

# 源自统计物理的数学论题 ——统计物理与概率的交叉

献给严士健教授 90 华诞

陈木法

北京师范大学数学科学学院, 北京 100875

E-mail: [mfchen@bnu.edu.cn](mailto:mfchen@bnu.edu.cn)

收稿日期: 2017-12-27; 接受日期: 2018-03-07; 网络出版日期: 2019-08-30

国家自然科学基金 (批准号: 11771046) 资助项目

**摘要** 从数学上讲, 本文介绍从 Markov 链到非平衡交互作用粒子系统的发展; 从物理上讲, 这是从平衡态统计力学到非平衡态统计物理的新进程. 本文介绍从中开发出来的几个典型成果和数学工具, 包括平衡态与非平衡态模型的判别, 非平衡态模型的数学构造、基本性质研究及概率度量和耦合方法等.

**关键词** Markov 链 交互作用粒子系统 反应扩散过程 平衡态与非平衡态

**MSC (2010) 主题分类** 60J27, 60K35

## 1 引言

本文是作者 2017 年 10 月在 “International Conference on Spatial Probability and Statistical Physics” (中国科学院数学与系统科学研究院) 的学术报告的第一部分.

故事发生在 40 多年前, 当时的数学界开始出现分流, 一部分人离开公理化路线, 开始回归自然. 作为代表, 有苏联的 Dobrushin 学派和美国的 Spitzer 学派. 前者于 1965 年开始探索统计力学的数学基础, 形成随机场的概率与统计物理的交叉学科; 后者于 1970 年左右提出交互作用粒子系统. 这两方面的研究均源于探索统计物理的数学基础. 稍晚些出现现已众所周知的几何与量子场论的交叉.

统计物理可分为平衡态和非平衡态两部分. 前者属经典, 后者大约初步形成于 1977 年.

下面的研究路线图 (见图 1) 大体上反应了我们最初二十年的研究历程.

本文主要介绍此图中的三部分:

- (1) 粒子系统、随机场  $\rightarrow$  平衡态、非平衡态统计物理;
- (2) Markov 链、跳过程  $\rightarrow$  存在唯一、遍历性  $\rightarrow$  反应扩散过程;

英文引用格式: Chen M F. Mathematical topics motivated from statistical physics—Interaction of statistical physics and probability theory (in Chinese). *Sci Sin Math*, 2020, 50: 1–12, doi: [10.1360/N012017-00270](https://doi.org/10.1360/N012017-00270)

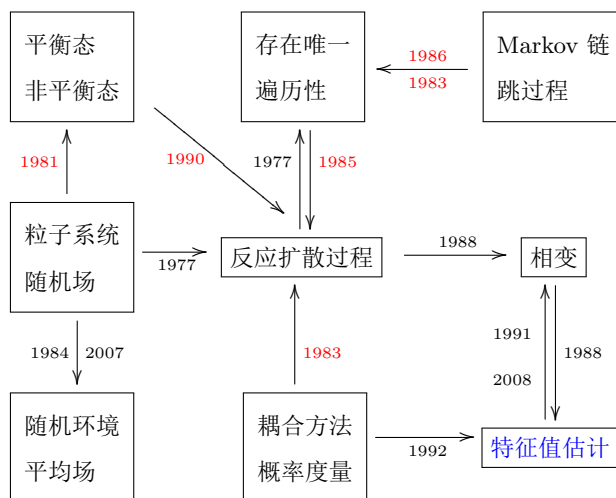


图 1 研究路线图

(3) 耦合方法、概率度量 → 反应扩散过程.

而将“耦合方法、概率度量 → 特征值估计 ↔ 相变”留到本文的姐妹篇 [1] 中讨论.

当年我们进入统计物理的背景有三个. 其一是非平衡统计物理的研究刚刚兴起, 北京师范大学一批老师成立了一个跨学科的“量子力学小组”, 严士健老师是成员之一. 他和李占柄的论文 [2] 是国内此方向最早的工作之一. 其二是美籍华裔学者钟开莱教授来华讲学时介绍了苏联概率论与平衡态统计物理交叉的新的研究方向. 其三是之前我跟随侯振挺老师学习 Markov 链, 有一些积累. 我们自然期望以概率论为工具, 探索非平衡统计物理的数学基础. 然而, 开局时就遇到大障碍. 对于局部系统, 其数学描述为有限维 Markov 链. 然而, 即使对于这种 Markov 链的唯一性的第一个门槛, 我们也经过五年之后才跨过. 这是本文后半部分的故事.

我们进入统计物理的第一步是从数学上区分平衡态与非平衡态.

## 2 平衡态与非平衡态—对称与非对称算子

描述一个物理系统常使用微分算子或矩阵算子. 物理上的平衡态在数学上对应于某个  $L^2(\mu)$  空间上的对称算子. 例如, 给定一个可数集  $E$  上的  $Q$  矩阵  $Q = (q_{ij} : i, j \in E)$  (非对角线元素非负, 行和为零). 当  $j \neq i$  时,  $q_{ij}$  表示从  $i$  跳到  $j$  的速率. 此时, 称  $Q$  在某个  $L^2(\mu)$  空间上对称 (自共轭), 如果关于其上的内积  $(\cdot, \cdot)$ , 有

$$(f, Qg) = (Qf, g), \quad f, g \in \mathcal{D}(Q) \subset L^2(\mu), \quad (2.1)$$

其中  $\mathcal{D}(Q)$  为算子  $Q$  的定义域. 若  $E$  有限, 则可忘掉这些技术性细节. 问题是, 如何判断一个矩阵  $Q$  可配称? 即存在一个正测度  $\mu$ , 使得  $Q$  在  $L^2(\mu)$  上对称 (等价地, 该物理系统属于平衡态). 本节的目的就是要回答这个问题.

