. Muller. "Extinction of Species by Periodic Comet Showers." *Nature* 308.5961 (1987) : 529-531.

4) : 715-717.

Filipovic, M. D. et al. "Mass Extinction and the Structure of the Milky Way." *Serbian Astronomical Journal* 87 (2013) : 43-52.

Grieve, Richard A. F., and Lauri J. Pesonen. "Terrestrial Impact Craters: Their Spatial and Temporal Distribution and Impacting Bodies." *Earth, Moon and Planets* 72.1-3 (1996) : 357-376.

Heisler, Julia, Scott Tremaine, and Charles Alcock. "The Frequency and Intensity of Comet Showers from the Oort Cloud." *Icarus* 70. 2 (1987) : 269-288.

Matese, J. "Periodic Modulation of the Oort Cloud Comet Flux by the Adiabatically Changing Galactic Tide." *Icarus* 116. 2 (1995) : 255-268.

Matese, J. J., K. A. Innanen, and M. J. Valtonen. "Variable Oort Cloud Flux due to the Galactic Tide." *Collisional Processes in the Solar System*. Ed. Mikhail Marov and Hans Rickman. Kluwer Academic Publishers, 2001. 91-102.

Melott, Adrian L., and Richard K. Bambach. "Nemesis Reconsidered." *Monthly Notices of the Royal Astronomical Society: Letters* 407.1 (2010) : L99-L102.

Napier, W. M. "Evidence for Cometary Bombardment Episodes." *Monthly Notices of the Royal Astronomical Society* 366. 3 (2006) : 977-982.

Nurmi, P., M. J. Valtonen, and J. Q. Zheng. "Periodic Variation of Oort Cloud Flux and Cometary Impacts on the Earth and Jupiter." *Monthly Notices of the Royal Astronomical Society* 327.4 (2001) : 1367–1376.

Rampino, M. R. "Disc Dark Matter in the Galaxy and Potential Cycles of Extraterrestrial Impacts, Mass Extinctions and Geological Events." *Monthly Notices of the Royal Astronomical Society* 448.2 (2015) : 1816–1820.

Rampino, M. R. "Galactic Triggering of Periodic Comet Showers." *Collisional Processes in the Solar System*. Ed. Mikhail Ya Marov and Hans Rickman. Kluwer Academic Publishers, 2001. 103–120.

Rampino, Michael, Bruce M. Haggerty, and Thomas C. Pagan. "A Unified Theory of Impact Crises and Mass Extinctions: Quantitative Tests." *Annals of the New York Academy of Sciences* 822.1 (1997) : 403–431.

Rampino, Michael R., and Richard B. Stothers. "Terrestrial Mass Extinctions, Cometary Impacts and the Sun's Motion Perpendicular to the Galactic Plane." *Nature* 308 (1984) : 709–712.

Schwartz, Richard D., and Philip B. James. "Periodic Mass Extinctions and the Sun's Oscillation about the Galactic Plane." *Nature* 308.5961 (1984) : 712–713.

Shoemaker, Eugene M. “Impact Cratering Through Geologic Time.” *Journal of the Royal Astronomical Society of Canada* 92 (1998) : 297–309.

Smoluchowski, R., J. M. Bahcall, and M. S. Matthews. *Galaxy and the Solar System*. University of Arizona Press, 1986.

Stothers, R. B. “Galactic Disc Dark Matter, Terrestrial Impact Cratering and the Law of Large Numbers.” *Monthly Notices of the Royal Astronomical Society* 300.4 (1998) : 1098–1104.

Swindle, T. D., D. A. Kring, and J. R. Weirich. “ $^{40}\text{Ar}/^{39}\text{Ar}$ Ages of Impacts Involving Ordinary Chondrite Meteorites.” *Geological Society, London, Special Publications* 378.1 (2013) : 333–347.

Torbett, Michael V. “Injection of Oort Cloud Comets to the Inner Solar System by Galactic Tidal Fields.” *Monthly Notices of the Royal Astronomical Society* 223 (1986) : 885–895.

Wickramasinghe, J. T., and W. M. Napier. “Impact Cratering and the Oort Cloud.” *Monthly Notices of the Royal Astronomical Society* 387.1 (2008) : 153–157.

Whitmire, Daniel P., and Albert A. Jackson. “Are Periodic Mass Extinctions Driven by a Distant Solar Companion?” *Nature* 308.5961 (1984) : 713–715.

Whitmire, Daniel P., and John J. Matese. “Periodic Comet Showers and Planet X.” *Nature* 313.5997 (1985) : 36–38.

Wickramasinghe, J. T., and W.M.Napier. “Impact Cratering and the Oort Cloud.” *Monthly Notices of the Royal Astronomical Society* 387.1 (2008) : 153–157.

