

XITONG SHENGWUXUE

系统生物学

—— 建模，分析，模拟

雷锦誌 著

上海科学技术出版社

科学出版社 中国科学院植物研究所 中国科学院微生物研究所

系统生物学

——建模、分析、模拟

王竹溪 著

科学出版社

ISBN 7-03-018888-8

系统生物学

——建模,分析,模拟

雷锦诤 著

上海科学技术出版社

图书在版编目 (C I P) 数据

系统生物学: 建模, 分析, 模拟 / 雷锦诒著. —上海:
上海科学技术出版社, 2010. 9
ISBN 978 -7 -5478 -0487 -2

I. ①系… II. ①雷… III. ①生物学 - 系统科学
IV. ①Q111

中国版本图书馆 CIP 数据核字 (2010) 第 157097 号

上海世纪出版股份有限公司 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)
新华书店上海发行所经销
常熟市兴达印刷有限公司印刷
开本 787 × 1092 1/16 印张: 13
字数: 283 千字
2010 年 9 月第 1 版 2010 年 9 月第 1 次印刷
ISBN 978 -7 -5478 -0487 -2/N · 5
定价: 50.00 元

本书如有缺页、错装或损坏等严重质量问题,
请向工厂联系调换

内 容 提 要

系统生物学是综合了生物、物理和数学等学科在内的交叉学科. 本书是系统生物学的入门教材, 结合作者近年的研究工作和国外一些新的研究进展, 介绍系统生物学这门新兴学科的研究对象和方法, 尽量去繁就简, 重点介绍对复杂的生物过程建立数学模型的方法, 并介绍相关的分析和模拟方法. 涉及基因表达、基因调控、生物振荡、状态切换、形态发生素梯度浓度、造血系统和霍奇金—赫胥黎方程(Hodgkin-Huxley equation)等生物学知识和随机过程、常微分方程、偏微分方程、时滞微分方程、随机微分方程、分岔分析等数学内容.

适用于生物、医学、数学、物理、化学、生物信息学、生物物理学等方向的学生、教师和相关研究人员.

前 言

在 20 世纪,生物学的研究内容和方法经历了由宏观到微观的发展过程,由形态及表型的描述逐步到生物体内的各种分子及其功能的动态演变. 20 世纪 50 年代克里克(F. Crick)和沃森(J. Watson)发现 DNA 的双螺旋结构,开启了分子生物学时代;从 70 年代起出现的基因工程技术加速了分子生物学的发展;1990 年人类基因组计划开启了生命科学的新时代;2003 年完成了人类及一些模式生物的基因测序工作,把生物学带入了系统生物学的时代. 根据美国科学院院士、首个系统生物学研究中心的创建者之一胡德(L. Hood)博士的定义,系统生物学是研究一个生物系统中所有的组成部分的构成,以及在特定的条件下这些组成部分相互关系的学科.

系统生物学的一个显著特点是其“系统”性. 所有生物现象,例如糖分子的代谢、人体心脏的跳动、对入侵病毒的免疫响应等等,都是复杂系统的综合行为的结果. 因此,为了了解这些复杂系统,我们必须把生物系统作为一个整体来研究,而不仅仅把研究重点放在单个的基因、蛋白质或者器官上. 在对生物系统的研究的过程中,仅通过实验方法很难了解复杂系统的行为. 而通过建立数学模型和计算机模拟的方法是了解复杂系统的行为,特别是某些动力学性质的重要手段,也是当前系统生物学的主要研究方法之一.

作者于 2004—2005 年期间分别访问了美国加州大学尔湾分校(UC Irvine)数学系和加拿大麦吉尔大学(McGill University)生理学系非线性动力学中心,开始从事系统生物学的研究工作. 从 2006 起作者在清华大学开设系统生物学选修课程以介绍自己所学习到的相关内容和研究方法. 本书作为这门课程的参考资料,主要目的是向读者介绍系统生物学这门新兴学科的研究对象和方法. 本书的主要内容结合作者的学习体会,从一个数学工作者的背景来介绍相关的生物学知识,尽量去繁就简,重点介绍对复杂的生物过程建立数学模型的过程,并介绍相关的分析和模拟方法. 每个看似简单的生物学过程,例如基因表达等,实际上都是非常复杂的过程. 本书在对生物学过程的介绍中尽量避免烦琐的过程描述和专业术语,而把重点放在生物学过程基本原理和数学的描述上. 这样的介绍更加突出建模的方法,而会略去一些细节的描述. 如果读者希望了解详细的生物学过程,建议读者参考专门的生物学方面的参考书. 另一方面,本书在介绍数学方

法时也有别于普通的数学参考书.在本书中将不会建立成体系的逻辑严密的数学理论,而是根据实际问题的需要介绍模型的分析方法.对相关数学理论感兴趣的读者请参考专业文献.