首先, 在 (2.1) 中取  $f$  和  $g$  为单点的示性函数, 立即得出

$$\mu_i q_{ij} = \mu_j q_{ji}, \quad j \neq i. \quad (2.2)$$

事实上, (2.2) 等价于 (2.1) (参见文献 [3, 第 6 章]). 由此看出, 对于对称性,  $Q$  必须满足

$$\text{共零性: } q_{ij} > 0 \Leftrightarrow q_{ji} > 0, \quad i \neq j. \tag{2.3}$$

为简单记, 当  $i \neq j$  时, 以  $i \rightarrow j$  表示  $q_{ij} > 0$ . 假定

$$i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_n,$$

称之为从  $i$  到  $j$  的一条路. 那么, 由 (2.2) 依次得出

$$\mu_{i_0} \frac{q_{i_0 i_1}}{q_{i_1 i_0}} = \mu_{i_1}, \quad \mu_{i_0} \frac{q_{i_0 i_1}}{q_{i_1 i_0}} \cdot \frac{q_{i_1 i_2}}{q_{i_2 i_1}} = \mu_{i_1} \frac{q_{i_1 i_2}}{q_{i_2 i_1}} = \mu_{i_2}, \quad \mu_{i_0} \frac{q_{i_0 i_1}}{q_{i_1 i_0}} \frac{q_{i_1 i_2}}{q_{i_2 i_1}} \cdots \frac{q_{i_{n-1} i_n}}{q_{i_n i_{n-1}}} = \mu_{i_n}. \tag{2.4}$$

如我们再向前走一步:  $i_n \rightarrow i_{n+1} = i_0$ , 则得出

$$\mu_{i_0} \frac{q_{i_0 i_1}}{q_{i_1 i_0}} \frac{q_{i_1 i_2}}{q_{i_2 i_1}} \cdots \frac{q_{i_{n-1} i_n}}{q_{i_n i_{n-1}}} \frac{q_{i_n i_0}}{q_{i_0 i_n}} = \mu_{i_0}.$$

两边消去  $\mu_{i_0}$ , 得到下述结果.

**定理 1** (Kolmogorov 圈形定理 [4]) 假定  $Q = (q_{ij} : i, j \in E)$  满足共零性 (2.2), 那么  $Q$  可配称当且仅当对于每一条闭路  $i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_n \rightarrow i_0$ , 都有

$$q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} q_{i_n i_0} = q_{i_0 i_n} q_{i_n i_{n-1}} \cdots q_{i_2 i_1} q_{i_1 i_0}. \tag{2.5}$$

余下是计算配称测度  $\mu$ . 由定义的“路”, 可赋予  $(E, Q)$  图结构. 然后使用共零性, 可将此图划分为若干联通子图. 我们只需在每一子图上算出其中的  $(\mu_k)$ . 为此, 任意选定子图上的参考点  $i_0$  和一条从  $i_0$  到  $k$  的路:  $i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_n = k$ , 不妨设  $\mu_{i_0} = 1$ , 那么  $\mu_k$  可由 (2.4) 的末式算出.

至此, 我们似乎已经回答了上述的可配称问题. 请看下面的五角星 (见图 2). 其闭路不会少于 25 条. 对于无限的  $E$ , 其闭路可能有无穷多. 可见问题没有这么简单. 为进一步深入, 我们将 (2.4) 末式改写成

$$\sum_{k=0}^{n-1} \log \frac{q_{i_k i_{k+1}}}{q_{i_{k+1} i_k}} = \log \frac{\mu_{i_n}}{\mu_{i_0}} = \log \mu_{i_n} - \log \mu_{i_0}. \tag{2.6}$$

此式有如下场论解释: 如将

$$\log \frac{q_{i_k i_{k+1}}}{q_{i_{k+1} i_k}}$$

定义为由  $Q$  生成的场沿  $i_k \rightarrow i_{k+1}$  所做的功, 那么 (2.6) 的左方就是这个场沿路  $i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_n$



图 2 五角星及其闭路

所做的功; 而 (2.6) 的右方则是在此路两端的位势之差. 这样, 我们找到了  $Q$  的可配称性的 (保守) 场论刻画. 由经典位势理论的启发, 我们容易得出下述结果.

**定理 2** [5]  $Q$  可配称当且仅当共零性成立, 而且它所生成的场是一个保守场 (等价地, 此场沿每一条最小闭路所做的功为零).

以上面的五角星为例, 最小闭路仅有五个角所构成的五个三角形及中间的正五边形. 因此, 由定理 2, 我们只需验证 6 个闭路条件. 随后将看到, 因为五个角具有对称性, 可能只需用一个条件表出.

下面以两个模型为例, 展示定理 2 的应用.

先考虑自旋系统 (spin system). 自旋空间为  $\{-1, +1\}$ . 对于给定的组态  $x = (x_u : u \in \mathbb{Z}^d)$ , 定义一个新的组态  ${}_u x$  如下:

$$({}_u x)_v = \begin{cases} -x_u, & \text{如果 } v = u, \\ x_v, & \text{如果 } v \neq u, \end{cases}$$

系统仅在单个位置  $u$  上自旋. 详言之, 系统的运动可描述为其自旋  $x \rightarrow {}_u x$  的速率为  $c(u, x) > 0$ . 这样, 描述此系统的算子可写成

$$\Omega f(x) = \sum_{u \in \mathbb{Z}^d} c(u, x)[f({}_u x) - f(x)].$$

注意, 每两个位置  $u$  和  $v$  ( $\neq u$ ) 的跳构成一个四边形 (见图 3). 关于这个算子的可配称性, 我们有如下简洁的判准.

