

第二章 信念传播方程

配分函数和自由能在平衡统计物理学中扮演中心的角色，而信念传播方程 (belief propagation, 简称BP) 在自旋玻璃理论、计算机科学、信息科学等领域有广泛的应用。本章在介绍了一般自旋玻璃模型的因子网络表示法后，将讨论配分函数的圈图展开公式，并由此推导出信念传播方程以及系统自由能的Bethe-Peierls近似表达式。Bethe-Peierls自由能以及信念传播方程构成自旋玻璃复本对称平均场空腔理论。本章也将从Bethe-Peierls近似的角度直观地理解信念传播方程，并讨论信念传播方程与Kikuchi自由能泛函之间的联系。

2.1 平衡统计物理基础

本节引入一般自旋玻璃模型以及因子网络表示法，并罗列平衡统计物理学理论的基本要点。

2.1.1 能量函数，因子网络表示

考虑一个统计物理系统，它包含 N 个粒子，这些粒子不能改变其位置，但有可变的内部微观状态。为方便起见，假定每个粒子 ($i = 1, 2, \dots, N$) 的微观状态是分立的且只有两个可能的态，用自旋 $\sigma_i = +1$ 和 $\sigma_i = -1$ 来表示。 N 个粒子自旋态的集合构成系统的一个微观构型， $(\sigma_1, \sigma_2, \dots, \sigma_N)$ 。系统的微观构型的总数目为 2^N 。在本书以后的讨论中常常会将系统的一个微观构型简记为 $\underline{\sigma}$ ，即 $\underline{\sigma} \equiv (\sigma_1, \sigma_2, \dots, \sigma_N)$ 。

每个粒子都可能受到外界力场的影响，而且粒子彼此之间也可能有相互作用。粒子 i 与外场的相互作用只与该粒子的微观状态 σ_i 有关，其能量记为 $E_i(\sigma_i)$ ；如果粒子 i 没有受到外界环境的直接影响，我们可以令外场能量为

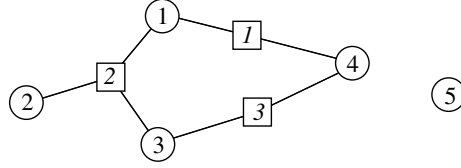


图 2.1: 一个包含5个粒子（圆点）和3个相互作用（方点）的模型系统的因子网络。该系统的内部相互作用能量函数的一般形式为 $E(\sigma_1, \sigma_2, \dots, \sigma_5) = E_1(\sigma_1, \sigma_4) + E_2(\sigma_1, \sigma_2, \sigma_3) + E_3(\sigma_3, \sigma_4)$ 。粒子5不和任何其它粒子相互作用，它在因子网络中是孤立的。

零，即 $E_i(\sigma_i) \equiv 0$ 。每一个内部相互作用 ($a = 1, 2, \dots, M$ ，其中 M 是内部相互作用的总数目) 涉及到两个或多个粒子。在以后的讨论中，相互作用 a 所涉及到的粒子构成的集合通常记为 ∂a 。内部相互作用 a 的能量记为 $E_a(\underline{\sigma}_{\partial a})$ ， $\underline{\sigma}_{\partial a}$ 表示集合 ∂a 中所有粒子的一个自旋构型。例如，如果相互作用 a 是两粒子 i 和 j 之间的自旋耦合，耦合常数为 J_{ij} ，那么 $\partial a = \{i, j\}$ ， $\underline{\sigma}_{\partial a} = \{\sigma_i, \sigma_j\}$ ， $E_a = -J_{ij}\sigma_i\sigma_j$ 。系统一个构型 $\underline{\sigma}$ 的能量 $E(\underline{\sigma})$ 是所有这些内部和外部相互作用能量之和：

$$E(\sigma_1, \sigma_2, \dots, \sigma_N) = \sum_{i=1}^N E_i(\sigma_i) + \sum_{a=1}^M E_a(\underline{\sigma}_{\partial a}). \quad (2.1)$$

在物理学、生物学、信息科学等研究领域中，有很多重要问题的能量函数都可以表达成 (2.1) 的加和形式。本书将发展一套统计物理学理论来研究由 (2.1) 所定义的一般模型的平衡性质。在这一多粒子复杂相互作用模型中，不同粒子的自旋状态并非独立，而是彼此影响彼此关联，常有很复杂的统计行为，而且可能出现奇异的集体性质。在信息科学研究中，人们常常用**因子网络** (factor graph) 来直观地描述一个多变量系统中的统计关联[10, 19]。由于本书所讨论的系统中，所有粒子都是不可移动的，我们也可以通过因子网络来描述统计物理系统 (2.1) 的所有微观相互作用。