本书的内容一共包括六章.第1章将介绍生物化学反应的数学描述.在分子生物学中,特别是基因调控和蛋白质相互作用等生物学过程,都可以分解为大量生化反应所组成的复杂系统.这一章中所介绍的相关数学描述方法是后面建立数学模型的基础.从第2章到第5章将按照从小到大的尺度介绍几类生物学问题,从问题出发介绍相关的数学分析方法.第2章主要介绍基因表达的数学描述.基因表达是生物系统的最基本过程.这一章将主要介绍基因表达的过程和相应的数学描述,重点了解基因表达过程中的随机性.第3章介绍几种基因调控关系的数学模型的建立和分析,包括双稳态、状态切换、生物振荡等.第4章介绍形态发生素浓度梯度形成的数学模型.形态发生素是发育生物学的重要概念,其浓度梯度的形成对于胚胎的发育至关重要.这一章介绍通过反应扩散方程建立描述形态发生素浓度梯度形成的数学模型.第5章介绍造血系统的数学模型的建立,参数估计与相关的分析方法,为研究动态血液病的致病机制建立基础.第6章介绍霍奇金—赫胥黎方程的建立.霍奇金—赫胥黎方程的建立是在生命科学领域实验与数学方法相结合的最成功典范之一.这一章将回顾霍奇金(A. L. Hodgkin)和赫胥黎(A. F. Huxley)的部分工作,沿着这两位大师揭开细胞兴奋性之谜的道路,在实验与理论相结合的过程中介绍建立霍奇金—赫胥黎方程的过程.

本书涉及的内容比较广,各章节的内容相对比较独立.建议读者在阅读时遇到不熟悉的内容可以大胆跳过,而需要理解详细过程时参考相关的专业文献.

本书的内容是在作者为清华大学研究生和高年级本科生开设系统生物学课程的教学讲稿的基础上修订而成.由于国内相关的教材非常缺乏,作者在开设本课程的过程中需要查阅很多文献资料.其中有些资料来源于最新发表的论文,还包含一些最新研究成果.书中的大部分公式都可以在有关的文献中找到,作者也重新绘制了相关的插图.限于作者水平,难免有不当或疏漏之处,恳切希望得到批评指正.

本书的相关内容还曾经在东北师范大学、哈尔滨工业大学、华南师范大学、北京师范大学、中山大学做过讲学.在此感谢所有听过我的课的学生,感谢范猛教授、刘胜强教授、翁佩萱教授、黎雄教授、袁荣教授和周天寿教授等的邀请.本课程的开设和讲义的编写还得到清华大学周培源应用数学研究中心的谢定裕、章梅荣、雍稳安、孙卫涛和来月英等同事的支持和帮助,在此深表谢意.最后,在本书即将出版之际,我对清华大学周培源应用数学研究中心名誉主任林家翘教