**定理 3** (参见文献 [3, 定理 11.2(1)]) 以有限集  $S$  代替  $\mathbb{Z}^d$ , 则此自旋系统可配称当且仅当下述四边形条件对于每一对互异的  $u, v \in S$  和  $x = (x_u : u \in S)$  成立:

$$c(u, x)c(v, {}_u x)c(u, {}_{uv}x)c(v, {}_v x) = c(v, x)c(u, {}_v x)c(v, {}_{vu}x)c(u, {}_u x).$$

注意对于有限集  $S$ , 乘积空间  $\{-1, 1\}^S$  中的闭路可能很多, 但这里只用一个统一的条件给出了判准, 这显示出定理 2 的效用. 由此出发, 补上自然条件  $c(u, \cdot)$  连续, 可证此定理的结论对于  $S = \mathbb{Z}^d$  依然是对的 (此时乘积空间  $\{-1, 1\}^S$  不可数), 参见文献 [3, 定理 11.10(1)].

自旋系统的最主要代表是 Ising 模型. 其状态空间为  $\{-1, +1\}^{\mathbb{Z}^d}$ , 自旋速率为

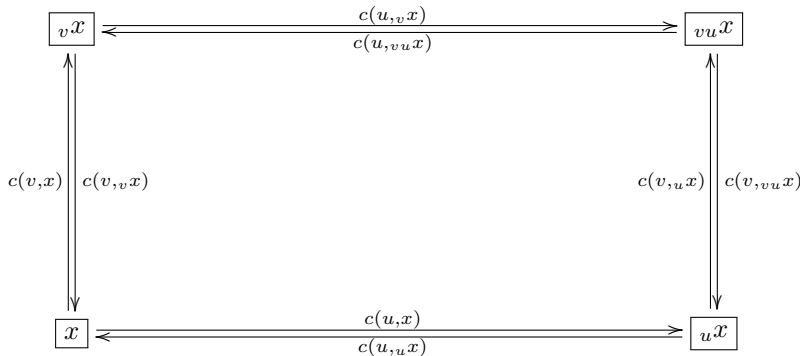


图 3 四边形闭路

$$c(u, x) = \exp \left[ -\beta \sum_{v:|u-v|=1} x_u x_v \right], \quad \beta \geq 0.$$

稍加计算, 可知它满足上述四边形条件, 因此属于平衡态系统. 之所以成为统计力学的典型模型, 是因为它具有相变. 以  $\mathcal{S}$  表示由  $\Omega$  生成的 Markov 过程的不变概率测度集, 则有如下结果: 当  $d = 2$  时,

$$|\mathcal{S}| = 1, \quad \text{如果 } \beta < \frac{1}{2} \log(1 + \sqrt{2}) =: \beta_c^{(2)} \approx 0.44;$$

$$|\mathcal{S}| > 1, \quad \text{如果 } \beta > \frac{1}{2} \log(1 + \sqrt{2}),$$

此处  $|\mathcal{S}|$  表示集合  $\mathcal{S}$  元素的个数. 当  $d \geq 3$ ,  $\beta_c^{(d)}$  的取值则是一个待解难题 (参见文献 [3, 第 10 章]).

可配称性研究是我们进入统计物理的开局一步. 定理 2 参见文献 [5] 和 [6, 第 6 章], 定理 3 源自文献 [7, 8]. 随后出现了一批成果, 如文献 [9-12] 等. 这些文献及更多信息均可从文献 [3] 中找到, 这里不赘述.

下面转入定理 2 的第二个应用: Schlögl 第二模型 [13]. 它由两部分构成: 反应部分和扩散部分. 反应部分是与位置无关的独立的具有生速  $b_k$  和死速  $a_k$  的生灭过程, 当然取值于非负整数集  $\mathbb{Z}_+$ :

$$b_k = \beta_0 + \beta_2 k(k-1), \quad a_k = \delta_1 k + \delta_3 k(k-1)(k-2), \quad \beta_0, \beta_2, \delta_1, \delta_3 > 0.$$

它表示在该位置上粒子的个数. 详言之, 如果以  $x = (x_u : u \in \mathbb{Z}^d)$  表示组态空间中的一点, 则  $x_u$  就是在位置  $u$  的粒子个数, 那么上面的  $b_k$  就应改写为  $b(x_u)$ . 系统的扩散部分乃是粒子从一个位置  $u$  到另一个位置  $v$  的运动. 也假定每个粒子的运动是独立的, 因此, 运动的速率关于  $x_u$  是线性的.

$$x \rightarrow x - e_u + e_v \text{ 的速率为 } x_u p(u, v),$$

其中  $e_u$  表示在  $u$  处为 1 而在其他地方为零的单位向量,  $(p(u, v))$  是  $\mathbb{Z}^d$  上的随机游动. 概括起来, 描述这个系统的算子可写成:

$$\Omega = \sum_u \text{在位置 } u \text{ 的生灭算子} + \sum_{u,v} \text{从 } u \text{ 到 } v \text{ 的扩散算子}.$$

取定一个  $u$ , 对于每一个  $v \neq u$ , 由在  $u$  处的生灭和在  $u$  与  $v$  之间的运动 (倘若  $p(u, v) > 0$ ), 构成了一个三角形 (见图 4). 这自然构成了这个系统的一条最小闭路. 留意此处的  $x + e_v$  来自  $(x + e_u) - e_u + e_v$ . 将 (2.5) 应用于此闭路, 立即写出如下的三角形条件:

$$b(x_u)(x_u + 1)p(u, v)a(x_v + 1) = b(x_v)(x_v + 1)p(v, u)a(x_u + 1), \quad \forall x, u, v, u \neq v.$$