在一幅因子网络中，圆点 ($i, j, k, \dots$) 代表系统中的粒子，称为**变量节点** (因为粒子的位置不能改变，但微观状态可变)；方点 ($a, b, c, \dots$) 代表内部相互作用，称为**功能节点**。因子网络中一条边 (i, a) 连接一个变量节点 i 和一个功能节点 a ，表示变量节点 i 所代表的粒子参与到了功能节点 a 所代表的内部相互作用。注意每条边连接一个变量节点和一个功能节点，两个变量节点之间或两个功能节点之间是没有边直接相连的。所有与功能节点 a 相连的变量

节点就构成集合 ∂a 。类似的，我们用记号 ∂i 表示所有与变量节点 i 相连的功能节点。如果一个变量节点 i 没有参与任何内部相互作用中，那么集合 ∂i 为空集。示例 2.1是一个简单模型系统所对应的因子网络。

在本书中，我们约定用字母 $i, j, k, \dots$ 来表示一个统计物理系统中的粒子及对应因子网络中的变量节点，而用 $a, b, c, \dots$ 来表示系统中的各种内部相互作用及对应因子网络中的功能节点。

对于只包含两体内部相互作用的系统，由于一个功能节点只连接两个变量节点，人们常常将功能节点省略，而用变量节点之间的边来表示系统中的两体相互作用。这样因子网络就退化为只包含变量节点和边的简单网络。

定义于 D 维晶格上的铁磁伊辛模型可以看成是模型 (2.1) 的特例，其能量函数为

$$E(\sigma_1, \sigma_2, \dots, \sigma_N) = - \sum_{(i,j)} J \sigma_i \sigma_j, \quad (2.2)$$

式中每一对最近邻格点 (i, j) 贡献一项相互作用能，耦合常数 $J > 0$ 。模型 (2.1) 的另一个重要特例是 D 维晶格上的Edwards-Anderson (EA) 自旋玻璃模型[9]，其能量函数为

$$E(\sigma_1, \sigma_2, \dots, \sigma_N) = - \sum_{(i,j)} J_{ij} \sigma_i \sigma_j, \quad (2.3)$$

其中边 (i, j) 的耦合常数是独立的随机参数，遵从同样的分布（例如 J_{ij} 有 $\frac{1}{2}$ 的概率为铁磁耦合， $J_{ij} = J$ ，有 $\frac{1}{2}$ 的概率为反铁磁耦合， $J_{ij} = -J$ ）。耦合常数的无序导致EA 模型有非平庸的低温统计物理性质。

本书研究的模型系统的能量函数都具有 (2.1) 的加和形式。值得强调的是，当系统中变量节点有超过两个微观状态，或者其微观状态需用连续参数来描述时，本书的大多数结论仍然是成立的。在以后的讨论中，我们将能量函数 (2.1) 所对应的因子网络记为 G 。

2.1.2 自由能，熵

系统处于温度为 T 的环境中，它的微观构型 $\underline{\sigma}$ 除了受到相互作用能量 (2.1) 的影响外，还受到环境热运动的影响。当系统达到平衡后，其宏观性质就不再随时间而改变，系统微观构型 $\underline{\sigma}$ 的平衡概率分布为玻耳兹曼分布，即

$$P_B(\underline{\sigma}) = \frac{1}{Z} \exp(-\beta E(\underline{\sigma})) = \frac{1}{Z} \prod_{i=1}^N \psi_i(\sigma_i) \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a}). \quad (2.4)$$

在上式中, β 称为逆温度, 它与温度呈反比关系, $\beta \equiv (k_B T)^{-1}$; $\psi_i \geq 0$ 是变量节点 i 感受到的外场所贡献的权重因子, 而 $\psi_a \geq 0$ 是内部相互作用 a 所贡献的权重因子, 它们的表达式分别为

$$\psi_i(\sigma_i) \equiv \exp[-\beta E_i(\sigma_i)], \quad (2.5a)$$

$$\psi_a(\underline{\sigma}_{\partial a}) \equiv \exp[-\beta E_a(\underline{\sigma}_{\partial a})]; \quad (2.5b)$$

Z 是玻耳兹曼分布的归一化系数, 称为配分函数,

$$Z \equiv \sum_{\underline{\sigma}} \exp(-\beta E(\underline{\sigma})) = \sum_{\underline{\sigma}} \prod_{i=1}^N \psi_i(\sigma_i) \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a}). \quad (2.6)$$