授致以真切的感谢。是林家翘教授的指导和鼓励使我从一名数学工作者走上系统生物学的研究道路的。林家翘教授对科学执着的追求和严谨的态度深深地影响着我此后的研究工作。

本书的部分研究内容得到自然科学基金(NSFC,10601029)和留学回国人员启动基金的资助,在此表示感谢。

最后感谢妻子和女儿对我的支持和帮助。没有她们的支持我不可能顺利完成对加州大学尔湾分校和麦吉尔大学的访问,开始对系统生物学的研究工作。

北京清华园

2010年3月

目 录

第 1 章 生物化学反应的数学描述	1
§ 1.1 生化反应系统	1
§ 1.2 化学主方程	3
1.2.1 方程的建立	3
1.2.2 方程的性质	4
1.2.3 吉莱斯皮算法	6
§ 1.3 化学速率方程	8
1.3.1 方程的建立	8
1.3.2 涨落耗散定理	10
§ 1.4 化学朗之万方程	13
1.4.1 方程的建立	13
1.4.2 随机积分的简单讨论	17
1.4.3 τ 跳跃算法	18
§ 1.5 福克尔—普朗克方程	19
§ 1.6 反应速率随时间变化的生化反应系统	20
1.6.1 外部噪声干扰下的反应速率	21
1.6.2 推广的化学朗之万方程	22
1.6.3 伊藤积分与斯特拉托诺维奇积分	23
1.6.4 有色噪声	24
§ 1.7 小结	24
习题	25
第 2 章 基因表达的数学描述	27
§ 2.1 遗传信息的传递与基因表达	27
§ 2.2 基因表达的内蕴随机性	29
2.2.1 模型的建立	29
2.2.2 平衡态	30
2.2.3 静态涨落	31
2.2.4 内蕴随机效应	33
§ 2.3 基因表达中的外部噪声	34
2.3.1 模型的建立	35

2.3.2 平衡态.....	37
2.3.3 静态涨落.....	38
2.3.4 外部噪声对基因表达的影响.....	40
§2.4 小结.....	42
习题	43
第3章 基因调控的数学模型	44
§3.1 数学基础.....	45
3.1.1 尺度分析.....	45
3.1.2 米氏函数和希尔函数.....	49
3.1.3 洛姆周期图.....	54
§3.2 正反馈调控与双稳态.....	55
3.2.1 乳糖操纵子.....	56
3.2.2 数学模型.....	57
3.2.3 平衡态分析.....	58
§3.3 噪声与细胞状态的切换.....	60
3.3.1 乳糖操纵子基因的状态切换.....	60
3.3.2 λ 噬菌体阻抑物基因的表达调控	62
3.3.3 内部噪声诱导的状态切换.....	68
§3.4 负反馈调控和生物振荡.....	71
3.4.1 阿特金森振子.....	72
3.4.2 随机激励振子.....	76
3.4.3 带时滞的负反馈调控.....	82
3.4.4 节律振荡.....	85
§3.5 小结.....	95
习题	97
第4章 信号分子浓度梯度形成的数学模型	98
§4.1 反应扩散方程的建立和模拟.....	98
4.1.1 一维守恒率方程.....	98
4.1.2 流的不同形式.....	99
4.1.3 初边值条件	100
4.1.4 高维守恒率方程	101
4.1.5 反应扩散方程的数值解	102
§4.2 形态发生素与胚胎的发育	103
4.2.1 形态发生素	103
4.2.2 Decapentaplegic 与果蝇翅膀的发育	104
§4.3 形态发生素的扩散与数学模型的建立	105
4.3.1 模型 A: 自由扩散和与受体的结合	106

4.3.2 模型 B: 自由扩散, 与受体结合, 信号分子的降解	108
4.3.3 模型 C: 配体与受体的结合和解离, 通过复合体的扩散	114
§4.4 小结	116
习题	117
第5章 造血系统的数学模型	119
§5.1 一些数据	119
§5.2 造血干细胞数量变化的数学模型	120
5.2.1 细胞周期与休眠期	120
5.2.2 数学模型的建立	121
5.2.3 反馈调控函数	123
5.2.4 参数估计	125
§5.3 干细胞模型的动力学分析	128
5.3.1 无量纲化方程	129
5.3.2 平衡态分析	129
§5.4 周期性白细胞减少症的动力学模型	131
5.4.1 模型介绍	132
5.4.2 参数估计	133
5.4.3 平衡点的存在性	134
5.4.4 平衡点稳定性与分岔分析	135
5.4.5 粒细胞集落刺激因子治疗	137
§5.5 小结	138
习题	139
第6章 霍奇金—赫胥黎方程	141
§6.1 离子通道与能斯特方程	141
§6.2 细胞膜模型	143
§6.3 离子通道的门控机制	145
6.3.1 门控机制的数学描述	145
6.3.2 莫里斯—莱卡尔模型	147
§6.4 霍奇金—赫胥黎方程	148
6.4.1 实验结果	149
6.4.2 离子通道的门控假设	149
6.4.3 方程的建立	150
§6.5 电缆方程和神经网络动力学方程	153
6.5.1 电缆方程	153
6.5.2 神经网络动力学方程	155
§6.6 小结	155
习题	156

