

复杂系统与复杂网络

Complex Systems and Complex Networks

何大韧 刘宗华 汪秉宏 编著



高等教育出版社
Higher Education Press

内容简介

复杂系统与复杂网络

Complex Systems and Complex Networks

何大韧 刘宗华 汪秉宏 编著

图在编录目(CIP)数

何大韧刘宗华汪秉宏

北京:高等教育出版社,2009

ISBN 978-7-04-025627

...①何...②刘...

—册价: 1.0414.2

中国版本图书馆CIP数据

薛气国金

集德

编著 王 王 王 王 王 王

版云版社 文新文

<http://www.hup.edu.cn>
<http://www.hup.com.cn>
<http://www.landisoo.com>
<http://www.landisoo.com.cn>
<http://www.wideth.com>

出版发行 高等教育出版社
 社 址 北京市西城区
 邮政编码 100120
 总 经 010-2888100
 经 理 北京理工大学
 印 刷 北京理工大学

2009年1月第1版
 2009年1月第1次印刷
 45.00元

开 本 787×1092 1/16
 印 张 19.25
 字 数 320 000

高等教育出版社

联系电话 010-5858100
 邮 政 25627-00

内容提要

本书从研究复杂系统的角度来介绍复杂网络。作为一本入门引导,前五章介绍一些复杂系统理论的基础知识,包括熵、计算机与信息、非平衡统计物理学、耗散结构与协同学、临界现象与自组织临界性、混沌、元胞自动机模型、复杂性的定义与量度、有关的统计物理学方法、博弈论、数理统计、图论等。第六、第七章介绍复杂网络的一些基础知识,包括描述网络拓扑结构的统计性质以及一些有影响的网络演化模型。在第八、第九、第十章中介绍了网络上的物理传输过程、生命网络和合作网络与合作-竞争网络。其中侧重介绍了作者群体的工作。另外,本书阐述了作者们对复杂网络研究前景的一些看法,特别是在第十一章中介绍了关于复杂网络动力学框架的一些最新研究。

本书可作为复杂系统与复杂网络研究方向的研究生教材,也可作为相关领域研究人员的参考书。

图书在版编目(CIP)数据

复杂系统与复杂网络/何大韧,刘宗华,汪秉宏编著.

北京:高等教育出版社,2009.1

ISBN 978-7-04-025627-7

I. 复… II. ①何…②刘…③汪… III. 统计物理学-研究 IV. O414.2

中国版本图书馆 CIP 数据核字(2008)第 185128 号

策划编辑 王超 责任编辑 李茜 封面设计 张楠 责任绘图 尹莉
版式设计 史新薇 责任校对 胡晓琪 责任印制 毛斯璐

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 北京机工印刷厂

开 本 787×1092 1/16
印 张 19.25
字 数 350 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
畅想教育 <http://www.widedu.com>

版 次 2009 年 1 月第 1 版
印 次 2009 年 1 月第 1 次印刷
定 价 42.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 25627-00

前 言

大约十年之前(1998年6月4日),Nature发表了两位年轻的物理学家(D. J. Watts和S. H. Strogatz)关于网络的一篇论文。一年多之后(1999年10月15日),Science又发表了另外两位年轻的物理学家(A. L. Barabasi和R. Albert)关于网络的另一篇论文。这两篇论文引发了关于复杂网络的研究热潮。这个潮流席卷全球,涉及数学、力学、物理学、计算科学、管理科学、系统科学、社会科学、金融经济科学等许多科学领域,以及交通运输、能源传输、通信工程、电子科学,甚至医学、烹饪等许多应用学科。至今(2008年3月),D. J. Watts和S. H. Strogatz的论文被SCI收录的论文引用5670次;A. L. Barabasi和R. Albert的论文被引用3275次。

人们把周围的许多系统(天然的或者人造的,例如交通网、电力网、人际关系网等)看作网络由来已久,运用数学的一个分支——“图论”对这些系统进行研究也已经有百年以上的历史。上述两篇文章的重要之处在于作者发现许多实际网络具有一些共同的拓扑统计性质,即“小世界性”和“无标度性”。这些性质既不同于规则网络,也不同于随机网络,正像近几十年来物理学家认为“复杂位于规则与随机之间”一样,所以大家把实际网络称为“复杂网络”。所谓小世界性是指实际网络具有比规则网络小得多的平均节点间距离和比随机网络大得多的平均集群系数(即邻点之间也相邻,形成紧密集团的比例);而无标度性则指实际网络中节点邻边数取一个定值的概率分布函数是幂函数(规则网的这个分布是 δ 函数,而随机网是正态分布)。这个幂函数标志基本单元与其邻居相互作用能力的极其不均匀分布。更加引人注目的是:论文的作者提出了解释这些独特规律的网络演化模型,而且运用统计物理学方法从这些模型解析地得出了这些独特规律。这些模型的思想简单明白、直观合理。产生小世界性的机制就是一部分基本单元之间相互作用的远程性、跳跃性和随机性;产生无标度性的机制就是基本单元建立相互作用的“优选”(或者称为“富者更富”)法则。这是第一次把统计物理学的思想和方法引进网络或者图论的研究,因此,若与传统的图论或网络理论比较,也许可以说当前的复杂网络研究的特征就是统计物理学的进入,所以应该把统计物理学列入复杂网络研究的基础知识之中。