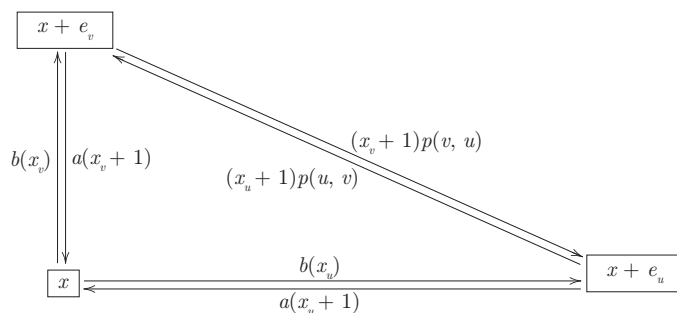


图 4 三角形闭路

现在, 对于每一对  $(u, v)$ ,  $u \neq v$ , 由三角形条件知,

$$x_u = x_v \Rightarrow p(u, v) = p(v, u).$$

这样, 若算子  $\Omega$  对称, 则

$$\text{转移矩阵 } (p(u, v)) \text{ 对称.} \tag{2.7}$$

由此进一步推出

$$\frac{\beta_k}{\delta_{k+1}} \equiv \text{常数 } \lambda > 0. \tag{2.8}$$

**定理 4** (参见文献 [14] 和 [3, 引理 14.16]) 假定  $(p(u, v) : u, v \in \mathbb{Z}^d)$  平移不变, 不可约且对每个  $u$ , 有  $p(u, u) = 0$ , 那么, Schlögl 第二模型对称当且仅当条件 (2.7) 和 (2.8) 同时成立.

特别地, 在 Schlögl 第二模型的参数空间  $\{\beta_0, \beta_2, \delta_1, \delta_3, p(u, v)\}$  上, 对称情形处于其低维子空间中. 此结果有重要物理意义. Schlögl 第二模型被设计为非平衡态统计物理的典型模型, 这与定理 4 相符. 该定理说的是, 仅在一个低维空间 (原参数空间的 Lebesgue 零测集) 上, 此模型会退化为平衡态.

新近找到定理 2 的新的应用, 我们将另文给出 [15].

### 3 非平衡态的典型模型—反应扩散过程

Schlögl 第二模型是反应扩散过程的主要代表. 我们先给出它与 Ising 模型的几种主要特征的比较 (见表 1).

从表 1 中易见, Schlögl 模型比 Ising 模型复杂得多. 例如, 因状态空间无局部紧性, 即使构造出半群, 也还得不到转移概率, 因为 Riesz 表现定理不适用. 我们知道, 描述统计物理的数学模型常是无穷维的, 因为有限维情形无相变. 当位置集  $S = \mathbb{Z}^d$  时, 状态空间  $E$  不可数, 若将  $\mathbb{Z}^d$  换成有限的  $S$ , 则  $E$  就是可数的, 相应的随机过程就是 Markov 链, 已有较多的积累. 有限位置之所以重要, 还在于它是走向无穷维模型的通路. 从有限维  $|S| < \infty$  出发, 通过极限过渡到无穷维有限维  $|S| = \infty$ , 是物理学家的常用手法, 也是我们的基本方法, 如同上面可配称性的处理.

原以为我们有足够本事处理有限维情形, 因为它是 Markov 链. 然而, 这恰恰是我们所遇到的第一道门槛. 当时, 我们所知道的是下述经典结果.

给定可数集  $E$  及其上的  $Q$  矩阵  $Q = (q_{ij} : i, j \in E)$ , 定义算子

$$\Omega f(i) = \sum_{j \in E} q_{ij} (f_j - f_i).$$

**表 1** 两种模型的比较

特征	模型	
	Ising 模型	Schlögl 模型
状态空间	$E = \{-1, +1\}^{\mathbb{Z}^d}$ , 紧	$E = \mathbb{Z}_+^{\mathbb{Z}^d}$ , 非局部紧
物理系统	平衡态	非平衡态
生成算子	局部有界	非局部有界且非线性
平稳分布	总存在, 有局部表示	存在性未知, 无局部表示

**定理 5** [16,17]  $Q$  过程唯一 (即由  $\Omega$  决定的 Markov 半群唯一) 当且仅当对于某一个 (等价地, 每一个)  $\lambda > 0$ , 下面方程只有零解:

$$\begin{cases} (\lambda I - \Omega)u(x) = 0, \\ 0 \leq u(x) \leq 1, \quad x \in \mathbb{Z}_+^S. \end{cases} \quad (3.1)$$

定理 5 有很多应用, 例如  $Q$  有界, 或者上述方程可解情形 (如生灭过程). 然而, 对于 Schlögl 第二模型, 我们希望对于任何有限维, 证明结论成立. 当然, 不能指望在有限维格子点  $\mathbb{Z}_+^{|S|}$  (可数集) 上直接解方程 (3.1). 在此, 读者可稍作停留, 想想如何克服这一难关.

