21 世纪 100 个科学难题，李喜先主编，吉林人民出版社 1998, 816–823

相变的数学理论

陈木法

(北京师范大学)

摘要. 本世纪初， D. Hilbert 在展望新世纪的著名演讲中提出：“使数学在其中起作用的那些物理科学公理化” (23 个问题中的第 6 个). 如果把“公理化”改为“严格化”，则本文所述的课题只不过是上述问题的一个子课题. 我们将尽可能通俗地介绍此课题的意义、少数有代表性的难题及若干研究现状. 本方向也许可视为新世纪主流数学 — 无穷维数学的代表性的一支.

1. **引言**. 什么是相变？相变在生活中处处可见. 水由液态变成固态或气态，铁磁体的磁化，都是相变.

研究相变的意义何在？只要想想超导体在工业应用中的重大价值便一清二楚. 超导现象也是一种相变.

相变现象的主要特征是什么？它是大量的具有交互作用的粒子所构成的系统的宏观发展形为. 这里的“大量”指的是：例如单位体积 (摩尔) 内气体分子数的数量级为 10^{23}. 这个天文数字宜视为无穷. 这样，这个研究对象在数学上属无穷维. 有限维和无穷维数学的差别好比“在常速下的牛顿力学”和“在光速下的爱因斯坦相对论”. 然而，无穷维数学目前依然是一个极待开垦的领域，它远比有限维数学广阔和艰深.

研究历史和现状如何？相变现象是统计物理的经久不衰的中心研究课题. 尽管统计物理与概率论有着天然的联系，但如同数学与物理的许多分支一样，它们长期分道扬镳. 直到六十年代初期，概率论和统计物理才重新汇合交融，逐步建立起统计物理的严格的数学基础. 其核心课题之一即是相变的数学理论. 至于目前的研究现状，正如 R. L. Dobrushin 所说：一方面，这里有若干数学理论的大陆，有一些数学成果的小岛屿；另一方面，这里有着未解决问题的汪洋大海.

在本文的余下部分，我们将从最简单的相变模型 — 渗流模型开始 (§2). 然后进入统计物理的代表—Ising 模型 (§3). 我们将提及若干著名难题. 文末 (§4) 介绍一种研究相变的新方法.

2. **渗流模型**. 让我们从简单模型谈起.

a) *边模型*. 考虑 $d\,(\geqslant 2)$ 维整点 $\mathbb{Z}^d=\{(i_1,\cdots,i_d): i_1,\cdots,i_d=0,\pm1,\pm2,\cdots\}$. 两紧邻点 (边长为 1) 构成一条边. 假定每条边独立地以概率 $p\in[0,1]$“开”；当然也就独立地以概率 $1-p$“闭”. 把开的边顺序地连接起来的无绕折线称为一条路，其长度定义为该路所含边的条数. 我们关心的是否有一条走向无穷 (即长度为无穷) 的路. 显然，当 $p=1$ 时，所有的边都开着，走向无穷的路很多；当 $p=0$ 时，所有的边都关闭着，没有走向无穷的路. 问题是：对于其它的 p，是否以概率 1 存在走向无穷的路？可见，这里的假设和问题容易说清楚. 显然，p 越大，越容易开，故存在一个最小值，叫 *临界值*，记为 p_c. 那么，容易猜出当 $d=2$ 时，$p_c=1/2$. 当 $p>1/2$ 时，总有一条走向无穷路；而当 $p<1/2$ 时则不行，存在走向无穷的路的概率为 0. 不过，我们强调此刻讲的是二维 ($d=2$) 情形.

b) *点模型*. 假如我们把上述模型稍许改一下，把“边”改为“点”. 即每点独立地以概率 p“开” (以概率 $1-p$“闭”). 如紧邻两点都开，则称连接这两点的边开. 于是可如上定义开路并提出同样的问题. 类似地，我们可以定义临界值 p_c. 对于二维情形，这个点模型的 p_c 是多少？与边模型作比较，我们可大致猜出 $p_c>1/2$. 但这个 p_c 究竟等于多少则是一个至今未能解决的著名难题.