在这十年中,不止一次地有人发表议论,说复杂网络研究已经差不多了,现在参加已经太晚;或者说复杂网络研究的前途只剩下在各个实际领域中的应用,基础研究已经基本到头等。然而,事实正好相反,国内外的复杂网络研究,包括

基础研究和可能的应用研究,都在快速发展。自从复杂网络研究兴起,我国物理学界就及时融入了世界科学的新潮流。2000—2001年间,胡进锬、熊诗杰、邹宪武、金准智、陈天仑、胡班比、胡岗、汪凯歌、高自友、朱陈平、汪小帆、陈关荣等人在SCI期刊Phys. Rev. E、Phys. Rev. B和Comm. Theo. Phys.上发表了早期工作。其中汪小帆、陈关荣等人2001—2002年的论文引起了较大的反响。与此大约同时的还有刘宗华、来颖诚等2002年在Phys. Lett. A,以及欧阳颀等人2002年在Phys. Rev. E发表的文章。2004年初夏在无锡召开全国第一次复杂网络学术论坛时,只有50人参加,20人发言。到了2006年初冬在武汉召开全国第二次复杂网络学术大会时,就有300人参加,200多人发言。在2007年11月上海的全国第三次复杂网络学术会议上,经过严格审稿最后收入会议论文集的论文就有249篇,参加会议的人数超过400。估计全国从事复杂网络研究的人数以千计,包括数学、力学、物理学、计算科学、管理科学、系统科学、社会科学、金融经济科学等许多科学领域,以及交通运输、能源传输、通信工程、电子科学,甚至医学、烹饪等许多应用学科的研究生和研究人员。正如香港城市大学的陈关荣教授为2006年武汉的全国第二次复杂网络学术大会文集所写的序言中说的:“从过去这几年国际上复杂网络的研究进展来看,发表的论文数量有增无减,专著陆续出版,论题的覆盖面不断拓宽,理论的探讨不断深入,并且渗透到越来越多的不同学科、特别是边缘和交叉学科里面去,可见其发展姿态依然是方兴未艾、形势喜人。展望不远的将来,可以预见,像复杂网络这样规模庞大、非线性强、复杂度高而且种类繁多的多节点连接巨型动态系统,其研究工作绝不会在短期内就取得比较完整的结论,其发展状态亦绝不会直线式地迅猛上升;但是也正因为这样,其发展趋势亦一定不会很快地衰落下来。毋庸置疑,复杂网络应该是相关领域里新一代研究生和青年科研工作者选题定向和专业发展最好的选择之一。”

为什么还需要深入地进行复杂网络的研究?复杂网络研究的最主要目标是什么?可能不同的人有不同的看法。我们3位作者作为物理学工作者,习惯于从物理学发展的角度看问题。

本书的读者们应该都熟悉经典物理学,尤其熟悉经典力学和经典电磁学,这是物理学中比较接近我们周围的世界,比较好懂,又特别能显示由牛顿及其追随者们创立、完善现代物理学方法论框架美妙魅力的部分。在经典力学和经典电磁学中,研究的对象——运动物质,被想象为可以被分割为无限多个无限小,又在空间中连续分布的基本单元的集合。这些基本单元(质点或者电荷元)被放在均匀空间中的各个规则格点位置上,因此,使用千年前数学家们创造的坐标系就可以完善地描述这个体系的运动。尽管各个基本单元的空间位置不同,而且一般来说在随时间变化,但是由于它们之间的相互作用遵从已经被认识的简

单的、普通的基本法则,因此运用几百年前创立的微积分工具(以及基本思想类似的一些现代数学工具)就可以非常简明地表示支配每一大类客观体系运动变化的普遍动力学规律,并且可以用来准确地预言这些体系未来的行为,为人类服务。以坐标的运动学描述和微积分(以及其他基本思路相似的现代数学方法)为基础的动力学描述成为现代物理学辉煌大厦的支柱。