配分函数 Z 是系统所有微观构型的加和, 每个构型 $\underline{\sigma}$ 贡献一个权重 $e^{-\beta E(\underline{\sigma})} = \prod_{i=1}^N \psi_i(\sigma_i) \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a})$ 。如果单个微观构型的能量越大, 它对配分函数的贡献就越小: 例如考虑一个总能量为 E 的构型 $\underline{\sigma}$ 及一个总能量为 $E' > E$ 的构型 $\underline{\sigma}'$, 构型 $\underline{\sigma}$ 对配分函数的贡献大于构型 $\underline{\sigma}'$ 的贡献。但另一方面, 系统中能量为 E' 的构型总数, 记为 $\Omega_{E'}$, 可能远远多于系统中能量为 E 的构型总数, Ω_E 。这样就有可能导致能量为 E' 的所有微观构型对配分函数的总贡献, $\Omega_{E'} e^{-\beta E'}$, 反而大于能量为 E 的所有微观构型对配分函数的总贡献, $\Omega_E e^{-\beta E}$ 。在这种情形下, 如果对系统的微观构型进行多次独立观测, 能量为 E' 的那些微观构型会比能量为 E 的那些微观构型有更多的概率被观测到。

上面的讨论表达了一个深刻的观点, 即系统的宏观性质是能量和微观构型数目相互竞争达到动态平衡的结果。能量为 E 的构型总数 Ω_E 的表达式为

$$\Omega_E = \sum_{\underline{\sigma}} \delta(E(\underline{\sigma}), E), \quad (2.7)$$

其中 $\delta(x, y)$ 是Kronecker记号: 如果 $x = y$, 那么 $\delta(x, y) = 1$; 否则 $\delta(x, y) = 0$ 。如果总能量为 E 的微观构型数不等于零 ($\Omega_E \geq 1$), 我们可以将它表示为 $\Omega_E = \exp[S(E)/k_B]$, 其中 $S(E)$ 是系统的熵:

$$S(E) = k_B \ln \Omega_E. \quad (2.8)$$

熵是系统微观状态数的度量, 它是系统能量 E 的函数: 在某个能量值 E 处的熵越大, 意味着系统在这个能量值处的微观构型越多。将对构型 $\underline{\sigma}$ 的求和转化为对能量 E 的求和, (2.6) 可以写成

$$Z = \sum_E \Omega_E e^{-\beta E} = \sum_E e^{-\beta[E - TS(E)]}. \quad (2.9)$$

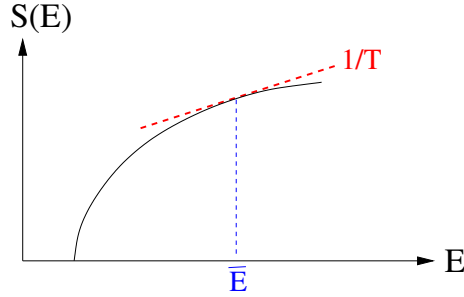


图 2.2: 环境温度决定宏观系统平均能量。给定温度 T ，系统的平均能量 $\bar{E}$ 由方程 (2.10) 决定，它对应于熵函数 $S(E)$ 对能量 E 的导数为 $1/T$ 处的能量值。

对通常的宏观统计物理系统而言，系统的粒子数 N 是很多的，系统的能量 E 和熵 $S(E)$ 都正比于粒子数 N 。那么配分函数的贡献绝大多数都来自能量值约为 $\bar{E}$ 的微观构型，其中 $\bar{E}$ 对应于函数 $E - TS(E)$ 的极小值。在 $N \rightarrow \infty$ 的极限下，假定 $S(E)$ 对能量 E 的导数存在，那么 $\bar{E}$ 的值由如下表达式决定

$$\left. \frac{dS(E)}{dE} \right|_{E=\bar{E}} = \frac{1}{T}. \quad (2.10)$$

能量 $\bar{E}$ 其实也是在温度 T 时对系统进行多次独立观测所获得的系统总能量的平均值。图2.2显示了温度 T 与系统平均能量 $\bar{E}$ 的相互关系。对通常的统计物理系统而言，系统的平均能量随着温度的减少而相应的减少。在某些特定温度处，平均能量作为温度的函数可能会出现一些奇异性，例如不连续或导数不连续，这对应于系统的宏观性质发生定性的改变，称为**相变**。在温度 T 趋向于零时，一个达到平衡的宏观系统的平均能量将趋向于系统的**基态能量**，记为 E_0 。基态能量是系统的所有微观状态所对应的能量的最小值，即

$$E_0 = \min_{\sigma} E(\sigma), \quad (2.11)$$

相应的微观状态称为系统的**基态**，基态的数目对应于**基态熵** $S_0 \equiv S(E_0)$ 。铁磁相互作用伊辛模型 (2.2) 只有两个基态，基态熵为 $k_B \ln 2$ 。我们将看到，对于很多更复杂的统计物理系统（例如包含多体相互作用的自旋玻璃系统），它们的基态熵可能正比于系统粒子数 N 。

