

PHYSICAL MODELS OF
LIVING SYSTEMS

生命系统的物理建模

[美] 菲利普·纳尔逊 (Philip Nelson) / 著

舒蛟根 黎明 / 译

上海科学技术出版社

$$E = H p,$$

raisonnement précédent.

$$P = \frac{H p}{S H p};$$

PHILIP NELSON'S PHYSICAL MODELS OF LIVING SYSTEMS

PHILIP NELSON'S PHYSICAL MODELS OF LIVING SYSTEMS

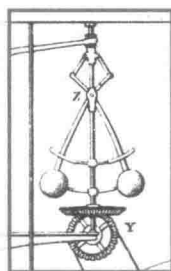
PHILIP NELSON'S PHYSICAL MODELS OF LIVING SYSTEMS

Physical Models of Living Systems

生命系统的物理建模

[美] 菲利普·纳尔逊 (Philip Nelson) / 著

舒咬根 黎 明 / 译



图书在版编目 (CIP) 数据

生命系统的物理建模 / (美) 菲利普·纳尔逊(Philip Nelson)著; 舒咬根, 黎明译. —上海: 上海科学技术出版社, 2018.8

ISBN 978-7-5478-3983-6

I. ①生... II. ①菲... ②舒... ③黎... III. ①生命系统理论-物理学-建模系统

IV. ①Q1-0

中国版本图书馆CIP数据核字(2018) 第081795号

Physical Models of Living Systems by Philip Nelson

First published in the United States

by

W. H. FREEMAN AND COMPANY, New York

Copyright © 2015 by W. H. FREEMAN AND COMPANY

All rights reserved.

上海市版权局著作权合同登记号 图字: 09-2015-357号

生命系统的物理建模

[美] 菲利普·纳尔逊(Philip Nelson) 著

舒咬根 黎明 译

上海世纪出版(集团)有限公司 出版、发行

上海科学技术出版社

(上海钦州南路71号 邮政编码200235 www.sstp.cn)

苏州望电印刷有限公司印刷

开本 787×1092 1/16 印张 24 插页 4

字数 450千字

2018年8月第1版 2018年8月第1次印刷

ISBN 978-7-5478-3983-6/Q·60

定价: 108.00元

本书如有缺页、错装或坏损等严重质量问题, 请向工厂联系调换

献给我的同学：Janice Enagonio, Feng Shechao, Andrew Lange.

来自落日的余晖，
来自大洋和清新的空气，
来自蓝天和人的心灵，
一种动力，一种精神，推动
一切有思想的东西，一切思想的对象，
穿过一切东西而运行。

——威廉·华兹华斯*(William Wordsworth)

*华兹华斯：浪漫主义诗人，英国文学史上最重要的诗人之一。以上选自其作品《延腾寺》（王佐良译本）。——译者注

中文版序

非常高兴中国读者能读到这本《生命系统的物理建模》。感谢舒咬根教授、黎明教授承担了繁杂的翻译工作，在整个翻译过程中我们进行了良好的沟通。我也感谢欧阳钟灿教授对本书出版的周密安排，感谢上海科学技术出版社对本书出版的大力支持。

在英文版面世两年之后，本书讨论的许多问题依然是学术界研究的热点，因此读者仍有很多机会做出重要发现！我建议读者多多探讨这些想法，同时思考如何将之应用于实际情况，这样就能从学习理解的层面逐渐过渡到创新思维的层面。