然而,自然界中存在大量不适于或者不能够用这种方法论(被称为“还原论”)讨论的系统。在第一类系统中,虽然我们原则上可以运用还原论进行它们演化的讨论,但是实际上,由于基本单元的相互作用涉及非常多的因素,这些因素又错综复杂地交连,所以不但进行动力学的解析讨论不可能,连不考虑近似的用“理想精确”数值“从头计算”也不大可能。与国计民生密切相关的气象系统、地震系统、足够复杂结构的各种材料系统都是这类系统的例子。可以说随着物理学研究的不断深化,几乎各个物理学分支的前沿都涉及这类复杂系统。物理学家越来越认识到不能继续被局限在还原论的框架内去讨论这类系统,必须探寻全新的思想方法论才有解决问题的希望。除此之外,另一类系统根本不能用还原论来处理,这是由于这类系统基本单元的“组织”会“涌现”许多种大量分立个体不会展示的性质,因此不可能仅仅依据单元个体性质来预言系统整体的丰富行为。这类系统的最典型代表是具有生命特征的所谓“自适应系统”(即基本单元具备根据外界信息进行预期、采取对策、改进自身及其与其他单元或环境的关系的系统),例如生物系统、生态系统、社会系统、经济系统等。近几十年来,由于各门科学,尤其是计算机科学和非线性科学的发展,许多科学家(包括物理工作者)认为突破还原论的限制,寻求描述复杂系统的概念和理论,把物理学的适用领域推广到复杂系统的任务已经提上日程。

改造世界必须先认识世界,而认识世界必须先描述世界。世界上的客观系统是由规则、均匀分布,全同、遵从简单普适规律的基本单元构成的,还是由高度不规则、不均匀分布,显示丰富多彩的种类、相互作用、有组织的基本单元构成的?这可能是把物理学推广向复杂系统时要回答的首要问题。这个问题正在引起物理学家(以及许多领域内的科学家)对复杂网络研究的空前热烈兴趣。

最简单地回顾一下生物学研究的历史可能是有益的。20世纪前,生物学基本上属于经验科学,没有多少真正可称为理论的东西。20世纪中叶以后,物理学、化学的仪器与方法大规模地进入生物学研究,同时带给生物学的还有物理学先分解、再综合的方法论。20世纪后期盛极一时的分子生物学体现了这些变化带来的辉煌成就。“分解”发现了蛋白质结构、基因以及遗传信息序列等。然而,“综合”却遇到了非常大的困难。生物系统远远不像经典物理学研究的力学或电学系统那样容易从基本单元综合起来,生物系统基本单元的种类太多,它们之间的作用太错综复杂,也许只有用网络(可能还是一大堆网络)来形容才比较

合适。这就是 21 世纪开始后系统生物学(包括所谓“生物网络”)兴起的原因。这个发展历程和形势吸引了大批物理学家试图进入生命科学研究领域。

前面所说的 Nature 和 Science 上发表的那两篇论文在物理学界引发了地震,使许多物理学家,尤其是统计物理学家认识到实际系统使用网络描述的重要性,统计物理学进入网络研究并大有作为的可能性,以及网络描述作为复杂系统研究工具的可能光辉前景。近十年来,复杂网络研究的论文车载斗量,然而,要达到物理学家寻求复杂系统动力学理论框架的目标,复杂网络研究还任重而道远。

首先,复杂网络研究依赖的主要数学基础理论——“图论”的主流还很难与物理学动力学描述的要求挂钩。从 1736 年诞生开始,图论研究就集中于位于某个平面(或二维曲面)上的一些基本单元之间的位置关系、这些位置之间的某种联系以及图上某种量的传播等问题。这种对系统基本单元及其相互作用的描述是“静态”和“平面”的,没有提供对相互作用丰富特性及其随时间演化的描述方法,更缺乏总结这些相互作用特征与规律,从而预言系统行为的手段。就我们所知,突破传统框架的图论研究已经开始发展,并带给物理学家极大鼓舞。此外,已经有一些从其他角度探索复杂网络动力学工具的研究。我们将根据自己的理解,选择这些研究成果的一部分在本书最后一章予以简介。然而,这些研究只能说是一个开始。

其次,把复杂网络作为研究对象对于统计物理学也是新问题。统计物理学还不能提供针对这类对象的最有效思想、方法、工具。目前的研究在借用一些多年来对许多类研究对象行之有效的办法,例如平均场理论、主方程、率方程、生成函数、最大似然估计等。本书将从物理角度简介这些方法。专门针对复杂网络研究的统计物理学方法的发展将是物理学工作者的一个重要探索方向。

还有不少原因可以被列举,来说明复杂网络研究还将不断地推出意义重大的研究成果并将热烈地延续很长时间。正像 R. Albert 和 A. L. Barabasi 在 2002 年 *Reviews of Modern Physics* 上发表的综述文章(被 SCI 收录论文引用 4231 次)中所说的:“我们相信,这些(已经获得的)结果仅仅是一杯冰淇淋的小尖。”我们同样相信,还会有很多年轻朋友们陆续进入这个前程远大、可能对人类具有重大影响的科学领域,还会有许多关于这个领域的研究专著出版。如果本书能作为一本入门引导,能被一些青年朋友喜欢,我们将感到无比欣慰。