第16章～第17章

Ahmed, Z et al. “Dark Matter Search Results from the CDMS II Experiment.” *Science* 327. 5973 (2010) : 1619–21.

Akerib, D. S. et al. “First Results from the LUX Dark Matter Experiment at the Sanford Underground Research Facility.” *Physical Review Letters* 112.9 (2014) : 091303.

Aprile, E. et al. “First Dark Matter Results from the XENON100 Experiment.” *Physical Review Letters* 105.13 (2010) .

Bergstrom, Lars. “Saas-Fee Lecture Notes: Multi-Messenger Astronomy and Dark Matter.” (2012) : 105.

Bertone, Gianfranco. “The Moment of Truth for WIMP Dark Matter.” *Nature* 468. 7322 (2010) : 389–393.

Bertone, Gianfranco. *Particle Dark Matter: Observations, Models and Searches*. Cambridge University Press, 2010.

Bertone, Gianfranco, and David Merritt. “Dark Matter Dynamics and Indirect Detection.” *Modern Physics Letters A* 20. 14 (2005) : 1021–1036.

Buckley, Matthew R., and Lisa Randall. "Xogenesis." *Journal of High Energy Physics*9 (2011) .

Cline, David B. "The Search for Dark Matter." *Scientific American*288. 3 (2003) : 50–59.

Cohen, Timothy et al. "Asymmetric Dark Matter from a GeV Hidden Sector." *Physical Review D*82. 5 (2010) .

Cohen, Timothy, and Kathryn M. Zurek. "Leptophilic Dark Matter from the Lepton Asymmetry." *Physical Review Letters*104. 10 (2010) .

Cui, Yanou, Lisa Randall, and Brian Shuve. "A WIMPy Baryogenesis Miracle." *Journal of High Energy Physics*4 (2012) : 75.

Cui, Yanou, Lisa Randall, and Brian Shuve. "Emergent Dark Matter, Baryon, and Lepton Numbers." *Journal of High Energy Physics*2011. 8 (2011) : 73.

Davoudiasl, Hooman et al. "Unified Origin for Baryonic Visible Matter and Antibaryonic Dark Matter." *Physical Review Letters*105. 21 (2010) .

Drukier, Andrzej K., Katherine Freese, and David N. Spergel. "Detecting cold dark-matter candidates." *Physical Review D*33.12 (1986) : 3495.

Freeman, Ken, and Geoff McNamara. *In Search of Dark Matter*. Springer, 2006.

Gaitskell, Richard J. “Direct Detection of Dark Matter.” *Annual Review of Nuclear and Particle Science* 54. 1 (2004) : 315–359.

Hooper, Dan, John March–Russell, and Stephen M. West. “Asymmetric Sneutrino Dark Matter and the Omega (b) /Omega (DM) Puzzle.” *Physics Letters B* 605. 3–4 (2005) : 228–236.

Jungman, Gerard, Marc Kamionkowski, and Kim Griest. “Supersymmetric Dark Matter.” *Physics Reports* 267. 5–6 (1996) : 195–373.

Kaplan, David B. “Single Explanation for Both Baryon and Dark Matter Densities.” *Physical Review Letters* 68. 6 (1992) : 741–743.

Kaplan, David E., Markus A. Luty, and Kathryn M. Zurek. “Asymmetric Dark Matter.” *Physical Review D* 79. 11 (2009) .

Napier, W. M. “Evidence for Cometary Bombardment Episodes.” *Monthly ; Notices of the Royal Astronomical Society* 366. 3 (2006) : 977–982.

“Neutralino Dark Matter.” http://www.picassoexperiment.ca/dm_neutralino.php.

Preskill, John, Mark B. Wise, and Frank Wilczek. “Cosmology of the Invisible Axion.” *Physics Letters B* 120. 1–3 (1983) : 127–132.

Profumo, Stefano. “TASI 2012 Lectures on Astrophysical Probes of Dark Matter.” (2013) : 41.

Shelton, Jessie, and Kathryn M. Zurek. “Darkogenesis: A Baryon Asymmetry from the Dark Matter Sector.” *Physical Review D*82.12 (2010) : 123512.

Thomas, Scott. “Baryons and Dark Matter from the Late Decay of a Supersymmetric Condensate.” *Physics Letters B*356. 2-3 (1995) : 256-263.

Turner, Michael S., and Frank Wilczek. “Inflationary Axion Cosmology.” *Physical Review Letters*66. 1 (1991) : 5-8.

Weinberg, Steven. “A New Light Boson?” *Physical Review Letters*40. 4 (1978) : 223-226.