菲利普·纳尔逊
2017 年于费城

译者序

本书是菲利普·纳尔逊教授继《生物物理学：能量、信息、生命》之后的又一部涉及定量生命科学的力作。前一著作着眼于生物系统分子层面（尤其是单分子层次）的物理学，主要论述分子相互作用的热力学与动力学；本书则强调细胞中各种动态现象的物理建模，主要论述随机过程以及动力系统这两类常用的建模方法。在本书的前半部分，作者介绍了随机过程在遗传突变现象（例如里程碑式的卢里亚-德尔布吕克实验）上的成功应用，同时也详细介绍与之紧密相关的统计推断方法（贝叶斯方法）及其在前沿技术（例如获 2014 年诺贝尔化学奖的超分辨显微镜）上的应用。在本书后半部分，作者用力学类比的方式引入了细胞基因调控中的负反馈、正反馈等概念，并介绍了如何建立确定性的动力学模型来解释细胞的内稳态、状态切换以及周期振荡等重要现象。上述这些概念和方法，对学习生物物理学以及合成生物学的读者来说具有很强的实用性。而作者通过对少数经典实例的透彻分析，详细展示了这些方法的流程和关键细节，为读者提供了一本容易上手的基础教程。

本书与我们之前翻译的《生物物理学：能量、信息、生命（修订版）》（2016，上海科学技术出版社）以及《细胞的物理生物学》（2013，科学出版社）在内容上几乎不重叠，可互为补充，作为从事定量生命科学的读者的系统教材。

本书的翻译得到了多方面的支持。译者要特别感谢欧阳钟灿院士的鼓励和支持；感谢国家重点基础研究计划（973 计划）“纳米生物机器的可控组装与多功能集成”（No.2013CB932804）、国家自然科学基金面上项目（No.11574329, 11774358, 11322543, 11105218, 11675180, 11421063, 11647601, 11675017）、中国科学院前沿重点专项（QYZDY-SSWSYS008）和国家自然科学基金委员会（NSFC）与以色列科学基金会（ISF）联合研究项目（The Joint NSFC-ISF Research Program No.51561145002）的资助；感谢卢禹锟博士参与了第5章和第6章的初译；感谢廖勇（重庆医科大学）、石汉平（首都医科大学）和童雪梅（上海交通大学医学院）三位教授提供了相关医学词汇的精准诠释。

由于本书涉及内容很广，译者的知识和水平所限，误译之处在所难免，敬请读者批评指正。

网页资源

本书英文网站<http://www.physics.upenn.edu/biophys/PMLS/index.html>或者<http://www.macmillanlearning.com/Catalog/product/physicalmodels-of-livingsystems-firstedition-nelson>, 包含下列资源链接:

- *Student's Guide* 介绍了一些可在计算机上运行的数学工具, 并给出了一些计算机实验练习。
- *Datasets* 包含习题使用到的数据集。在正文中的引用方式形如 Dataset 1, 这里的数字 1 就是本网站列表中的序号。
- *Media* 给出了外部媒体(图像、录音和视频)的链接。在正文中的引用方式形如 Media 2, 数字 2 即是指本网站列表中的序号。
- 最后, *Errata* 是勘误表。

致学生

科学教育我们不可迷信专家。

——理查德·费恩曼*(Richard Feynman)

这是一本有关生命系统物理建模的书。如果你能仔细研读本书，那你就能掌握某些建模技巧，并且也能更好地评估各种科学主张，而无需依赖专家。

我们研究的生命系统涉及的尺度从单个大分子一直到完整的有机体。如果你习惯了物理学研究，则初始时会感觉有机体在每个层次的固有复杂程度都是异乎寻常的。例如，相比于单个细胞作出决策时所需要的分子相互作用，描述月球轨道的牛顿方程则简单得不值一提。然而，仔细观察发现月球运动也是复杂的，存在潮汐相互作用、锁相和进动等。为了研究任何复杂系统，我们必须首先建立物理模型，即通过一系列理想化近似把我们的焦点集中在最重要的特征上，从而使问题变得易于处理。

物理建模通常会拿我们理解得比较透彻的系统进行类比。通过类比，我们惊讶地发现，少数几个基本概念就能让我们理解生命科学和物理科学中的大量问题。

物理建模追求的是定量阐明实验数据，目的不只是简单地总结一下，而是通过检验各种竞争模型给出的不同预测来揭示系统隐含的机制。我们之所以坚持定量预测是因为我们通常设想出的卡通图像（无论是草图或仅是文字描述）多半是听起来合理但定量上是失败的。反之，如果模型的数值预测被详细证实，则卡通图像（模型）不太可能是一种侥幸。有时候某些模型能明确预言每次会发生什么，这样的模型可由单个试验检验。可是，通常情况下模型的输出具有概率的特性。本书将介绍概率论的一些关键概念，使我们能够对模型预测及其与实际数据的符合程度作出精确陈述。