在我们的知识范围内,虽然国外关于复杂网络的科学专著已经出版过几十本,还没有一本特别考虑初学者的需求。国内已经出版了两本关于复杂网络的高水平专著。第一本是《复杂网络》,由郭雷、许晓鸣等(包括本书作者之一汪秉宏)主编(上海科技教育出版社,2006)。它由 31 位作者(包括本书作者之一何大韧)分工写作 13 章。每一章由熟悉该方向的专家综述其前沿研究,并兼顾他们自己的工作汇报。内容全面,质量高,但是部分初学者阅读可能有一定困难。

第二本是《复杂网络:理论及其应用》,由汪小帆、李翔、陈关荣著(清华大学出版社,2006),它由这3位长期合作的高水平专家撰写,有很好的系统性和前沿水平。该书作者具有开设研究生课的基础,因此也兼顾了与本科生水平的接轨。然而,3位作者都具有工程和数学的学术背景,所以他们的大作与本书的风格有差异。另外,他们的研究水平很高,强调学术水平,不像本书这样以作者6年来为研究生授课的讲义为基础,特别强调介绍基础知识。本书共分为11章,我们使用其中的5章,分别简单介绍进行复杂网络研究所需的5个科学领域的基础知识,包括复杂性与复杂系统、有关的统计物理学方法、博弈论、数理统计、图论,并且推荐一些比较容易阅读的参考文献。我们期望读者们仍旧会感到本书有用。

本书的3位作者进行亲密的研究合作已经接近20年,现在我们领导的集体(中国科技大学理论物理研究所、华东师范大学理论物理研究所、扬州大学复杂性科学研究中心)又共同承担国家自然科学基金的重点项目“基于复杂网络的复杂系统动力学与统计行为的研究”(项目号:10635040)。本书是由这个重点项目资助获得的一个重要成果,也是我们合作的一个重要纪念。本书包含3个研究集体6年来一百多人的辛勤研究成果,无法在此一一列举,仅此表示感谢。另外,应我们的特别邀请,南京航空航天大学的朱陈平副教授为本书编写了第7章第11节,扬州大学的官山副教授为本书编写了第10章第3节。在此我们要特别感谢朱陈平副教授,他不但非常仔细地阅读了本书全稿,提出了很多极有价值的修改意见,而且也积极参加我们基金重点项目(10635040)的研究工作,贡献了不少高质量的论文。同时,本书的出版得到扬州大学出版基金和高等教育出版社的大力支持与资助,在此致以深深的谢意。

目 录

第一章 漫谈复杂性与复杂系统	1
1.1 熵	1
1.2 计算机与信息	5
1.3 算法复杂性	7
1.4 非平衡统计物理学、耗散结构与协同学	8
1.5 临界现象与自组织临界现象	11
1.6 混沌	21
1.7 原胞自动机	25
1.8 描述复杂性与统计复杂性	29
第一章参考文献	40
第二章 一些有关复杂网络研究的统计物理学方法	43
2.1 连续相变的平均场理论	43
2.2 自组织临界现象的平均场理论	46
2.3 流行病传播的平均场理论简介	50
2.4 主方程	52
2.5 生成函数	55
2.6 率方程	57
第二章参考文献	57
第三章 博弈论及演化网络博弈	59
3.1 基本概念	60
3.2 完全信息静态博弈与纳什均衡	63
3.3 完全信息动态博弈与子博弈精炼纳什均衡	66
3.4 不完全信息静态博弈与贝叶斯纳什均衡	68
3.5 不完全信息动态博弈与精炼贝叶斯纳什均衡	70
3.6 合作博弈	72
3.7 演化网络博弈	73
3.8 城市公交网络的网络操纵者博弈模型	82
第三章参考文献	85
第四章 数理统计简介	87
4.1 一些基本概念	87
4.2 统计假设及其检验	90
4.3 一元线性回归	92
4.4 回归的一些问题	95

4.5	漫谈数据的采集与处理	103
	第四章参考文献	105
	附录	105
第五章	图论简介	107
5.1	一些基本概念	110
5.2	图的连通性	114
5.3	树图	115
5.4	最短道路问题	116
5.5	图的矩阵描述	117
5.6	有向图	119
5.7	二分图	121
5.8	网络流	122
	第五章参考文献	124
第六章	复杂网络的统计描述	126
6.1	平均距离、谐平均距离、效率与脆弱性	126
6.2	集群系数、圈系数、富人集团系数、集团度	127
6.3	度、度分布、度相关性	130
6.4	边权网及边权的一些统计性质	131
6.5	二分图的二分度	132
6.6	中心度与中心化	132
6.7	谱分析	136
6.8	模体	137
6.9	群落、派系与层次	138
6.10	度分布熵、目标熵以及不同的网络信息熵	141
6.11	多标度分形的分数维谱	143
6.12	漫谈复杂网络的统计描述	145
	第六章参考文献	145
第七章	一些网络演化模型	147
7.1	ER 随机网模型	147
7.2	WS 小世界网模型	150
7.3	BA 无标度网模型	155
7.4	BA 无标度网模型的主方程解	159
7.5	BA 无标度网模型的率方程解	161
7.6	部分优选、部分随机选择模型	162
7.7	局域世界模型	164
7.8	赋权演化网络的 BBV 模型	165
7.9	可调集群系数的 HK 模型及其改进模型	167
7.10	JGN 社会网络模型	169