由 (2.6) 定义的配分函数之所以在平衡统计物理学中起着非常重要的作用，是因为它与系统的平衡**自由能** F 有如下的关系

$$F = -k_B T \ln Z. \quad (2.12)$$

这一关系是平衡统计物理学的一个基本方程，它将系统的宏观性质(自由能)与系统的微观构型(配分函数)建立了定量的联系，为从微观相互作用出发定量理解系统宏观性质提供了理论基础。平衡自由能 F 是逆温度 β 以及其它环境参数(例如作用于各变量节点的外场)的函数。由(2.9)可知当 $N \rightarrow \infty$ 时，在给定的温度 T ，自由能 F 的值为

$$F = \bar{E} - TS(\bar{E}) \quad (N \rightarrow \infty), \quad (2.13)$$

其中 $\bar{E}$ 作为 T 的函数由(2.10)给出。对于一般的有限 N 值，与上式对应的自由能表达式可以由(2.12)导出为

$$\begin{aligned} F &= \sum_{\underline{\sigma}} P_B(\underline{\sigma}) E(\underline{\sigma}) - T \left[- \sum_{\underline{\sigma}} P_B(\underline{\sigma}) \ln P_B(\underline{\sigma}) \right] \\ &= \langle E(\underline{\sigma}) \rangle - TS, \end{aligned} \quad (2.14)$$

其中 $\langle E(\underline{\sigma}) \rangle$ 和 S 分别是系统处于温度为 T 的环境中所表现出的平均能量和熵，

$$\langle E(\underline{\sigma}) \rangle = \sum_{\underline{\sigma}} P_B(\underline{\sigma}) E(\underline{\sigma}) = \frac{\partial(\beta F)}{\partial \beta}, \quad (2.15)$$

$$S = -k_B \sum_{\underline{\sigma}} P_B(\underline{\sigma}) \ln P_B(\underline{\sigma}) = -\frac{\partial F}{\partial T}. \quad (2.16)$$

由(2.16)可以得出两个重要结论：(1) 系统的平衡熵 S 是非负的， $S \geq 0$ ；(2) 平衡自由能 F 是温度 T 的非增函数，即随着温度减少， F 会增加或至少保持不变。另外，平衡自由能 F 还是温度 T 的凹(concave)函数，这是因为 F 对 T 的二阶导数为非正：

$$\frac{d^2 F}{dT^2} \equiv -\frac{1}{k_B T^3} \left[\langle E(\underline{\sigma})^2 \rangle - \langle E(\underline{\sigma}) \rangle^2 \right] \leq 0. \quad (2.17)$$

上式中， $\langle A \rangle$ 表示物理学量 A 在玻耳兹曼分布下的平均值，即

$$\langle A \rangle \equiv \sum_{\underline{\sigma}} P_B(\underline{\sigma}) A(\underline{\sigma}).$$

由(2.12)所定义的自由能适用于描述一个统计物理系统的平衡性质。当系统处于非平衡状态时，对其微观构型进行观测，所得到的概率分布将不同于玻耳兹曼分布，而且该非平衡分布通常是与时间 t 有关的，让我们记之为 $P(\underline{\sigma}; t)$ 。获得概率分布 $P(\underline{\sigma}; t)$ 原则上需要对一个系统在相同的外部条件下进行多次独立测量。可以考虑如下的思想实验：将系统的环境温度升至足

够高并等待足够长时间使系统忘记历史达到平衡，然后将温度突降为实验温度 T ，等待一段时间 t 后记录系统在 t 时刻的微观构型 $\underline{\sigma}^1$ ；重复这一升温-降温过程多次，获得微观构型 $\underline{\sigma}^2, \underline{\sigma}^3, \dots$ ；最后，根据这些抽取的微观构型样本来构造概率分布 $P(\underline{\sigma}; t)$ 。我们针对非平衡分布也可以定义一个自由能 $F[P(\underline{\sigma}; t)]$ ，它是 $P(\underline{\sigma}; t)$ 的泛函：

$$F[P(\underline{\sigma}; t)] \equiv \sum_{\underline{\sigma}} P(\underline{\sigma}; t) E(\underline{\sigma}) + k_B T \sum_{\underline{\sigma}} P(\underline{\sigma}; t) \ln P(\underline{\sigma}; t). \quad (2.18)$$

自由能泛函 $F[P(\underline{\sigma}; t)]$ 也是由两方面的贡献叠加而成的，即概率分布 $P(\underline{\sigma}; t)$ 所对应的平均能量，及其蕴含的熵 $S[P(\underline{\sigma}; t)]$ 所导致的自由能贡献。一个概率分布的熵，又称为**信息熵**，可由如下表达式计算出来[8]：

$$S[P(\underline{\sigma}; t)] \equiv -k_B \sum_{\underline{\sigma}} P(\underline{\sigma}; t) \ln P(\underline{\sigma}; t). \quad (2.19)$$