好的模型不仅能解释已有数据，而且能外推预言新数据及其走势，这或许是实际建模中最重要的一点。例如，为了揭示模型的缺陷，模型本身就可能提示我们需要对这些数据采取哪些定量的物理干预。一旦找出这些缺陷，该模型还能进一步提示我们如何改进它以便更细致地考虑系统细节或者更真实地描述系统。一旦模型经过足够多的证伪实验的检验，那么这个模型最终就会被认为是“有望成功的”，甚至某天被人们广泛认可。

*又常译为费曼。——译者注

本书将展示一些建模实例，在某些情况下，对于定量数据的物理建模就能使得科学家推断出背后的机制，而其中某些关键的分子在当时是完全未知的。这些案例的研究值得学习，在你进行探索研究时就可以模仿这种模式。

技能

科学不仅仅是供你记忆的一堆事实。你当然需要知道很多事实，本书也将提供一些事实作为案例研究的背景材料，但是你尤其需要技能。你仅仅阅读本书（或任何其他书）是不会获得这种技能的，至少需要完成每章末尾和正文中的一些练习。

本书特别强调：

- 模型构建技能：重点是确定一个恰当的描述层次并写出在该层次上能自圆其说的表达式。（随机性可能是该系统的基本特征吗？模型在量纲分析层次被验证了吗？）在了解其他人的工作时，我们也要重视他们的模型作了哪些假设和哪方面的近似等。
- 融会贯通技能：物理模型通常能揭示出隐藏在看似无关的话题背后的相似性，从而将这些相似性联系起来。一旦人们发现了这种共性，就会导致科学上的重大突破。
- 批判的技能：有时候一个心爱的物理模型被证明是……错的。例如，亚里士多德就曾经主张人脑的主要功能是冷却血液。为了评估更加现代的各种假说，你必须了解原始数据如何向我们传递信息、并进而转化为可理解的知识。
- 计算机技能：特别在研究生物系统时，通常需要进行多次试验，每次试验的结果略有不同。实验数据很快超越了我们数学课程教授的分析工具的处理能力。前不久，当一本类似的教材列举出前人做过的工作时，你还无法亲自动手对它们进行分析，因为当时的计算机尚不具备这样的计算能力。但现在你可以在个人电脑上工业强度级的分析。
- 沟通技能：如果不能有效地传播，即便最伟大的发现也会变得毫无用处。为此，你需要提升沟通技巧。在解答书中习题时，你可以设想自己是在为持怀疑态度的读者准备同行评审报告。你能再花几分钟时间使别人更容易明白你做了什么及为什么要这么做吗？为提高可读性，能否将坐标轴标示得更好些？对代码能否作一些注解？或者能否增加一步证明？你能够预料某些质疑吗？

你确实需要上述这些技能，无论是读原始研究文献，还是解释自己的实验数据，甚至是评估你在报纸上看到的那些由统计和庸统计做出的主张。

还有一项技能值得单独提及，本书的一些习题也许听起来可疑，例如，“评论……”，他们故意这样写其实是让你问“什么是有趣点？值得在这里评论

吗？”对这类问题可能存在多种“正确”的答案，因为也许有趣的事情不止一件。在你自己的科学研究中没有人会告诉你这些问题，因此要养成自己提出这些问题的好习惯。

这些技能将大大增强我们的能力。例如，本书的许多最为有趣的图片在其他地方都没有出现过，你可以利用网站上的数据自己构建这些图片。

计算机能做的

所谓模型其实就是你针对某些观察事实提出的机制。你可以在纸上勾画一个草图来表达这些想法。这类草图可以帮你进一步厘清头绪，更方便地解释给他人听，并做出可被实验验证的预测。