我们使用数学上常用的比较方法, 即将高维过程与一维过程比较. 例如, 研究唯一性时, 我们所关心的是非爆炸, 即不会跑到无穷远. 注意要跑向无穷远, 必定跨过所有水平集  $E_n := \{x \in E : \sum_{u \in S} x_u = n\}$  所形成的阶梯. 所有阶梯可用非负整数  $\mathbb{Z}_+$  编号. 为控制原过程, 我们要构造  $\mathbb{Z}_+$  上的 Markov 链, 它跑出高水平集的速率不小于原过程的跑出速率. 详言之, 设  $y \in E_2$ , 将它跑到某  $z \in E_3$  的速率记为  $\text{Rate}(y \rightarrow z)$ . 这样, 原过程从  $y$  跑到水平集  $E_3$  的速率不超过

$$\sum_{z \in E_3} \text{Rate}(y \rightarrow z).$$

进一步, 原过程从水平集  $E_2$  跑到水平集  $E_3$  的速率不会超过

$$\max_{y \in E_2} \sum_{z \in E_3} \text{Rate}(y \rightarrow z).$$

我们将这个量定义为新的一维 Markov 链的转移速率  $q_{23}$ . 类似地, 我们可定义其他的  $q_{ij}$  ( $j \neq i$ ), 如图 5 所示.

十分幸运, 对于我们所关心的非平衡态统计物理的 16 个模型, 我们只需要用到单生过程, 即

$$\text{当 } j > i \text{ 时, } q_{ij} > 0 \text{ 当且仅当 } j = i + 1.$$

这引导我们对单生过程作了相当系统的研究, 如文献 [18] 及所引文献. 进而完成了有限维反应扩散过程的基础性研究. 这项工作完成于 1983 年, 最早发表于文献 [19] (又可参见文献 [3, 第 3 和 4 章]).

自然想到, 上述水平集划分会会有很多变种. 由此前进, 我们得到了抽象空间中 Markov 跳过程的非常一般的充分性条件. 今就可数情形重述为下述定理.

**定理 6** [20,21] 给定可数集  $E$  上的  $Q$  矩阵  $Q = (q_{ij})$ . 假定存在序列  $\{E_n\}_1^\infty$ 、常数  $c \in \mathbb{R}$  和非负函数  $\varphi$  使得

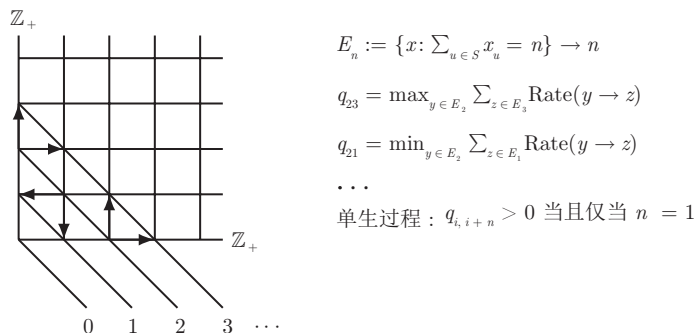


图 5 化高维为一维单生过程

- (1)  $E_n \uparrow E$ , 对于每个  $n \geq 1$ ,  $\sup_{i \in E_n} (-q_{ii}) < \infty$ ;
- (2)  $\lim_{n \rightarrow \infty} \inf_{i \notin E_n} \varphi_i = \infty$ ;
- (3) 对于每个  $i \in E$ ,  $\sum_j q_{ij}(\varphi_j - \varphi_i) \leq c\varphi_i$ ,

则  $Q$  过程唯一.

为展示此定理的效力, 对于 Schlögl 第二模型, 只需取  $\varphi(x) = c(1 + \sum_{u \in S} x_u)$ , 其中  $E_n = \{x : \sum_{u \in S} x_u \leq n\}$ ,  $c$  为常数. 因为死的速率比生的速率高一阶, 所以, 显然过程唯一. 有趣的是, 新近已证, 对于可数空间, 定理 6 的条件也是必要的, 从而成为唯一性的一个判别准则, 详见文献 [22] 和综述报告 [23].

#### 4 概率度量与耦合方法

在完成了有限维的基本工作之后, 便可进入无穷维过程的构造. 为此, 先讨论我们的主要工具: 概率度量和耦合方法. 在我们的工作之前, 苏联学派只用名称“概率度量”, 而美国学派只用“耦合方法”. 两者虽有区别, 但实质上紧密联系. 我们先复习一下概率论中常见的几种收敛性. 以下是笔者于 1986 年领悟到的蕴含关系图 (见图 6).

熟知, 图 6 中的依  $L^p$  收敛、a.s. 收敛和依概率收敛依赖于概率空间这一参考标架; 但其余三种收敛性不依赖于这一参考标架. 依分布收敛与淡收敛等价, 参见文献 [24, 定理 4.4.5 和 4.4.2]. 另一方面, 由 Skorohod 定理 (参见文献 [25, 定理 2.7]) 知, 对于适当选择的参考标架, 由弱收敛可推出 a.s. 收敛. 这样, 除  $L^p$  收敛之外, 其余的收敛性就内在性而言是一样的. 因此, 为寻找本质上不同的收敛性, 应当考虑  $L^p$  收敛的类似物.

设  $\xi_1$  和  $\xi_2$  是定义在概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上、取值于可测距离空间  $(E, \rho, \mathcal{E})$  的随机元, 那么, 通常的  $L^p$  度量是  $\|\xi_1 - \xi_2\|_p = \{\mathbb{E}[\rho(\xi_1, \xi_2)^p]\}^{1/p}$ . 再设  $\xi_i$  有分布  $P_i$ , 记作  $\xi_i \sim P_i, i = 1, 2, (\xi_1, \xi_2) \sim \tilde{P}$ , 则

$$\|\xi_1 - \xi_2\|_p = \left[ \int_{E \times E} \rho(x_1, x_2)^p \tilde{P}(dx_1, dx_2) \right]^{1/p}.$$