Wilczek, F. “Problem of Strong P and T Invariance in the Presence of Instantons.” *Physical Review Letters*40. 5 (1978) : 279-282.

第18章

Ackerman, Lotty et al. “Dark Matter and Dark Radiation.” *Physical Review D*79. 2 (2009) : 023519.

Bovy, Jo, Hans-Walter Rix, and David W. Hogg. “The Milky Way Has No Distinct Thick Disk.” *The Astrophysical Journal*75

1.2 (2012) : 131.

Buckley, Matthew R., and Patrick J. Fox. "Dark Matter Self-Interactions and Light Force Carriers." *Physical Review D* 81.8 (2010) .

De Blok, W. J G. "The Core-Cusp Problem." *Advances in Astronomy* (2010) .

Faber, S. M., and R.E. Jackson. "Velocity Dispersions and Mass-to-Light Ratios for Elliptical Galaxies." *The Astrophysical Journal* 204 (1976) : 668-683.

"First Signs of Self-Interacting Dark Matter?" ESO Press Release, European Southern Observatory. <http://www.eso.org/public/news/esol1514/>

Goldberg, Haim, and Lawrence J. Hall. "A New Candidate for Dark Matter." *Physics Letters B* 174.2 (1986) : 151-155.

Governato, F et al. "Bulgeless Dwarf Galaxies and Dark Matter Cores from Supernova-Driven Outflows." *Nature* 463. 7278 (2010) : 203-6.

Holmberg, Johan, and Chris Flynn. "The Local Surface Density of Disc Matter Mapped by Hipparcos." *Monthly Notices of the Royal Astronomical Society* 352. 2 (2004) : 440-446.

Kuijken, Konrad, and Gerard Gilmore. "The Galactic Disk Surface Mass Density and the Galactic Force $K(z)$ at $Z=1.1$ Kiloparsecs." *The Astrophysical Journal* 367 (1991) : L9-L13.

Langdale, Jonathan. “Could There Be a Larger Dark World with Dark Interactions? There Is More Dark Matter than Visible.” (2013). <https://plus.google.com/+JonathanLangdale/posts/Es7M9VhiFNp>.

Markevitch, M. et al. “Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657–56.” *The Astrophysical Journal* 606.2 (2004) : 819–824.

Moore, Ben et al. “Dark Matter Substructure within Galactic Halos.” *The Astrophysical Journal* 524. 1 (1999) : L19–L22.

Oort, J. H. “The Force Exerted by the Stellar System in the Direction Perpendicular to the Galactic Plane and Some Related Problems.” *Bulletin of the Astronomical Institutes of the Netherlands* 6 (1932) : 249–287.

Oort, J. H. “Note on the determination of K_z and on the mass density near the Sun.” *Bulletin of the Astronomical Institutes of the Netherlands* 15 (1960) : 45.

Read, J I. “The Local Dark Matter Density.” *Journal of Physics G: Nuclear and Particle Physics* 41. 6 (2014) : 063101.

Salucci, Paolo, and Annamaria Borriello. “The Intriguing Distribution of Dark Matter in Galaxies.” *Particle Physics in the New Millennium* 616 (2003) : 66–77.

Spergel, David N., and Paul J. Steinhardt. “Observational Evidence for Self-Interacting Cold Dark Matter.” *Physical Rev*

*iew Letters*84.17 (2000) : 3760–3763.

Weinberg, David H. et al. “Cold Dark Matter: Controversies on Small Scales.” *Proceedings of the National Academy of Sciences* (2015) : <http://arxiv.org/abs/1306.0913>.

Weniger, Christoph. “A Tentative Gamma-Ray Line from Dark Matter Annihilation at the Fermi Large Area Telescope.” *Journal of Cosmology and Astroparticle Physics*2012. 08 (2012) .

Zhang, Lan, et al. “The gravitational potential near the sun from SEGUE K-dwarf kinematics.” *The Astrophysical Journal* 772. 2 (2013) : 108.

第19章

Cline, James M., Zuowei Liu, and Wei Xue. “Millicharged Atomic Dark Matter.” *Physical Review D*85. 10 (2012) : 101302.

Cooper, A. P. et al. “Galactic Stellar Haloes in the CDM Model.” *Monthly Notices of the Royal Astronomical Society*406.2 (2010) : 744–766.

Dienes, Keith R., and Brooks Thomas. “Dynamical Dark Matter: A New Framework for Dark-Matter Physics.” *Workshop on Dark Matter, Unification and Neutrino Physics: CETUP*2012. Vol.1 534. AIP Publishing, 2013. 57–77.

Fan, JiJi et al. “Dark-Disk Universe.” *Physical Review Letters*110. 21 (2013) : 211302.