附录 A 常微分方程	158
§ 1.1 常微分方程模型	158
§ 1.2 二阶微分方程	160
§ 1.3 二阶常微分方程边值问题的数学基础	161
附录 B 随机微分方程	168
§ 2.1 随机微分方程与随机积分	168
§ 2.2 伊藤公式	170
§ 2.3 福克尔—普朗克方程	171
§ 2.4 随机微分方程数值方法	174
2.4.1 1.0 阶差分格式	174
2.4.2 玛尔萨利亚随机数发生器	175
附录 C XPPAUT 软件使用介绍	177
§ 3.1 建立 ODE 文件	177
§ 3.2 运行和退出程序	180
§ 3.3 保存结果	180
§ 3.4 相平面分析	181
§ 3.5 分岔分析	182
§ 3.6 通过脚本语言运行 XPPAUT	183
参考文献	186
索引	193

第 1 章 生物化学反应的数学描述

系统生物学主要在分子层次研究生命行为的机制. 这些分子行为包括蛋白质 (protein)—蛋白质相互作用、蛋白质—DNA 相互作用等, 构成生命行为的最基本元素. 例如, 基因表达这一最基本的生命现象就包括基因 (DNA 片段) 转录成信使 RNA (mRNA) 和 mRNA 翻译成蛋白质等过程. 而在分子层次, 生物化学反应是最基本的表现方式. 这一章将介绍生物化学反应的几种基本数学描述方法. 这一章所涉及的数学内容比较抽象, 不熟悉相关内容的读者只需要了解不同数学描述的形式, 而忽略细节的推导过程.

常微分方程是描述化学反应的常用数学工具. 一般地, 普通化学反应包含的分子的个数非常多 (约 10^{23}), 因此使用常微分方程可以得到很准确的描述. 然而, 对于生命系统中所发生的生物化学反应, 常微分方程的描述有时候是不合适的. 这是因为生命行为中的生物化学反应所涉及的分子的个数通常都非常少. 例如, 在基因表达中, 同一种蛋白质所对应的基因的个数通常只有 1 个, 而对应的 mRNA 分子的数目通常只有几十个. 在蛋白质相互作用中, 作为反应物的同一种蛋白质的个数也不过几千个, 远远小于普通化学反应的分子个数的量级. 另一方面, 生物体内发生化学反应所需的时间都比较长. 例如, 基因的转录过程通常需要几分钟才能完成. 而因为分子的个数少而且反应比较慢, 发生化学反应的随机性就非常明显. 这些随机性是由两方面的原因造成的, 一方面是因为反应物碰撞后才可能发生反应. 而当分子个数很小 (浓度很低) 时, 反应物的碰撞概率是非常小的. 另一方面是热力学涨落. 即使参加反应的反应物碰撞在一起后, 也需要有足够大的活化能才能发生相应的反应. 而活化能是受热涨落的影响的, 具有显著的随机性. 因此, 随机过程是描述生物化学反应的重要数学手段.

§ 1.1 生化反应系统

考虑有 $N \geq 1$ 种分子 $\{S_1, \dots, S_N\}$ 参加的化学反应. 假设系统的温度恒定, 所有分子充分混合在一个体积 (记为 Ω) 固定的容器中, 并且所有反应一旦发生都可以在瞬间完成. 以 $X(t) = (X_1(t), \dots, X_N(t))$ 表示该系统在时刻 t 的状态, 其中

$X_i(t)$ 表示系统中分子 S_i 在时刻 t 的个数 ($i = 1, \dots, N$).

显然地, 系统的状态 $X(t)$ 是一个随机过程, 这是因为每个化学反应实际发生的时刻是随机的, 受热涨落的影响.

假设所考虑的系统中共有 $M \geq 1$ 个基本反应通道 $\{R_1, \dots, R_M\}$. 每个反应通道 R_j 对应于一个趋向性函数 (propensity function) a_j , 定义为

$a_j(x)dt$ 表示给定 $X(t) = x$ 时, 反应 R_j 在时间区间 $[t, t+dt)$ 内, 在容器中某处发生一

次的概率.

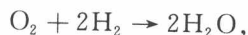
每次化学反应都会引起分子个数的改变. 以状态改变向量 (state-change vector) $v_j = (v_{j1}, \dots, v_{jN})$ 表示反应 R_j 所引起的系统状态的改变:

v_{ji} 表示分子 S_i 因为反应 R_j 所引起的改变量 ($j = 1, \dots, M; i = 1, \dots, N$). 即反应 R_j 发生后, 分子 S_i 的个数由 $X_i(t)$ 变为 $X_i(t) + v_{ji}$.