自由能泛函 $F[P(\underline{\sigma}; t)]$ 与系统的平衡自由能 (2.12) 有如下关系，即

$$F[P(\underline{\sigma}; t)] = -k_B T \ln Z + k_B T \sum_{\underline{\sigma}} P(\underline{\sigma}; t) \ln \left[\frac{P(\underline{\sigma}; t)}{P_B(\underline{\sigma})} \right]. \quad (2.20)$$

由此可知 $F[P(\underline{\sigma}; t)]$ 总是大于系统的平衡自由能。随着时间 t 的增加，系统的概率分布 $P(\underline{\sigma}; t)$ 将向平衡玻耳兹曼分布 $P_B(\underline{\sigma})$ 演化，伴随有 $F[P(\underline{\sigma}; t)]$ 的递减。当系统达到平衡后，自由能泛函 $F[P(\underline{\sigma}; t)]$ 也将达到其唯一的极小值，即平衡自由能。换句话说，在给定温度 T 下，系统的自由能在平衡时达到极小值。

2.2 配分函数展开

用解析方法严格求解配分函数和平衡自由能只在为数不多的模型系统可以实现，这方面最著名的工作当然是Onsager的二维伊辛模型精确解。对大多数系统而言，在温度比较高因而系统内部的关联长度较短的情形，可以用高温展开的办法计算配分函数和自由能，常用的方法有Mayer的低密度气体集团展开法[11, 30, 1]、Brout的格点模型圈图展开法[5]、Kikuchi的团簇变分(cluster variation) 法[16, 3, 26, 33]、Plefka 展开法及其推广[34, 12, 36, 35]等等。

Chertkov和Chernyak[6, 7] 发现由 (2.1) 定义的模型系统，如果所有变量节点的状态都是两分量自旋（即 $\sigma_i = \pm 1$ ），那么系统的配分函数可以表达为有限个圈图贡献之和。文献[42]进一步提出了一般因子网络系统的配分函

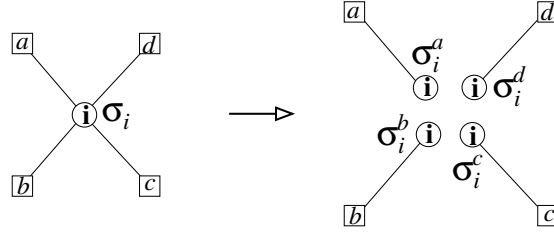


图 2.3: 一个连通度为 k_i 的变量节点 i 与 k_i 个功能节点相连 (这些功能节点的集合记为 ∂i)。我们将 i 想象为有 k_i 个化身, 每个化身连到一个不同的功能节点 $a \in \partial i$, 且具有自旋态 σ_i^a 。在本例中, $\partial i = \{a, b, c, d\}$, $k_i = 4$ 。

数圈图展开方法, 该方法很简单, 且不要求变量节点的状态为二分量自旋 (它可以是多分量离散变量, 也可以是连续变量, 甚至可以是多维变量或函数)。

我们采用文献[42]的方法来推导两分量自旋模型配分函数的圈图展开公式。在下一章我们将用同样的方法推导出广义配分函数的圈图展开公式。

在能量 (2.1) 所对应的因子网络 G 中, 变量节点 i 与一些功能节点 a 相连, 这些功能节点组成集合 ∂i , 该集合包含的功能节点数称为节点 i 的连通度, 记为 k_i ($\equiv |\partial i|$)。由于变量节点 i 参与了 k_i 个相互作用中, 我们将 i 想象成具有 k_i 个“化身”, 每个化身与一个不同的功能节点 $a \in \partial i$ 相连, 且有自己的自旋态 σ_i^a 。当然, 由于这 k_i 个化身实际上都是节点 i , 所以这些自旋值 σ_i^a 都等于节点 i 的自旋 σ_i (见图2.2)。假设因子网络 G 中功能节点 a 所连的变量节点集合为 $\partial a = \{i, j, \dots, l\}$ 。在对每个变量节点都引入化身后, 从节点 a 所“看到”的变量节点自旋构型 $\underline{\sigma}_{\partial a}$ 就是 $\underline{\sigma}_{\partial a} = \{\sigma_i^a, \sigma_j^a, \dots, \sigma_l^a\}$ 。通过引入变量节点的化身状态, 配分函数 (2.6) 可以改写为

$$Z = \sum_{\underline{\sigma}} \prod_{i=1}^N \psi_i(\sigma_i) \sum_{\{\underline{\sigma}_{\partial a} | a \in G\}} \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a}) \prod_{(j,b) \in G} \delta(\sigma_j^b, \sigma_j). \quad (2.21)$$