Fan, JiJi et al. “Double-Disk Dark Matter.” *Physics of the Dark Universe*2. 3 (2013) : 139–156.

Foot, R. “Mirror Dark Matter: Cosmology, Galaxy Structure and Direct Detection.” *International Journal of Modern Physics A*29. 11n12 (2014) : 1430013.

Foot, R., H. Lew, and R.R.Volkas. “A Model with Fundamental Improper Spacetime Symmetries.” *Physics Letters B*272.1–2 (1991) : 67–70.

Kaplan, David E et al. “Atomic Dark Matter.” *Journal of Cosmology and Astroparticle Physics*05 (2010) . 21.

Kaplan, David E et al. “Dark Atoms: Asymmetry and Direct Detection.” *Journal of Cosmology and Astroparticle Physics*10 (2011) : 19.

Pillepich, Annalisa et al. “The Distribution of Dark Matter in the Milky Way’s Disk.” *eprint arXiv: 1308. 1703* (2013) .

Powell, Corey S. “Inside the Hunt for Dark Matter.” *Popular Science*. (2013) . <http://www.popsci.com/article/science/inside-hunt-dark-matter>.

Powell, Corey S. “The Possible Parallel Universe of Dark Matter.” *Discover Magazine. com*. (2013) .<http://discovermagazine.com>

zine.com/2013/julyaug/21-the-possible-parallel-universe-of-dark-matter.

Purcell, Chris W., James S. Bullock, and Manoj Kaplinghat. “The Dark Disk of the Milky Way.” *The Astrophysical Journal* 703.2 (2009) : 2275–2284.

Read, J. I. et al. “Thin, Thick and Dark Discs in Λ CDM.” *Monthly Notices of the Royal Astronomical Society* 389.3 (2008) : 1041–1057.

Rosen, Len. “Is There Only One Type of Dark Matter?” (2013) . <http://www.21stcentech.com/type-dark-matter/>.

第20章

Bovy, Jo, and Hans-Walter Rix. “A Direct Dynamical Measurement of the Milky Way’s Disk Surface Density Profile, Disk Scale Length, and Dark Matter Profile at $4 \text{ Kpc} \leq R \leq 9 \text{ Kpc}$.” *The Astrophysical Journal* 779.2 (2013) : 1–30.

Bovy, Jo, and Scott Tremaine. “On the Local Dark Matter Density.” *The Astrophysical Journal* 756. 1 (2012) : 89.

Bruch, Tobias et al. “Dark Matter Disc Enhanced Neutrino Fluxes from the Sun and Earth.” *Physics Letters B* 674. 4–5 (2009) : 250–256.

Bruch, Tobias et al. “Detecting the Milky Way’s Dark Disk.” *The Astrophysical Journal* 696. 1 (2009) : 920–923.

Buckley, Matthew R. et al. “Scattering, Damping, and Acoustic Oscillations: Simulating the Structure of Dark Matter Halos with Relativistic Force Carriers.” *Physical Review D* 90.4 (2014) : 043524.

Cartlidge, Edwin. “Do Dark-Matter Discs Envelop Galaxies?” *PhysicsWorld.com*. (2013) .<http://physicsworld.com/cws/article/news/2013/jun/03/do-dark-matter-discs-envelop-galaxies>.

Cyr-Racine, Francis-Yan et al. “Constraints on Large-Scale Dark Acoustic Oscillations from Cosmology.” *Physical Review D* 89. 6 (2014) .

Cyr-Racine, Francis-Yan, and Kris Sigurdson. “Cosmology of Atomic Dark Matter.” *Physical Review D* 87. 10 (2013) .

Holmberg, J, and C Flynn. “The Local Density of Matter Mapped by Hipparcos.” *Monthly Notices of the Royal Astronomical Society* 313. 2 (2000) : 209–216.

Kuijken, K., and G. Gilmore. “The Mass Distribution in the Galactic Disc—II—Determination of the Surface Mass Density of the Galactic Disc Near the Sun.” *Monthly Notices of the Royal Astronomical Society* 239 (1989) : 605–649.

Kuijken, Konrad, and Gerard Gilmore. “The Mass Distribution in the Galactic Disc. I—A Technique to Determine the Integ

ral Surface Mass Density of the Disc near the Sun.” *Monthly Notices of the Royal Astronomical Society* 239 (1989) : 571–603.

March–Russell, John, Christopher McCabe, and Matthew McCullough. “Inelastic Dark Matter, Non–Standard Halos and the DAMA/LIBRA Results.” *Journal of High Energy Physics* 2009. 05 (2009) .

McCullough, Matthew, and Lisa Randall. “Exothermic Double–Disk Dark Matter.” *Journal of Cosmology and Astroparticle Physics* 2013. 10 (2013) : 58.