为摆脱参考标架, 令  $\tilde{P}$  跑遍  $\xi_1$  和  $\xi_2$  的一切联合分布. 这引导出耦合概念. 称  $E_1 \times E_2$  上的概率测度  $\tilde{P}$  为  $E_1$  上的概率测度  $P_1$  与  $E_2$  上的概率测度  $P_2$  的耦合, 如果下述边缘性条件满足:

$$\tilde{P}(dx_1, E_2) = P_1(dx_1), \quad \tilde{P}(E_1, dx_2) = P_2(dx_2).$$

然后我们可定义  $p$  阶 Wasserstein 距离如下:

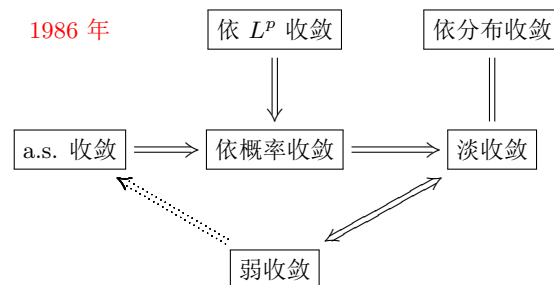


图 6 各种收敛性的蕴含关系图



$$W_p(P_1, P_2) = \inf_{\tilde{P}} \left[ \int_{E \times E} \rho(x_1, x_2)^p \tilde{P}(dx_1, dx_2) \right]^{1/p}, \quad p \geq 1, \quad (4.1)$$

其中  $\tilde{P}$  跑遍  $P_1$  和  $P_2$  的一切耦合. 此距离是文献 [26] 首先引进的.

一般地讲, 很难算出此距离的精确解. 然而, 任一耦合都提供此距离的上界估计 (这正是实际应用所需要的):

$$W_p(P_1, P_2) \leq \left[ \int_{E \times E} \rho(x_1, x_2)^p \tilde{P}(dx_1, dx_2) \right]^{1/p}.$$

在应用中, 常可找出有效或高效的  $\tilde{P}$ . 从这里已可看出概率度量与耦合方法的姐妹关系.

与许多概率度量 (如弱收敛拓扑) 相比, 这里的度量有重要几何意义:

$$\text{如果 } \xi \sim P_1, \quad \xi + \theta \sim P_2, \quad \text{则 } W_p(P_1, P_2) = |\theta|.$$

详言之, 考虑 Euclid 空间中的一个随机变量和它的常值推移. 那么, 这两个概率分布的  $W_p$  距离就等于其推移的长度.

因为我们要研究的是 Markov 过程, 自然需要将概率测度的耦合推广到转移概率的耦合. 称  $(E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$  上的转移概率  $\tilde{P}(x_1, x_2; dy_1, dy_2)$  为  $(E_1, \mathcal{E}_1)$  上的转移概率  $P_1(x_1, dy_1)$  与  $(E_2, \mathcal{E}_2)$  上的转移概率  $P_2(x_2, dy_2)$  的耦合, 如果下述边缘性条件成立:

$$\begin{aligned} \tilde{P}(x_1, x_2; dy_1, E_2) &= P_1(x_1, dy_1), \\ \tilde{P}(x_1, x_2; E_1, dy_2) &= P_2(x_2, dy_2), \quad x_k \in E_k. \end{aligned}$$

等价地, 若将单变量函数  $f(x_k)$  视为双变量函数  $\tilde{f}(x_1, x_2)$ , 则上述 Markov 耦合的边缘性可改写为

$$\tilde{P}\tilde{f}(\cdot, x_2) = P_1 f, \quad \tilde{P}\tilde{f}(x_1, \cdot) = P_2 f, \quad x_k \in E_k.$$

对于连续时间的 Markov 过程, 应当处理其 (形式生成) 算子, 而不能将条件施加到无法操作的转移概率上. 给定两个 Markov 算子  $\Omega_k$  ( $k = 1, 2$ ), 乘积空间  $(E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$  上的 Markov 算子  $\tilde{\Omega}$  称为  $\Omega_1$  和  $\Omega_2$  的耦合, 倘若

$$\tilde{\Omega}\tilde{f}(\cdot, x_2) = \Omega_1 f, \quad \tilde{\Omega}\tilde{f}(x_1, \cdot) = \Omega_2 f,$$

此处  $\tilde{f}$  (如上述) 为  $f$  的提升.

当然, 对于给定的边缘算子, 耦合算子有无穷多. 幸运地, 我们有如下结果.

**定理 7** (耦合基本定理 [3, 20, 21, 27]) 对于 Markov 跳过程 (以 Markov 链为其特例), 若某个耦合过程非爆炸 (过程唯一), 则其两个边缘过程也都非爆炸. 反之, 若两个边缘过程都非爆炸, 则它们的任意耦合过程也非爆炸.

对于不同的目的, 有效的耦合是不同的. 这就提出了耦合方法的分类和优化问题. 例如, 对于收敛速度, 何种耦合最好, 这曾经让我们摸索了 6 年时间. 结果答案竟来自 Wasserstein 距离:  $W = W_1$ . 受其启发, 我们找到如下概念 (参见文献 [28]).

**定义 8** 称耦合算子  $\bar{\Omega}$  是  $\rho$  最优的, 倘若

$$\bar{\Omega}\rho(x_1, x_2) = \inf_{\tilde{\Omega}} \tilde{\Omega}\rho(x_1, x_2),$$

其中  $\tilde{\Omega}$  跑遍所有耦合算子.

至此, 我们完成了耦合故事三部曲中的前两步 (详情参见文献 [27, 29]):

- (1) Markov 耦合基本定理 (定理 7),
- (2)  $\rho$  最优耦合 (定义 8),
- (3) 耦合关于距离的优化.