在上式中, 变量节点 i 的不同化身的自旋状态 $\sigma_i^a, \sigma_i^b, \dots$ 被看成是相互独立变量。通过在因子网络 G 的每一条边 (i, a) 上引入 Kronecker 记号 $\delta(\sigma_i^a, \sigma_i)$, 保证了只有当节点 i 的 k_i 个化身状态都取同一个值 σ_i 时才能真正对配分函数有贡献。

现在对每一条边 (i, a) 引入一个辅助概率函数 $q_{i \rightarrow a}(\sigma_i^a)$, 它是非负的, 且

满足归一化条件 $\sum_{\sigma_i^a} q_{i \rightarrow a}(\sigma_i^a) = 1$ 。表达式 (2.21) 被改写成

$$Z = \sum_{\underline{\sigma}} \prod_{i=1}^N \psi_i(\sigma_i) \sum_{\{\underline{\sigma}_{\partial a} | a \in G\}} \prod_{a=1}^M \left[\psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j^a) \right] \prod_{(k,b) \in G} \frac{\delta(\sigma_k^b, \sigma_k)}{q_{k \rightarrow b}(\sigma_k^b)}. \quad (2.22)$$

在上式中，如果不考虑Kronecker记号带来的对变量节点化身状态的约束¹，每个功能节点 a 将对配分函数贡献一个乘积因子 Z_a ：

$$Z_a \equiv \sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i^a). \quad (2.23)$$

我们来看看 Z_a 的物理意义。定义一个只包含功能节点 a 以及和 a 相连的所有 k_a 个变量节点的子系统 ($k_a = |\partial a|$)，并要求该子系统的能量为

$$- \sum_{i \in \partial a} h_{i \rightarrow a} \sigma_i^a + E_a(\underline{\sigma}_{\partial a}),$$

其中磁场 $h_{i \rightarrow a}$ 的值由概率分布函数 $q_{i \rightarrow a}(\sigma_i^a)$ 决定：

$$h_{i \rightarrow a} \equiv \frac{1}{2\beta} \ln \frac{q_{i \rightarrow a}(+1)}{q_{i \rightarrow a}(-1)}. \quad (2.24)$$

如果不考虑相互作用能量 E_a ，这个子系统就是 k_a 个相互独立的节点，其配分函数为

$$\sum_{\underline{\sigma}_{\partial a}} \prod_{i \in \partial a} e^{\beta h_{i \rightarrow a} \sigma_i^a} = \prod_{i \in \partial a} [2 \cosh(\beta h_{i \rightarrow a})]. \quad (2.25)$$

当相互作用能量 E_a 被考虑进来后，子系统的配分函数变为

$$\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} e^{\beta h_{i \rightarrow a} \sigma_i^a} = \prod_{i \in \partial a} [2 \cosh(\beta h_{i \rightarrow a})] \sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j^a). \quad (2.26)$$

注意到 Z_a 是 (2.26) 与 (2.25) 之比。因此 Z_a 所对应的自由能

$$f_a \equiv -k_B T \ln Z_a = -k_B T \ln \left[\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i^a) \right] \quad (2.27)$$

是相互作用能量 E_a 所导致的这个子系统自由能的变化量。

为了对每个变量节点 i 也得到一个和 Z_a 类似的配分函数因子，可以对每条边 (i, a) 引入另外一个概率分布函数 $p_{a \rightarrow i}(\sigma_i)$ ，它同样是非负的，满足归一

¹ 这等于将 (2.21) 等式右侧每条边 (k, b) 的权重因子 $[\delta(\sigma_k^b, \sigma_k)/q_{k \rightarrow b}(\sigma_k^b)]$ 假想为1。

化条件 $\sum_{\sigma_i} p_{a \rightarrow i}(\sigma_i) = 1$ 。表达式 (2.22) 被进一步改写为

$$Z = \sum_{\{\sigma_i | i \in G\}} \prod_{i=1}^N \left[\psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i) \right] \sum_{\{\sigma_{\partial b} | b \in G\}} \prod_{b=1}^M \left[\psi_b(\sigma_{\partial b}) \prod_{j \in \partial b} q_{j \rightarrow b}(\sigma_j^b) \right] \\ \times \prod_{(k,c) \in G} \frac{\delta(\sigma_k^c, \sigma_k)}{p_{c \rightarrow k}(\sigma_k) q_{k \rightarrow c}(\sigma_k^c)}. \quad (2.28)$$

在上式中, 如果同样忽略Kronecker记号带来的约束, 每个变量节点 i 将对配分函数贡献一个乘积因子 Z_i :

$$Z_i = \sum_{\sigma_i} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i). \quad (2.29)$$

Z_i 的物理意义也很容易探讨。考虑一个只包含变量节点 i 的 k_i 个化身的子系统 ($k_i = |\partial i|$), 并在每个化身节点上引入一个磁场 $u_{a \rightarrow i}$, 它的值由概率分布函数 $u_{a \rightarrow i}(\sigma_i)$ 决定:

$$u_{a \rightarrow i} \equiv \frac{1}{2\beta} \ln \frac{p_{a \rightarrow i}(+1)}{p_{a \rightarrow i}(-1)}. \quad (2.30)$$