7.11 自组织耦合演化模型	171
7.12 其他运用统计物理学方法的模型研究	181
第七章参考文献	183
第八章 复杂网络上的物理传输过程	187
8.1 流行病传播的基本模型	187
8.2 复杂网络上的流行病传播	190
8.3 复杂网络上的舆论传播	194
8.4 群落网结构对流行病传播的影响	198
8.5 动态群落网上的流行病传播	203
8.6 因特网上的信息包传递	208
8.7 因特网上交通堵塞的控制	210
8.8 交通数据的去趋势涨落分析	216
8.9 复杂网络上的粒子输运	220
8.10 粒子输运的平均场方法	221
8.11 加权复杂网络上的粒子输运	223
8.12 简单网络上能量输运	225
8.13 复杂网络上能量输运	229
展望	234
第八章参考文献	234
第九章 一些生命网络的研究	239
9.1 大脑功能网络	239
9.2 两态小动物群体网络	243
9.3 生物分子网络	250
第九章参考文献	255
第十章 合作网络与合作 - 竞争网络	258
10.1 简介	258
10.2 比较早期的合作网实证研究	260
10.3 合作网的项目大小分布和项目度分布	264
10.4 合作网的同类性与项目度分布的相关性	271
10.5 二分图投影的资源分配方法	274
10.6 近期关于合作网络的实证研究	276
10.7 关于合作 - 竞争网络的研究	279
第十章参考文献	282
第十一章 网络动力学的一些探索	285
11.1 布尔网络、信息距离及一些复杂网络的非线性动力学	285
11.2 最小作用量原理与网络形态的自然选择	289
11.3 图的动力学谱分析	292
第十一章参考文献	294

第一章 漫谈复杂性与复杂系统

从 20 世纪末以来,不少物理学工作者一直在寻求描述复杂系统的概念和理论,力图把物理学的适用领域推广到复杂系统^[1-7]。近十年来,复杂网络成为被寄予希望的一种描述工具。然而,要把复杂网络与复杂系统的研究很好地结合,首先要知道什么是复杂,什么是复杂系统、简单系统以及复杂性。这可能也是古往今来许多智者反复思考过的问题。除了定性的回答之外,几十年来,许多科学家致力于建立定量的定义,希望利用这样的定义来定量计算各种系统的复杂程度,从而比较不同系统复杂程度的大小。他们的成果尽管都还没有得到公认,但是很可能是建立复杂系统理论的必经阶段之一。这类研究论文数量相当多,已经提出的定义五花八门,有些很难搞懂,更难计算,要在本章中全面介绍不大可能,然而,这些研究发展的大趋势是一致的,都是从 20 世纪中叶以来物理学及一些别的科学分支的一些大进展延拓开来的。我们认为与复杂网络研究直接关联的正是这些大进展带来的物理学新理解,而不是个别的复杂性定义。因此,本章将主要介绍这些大进展,对复杂性定义仅按照我们的看法选择一小部分进行简介。知道这些知识对读者们理解本书以后介绍的内容很可能是必要的。由于本章各节的内容都涉及一门大学科,这里不可避免地只能作科普性的介绍,就当是在以后各章理论内容之前的一段比较轻松的阅读吧。

1.1 熵

物理学中对“复杂”的最初考虑和争论也许要从统计物理学的建立,也就是概率论的描述和统计方法进入物理学算起^[8]。在一百多年前,以玻耳兹曼为代表的物理学家倡导统计物理学的时候,一切并不那么容易。

为了说明这一点,我们有必要简单介绍牛顿(图 1.1)和其他几位伟大的力学家的贡献。牛顿在 1661 年(19 岁)进入剑桥大学,靠为学院做杂务支付学费。教授伊萨克·巴罗独具慧眼,在 1664 年,把牛顿选为自己的助手。1669 年,巴罗为了提携牛顿而辞去了教授之职,26 岁的牛顿晋升为数学教授。微积分的创立和经典力学完整、严密体系的建立是牛

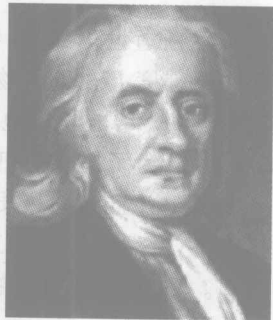


图 1.1 牛顿(1642—1727)