Motl, Luboš. “Exothermic Double–Disk Dark Matter.” (2013) . <http://motls.blogspot.com/2013/07/exothermic-double-disk-dark-matter.html>.

Nesti, Fabrizio, and Paolo Salucci. “The Dark Matter Halo of the Milky Way, AD 2013.” *Journal of Cosmology and Astroparticle Physics* 2013. 07 (2013) : 16.

Randall, Lisa, and Jakub Scholtz. “Dissipative Dark Matter and the Andromeda Plane of Satellites.” (2014) . <http://arxiv.org/abs/1412.1839>.

Rix, Hans–Walter, and Jo Bovy. “The Milky Way’s Stellar Disk.” *The Astronomy and Astrophysics Review* 21. 1 (2013) : 61.

第21章

Aron, Jacob. “Did Dark Matter Kill the Dinosaurs? Maybe……” *New Scientist*. (2014) . <http://www.newscientist.com/article/dn25177-did-dark-matter-kill-the-dinosaurs-maybe.html#.VVY1fv1VhBc>.

Choi, Charles Q. “Dark Matter Could Send Asteroids Crashing Into Earth: New Theory.” (2014) . <http://www.space.com/25657-dark-matter-asteroid-impacts-earth-theory.html>

Gibney, Elizabeth. “Did Dark Matter Kill the Dinosaurs?” *Nature*. (2014) . <http://www.nature.com/news/did-dark-matter-kill-the-dinosaurs-1.14839>.

Nagai, Daisuke. “Viewpoint: Dark Matter May Play Role in Extinctions.” *Physical Review Letters Physics*7 (2014) : 41.

Nair, Unni K. “Dinosaurs Extinction from Dark Matter?” (2014) . <http://guardianlv.com/2014/03/dinosaurs-extinction-from-dark-matter/>.

Piggott, Mark. “Were Dinosaurs Killed by Disc of Dark Matter?” (2014) . <http://www.ibtimes.co.uk/were-dinosaurs-killed-by-disc-dark-matter-1439500>.

Randall, Lisa, and Matthew Reece. “Dark Matter as a Trigger for Periodic Comet Impacts.” *Physical Review Letters*112. 16 (2014) : 161301.

Sharwood, Simon. “Dark Matter Killed the Dinosaurs, Boffns Suggest” *The Register*. 5 Mar. 2014. http://www.theregister.co.uk/2014/03/05/dark_matter_killed_the_dinosaurs_boffns_suggest/.

结 语

Bettencourt, Luís M A et al. “Growth, Innovation, Scaling, and the Pace of Life in Cities.” *Proceedings of the National Academy of Sciences of the United States of America* 104.17 (2007) : 7301–6.

Brynjolfsson, Erik, and Andrew McAfee. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W.Norton, 2014.

“Geoffrey West.” (2015) . <http://www.santafe.edu/about/people/profile/Geoffrey%20West>.

“On Care for Our Common Home.” Encyclical Letter *Laudato Si'* of the Holy Father Francis (2015) . http://w2.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html.

Weisman, Alan. *The World Without Us*. Reprint Edition. Picador, 2008.

West, Geoffrey. “Why Cities Keep Growing, Corporations and People Always Die, and Life Gets Faster.” *The Edge* (2011) .
<http://edge.org/conversation/geoffrey-west>.

湛庐，与思想有关……

如何阅读商业图书

商业图书与其他类型的图书，由于阅读目的和方式的不同，因此有其特定的阅读原则和阅读方法，先从一本书开始尝试，再熟练应用。

阅读原则1 二八原则

对商业图书来说，80%的精华价值可能仅占20%的页码。要根据自己的阅读能力，进行阅读时间的分配。

阅读原则2 集中优势精力原则

在一个特定的时间段内，集中突破20%的精华内容。也可以在一个时间段内，集中攻克一个主题的阅读。

阅读原则3 递进原则

高效率的阅读并不一定要按照页码顺序展开，可以挑选自己感兴趣的部分阅读，再从兴趣点扩展到其他部分。阅读商业图书切忌贪多，从一个小主题开始，先培养自己的阅读能力，了解文字风格、观点阐述以及案例描述的方法，目的在于对方法的掌握，这才是最重要的。