如果这 k_i 个节点彼此之间没有相互作用, 那么该子系统的配分函数为

$$\prod_{a \in \partial i} \left[\sum_{\sigma_i^a} e^{\beta u_{a \rightarrow i} \sigma_i^a} \right] = \prod_{a \in \partial i} [2 \cosh(\beta u_{a \rightarrow i})]. \quad (2.31)$$

由于这 k_i 个节点都是变量节点 i 的化身, 故要求它们的自旋态必须都相同。这一约束以及和外场有关的权重因子 $\psi_i(\sigma_i)$ 导致子系统的配分函数变为

$$\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{a \in \partial i} \left[\sum_{\sigma_i^a} e^{\beta u_{a \rightarrow i} \sigma_i^a} \delta(\sigma_i^a, \sigma_i) \right] \\ = \prod_{a \in \partial i} [2 \cosh(\beta u_{a \rightarrow i})] \sum_{\sigma_i} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i). \quad (2.32)$$

Z_i 等于配分函数 (2.32) 与配分函数 (2.31) 的比值。因此, Z_i 所对应的自由能

$$f_i \equiv -k_B T \ln Z_i = -k_B T \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i) \right] \quad (2.33)$$

就是外场能量 $E_i(\sigma_i)$ 及变量节点 i 的所有化身节点具有相同的自旋态这一约束所导致的子系统自由能的变化量。

类似于 Z_a 和 Z_i ，我们引入每一条边 (i, a) 的配分函数因子 $Z_{(i,a)}$ 为

$$Z_{(i,a)} = \sum_{\sigma_i} q_{i \rightarrow a}(\sigma_i) p_{a \rightarrow i}(\sigma_i), \quad (2.34)$$

以及相应的自由能 $f_{(i,a)}$

$$f_{(i,a)} \equiv -k_B T \ln Z_{(i,a)} = -k_B T \ln \left[\sum_{\sigma_i} q_{i \rightarrow a}(\sigma_i) p_{a \rightarrow i}(\sigma_i) \right]. \quad (2.35)$$

配分函数 $Z_{(i,a)}$ 对应着由变量节点 i 及它的一个化身组成的子系统，它们分别受到磁场 $h_{i \rightarrow a}$ 和磁场 $u_{a \rightarrow i}$ 的作用。子系统受到的约束是节点 i 和它的化身自旋值相同。这一约束所导致的子系统自由能变化量就是 $f_{(i,a)}$ 。

有了这些准备，表达式(2.28)可以被进一步改写为

$$\begin{aligned} Z = & \sum_{\{\sigma_i | i \in G\}} \prod_{i=1}^N \left[\psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i) \right] \sum_{\{\sigma_{\partial b} | b \in G\}} \prod_{b=1}^M \left[\psi_b(\sigma_{\partial b}) \prod_{j \in \partial b} q_{j \rightarrow b}(\sigma_j^b) \right] \\ & \times \prod_{(k,c) \in G} \left[\frac{1 + \Delta_{(k,c)}(\sigma_k^c, \sigma_k)}{Z_{(k,c)}} \right], \end{aligned} \quad (2.36)$$

其中 $\Delta_{(i,a)}$ 是一个修正因子，其表达式为

$$\Delta_{(i,a)}(\sigma_i^a, \sigma_i) = \frac{Z_{(i,a)}}{q_{i \rightarrow a}(\sigma_i^a) p_{a \rightarrow i}(\sigma_i)} \delta(\sigma_i^a, \sigma_i) - 1. \quad (2.37)$$

我们之所以将修正因子 $\Delta_{(i,a)}$ 构造成(2.37)的形式，是希望它的值能够尽可能接近于零，或至少在某种平均的意义下其值接近于零。修正因子 $\Delta_{(i,a)}$ 有点类似于Mayer集团展开法中的Mayer函数[30, 1]。将 $\Delta_{(i,a)}$ 视为小量，因而对(2.36)式中的边连乘项进行展开，就得到

$$\begin{aligned} \prod_{(i,a)} [1 + \Delta_{(i,a)}] &= 1 + \sum_{(i,a)} \Delta_{(i,a)} + \sum_{(i,a), (j,b)} \Delta_{(i,a)} \Delta_{(j,b)} + \dots \\ &= 1 + \sum_{g \subseteq G} \prod_{(i,a) \in g} \Delta_{(i,a)}. \end{aligned} \quad (2.38)$$