尽管上述传统做法是可行的，但你通常还必须做一些计算，才能给出足够详细的预言，从而检验你的模型。某些时候，一个计算器、一张纸和一支笔就可以轻易完成这些计算步骤。可是，更多的时候，在某一步你需要一个非常快速和精确的助手，计算机可以担当这个角色。

此外，模型做统计预测需要大量的实验数据来验证；有时候你提出的机制中可能包含大量变量，因此需要大量计算；有时候模型验证还涉及模拟系统包含的每个随机变量；为了寻找最佳的拟合值，有时候还必须做多次模拟，每次选用不同的未知参数值。而所有这些事只有计算机才能非常快速地完成。

为了更可靠地使用计算机，你也需要了解一点后台的计算进程，这就要求你自行编写简单的分析程序。本书中有很多习题就是要训练你的编程技能。

最后，你需要理解计算结果，并将结果传达给别人。数据的可视化是一门艺术，它可以忠实和直观地展现数据中蕴含的定量信息。从最简单的 xy 二维曲线到最炫的交互式 3D 图像。计算机比以往任何时候都更快捷和方便地实现了数据的可视化。

本书没有专门安排计算机编程和数据可视化的章节。学生指南包含一个简要介绍，指导教师会帮你寻找适合你使用的平台和其他资源。

计算机做不了的

首先，计算机不擅长提出富有想象力的模型；其次，计算机没有直觉，而这种直觉是建立在与以往的经验作类比的基础上的，这些经验帮助我们鉴别主要因素及其相互作用；再次，计算机不知道哪种预测容易在实验室测量；最后，计算机也不能选择哪种可视化模式能帮你最好地传布你的结论。

最重要的是，计算机不能告诉你何时应该用计算机进行计算，何时用纸笔计算会更好。计算机也不会告知你某种可视化方式会误导别人，或者会被一些不相关的信息所干扰。这些洞察力属于你自己的工作。

本书的结构和特点

- 每章都有“思考题”，这些问题普遍是简单和容易的（尽管不全是）。可是，除了这些简明的问题，大部分公式是前面提到过而又必须由你自己推导的。这么做会大大提高你对问题的理解，尤其是考试遭遇到这类问题时你会轻松很多。
- 很多章内容都以“拓展”节结束，该节主要是为比较优秀的学生准备的，部分内容比正文材料假设了更多的背景知识（有些内容只是更详细而已）。类似的情况发生在注有 $\mathcal{T}$ 字形的脚注和习题中。
- 附录 A 总结了本书一直沿用的数学符号和关键字符。附录 B 讨论了一些有用的解题工具。附录 C 收集了几个供参考的物理常量。
- 很多方程和重要观点都被提出来并给予参考编号。符号“方程x.y”和“要点x.y”均指同一个系列的编号。

其他书籍

本书是帮助那些有志于成为科学家的学者在生命系统的物理建模方面获得一些技能和框架。《生物物理学：能量、信息、生命》介绍了其他方面的内容，这些主题涉及力学与流体力学、熵与熵力、生物电与神经冲动以及力学化学的能量转换。

其他书籍对生物物理的介绍更全面，对本书也是很好的补充。关联较强的书籍包括：

普及类：Ahlborn, 2014; Franklin *et al.*, 2010; Nordlund, 2011.

细胞生物学/生化背景知识：Alberts *et al.*, 2014; Berg *et al.*, 2012; Karp, 2013; Lodish *et al.*, 2012.

医学/生理学：Amador Kane, 2009; Dillon, 2012; Herman, 2007; Hobbie & Roth, 2007; McCall, 2010.

网络：Alon, 2006; Cosentino & Bates, 2012; Vecchio & Murray, 2014; Voit, 2013.

数学背景知识：Otto & Day, 2007; Shankar, 1995.

生物学和物理学中的概率：Denny & Gaines, 2000; Linden *et al.*, 2014.