顿最卓越的成就。英国的政府与社会也给了他丰厚的回报。1705年(63岁)牛顿被封为贵族,非常富有,担任英国皇家学会会长直到他1727年(85岁)逝世。死后被埋葬在去世的著名人物汇集的威斯敏斯特教堂。他的墓碑上镌刻着:“让人们欢呼这样一位多么伟大的人类荣耀曾经在世界上存在(Let mortals rejoice that there has existed such and so great an ornament of Nature)。”

在牛顿之后,一大批力学家、数学家继承牛顿的研究,把经典力学提升到了新的高度。其中最值得提及的是拉格朗日(1736年生于意大利的都灵,19岁当上教授,1813年卒于巴黎)和哈密顿(1805年生于爱尔兰都柏林,1827年22岁时被任命为教授,1865年卒于都柏林),如图1.2和图1.3。以他们的名字命名的方程和动力学理论与方法至今仍是物理系大学生的必修内容。经典力学的完美理论框架和它在各个科学技术领域(天文学、机械工程学等)取得的伟大成就使科学界普遍地接受了确定论。大家相信,只要精确地知道一个系统演化的方程和初值,就可以精确预言它任何时刻的运动,而且这个原则适用于宇宙间的万物。那些列不出方程或者解不出方程的相对复杂的系统的行为只是暂时不能精确预言,当人类的认识能力和解析技巧不断改进之后,总有一天它们对人类再无神秘可言。



图 1.2 拉格朗日(1736—1813)

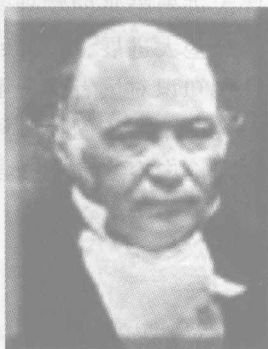


图 1.3 哈密顿(1805—1865)

在牛顿逝世不到50年之后,蒸汽机的发明带来了现代化的工业组织形式,也在一定程度上改变了物理学研究的方向。1785年蒸汽机被成功地应用于纺织工业,1807年被成功地应用于轮船和火车。然而,当时的蒸汽机的效率只有5%~8%。为了提高热机的效率,许多物理学家致力于研究热机的理论。1824年卡诺提出并且利用热质说(已经被证明是不正确的)证明了著名的卡诺定理,即相同高、低温热源之间工作的一切可逆热机效率相等,并大于一切不可逆热机的效率。不久,克劳修斯提出热机循环中的守恒量不应该是热量,而应该是热量与温度的比, Q/T ,从而提出了著名的克劳修斯不等式: $\oint dQ/T \leq 0$,其中等号对

应于可逆循环,小于号对应于不可逆循环。这个不变量被命名为克劳修斯熵。这是熵增加原理的最初形式^[8]。这只是一种经验总结,或者可以称为“唯象”的理论,虽然实际上表示了对一个全新的普遍自然规律的认识,但是还远远没有表述成能够阐明这个自然规律深刻内涵的形式。物理学历史上一再出现这种情况,科学实验已经积累了足够的素材,只等待一位有勇气独立思考的天才实现一次划时代的认识上的飞跃。玻耳兹曼(图 1.4)正好出现在这个时刻,然而他的命运和牛顿大不相同。在牛顿、拉格朗日、哈密顿等一大批力学家建筑起来的经典力学的辉煌大厦面前,玻耳兹曼和他倡导的统计物理学似乎显得非常渺小甚至荒唐,尽管历史最终证明玻耳兹曼提出的确实是科学真理,他和牛顿同样伟大,玻耳兹曼在他生前却历尽艰辛。



图 1.4 玻耳兹曼
(1844—1906)

玻耳兹曼 1844 年生于维也纳,家境极端困难,1863 年进入维也纳大学物理学和数学专业。1866 年 2 月 6 日他不满 22 岁就拿到博士学位,开始把统计学的思想引入分子运动论。1869 年,他在担任奥地利的格拉茨大学讲师期间提出了著名的熵和热力学第二定律的微观解释,也就是著名的玻耳兹曼公式: $S = k \ln \Omega$,其中 Ω 表示系统“包含微观状态的数目”, k 就是如今物理系学生耳熟能详的玻耳兹曼常量^[8]。然而,当时的科学界远远没有能够像今天这样深刻认识到必须极其小心地保护新生的科学幼苗。玻耳兹曼在维也纳大学任教时受到了以马赫为代表的实证论者的强烈批评及指责。他提出的熵与概率之间的联系遭到绝大多数物理学前辈的责难。在他任教的学校中,马赫开设的“归纳科学的历史与哲学”讲座吸引了大批优秀学生,包括玻耳兹曼的研究生在内。大多数学生仅仅选择玻耳兹曼为第二指导老师,这种精神压抑深深地刺痛了玻耳兹曼的自尊心。为了改善自己的环境,玻耳兹曼在 1900 年转向德国莱比锡大学,担任了理论物理学教授职务,但在那里,他同以奥斯特瓦尔德为代表的唯能论者之间的观点对立使他并不比在格拉茨大学更轻松。1902 年,当马赫因病辞去科学哲学教授之后,玻耳兹曼再次回到维也纳大学,可是情况并没有改善。从 1903 年到 1906 年,玻耳兹曼开设的讲座的听讲学生人数逐年减少。这位哲学争论中的孤独者于 1906 年 9 月 5 日(62 岁)在意大利一个度假店里上吊自杀。他被亲友埋葬在一个公园的小径旁,长时间被荒草淹埋的墓碑上只刻着他的著名公式: $S = k \ln \Omega$ 。