其实, 最后两步来自文章 [1] 谈到的特征值估计.

有了这些准备之后, 我们可以介绍构造无穷维反应扩散过程的要点. 如前所述, 由于状态空间  $\mathbb{Z}_+^d$  无局部紧性, 即使构造出半群, 依然得不到转移概率. 因为通常的 Riesz 表现定理不能用. 可以想像, 这种 Markov 过程的轨道空间也不会易于处理. 我们的手法还是使用有限维的 Markov 链逼近. 为此, 取定有限集序列  $\Lambda_n \uparrow \mathbb{Z}^d$ . 设  $n < m$ . 我们从  $\mathbb{Z}_+^{\Lambda_n}$  上的转移概率所生成的半群  $P_n(t)$  出发, 可将  $P_n(t)$  视为  $\mathbb{Z}_+^{\Lambda_m}$  上的半群. 我们的关键在于证明, 对于固定的  $t \geq 0, x_1, x_2 \in \mathbb{Z}_+^{\Lambda_m}$ , 有

$$W(P_n(t, x_1, \cdot), P_m(t, x_2, \cdot)) \rightarrow 0, \quad n, m \rightarrow \infty.$$

这一步需要构造较好的耦合算子, 使得关于  $W$  距离有较好的上界估计以保证收敛性. 倘若如此, 我们就证得  $\{P_n(t)\}_{n \geq 1}$  是关于  $W$  距离的 Cauchy 列. 再由  $W$  空间的完备性得到一个极限点. 这只是对于固定的时间  $t$  和固定的出发点  $x$  而言的极限. 还需要证明所获得的极限实际上是转移概率函数. 可以想象, 整个证明是有难度的.

无穷维反应扩散过程的构造完成于文献 [30], 这类 Markov 过程也是在该文献中命名. 笔者关于此文的后续工作有

- (1) 一般非平衡粒子系统的存在定理 (参见文献 [31]),
- (2) 非紧空间非平衡粒子系统的平稳分布 (参见文献 [32]),
- (3) 反应扩散过程的遍历性 (参见文献 [33, 34]),
- (4) 反应扩散过程的唯一性 (参见文献 [35]).

此外, 书 [27, 第 9 章] 乃是反应扩散过程的专题综述. 关于此方向, 已积累不少文献. 例如, 我们研究团队的文献 [14, 36–42] 和 [43–50] 等. 关于反应扩散过程的简化模型及共存现象的系统研究, 参见文献 [51–53]. 事实上, 如同文献 [3] 的书名所示, 全书以非平衡粒子系统为目标和导向. 一方面, 这开拓了新的数学研究领域; 另一方面, 使我们有机会重新考察已有的数学工具的有效性, 发展已有方法或探索新的数学工具.

致谢 本文回顾了严士健老师带领下, 我们进入统计物理交叉研究领域的几个片段. 应当说, 历程充满艰辛. 作为纪念, 献给严老师的 90 华诞.

## 参考文献

- 1 陈木法. 源自统计物理的数学论题—相变现象与数学的交叉. 中国科学: 数学, 2019, 49: 327–338
- 2 严士健, 李占柄. 非平衡系统的概率模型以及 Master 方程的建立. 物理学报, 1980, 29: 139–152
- 3 Chen M F. From Markov Chains to Nonequilibrium Particle Systems, 2nd ed. Singapore: World Scientific Publishing, 2004
- 4 Kolmogoroff A. Zur theorie der Markoffschen Ketten. *Math Ann*, 1936, 112: 155–160
- 5 侯振挺, 陈木法. 马尔可夫过程与场论. 科学通报, 1980, 25: 807–811 [收入文献 [6, 第 6 章第 194–242 页]]
- 6 钱敏, 侯振挺, 陈木法. 可逆马尔可夫过程. 长沙: 湖南科技出版社, 1979
- 7 Yan S J, Chen M F, Ding W D. Potentiality and reversibility for general speed functions (I). *Chin Ann Math Ser B*, 1982, 3: 572–586
- 8 Yan S J, Chen M F, Ding W D. Potentiality and reversibility for general speed functions (II). *Chin Ann Math Ser B*, 1982, 3: 705–720
- 9 唐守正. 自旋过程的可逆性. 数学学报, 1982, 25: 306–314