细胞和分子生物物理学：Boal, 2012; Phillips *et al.*, 2012; Schiessel, 2013.

生物物理化学：Atkins & de Paula, 2011; Dill & Bromberg, 2010.

实验方法：Leake, 2013; Nadeau, 2012.

计算方法：算法：DeVries & Hasbun, 2011; Newman, 2013. 其他计算机技能：Haddock & Dunn, 2011.

纸质书不能像网上资源那样可以随时更新，维基百科(Wikipedia)之类的通用网站拥有很多有益的文献。你也可以从<http://bionumbers.hms.harvard.edu/>获得大量参数的测量值，这些数值在生命系统物理建模时经常用到。

致指导教师

物理学家：“我想研究大脑。告诉我一点有用的知识吧。”

生物学家：“好的，首先，大脑分为两半……”

物理学家：“行了，这些就足够了！”

——阿德里安·帕西金(V. Adrian Parsegian)

本书是我在宾夕法尼亚大学授课数年的讲义，学员主要是大学 2~3 年级理工科学生，他们至少受过一年的物理学和相关数学课程的训练，并对合成生物学、超高分辨显微镜等有所了解，且希望有所作为。

最近很多文献强调未来生命科学和医学的突破有赖于拥有定量研究背景和系统分析经验的研究人员。作为呼应，一系列“数学生物学”、“系统生物学”和“生物信息学”之类的书相继面世。但是，这些书都没有强调物理建模的重要性，有的甚至故意贬低物理建模的意义。本书试图用一些案例来澄清这一点。

本书恪守下列原则*：

- 即便未来不想从事生物物理研究，物理专业的学生也应该对生命体的研究充满激情，因为在此我们将获得许多基本的物理概念；
- 即便未来不想从事生物物理研究，生命科学专业的学生也应该学习基本物理概念，因为这些概念有助于我们阐明生命体的构造和功能及相应的研究仪器。

总之，这是一本物理与生命两学科间如何交叉而又相互促进的教科书。

我们还将达成如下共识：

- 我们尽可能将基本概念与现有经验紧密结合；
- 为了了解现在很多领域使用的各种方法，所有理科学生需要具备一些对概率和推理的直觉，包括似然最大化和贝叶斯建模，还有在本科教育大纲中经常被忽略的其他重要内容，包括：卷积、长尾分布、反馈控制和泊松过程（和其他马尔可夫过程）的概念。

*见《致学生》。

- 算法思维不同于纸笔分析，很多学生在他们目前的学习、研究阶段还没有遇到这类问题，但算法思维对于每个分支学科的日常实践至关重要。最新的报道已经注意到现有课程在这方面的严重不足并提出了改进建议（如：Pevzner & Shamir, 2009; 美国国家研究委员会, 2003），学生越早习惯这一思维模式越好。
- 学生需要在具体案例学习中明确了解“理论的出处”。

本书当然不是要对庞杂的生物物理学领域给出全面深入的阐述。我的目的仅仅是让学生掌握一些在自然科学、工程学、应用数学等方面都能用得上的基本技巧和思考原则，以便理解生物体如何运用它们的某些卓越能力。我将向学生们讲解数量有限但细节充实的一些案例，使他们能自行开展科研水平的分析。案例的选择保持了前后一致的风格，我认为它们为理解当前的研究进展提供了最佳入口。这些案例也适合学生们从头到尾完成所有计算，从而避免在文中出现“某人已经证明……”这类字眼。

选修本课程的学生拥有不同的专业背景，学科跨度很大。对授课老师既是挑战也是机遇，在学科交叉之际就会出现本节题词中的那一幕。我发现只要略施小计就能将不同背景和能力学生集合在一起，开展学科交互，并逐渐形成一种习惯。

本书的使用方法

大多数章节都以“拓展”节结束，某些内容适合于掌握更高级背景知识的学生；另一些是本科层次的讨论，但本书后续内容又不需要以此为基础，根据你和学生们的兴趣可以给予单独讨论。另外，教师指南包括许多其他参考书目，其中一些可能有助于启动基于原始文献的项目。