确定论规律先于深藏在大量随机现象深处的统计规律被发现,确定论科学比统计性科学容易得到承认。这也许不是历史的偶然,而包含发人深省的深刻

道理。各种人都希望听到确定的预言,哪怕这种预言来自只能勉强和我们交流的其他生物也好。这种一厢情愿的希望科学中的反映也许是非常深远的。在20世纪我国从60到90年代理工科常用的一些大学物理教材中长期宣传这样的观点:在分子运动论中采用的统计物理方法不是必需的,不是因为其科学方法和思路与经典力学根本不同,而是由于分子数太多,求解牛顿方程太麻烦,不得已而为之,即所谓“求解这么多的牛顿方程既不可能,也不必要”。几十届物理学科和其他学科的学生都接受了这种说法。实际上,人们容易想起一个理所当然的问题:力学微分方程都是可逆的,如果统计物理学没有引入新的原则,热力学系统的不可逆性如何从可逆的力学微分方程中得出?这个问题必然引向这样的结论,即大量分子的随机运动会产生新的、与少量分子确定性运动根本不同的性质!实际上,在大学物理中一再讲授的玻耳兹曼热力学第二定律的概率解释完全能回答这个问题。它告诉我们,大量随机事件构成的系统会自动趋向于它的大概率状态。反过来,系统趋向于它的小概率状态就必需外界的干预(输入能量、信息等),绝对不会自动进行^[8]。也许还可以追问一下,为什么对这类系统的少量次数行为的预言几乎完全被随机性支配(而不能从牛顿定律求解精确预言),只有大量多次的重复测量才会展示统计规律?用大家熟悉的投掷硬币作为例子,提高投掷过程的控制、测量精度可以排除对少量投掷结果预测的随机性吗?显然不能。设想我们投掷的不是硬币,而是不倒翁,那么行为预言的随机性就不复存在。看来关键在于系统演化结果对于扰动的敏感性。由于干扰总是存在,而且很难找到它们之间的关联和支配他们的规律,如果太敏感,运动就不可预言。任何一个事物运动遵循的压倒性的少数道理才是道理,才会导致可预言的秩序。大量不相关联的道理近于根本没有道理,运动实际上就没有秩序和可预言性。

也许还可以再问一个问题:既然大概率事件导致统计规律,而且统计量越大,涨落(小概率、随机性事件)越小,小概率事件越不重要,那么我们身边的绝大多数宏观系统的整体运动就应该基本上不依赖于涨落。因此,统计规律和概率描述毕竟远远不如确定性描述重要。对吗?我们认为此说法不妥。现代科学表明:定性、总体来看,系统越复杂,小概率事件越重要。例如你现在正在看我们的这本书,这个事件发生的概率其实小到了极点。首先地球必须在宇宙大爆炸以后的一个适当时间、地点产生,地球必须提供很合适的条件,使得它上面的物质微粒能够逐渐结合成足够大的分子,而且在一个阈值上产生生命,我们的父亲和母亲都必须在茫茫人海中相遇并且互相赏识,等等。这个小概率事件对我们当然太重要了。最近的一本畅销书《黑天鹅》(The Black Swan: The impact of the highly improbable)对此有较详细的描述。它说得是在发现澳大利亚的黑天鹅之前,欧洲人认为天鹅都是白色的,“黑天鹅”曾经是他们言谈与写作中的惯用语,

用来指不可能存在的事物,但这个不可动摇的信念随着第一只“黑天鹅”的出现而崩溃。从次贷危机到东南亚海啸,从“泰坦尼克号”的沉没到“9·11”事件,“黑天鹅”存在于各个领域。“黑天鹅”就是我们这里说的小概率事件。因此小概率事件的重要性也许可以作为复杂性度量的标准之一。我们将在本章后面再回到这个问题。

1.2 计算机与信息

另一个需要介绍的概念是信息,尤其是与计算机相关的信息概念^[9,12]。信息的产生、传递和处理可能是贯穿人类历史的重要问题之一。从古代的结绳记事、烽火台,到后来的纸笔信件、印刷术、邮递员,到今天的无线电话、电视、电子邮件、因特网等,人类经历了巨大的进步。借助于人造的计算装置来传递、处理信息也已经有相当长的历史。打字机、算盘、机械计算机都可以看作对原始信息进行某种处理,以方便传递或其他目的的装置。再后来出现的电报、电话等已经是很具备现代味道至今还在使用的设备和方式。