- 10 李世取. 一类混合无穷粒子系统的有势性和可逆性. 数学年刊 A 辑, 1983, 4: 773–780
- 11 曾文曲. 两类无穷粒子系统的可逆性. 数学年刊 A 辑, 1983, 4: 763–772
- 12 李勇. 关于有势排它过程的正可逆测度. 应用概率统计, 1990, 6: 121–126
- 13 Schlögl F. Chemical reaction models for non-equilibrium phase transitions. *Z Phys*, 1972, 253: 147–161
- 14 Ding W D, Durrett R, Liggett T M. Ergodicity of reversible reaction diffusion processes. *Probab Theory Related Fields*, 1990, 85: 13–26
- 15 Chen M F. Hermitizable, isospectral complex matrices or differential operators. *Front Math China*, 2018, 13: 1267–1311
- 16 Feller W. On boundaries and lateral conditions for the Kolmogorov differential equations. *Ann of Math (2)*, 1957, 65: 527–570
- 17 Reuter G E H. Denumerable Markov processes and the associated contraction semigroups on  $l$ . *Acta Math*, 1957, 97: 1–46
- 18 Chen M F, Zhang Y H. Unified representation of formulas for single birth processes. *Front Math China*, 2014, 9: 761–796
- 19 Yan S J, Chen M F. Multidimensional  $Q$ -processes. *Chin Ann Math Ser B*, 1986, 7: 90–110
- 20 陈木法. 跳过程与粒子系统. 北京: 北京师范大学出版社, 1986
- 21 Chen M F. Coupling for jump processes. *Acta Math Sin Engl Ser*, 1986, 2: 123–136
- 22 Spieksma F M. Countable state Markov processes: Non-explosiveness and moment function. *Probab Engrg Inform Sci*, 2015, 29: 623–637
- 23 Chen M F. Practical criterion for uniqueness of  $Q$ -processes. *Chinese J Appl Probab Statist*, 2015, 31: 213–224
- 24 Chung K L. *A Course in Probability Theory*, 3rd ed. Amsterdam: Academic Press, 2001
- 25 Ikeda N, Watanabe S. *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. Amsterdam: Holland-Kodansha, 1989
- 26 Wasserstein L N. Markov processes on a countable product space, describing large systems of automata. *Problemy Peredachi Informatsii*, 1969, 5: 64–73
- 27 Chen M F. *Eigenvalues, Inequalities, and Ergodic Theory*. New York: Springer, 2005
- 28 Chen M F. Optimal Markovian couplings and applications. *Acta Math Sin Engl Ser*, 1994, 10: 260–275
- 29 Chen M F. Trilogy of couplings and general formulas for lower bound of spectral gap. In: *Probability Towards 2000. Lecture Notes in Statistics*, vol. 128. New York: Springer, 1998, 123–136
- 30 Chen M F. Infinite dimensional reaction-diffusion processes. *Acta Math Sin Engl Ser*, 1985, 1: 261–273
- 31 Chen M F. Existence theorems for interacting particle systems with non-compact state spaces. *Sci China Ser A*, 1987, 30: 148–156
- 32 Chen M F. Stationary distributions of infinite particle systems with noncompact state space. *Acta Math Sci Engl Ser*, 1989, 9: 7–19
- 33 Chen M F. Ergodic theorems for reaction-diffusion processes. *J Stat Phys*, 1990, 58: 939–966
- 34 Chen M F. On the ergodic region of Schlögl's model. In: *Proceedings of International Conference on Dirichlet Forms and Stochastic Processes*. Berlin: Walter de Gruyter, 87–102
- 35 Chen M F. Uniqueness of reaction-diffusion processes. *Chinese Sci Bull*, 36: 969–973
- 36 唐守正. 扩散速度有限的多物种反应扩散过程的存在性及唯一性. 应用概率统计, 1985, 1: 11–22
- 37 黄力平. 一类无穷质点马尔可夫过程的平稳分布存在性. 应用概率统计, 1987, 3: 152–158
- 38 Zheng X G, Ding W D. Existence theorems for linear growth processes with diffusion. *Acta Math Sin New Ser*, 1987, 7: 25–42
- 39 Ding W D, Zheng X G. Ergodic theorems for linear growth processes with diffusion. *Chin Ann Math Ser B*, 1989, 10: 386–402
- 40 Chen M F, Ding W D, Zhu D G. Ergodicity of reversible reaction diffusion processes with general reaction rates. *Acta Math Sin Engl Ser*, 1994, 10: 99–112
- 41 韩东. 多物种无穷维反应扩散粒子系统鞅解的存在性. 应用概率统计, 1990, 6: 265–278
- 42 韩东. 一维 Brusselator 模型的遍历性. 新疆大学学报: 自然科学版, 1991, 8: 37–40
- 43 韩东. 多物种无穷维反应扩散粒子系统鞅解的唯一性. 数学年刊 A 辑, 1992, 13: 271–277
- 44 韩东. 多物种无穷维反应扩散过程的唯一性. 数学年刊 A 辑, 1995, 16: 572–578
- 45 李勇. 无穷维反应扩散过程的唯一性. 科学通报, 1991, 22: 1681–1684
- 46 李勇. 一类平移不变无穷粒子反应扩散过程的遍历性. 数学年刊 A 辑, 1995, 16: 223–229
- 47 Feng S, Zheng X G. Solutions of a class of nonlinear master equations. *Stochastic Process Appl*, 1992, 43: 65–84
- 48 陈金文. 有限维 Brussel 模型的常返性. 数学物理学报, 1995, 15: 121–125
- 49 Chen J W. Large deviations for infinite dimensional and reversible reaction-diffusion processes. *Acta Math Appl Sin*, 1996, 12: 300–307

- 50 Chen J W. A note on the large deviations for reversible reaction-diffusion processes. *Chinese J Contemp Math*, 1996, 17: 397–404
- 51 Durrett R, Levin S. The importance of being discrete (and spatial). *Theor Popul Biol*, 1994, 46: 363–394
- 52 Durrett R. Ten lectures on particle systems. In: *Lectures on Probability Theory. Lecture Notes in Mathematics*, vol. 1608. Berlin: Springer, 1995, 97–201
- 53 Durrett R, Neuhauser C. Particle systems and reaction-diffusion equations. *Ann Probab*, 1994, 22: 289–333

## Mathematical topics motivated from statistical physics —Interaction of statistical physics and probability theory

Mufa Chen

**Abstract** In mathematical language, the paper introduces the development from Markov chains to the nonequilibrium interacting particle systems. In physical language, it is a way motivated from equilibrium statistical physics to nonequilibrium. Mainly, some typical results and mathematical tools developed in period are introduced, including the criteria of equilibrium/nonequilibrium, the construction and further property of a typical class of nonequilibrium system—the reaction-diffusion processes, as well as the important tools—probability metrics and coupling.

**Keywords** Markov chain, interacting particle system, reaction-diffusion process, equilibrium/nonequilibrium

**MSC(2010)** 60J27, 60K35

**doi:** 10.1360/N012017-00270