阅读原则4 好为人师原则

在朋友圈中主导、控制话题，引导话题向自己设计的方向去发展，可以让读书收获更加扎实、实用、有效。

阅读方法与阅读习惯的养成

- (1) 回想。阅读商业图书常常不会一口气读完，第二次拿起书时，至少用15分钟回想上次阅读的内容，不要翻看，实在想不起来再翻看。严格训练自己，一定要回想，坚持50次，会逐渐养成习惯。
- (2) 做笔记。不要试图让笔记具有很强的逻辑性和系统性，不需要有深刻的见解和思想，只要是文字，就是对大脑的锻炼。在空白处多写多画，随笔、符号、涂色、书签、便签、折页，甚至拆书都可以。
- (3) 读后感和PPT。坚持写读后感可以大幅度提高阅读能力，做PPT可以提高逻辑分析能力。从写读后感开始，写上5篇以后，再尝试做PPT。连续做上5个PPT，再重复写三次读后感。如此坚持，阅读能力将会大幅度提高。
- (4) 思想的超越。要养成上述阅读习惯，通常需要6个月的严格训练，至少完成4本书的阅读。你会慢慢发现，自己的思想开始跳脱出来，开始有了超越作者的感觉。比拟作者、超越作者、试图凌驾于作者之上思考问题，是阅读能力提高的必然结果。



扫码关注湛庐文化，
回复“阅读”
这5种方法，让读过的书变成你的影子

[特别感谢：营销及销售行为专家 孙路弘 智慧支持！]

老 我们出版的所有图书，封底和前勒口都有“湛庐文化”的标志



并归于两个品牌



老 找“小红帽”

为了便于读者在浩如烟海的书架陈列中清楚地找到湛庐，我们在每本图书的封面左上角，以及书脊上部 47mm 处，以红色作为标记——称之为“小红帽”。同时，封面左上角标记“湛庐文化 Slogan”，书脊上标记“湛庐文化 Logo”，且下方标注图书所属品牌。

湛庐文化主力打造两个品牌：**财富汇**，致力于为商界人士提供国内外优秀的经济管理类图书；**心视界**，旨在通过心理学大师、心灵导师的专业指导为读者提供改善生活和心境的通路。



老 阅读的最大成本

读者在选购图书的时候，往往把成本支出的焦点放在书价上，其实不然。

时间才是读者付出的最大阅读成本。

阅读的时间成本=选择花费的时间+阅读花费的时间+误读浪费的时间

湛庐希望成为一个“与思想有关”的组织，成为中国与世界思想交汇的聚集地。通过我们的工作和努力，潜移默化地改变中国人、商业组织的思维方式，与世界先进的理念接轨，帮助国内的企业和经理人，融入世界，这是我们的使命和价值。

我们知道，这项工作就像跑马拉松，是极其漫长和艰苦的。但是我们有决心和毅力去不断推动，在朝着我们目标前进的道路上，所有人都是同行者和推动者。希望更多的专家、学者、读者一起来加入我们的队伍，在当下改变未来。



湛庐文化获奖书目



《大数据时代》

国家图书馆“第九届文津奖”十本获奖图书之一
CCTV

《光明日报》2013年度“九图十书”八选图书

《第一财经日报》2013年第一财经金融价值榜“推荐财经图书奖”

2013年度和讯华文财经图书大奖

2013亚马逊年度图书排行榜经济管理类图书榜首

《中国企业家》年度好书经管类TOP10

《创业家》“5年来最值得创业者读的10本书”

《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”

《中国新闻出版报》2013年度好书20本之一

2013百道网·中国好书榜·财经类TOP100榜首

2013蓝狮子·腾讯文学十大最佳商业图书和最受欢迎的数字阅读出版物

2013京东经管图书年度畅销榜上榜图书,综合排名第一,经济类榜榜首



《牛奶可乐经济学》

国家图书馆“第四届文津奖”十本获奖图书之一
搜狐、《第一财经日报》2008年十本最佳商业图书



《影响力》(经典版)

《商学院》“2013经理人阅读趣味年报·心理学和行为科学类最受关注图书”

2013亚马逊年度图书分类榜心理励志图书第八名

《财富》鼎力推荐的75本商业必读书之一



《人人时代》(原名《未来是湿的》)

CCTV《子午书简》·《中国图书商报》2009年度最值得一读的30本好书之“年度最佳财经图书”

《第一财经周刊》·蓝狮子读书会·新浪网2009年度十佳商业图书TOP5



《认知盈余》

《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”

2011年度和讯华文财经图书大奖



《大而不倒》

《金融时报》·高盛2010年度最佳商业图书入选作品

美国《外交政策》杂志评选的全球思想家正在阅读的20本书之一

蓝狮子·新浪2010年度十大最佳商业图书,《智囊悦读》2010年度十大最具价值经管图书

《第一大亨》

普利策传记奖,美国国家图书奖

2013中国好书榜·财经类TOP100

《真实的幸福》

《第一财经周刊》2014年度商业图书TOP10

《职场》2010年度最具阅读价值的10本职场书籍

《星际穿越》

国家图书馆“第十一届文津奖”十本获奖图书之一

2015年全国优秀科普作品三等奖

《环球科学》2015最美科学阅读TOP10



《翻转课堂的可汗学院》

《中国教师报》2014年度“影响教师的100本书”TOP10

《第一财经周刊》2014年度商业图书TOP10

湛庐文化获奖书目

《爱哭鬼小华》

国家图

《新京报》2013年度重书

《中国教育报》2013年度教师推荐的10大童书

新阅读研究所“2013年度最佳童书”