说起现代计算机的发明历史,可能要提到1896年赫曼·霍列瑞斯报表公司为了完成一次人口普查而首先制造、使用的穿孔纸带或卡片输入的机械计算机。到了1940年前后,美国政府机构为了进行设计原子弹所需的大量计算,租用了IBM公司的穿孔卡片输入式电子真空管计算机,这台计算机的运算速度是每7秒完成一次乘法。1943年,25岁的研究生埃克特领导火炮弹道计算研究项目,1946年研制成功了被命名为埃尼亚克、用电子管和继电器完成计算逻辑控制的计算机,进行一次乘法的时间缩短到0.003 s。计算方法理论和计算机理论的创始人之一冯·诺伊曼就是在这期间开始参加计算机的改进,提出、奠定了现代的程序存储计算机设计理论的基础。1949年世界上第一台程序存储计算机在美国诞生,包含3600个真空管、10000个锗二极管,主频为1 MHz。它在1952年正式开始运行,1962年光荣退休。值得一提的是,中国科学院计算机研究所研制的“中国104计算机”在1959年制成,包含4200个真空管、4000个二极管,每秒完成10000次运算,内存1048字节(浮点40二进制位)。使用一台磁带机作为外存,耗电70 kW。使用一台专用发电机组,占用一座不太大的楼房。由此可见,我国当时与世界先进水平的差距并不十分大^[9]。

冯·诺伊曼(图1.5)的历史可能给我们新的启发。有趣的是他并不是一个工程师,而是一个数学家。他是美籍匈牙利人,1903年生于匈牙利的布达佩斯。1911—1921年,冯·诺伊曼在布达佩斯的卢瑟伦中学读书期间就发表了第一篇数学论文,当时冯·诺伊曼还不到18岁。1930年接受了普林斯顿大学客座教授职位,1931年成为该校终身教授。1951年至1953年任美国数学会主席。

1957年在华盛顿去世,终年54岁。

冯·诺伊曼在算子理论、集合论、群论、算子代数(冯·诺伊曼代数)、博弈论(创立者)、格论、连续几何、理论物理、动力学、连续介质力学、气象计算、原子能和经济学等诸多领域都进行了开创性工作,并作出了重大贡献。但他对人类的最大贡献仍旧是对计算机科学、计算机技术和数值分析的开拓性工作。1945年,他和埃尼亚克机研制小组的富有创新精神的年轻科技人员们发表了一个全新的“存储程序通用电子计算机方案”,方案中新机器由五个部分组成,包括运算器、逻辑控制装置、存储器、输入和输出设备,程序和指令采用了二进制。最为重要的是指令和数据一起存储的程序存储方式。这种综合设计思想导致了上述的第一台程序存储计算机的诞生以及著名的“冯·诺伊曼计算机”的基本概念,被誉为“计算机发展史上的一个里程碑”。



图 1.5 “计算机之父”

冯·诺伊曼(1903—1957)

这些故事告诉我们掌握理论知识,同时对实际问题满怀兴趣,时刻准备向实践专家学习的重要性。改变人类命运的重大发现大多数是由这样的人完成的。

现代计算机的出现和不断加速的性能改进、成本降低已经、正在、将要本质性地改变包括物理学在内的一切科学。计算机使得以前许多难于直接观测的自然现象可以借助于计算机的数据处理或者模拟计算而直接或间接地得到观测,许多难于或根本不可能解析地预测的现象可以通过快速数值计算来预测,许多原来淹没于噪声的现象得以分辨,许多复杂的模型和科学猜测得以验证。并且,许多概念很可能由于计算机的普遍使用和不断改进而发生根本性地改变。伴随着计算机飞速发展的通信技术自然地提出了量度、刻画“信息”的要求。

信息的科学定义几乎与现代计算机的出现同步,而且密切相关。与熵的概念不同,到现在为止,信息的科学定义还没有统一。更多地得到公认的定义是所谓的“香农定义”。

香农(图 1.6)1916年诞生于美国密歇根州。他的大部分时间是在贝尔实验室和麻省理工学院度过的。在二次世界大战时,香农是一位著名的密码破译者。1948年香农长达数十页的论文《通信的数学理论》成了信息论正式诞生的里程碑。2001年香农在马里兰州去世,享年85岁。贝尔实验室和麻省

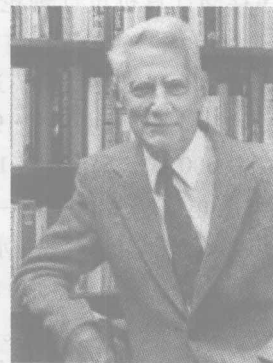


图 1.6 香农(Claude Elwood Shannon)(1916—2001)