进一步，如果把 $d=2$ 换成 $d\geqslant 3$. 那么无论是边模型或点模型，其临界值都是未知的，目前仅知道一个大概的范围. 以上这些以及本学科的大量的未解决问题均具备好数学问题的特征：“易于表述而又难于解决”. 渗流模型是数学上描述相变现象的简化模型，正因为模型简单，所以也成为研究其它问题的基本数学工具. 现在，渗流理论是概率论和数学物理的重要研究方向. 详见 H. Kesten[5]， G. Grimmett[6], [7].

3. **Ising 模型**.

a) *有限维情形*. 还是从简单情形谈起. 假定只考虑一个粒子, 它只取两个值 $\{-1,+1\}$. 从 -1 转移到 $+1$ 的速率为 $b>0$, 而从 $+1$ 转移到 -1 的速率为 $a>0$. 于是有 Q 矩阵

$$Q=\begin{pmatrix} -b & b \\ a & -a \end{pmatrix}$$

要点是每一行和等于零. 由 Q 导出转移概率半群

$$\begin{aligned} P(t)=(p_{ij}(t))=e^{tQ}&=\sum_{n=0}^{\infty}\frac{t^n}{n!}Q^n \\ &=\frac{1}{a+b}\begin{pmatrix} a+be^{-\lambda_1 t} & b[1-e^{-\lambda_1 t}] \\ a[1-e^{-\lambda_1 t}] & b+ae^{-\lambda_1 t} \end{pmatrix} \end{aligned} \tag{3.1}$$

命 $t\to\infty$ 得出

$$P(t)\longrightarrow\begin{pmatrix}\pi \\ \pi\end{pmatrix}=\begin{pmatrix}\pi_{-1} & \pi_{+1} \\ \pi_{-1} & \pi_{+1}\end{pmatrix}$$

其中 $\pi_{-1}=\frac{a}{a+b}$, $\pi_{+1}=\frac{b}{a+b}$. 由 (3.1) 易证 $\pi=(\pi_{-1},\pi_{+1})$ 满足如下方程

$$\pi_j=\sum_i \pi_i p_{ij}(t),\qquad j\in\{-1,+1\},\quad t\geqslant 0. \tag{3.2}$$

右方表示从 π 出发, 相应于 $P(t)$ 的马氏链在时刻 t 的概率分布. 此式表明, 分布 π 与 t 无关, 故称为 $P(t)$ 的 *平稳分布*, 在物理上称为 *平衡态*. 由 (3.1) 和 (3.2) 导出如下结论:

i) 平衡态 π 存在且唯一.

ii) 收敛于平衡态的速度为 $e^{-\lambda_1 t}$, 即

$$|p_{ij}(t)-\pi_j|=O\big(e^{-\lambda_1 t}\big),\qquad t\to\infty.$$

其中 $\lambda_1=a+b$ 为矩阵 $-Q$ 的非平凡特征值.

往前走一步, 考虑有限乘积空间. 即把 $\{-1,+1\}$ 换成 $\{-1,+1\}^S$ (S 有限). 那么上述两条结论依然成立.

b) *无限维情形*. 再往前走一步, 考虑无限 (可数) 乘积空间 $E=\{-1,+1\}^{\mathbb{Z}^d}$. 简单理解如下: 在每一 d 维整点 ($\mathbb{Z}^d$) 上放一磁单子, 它有正、负极之分. 整个系统代表一个铁磁体. 我们所关心的基本问题是何时铁磁体被磁化, 即从 “无序” 变为 “有序”. 现在, 我们给出这个系统的数学描述. 用 $x=(x_u: u\in\mathbb{Z}^d)$ 来表示 $E=\{-1,+1\}^{\mathbb{Z}^d}$ 中点, 把 $u\in\mathbb{Z}^d$ 称为 *位置* 或 *坐标*. x 在位置 u 处的值 x_u 仅取 ±1 两个值. 给定 $x\in E$ 和 $u\in\mathbb{Z}^d$. 以 ${}_ux$ 表 E 中的一点: 它与 x 的差别只在 u 处变号, 即 $({}_ux)_v=x_v$, 如 $v\neq u$; 而 $({}_ux)_u=-x_u$. 系统由 x 到 ${}_ux$ 的转移速率取为

$$c(u,x)=\exp\Bigg[-\beta\sum_{v:|v-u|=1}x_u x_v\Bigg],$$

此处 $\beta>0$ 称为 *反温度*. 我们不考虑其它方式的转移. 速度函数的这种取法有特定的物理意义而在此处不能详细解释. 代替原先的矩阵 Q (因为此时状态空间 E 不可数), 我们使用算子

$$\Omega f(x)=\sum_{u\in\mathbb{Z}^d}c(u,x)[f({}_ux)-f(x)],\qquad x\in E \tag{3.3}$$