这里 $v_{ji} > 0$ 表示反应 R_j 产生分子 S_i , $v_{ji} < 0$ 表示反应 R_j 消耗分子 S_i .

由反应趋向性函数 a_j 和状态改变向量 v_j 一起就可以给出反应通道 R_j 的完整描述. 根据这些量, 可以建立数学模型, 用于描述所研究的生化反应系统.

一般地, 状态改变向量 v_j 可以通过化学反应本身得到, 例如对于氢气燃烧生成水的化学反应



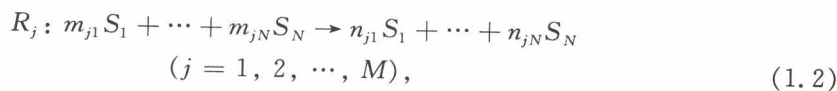
每次反应消耗一个氧原子、两个氢原子, 生成两个水分子. 因此此反应对应的状态改变向量是 $v = (-1, -2, 2)$.

反应趋向性函数 a_j 的准确描述比较困难, 一般情况下只能得到近似的描述. 函数 a_j 通常包括两部分的乘积: 所有参加反应的反应物按反应 R_j 所需比例碰撞在一起的概率乘以反应物克服化学势垒发生化学反应的概率. 前一部分可以通过组合的方法得到, 而后一部分由参加反应的反应物发生反应需要克服的化学势所定义. 例如, 对于上面的水合成的反应, 如果系统中有 x_1 个氧分子和 x_2 个氢分子, 并假设这些分子充分混合. 则一个氧分子和两个氢分子碰撞在一起的概率与从系统中任意取三个分子, 组合为一个氧分子和两个氢分子的组合数, 即 $x_1 x_2 (x_2 - 1)/2$ 成正比. 而当两个氢分子和一个氧分子发生碰撞后, 若要发生化学反应生成水, 还需要打开原有的共价键和产生新的共价键. 这一过程需要克服一定的能量势垒. 如果所需要克服的能量势垒为 $\Delta\mu$, 则克服能量势垒发生化学反应的概率与 $e^{-\Delta\mu/(k_B T)}$ 成正比, 这里 k_B 为玻耳兹曼常数, T 表示绝对温度. 这两部分的乘积就给出了由系统状态定义的该反应所对应的反应趋向性函数 $a(x) = c x_1 x_2 (x_2 - 1)/2$, 其中 c 为常数, 通常与温度和体积 Ω 有关. 当系统的温度随时间变化时, 比例系数也是时变的. 一般地, 趋向性函数 $a_j(x)$ 可以按以下形式定义:

$$a_j(x) = c_j h_j(x). \quad (1.1)$$

其中 c_j 是反应通道 R_j 的化学反应常数, 函数 $h_j(x)$ 表示在状态 $X(t) = x$ 并充分混合时, 反应物按反应通道 R_j 所需的比例进行组合的可能组合数.

一般地, 对于化学反应系统



有

$$v_{ji} = n_{ji} - m_{ji}, \quad a_j(x) \propto \prod_{k=1}^N C_{x_k}^{m_{jk}}, \quad (1.3)$$

其中

$$C_{x_k}^{m_k} = \frac{x_k!}{m_{j_k}!(x_k - m_{j_k})!}$$

表示从 x_k 个 S_k 分子中随机选取 m_{j_k} 个分子的可能的组合数. 如果对任意的 k 和 j , 都有 $x_k \gg m_{j_k}$, 可以近似地取

$$a_j(x) = c_j \prod_{k=1}^N x_k^{m_{j_k}}.$$

这里的常数 c_j 只能通过实验方法来确定. 对于很多复杂的生化反应, 如何确定对应的化学反应常数还是非常困难的. 需要注意的是, 在这里以分子个数表示系统的状态, 如果以浓度或者摩尔数表示系统的状态, 化学反应常数需要做相应的修正.

对于生物化学反应系统, 时刻都可能有反应发生, 因此系统状态的改变是一个复杂的动态过程. 下面首先引入描写该过程的最基本和最本质的数学描述——化学主方程 (chemical master equation), 然后在不同假设条件下分别引入化学速率方程 (chemical rate equation), 化学朗之万方程 (chemical Langevin equation) 和福克尔—普朗克方程 (Fokker-Planck equation).