本书是现代生物物理学的基础教材，也可作为诸如物理学、生物物理学、各类工程学和应用数学等专业课的补充教材。尽管正文是一门本科课程，但它的很多内容超出了一般本科物理课程的范围。因此，只要添加部分或全部“拓展”小节的内容和你自己专业的一些文献（或在教师指南中引用的工作），本书很容易就成了研究生的基础教材。

本书不是我先前那本《生物物理学：能量、信息、生命》的续集，两书几乎没有交集，这部分解释了为什么某些主题不在此处涵盖。还有一些其他主题将出现在即将出版的关于光、成像和视觉的书中。近期有些重叠内容的书，部分列在“致学生”中，另一部分则出现在章节的结尾。

书中材料的编排方案可以有很多种，例如根据生物体系类别或大小尺度。我的思路是通过各种课题的介绍，逐渐建立起一个完整的分析框架，最终能够让读者理解第 11 章将讨论的重要且典型的生物现象。

基于计算机的作业

没有习题的教材与带有习题的教材之间的差别，
类似于学习语言时读和说的差别。

——弗里曼·戴森(Freeman Dyson)

书中的所有习题都已经通过了学生的亲身实践。很多习题要求学生使用计算机。当然，不用计算机也能学到些知识，但我依然认为学生应该学会从头开始写短程序。最好不是在没有相关内容的编程课中学习写短程序，而是在遭遇某些诸如生物物理等具有独特科学意义的问题时去实践。本书的网站收集了对应于习题的实验数据集，很多文献强调了学生使用这类实验数据的重要性（参见美国国家研究委员会，2003）。

为了做研究，学生需要学会数据可视化、随机变量模拟和数据处理，本书习题会涉及所有这些训练。几个通用的编程环境都能满足上述要求，具体取决于个人的喜好，如：Mathematica[®]，MATLAB[®]，Octave，Python，R 或 Sage，有些是免费开放的。当一个完美的数据拟合呈现时是很激动人心的，而早期经常获得这种经历对学生来说很重要。

在我自己的教学班里，许多学生没有编程经验。学生指南提供了一些计算机实习及如何入门的建议。教师指南给出了这些实习的答案，同时也解答了本书的习题和“思考题”。记住：对初学者来说，编程是非常耗时的。在整个学期中，你可能只分派了几个较长的习题，你的学生则可能需要大量的支持。

课堂演示

体验式学习几乎是物理学课程所独有：我们将一套仪器带入教室，给学生展示一些令人惊讶的真实的现象，既不是模拟也不是隐喻。教师指南建议了可以对哪些内容进行课堂演示。

标准申明

这是一本教科书而不是专著。很多精妙的论点被有意放置在“拓展”小节或“教师指南”中，有的甚至干脆没有提及。选择书中描述的实验仅仅是为了服务于我想阐明的观点。书中对原始文献的引用比较随意，书中的任何内容都并非我个人的原创。我也不保证书中所述历史的完整性。

目 录

前言：HIV研究的突破得益于学科交叉	1
I 预备知识	7
1 病毒动力学	9
1.1 导读	9
1.2 HIV 感染过程建模	9
1.2.1 生物背景	10
1.2.2 恰当的图表有利于揭示数据的关键特征	11
1.2.3 鉴别系统主因及其主要相互作用是物理建模的第一步	12
1.2.4 数学分析可以预测一系列行为	14
1.2.5 大部分模型都必须适用于数据拟合	15
1.2.6 过约束与过拟合	16
1.3 有关建模的几句忠告	17
总结	18
拓展	21
习题	23
2 物理学与生物学	27
2.1 导读	27
2.2 交叉	28
2.3 量纲分析	29
总结	29
习题	31
II 生物学的随机性	33
3 离散型随机性	35
3.1 导读	35
3.2 随机性事例	36