此处 f 属于某个 “好” 的函数类. 由此出发, 可导出半群 $P(t)$, 形式上写成 $P(t)=e^{t\Omega}$. 进而有平衡态 π:

$$\pi=\pi P(t),\qquad t\geqslant 0.$$

当 $d=1$ 时，平衡态依然唯一. 但当 $d\geqslant 2$ 时，出现了十分有趣的现象：存在临界值 $\beta_c^{(d)}$ 使得

$$\begin{aligned}&\text{如 } \beta<\beta_c^{(d)}\,(<\infty)\text{，则平衡态唯一；}\\&\text{如 } \beta>\beta_c^{(d)}\,(>0)\text{，则平衡态非唯一.}\end{aligned}\tag{3.4}$$

这给出了相变的一种数学刻画. 回顾前面所说，有限维 (即有限乘积) 空间上平衡态唯一，因而无相变. 这说明我们不得不要研究无穷维. 上述例子是统计物理的典型模型，叫做 Ising 模型 (1925)，是每一本统计物理教材都要讲到的 (当然，使用的是物理语言). 关于二维情形，临界值已由 L. Onsager (1944) 定出：

$$\beta_c^{(2)}=\frac{1}{2}\log\left(1+\sqrt{2}\right)\approx 0.44.$$

值得一提的是李政道、杨振宁 (1952) 曾有杰出贡献. 当 $d\geqslant 3$ 时，$\beta_c^{(d)}$ 却是未知的，这是数学界和物理学界普遍了解的经典难题. 数学上的兴趣并不完全在于确定这些临界值，还在于它们是检验新理论、新方法的重要标尺.

由上述讨论自然想到要研究更多的模型、更多的问题 (例如收敛于平衡态的速度等). 这形成了当代的一个重要的数学分支 — 无穷维交互作用粒子系统,并已出版了一批总结性论著,例如 T. M. Liggett[8]，R. Durrett[3]，严士健 [12] 和笔者的 [1] 和 [2]. 作为姐妹学科，有另一重要数学分支 — 随机场. 可参考 H. O. Georgii [4].

4. **数学方法**. a) 概述. 我们的研究对象属于无穷维数学. 因为是无穷维，许多有限维的数学工具用不上，变得极为艰难. 有些人觉得这里缺乏“系统的”理论，其实是一种误解. 无穷维数学远比有限维数学丰富多彩. 即使将来形成象有限维数学那么“完整”的理论，恐怕其名目也会极其繁多. 正是为研究无穷维数学，人们不断地寻求新方法、新工具. 卅多年来，也已发展出一大批的方法和工具. 如 1) 耦合方法 [1], [2], [8], [12]，2) 对偶方法 [8], [12]，3) 图表示与渗流理论 [3], [7]，4) Peirels 方法与 Pirogov-Sinai 理论 [11]，5) 聚团展开方法 [9], [10], 以及新近的 6) 谱隙估计方法等. 此处我们只能挂一漏万，对末一方法略作介绍.

b) 谱隙. 让我们从近乎平凡的情形开始，对于 (3.1) 的矩阵 Q，$-Q$ 有平凡特征值 $\lambda_0=0$ 及 $\lambda_1=a+b$. 平稳分布 $\pi=(\pi_{-1},\pi_{+1})$ 决定了一个实平方可积函数空间 $L^2(\pi)$. 容易验证：$P(t)$ 在 $L^2(\pi)$ 上为对称算子：

$$(f,P(t)g)=(g,P(t)f),\qquad f,g\in L^2(\pi),\quad t\geqslant 0.$$

此处 $(\cdot,\cdot)$ 表 L^2 内积. 这等价于

$$(f,Qg)=(g,Qf),\qquad f,g\in L^2(\pi).$$

以 $\|\cdot\|$ 表 L^2 范数，则可证明 L^2 指数式收敛：

$$\|P(t)f-\pi(f)\|\leqslant\|f-\pi(f)\|e^{-\varepsilon t},\qquad f\in L^2(\pi),\quad t\geqslant 0.\tag{3.5}$$