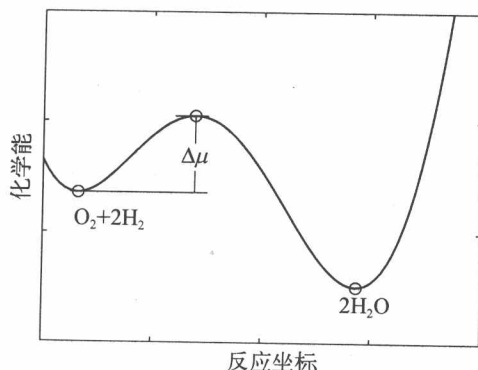


图 1.1 氢气燃烧生成水的化学反应的化学势

§ 1.2 化学主方程

1.2.1 方程的建立

生物化学反应是一个随机过程. 在任一时刻, 系统的状态决定了该系统下一时刻发生某个化学反应的概率, 而每个化学反应的发生都可以改变系统的状态. 因此, 从数学上可以把生物化学反应看成是状态随时间变化的马氏过程.

假设系统在时刻 t_0 的状态为 $X(t_0) = x_0$, 定义条件概率密度函数 $P(x, t | x_0, t_0)$ 为系统在 t 时刻状态为 $X(t) = x$ 的概率, 即

$$P(x, t | x_0, t_0) = \text{Prob}\{X(t) = x, \text{如果 } X(t_0) = x_0\}. \quad (1.4)$$

为了理解概率密度函数, 假设有许多独立的反应系统 (也称为系综), 每个反应系统都包含相同的分子, 并且体积、温度和初始状态都相同. 则在时刻 t , 这些系统的状态 (分子的个数) 可以是任意值. 概率密度函数 $P(x, t | x_0, t_0)$ 由在时刻 t 状态为 x 的系统的个数所占的百分比给出.

如果已知系统在时刻 t_0 的状态和任意时刻的概率密度函数 $P(x, t | x_0, t_0)$, 可以得到整个反应系统在任意时刻的统计性质. 例如, 系统的状态在任何时刻 t 的平均值 (期望) 由

$$\langle X \rangle = \sum_{x \in \mathcal{X}} x P(x, t | x_0, t_0)$$

给出, 其中 $\mathcal{X}$ 为系统的所有可能状态的集合. 需要特别指出的是, 在实验中所看到的通常是

系统的统计性质,而不是单独某个样本的行为.

下面,来推导 $P(x, t | x_0, t_0)$ 随时间演化所满足的方程.

对给定的 (x_0, t_0) , 把 $P(x, t | x_0, t_0)$ 看成是 (x, t) 空间中的概率流, 下面来推导这个概率流随时间的演化方程. 取 dt 充分小使得在时间 dt 内发生两次或者更多次化学反应的概率可以忽略. 则在 dt 时间区间内系统的概率密度函数的改变量主要是因为发生(一次)化学反应而导致系统状态改变引起的. 因此, 函数 $P(x, t | x_0, t_0)$ 随时间的变化可以通过下面关系来刻画:

$$P(x, t + dt | x_0, t_0) - P(x, t | x_0, t_0) = \text{概率流在}(t, t + dt)\text{内流入}(x, t)$$

$$- \text{概率流在}(t, t + dt)\text{内从}(x, t)\text{流出.}$$

特别地, 如果某个系统在 t 时刻的状态为 $(x - v_j)$, 并且在 t 到 $t + dt$ 时间内发生了一第 j 个反应, 则系统的状态在 $(t + dt)$ 时刻就变成 x , 相应地, 系综在 $(t + dt)$ 时刻的状态为 x 的概率增加了. 由前面给出的反应趋向性函数和状态改变向量的定义可以得到

$$\text{概率流在}(t, t + dt)\text{内流入}(x, t) = \sum_{j=1}^M P(x - v_j, t | x_0, t_0) a_j(x - v_j) dt.$$