《群体性孤独》

国家图书馆“第十届文津奖”十本获奖图书之一

2014“腾讯网·啖书局”TMT十大最佳图书

《用心教养》

国家新闻出版广电总局2014年度“大众喜爱的50种图书”生活与科普类TOP6

《正能量》

《新智囊》2012年经管类十大图书, 京东2012好书榜年度新书

《正义之心》

《第一财经周刊》2014年度商业图书TOP10

《神话的力量》

《心理月刊》2011年度最佳图书奖

《当音乐停止之后》

《中欧商业评论》2014年度经管好书榜·经济金融类

《富足》

《哈佛商业评论》2015年最值得读的八本好书

2014“腾讯网·啖书局”TMT十大最佳图书

《稀缺》

《第一财经周刊》2014年度商业图书TOP10

《中欧商业评论》2014年度经管好书榜·企业管理类

《大爆炸式创新》

《中欧商业评论》2014年度经管好书榜·企业管理类

《技术的本质》

2014“腾讯网·啖书局”TMT十大最佳图书

《社交网络改变世界》

新华网、中国出版传媒2013年度中国影响力图书

《孵化Twitter》

2013年11月亚马逊(美国)月度最佳图书

《第一财经周刊》2014年度商业图书TOP10

《谁是谷歌想要的人才?》

《出版商务周报》2013年度风云图书·励志类上榜书籍

《卡普新生儿安抚法》(最快乐的宝宝1·0~1岁)

2013新浪“养育有道”年度论坛养育类图书推荐奖



延伸阅读

《叩响天堂之门》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂宇宙求索的漫漫历程。
- ◎ 宇宙如何起源？为什么我们要耗资巨额，建造史上最大型的科学仪器——大型强子对撞机？宇宙万物的真相又如何向我们徐徐展开？
- ◎ 科学小白与科学大V都不可错过的年度最佳科普巨作，韩涛、张双楠、陈学雷、朱进、苟利军、吴岩、万维钢、郝景芳等众多顶尖科学家与科学达人挚爱推荐。



扫码直达本书购买链接



《弯曲的旅行》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂神秘的额外维度。
- ◎ 我们了解宇宙吗？宇宙有哪些奥秘？宇宙隐藏着与我们想象中完全不同的维度吗？我们将怎样证实这些维度的存在？



扫码直达本书购买链接



《暗物质与恐龙》

- ◎ 理论物理学大师丽莎·兰道尔“宇宙三部曲”——一本书读懂暗物质以及恐龙灭绝背后的秘密。
- ◎ 暗物质是什么？它是如何让昔日的地球霸主毁灭的？宇宙万物又是如何在看似无关的情况下联系在一起，从而改变了世界的发展的？



扫码直达本书购买链接



《星际穿越》

- ◎ 天体物理学巨擘，引力波领域大师，同名电影科学顾问基普·索恩巨著，媲美霍金《时间简史》。
- ◎ 国家天文台8位天体物理学科学家权威翻译。
- ◎ 国家图书馆“第十一届文津奖”科普奖获奖图书。



扫码直达本书购买链接



[1] 介于普通物质和通常的暗物质之间。——译者注

[2] 电影《星际迷航》中的飞船。——译者注

[3] 一般认为，这个比喻到此为止。不像那些被流放了的不满意的重元素（曾经在宇宙中分布），它们不会进一步加剧混乱。更好的是，它们还对恒星系统甚至生命的形成有贡献。

[4] 事实上，这个物理概念（尽管被广泛接受）在专业人士中存在争议。一方面，对白矮星爆发所预言的光谱和光变曲线和观测吻合得很好。另一方面，预期中白矮星的伴星从未有人看到过。天文学家由此认为，也可能是两颗白矮星合并而导致了爆发，有些数据是支持这个结论的（大部分是关于双星形成和爆发的时间差的测量）。但是关于单个白矮星爆发的方案，其预言细节还有待证实，所以这个问题还未完全解决。