(其中 $\pi(f)=\int f\mathrm{d}\pi$) 的最大速度 $\varepsilon_{\max}=\lambda_1$. 这刻画了 L^2 指数式收敛性与谱隙 $(=\lambda_1-\lambda_0=\lambda_1$, 即第一 (非平凡) 特征值) 之间的关系.

c) 谱隙与相变. 对于 (3.3) 所定义的算子 Ω. 我们也有平凡特征值 $\lambda_0=0$, 对应于常值特征函数 1. (3.5) 式对于一般的 L^2 半群都成立且 $\varepsilon_{\max}=\lambda_1$. 例如当 $d=1$ 时，对于 Ising 模型，我们有 $\lambda_1>0$. 当 $d\geqslant 2$ 时，在遍历性区域，平稳分布 π 唯一，因而依然有 $\lambda_1=\lambda_1(\beta)>0$. 至少对于充分小的 β (高温区) 已获证明. 但当 β 充分大时，因为有不同的平衡态，不可能指望 $\lambda_1>0$. 人们使用有限方体 Λ 代替 $\mathbb{Z}^d$. 即考虑 $\{-1,+1\}^\Lambda$ 上的马氏链，此时有 $\lambda_1(\Lambda)>0$. 然后证明对于充分大的 β, 当 $\Lambda\uparrow\mathbb{Z}^d$ 时，$\lambda_1(\Lambda)\downarrow 0$. 事实上，已证得：$\lambda_1(\Lambda)$ 指数式下降于零，其速率与 Λ 的低一维的体积成正比. 总之，我们已描绘出一个图像：随着 β 增大，$\lambda_1=\lambda_1(\beta)$ 从正退化为零. 这提供了描述相变的一种数学方式，它有别于 (3.4). 这些不同的描述方式体现出相变的数学理论的丰富多彩.

d) 对有限维数学的反作用. 当然，人们不会怀疑这种根本性新思想所产生的重大价值. 但本学科研究中所提出的新方法、新工具对有限维数学的一些重要应用则使不少人感到相当的意外. 有兴趣的读者可参考笔者新近的综述报告 “耦合、谱隙估计及相关课题”，《科学通报》， 1997.

5. **结束语**. 相变理论的迷人之处和广阔前景主要在于它有太多的未解决的重要问题，面对它们，人们还常感到手无寸铁. 在这里，既需要建造象牙之塔的能工巧匠，也需要一大批勇于献身的垦荒者. 一方面，科学难题的解决常带来深远的影响；另一方面，对于未知的基础领域的探索，也同样具有强大的生命力. 例如本世纪的计算机革命，并非导源于经典难题. 我们不难造就一批攻克难题的能手，但却不易培养出统帅千军的帅才. 科学的价值在于造福于人类，它有无限的内函，而非“功利主义”那种小尺子可以测量的.

参考文献

[1] 陈木法，跳过程与粒子系统，北京师范大学出版社， 1986.
[2] Chen, M. F., From Markov Chains to Non-Equilibrium Particle Systems, World Scientific, Singapore, 1992.
[3] Durrett, R., Lecture Notes on Particle Systems and Percolation, Wadsworth and Brooks /Cole, Pacific Grove, Calif., 1988.
[4] Georgii, H., Gibbs Measures and Phase Transitions, de Gruyter Studies in Math., vol. 9, Water de Gruyter, 1988.
[5] Kesten, H., Percolation Theory for Mathematicians, Birkhäuser, Boston, 1982.
[6] Grimmett, G., Percolation, Springer-Verlag, 1989.
[7] Grimmett, G., Percolation and Disordered Systems, to appear in LNM, 1998.
[8] Liggett, T. M., Interacting Particle Systems, Springer-Verlag, 1985.
[9] Malyshev, V. A. and Minlos, R. A., Gibbs Random Fields, Cluster Expansions, Kluwer Acad. Publ., 1991.
[10] Malyshev, V. A. and Minlos, R. A., Linear Infinite-Particle Operators, Transl. Math. Monographs, vol. 143, Am. Math. Soc., 1994.
[11] Sinai, Ya. G., Thoery of Phase Transitions: Rigorous Results, Pergamon Press, 1982.
[12] 严士键，无穷粒子马尔可夫过程引论，北京师范大学出版社， 1989.