类似地, 如果某个系统在 t 时刻的状态为 x , 并且在 t 到 $t + dt$ 时间内发生了一第 j 个反应, 则系统的状态在 $(t + dt)$ 时刻就变成 $x + v_j$, 相应地, 系综在 $(t + dt)$ 时刻的状态为 x 的概率减少了. 因此, 可以得到

$$\text{概率流在}(t, t + dt)\text{内从}(x, t)\text{流出} = P(x, t | x_0, t_0) \sum_{j=1}^M a_j(x) dt.$$

这样, 可以得到下面关系:

$$\begin{aligned} P(x, t + dt | x_0, t_0) - P(x, t | x_0, t_0) &= \sum_{j=1}^M P(x - v_j, t | x_0, t_0) a_j(x - v_j) dt \\ &\quad - P(x, t | x_0, t_0) \sum_{j=1}^M a_j(x) dt. \end{aligned}$$

两边除以 dt , 并令 $dt \rightarrow 0$, 就得到化学主方程

$$\frac{\partial}{\partial t} P(x, t | x_0, t_0) = \sum_{j=1}^M [a_j(x - v_j) P(x - v_j, t | x_0, t_0) - a_j(x) P(x, t | x_0, t_0)]. \quad (1.5)$$

1.2.2 方程的性质

方程(1.5)从本质上反映了我们所研究的化学反应系统. 如果可以求解出 P , 则可以完整地刻画随机过程 $X(t)$. 事实上, 方程(1.5)是一组线性方程组. 如果省略初始状态 (x_0, t_0) , 并且记 $P(x, t)$ 为 $P_x(t)$, 其中 x 的取值范围为系统所有可能的状态的集合, 记为 $x \in \mathcal{X}$, 则方程(1.5)可以改写为线性方程组

$$\frac{dP_x(t)}{dt} = \sum_{j=1}^M [a_j(x - v_j) P_{x-v_j}(t) - a_j(x) P_x(t)] \quad (x \in \mathcal{X}). \quad (1.6)$$

注意到如果反应趋向性函数 a_j 不依赖于时间 t , 则式(1.6)是一组常系数线性方程组. 因此, 常系数线性微分方程组的研究方法和结果都可以用于研究化学主方程(1.6). 特别地, 令 A 表示系数矩阵, 即

$$A_{x, x-v_j} = a_j(x - v_j), \quad A_{x, x} = -\sum_{j=1}^M a_j(x) \quad (\forall x \in \mathcal{X}).$$

以上方程组可以改写为

$$\frac{dP_x(t)}{dt} = AP_x(t) \quad (x \in \mathcal{X}). \quad (1.7)$$

有时也称矩阵 A 为状态转移矩阵.

这里关心的是方程(1.7)的作为概率密度函数的解, 即满足条件

$$P_x(t) \geq 0, \quad \sum_{x \in \mathcal{X}} P_x(t) = 1 \quad (\forall t > 0) \quad (1.8)$$

的解. 根据条件(1.8)容易得到 $P_x(t) \leq 1$ 对任意 $t > 0$ 都成立, 即解总是有界的.

如果 $\mathcal{X}$ 是有限集, 则方程(1.7)的解可以表示为

$$P_x(t) = e^{A(t-t_0)} P_x(t_0) \quad (t > t_0), \quad (1.9)$$

其中初值满足条件(1.8). 特别地, 如果 $X(t_0) = x_0$, 则有

$$P_x(t_0) = \delta_{x, x_0} = \begin{cases} 0, & x \neq x_0 \\ 1, & x = x_0 \end{cases}.$$

容易看到如果初值满足条件(1.8), 则方程(1.7)的解在任何时刻均满足式(1.8).

根据线性代数中著名的若尔当标准型定理, 对于有限阶矩阵 A , 存在一个非奇异矩阵 Q , 使 $Q^{-1}AQ = J$ 为若尔当标准型, 即 J 为下面的分块对角矩阵

$$J = \text{diag}(J_1, \dots, J_s) := \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix}.$$

其中

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \ddots & 0 \\ 0 & 0 & \lambda_k & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_k \end{pmatrix}$$

为 $n_k \times n_k$ 阶矩阵. 这里 λ_k 是矩阵 A 的特征值, 重数为 n_k . 因为 $A = QJQ^{-1}$, 所以

$$e^{A(t-t_0)} = Qe^{J(t-t_0)}Q^{-1} = Q\text{diag}(e^{J_1(t-t_0)}, \dots, e^{J_s(t-t_0)})Q^{-1}.$$