[5] Λ 指的是希腊字母Lambda，CDM是“cold dark matter”（冷暗物质）的缩写。

[6] 因为洛杉矶的居民比较注重美。——译者注

[7] 1开尔文等于-273.15°C，等于-459.67° F。

[8] 当时他们所在的贝尔实验室就位于纽约市附近的新泽西州。——译者注

[9] 这里指的是威廉·布莱克的诗作《天真的预言》中的一句诗：“一沙一世界，一花一天国。”——编者注

[10] 赫歇尔是天王星的发现者，对于早期的观测者而言，低分辨的望远镜让这些星云看起来和行星很像。——译者注

[11] 1984年，美国联邦法律规定，21岁是饮酒的最小年龄。而在此之前，各个州的饮酒最小年龄各不相同。——译者注

[12] 天文单位（AU），即日地平均距离。

[13] 此探测器已经于2016年9月8日发射，按计划，将在2018年抵达小行星贝努，在2023年返回样本。——译者注

[14] 电影讲述了即将撞上地球的小行星。——译者注

[15] 如果森林里的一棵树倒下，但是附近没人听到它倒下，那么它究竟是发出声音了还是没有？作者此处想借此说明观察和现实之间的关系。——译者注

[16] 请注意，温度和压力都会起作用，因为熔点和沸点在不同压力下是不同的。任何曾试图在海拔2 400米的地方煮面条的人都知道这一点。

[17] 这里的“暗”指通常意义上的是否吸收光，而非“暗物质”的“暗”。

[18] 英文电影Deep Impact的中文名字被翻译成了《天地大冲撞》。——译者注

[19] 当然比柯伊伯最初所预期的还是要小很多。——译者注

[20] 安·德鲁彦与卡尔·萨根因此而喜结连理。——译者注

[21] 原句改编自“如果大山不走向穆罕默德，穆罕默德就会走向大山”。这里意指“如果我们去不了小行星上，那么小行星将会造访地球”。

[22] 阿波菲斯，埃及神话中的灾难和破坏之神，是破坏、混沌、黑暗的化身，因此作者在此称为“魔鬼”。——编者注

[23] 译者认为，这里作者写错了，应该是几百米而非几百公里。从上文的“直径为1000米的天体”等内容判断，此处也应为“几百米”。——译者注

[24] 作者在英文原文中用了不同的词语，但实际上，上下文中的“撞击坑”和“陨石坑”意义是一样的。为遵从原文，原文为“impact crater”的，仍译为了“撞击坑”。——译者注

[25] 这些天体资料已被很好地归档，并保存在地球撞击数据库（Earth Impact Data Base）中，你可以从网上查到相关信息。这些观测资料对于思考暗物质在引发流星体撞击所起的作用时至关重要。

[26] 水星的大气层极其稀薄且不稳定，也无法提供相应的保护。——译者注

[27] 国际地层委员会（ICS）负责命名这些时间框架，也试图消除第四个时间部分——第四纪，但国际第四纪研究联盟表示反对。因此，在2009年，ICS恢复了此术语。第三纪，由于没有那么多热切的捍卫者，因此不再是官方用语，这也是K-T（白垩纪-第三纪）被K-Pg（白垩纪-古近纪）取代的原因。

[28] 这一族群包含了虚幻龙，这是雷龙最初的名字，也许是内雷龙的首选名字。关于雷龙名字的辩论，可以和关于冥王星是否为行星的辩论相媲美。

[29] 这里作者引用的是《金发女孩和三只熊》的故事。有兴趣的读者可以阅读一下这个意义非凡的故事。——译者注

[30] 《爱丽丝梦游仙境》中的兔子窝，指的是各种困惑麻烦的源泉。——译者注

[31] 有关大型强子对撞子的内容，我在《叩响天堂之门》中作了详细介绍。

[32] 从技术上讲，从其他过程中得到的精确测量会给出一些限制，但是这些限制一般在关于希格斯玻色子直接探测的报告中被忽略掉了。

[33] 内梅西斯，希腊神话中复仇女神的名字。以前认为是这颗星的运动引发了彗星和恐龙灭绝之间的关联，但一直没有找到这颗星。

[34] googol，等于10的100次方。——译者注

[35] 此处原书写的是中子，但估计是排版错误。——译者注

[36] 鲁布·戈德堡，美国犹太人漫画家，画了许多用复杂方法做简单小事的漫画，赢得了许多读者。——译者注

[37] 玛莎·斯图尔特，美国企业家，几经挫折最终跻身全美第二女富豪。她在烹饪、园艺和室内装饰等方面颇具才华。——译者注

[38] 即关于地球和太阳谁是宇宙中心的问题。——译者注

[39] 粒子物理学标准模型上缺失的一片，近期刚被发现的质量。

[40] 双黑洞合并所辐射的引力波已经在2016年被发现。——译者注

[41] 就是学术界常用的arXiv.org网站。——译者注