在上式中， g 代表因子网络 G 的任意一个子网络，它包含 G 的一部分边以及这些边两端的变量节点和功能节点。将(2.38)代入(2.36)，就得到配分函数 Z 的展开表达式

$$Z = Z_0 \left(1 + \sum_{g \subseteq G} L_g \right). \quad (2.39)$$

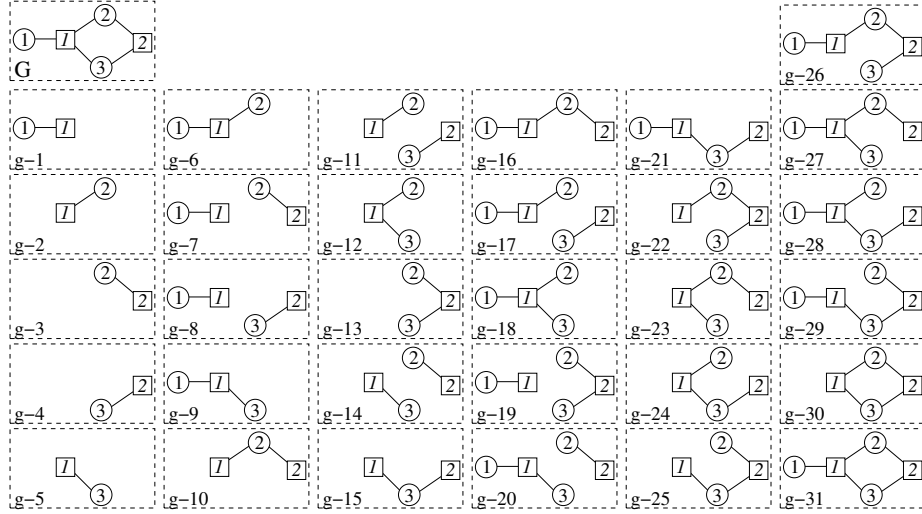


图 2.4: 因子网络 G 及其子网络集合。作为示例, 网络 G 只包含三个变量节点和两个功能节点, 它的子网络数目有31个: 有5个子网络只包含一条边($g-1, g-2, \dots, g-5$), 有10个子网络只包含两条边($g-6, g-7, \dots, g-15$), 有10个子网络包含三条边($g-16, g-17, \dots, g-25$), 有5个子网络包含四条边($g-26, g-27, \dots, g-30$), 还有一个子网络 $g-31$ 包含网络 G 的所有边。

在上式中, Z_0 的表达式为

$$Z_0 = \frac{\prod_{i \in G} Z_i \prod_{a \in G} Z_a}{\prod_{(i,a) \in G} Z_{(i,a)}}; \quad (2.40)$$

子网络修正项 L_g 的表达式为

$$L_g = \sum_{\{\sigma_i | i \in g\}} \sum_{\{\sigma_{\partial a} | a \in g\}} \prod_{i \in g} \omega_i(\sigma_i) \prod_{a \in g} \omega_a(\sigma_{\partial a}) \prod_{(j,b) \in g} \Delta_{(j,b)}(\sigma_j^b, \sigma_j), \quad (2.41)$$

其中概率分布 $\omega_i(\sigma_i)$ 和 $\omega_a(\sigma_{\partial a})$ 分别表达为

$$\omega_i(\sigma_i) = \frac{1}{Z_i} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i), \quad (2.42a)$$

$$\omega_a(\sigma_{\partial a}) = \frac{1}{Z_a} \psi_a(\sigma_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i^a). \quad (2.42b)$$

对一个包含 $\mathcal{M}$ 条边的因子网络 G , 其子网络的总数目为 $2^{\mathcal{M}} - 1$, 所以(2.39)中的求和中包含的项在 $\mathcal{M} \gg 1$ 的情况下是很多的。在图2.4中, 我们对一个含

有三个变量节点和两个功能节点的因子网络 G 的所有子网络都列出来了。在下节将看到，通过适当选择辅助概率分布函数 $\{q_{i \rightarrow a}(\sigma_i^a), p_{a \rightarrow i}(\sigma_i)\}$ ，我们可以使许多子网络 g 对配分函数 Z 的修正贡献 $L_g = 0$ ，从而使求和项的数目大为减少。

2.3 信念传播方程

展开式 (2.39) 形式上将配分函数写成了两项的乘积，一项是 Z_0 ，另一项是所有子网络修正贡献之和。该表达式对于任意选取的辅助概率函数都成立。现在需要考察如何优化辅助概率函数 $\{q_{i \rightarrow a}(\sigma_i^a), p_{a \rightarrow i}(\sigma_i)\}$ 的选择。

图2.4列出了一个小的因子网络 G 的所有子网络。注意到在很多子网络中都有一个或多个节点只连着一边。例如子网络 g_{-31} 中，变量节点1的连通度为1；在子网络 g_{-27} 中，功能节点2的连通度为1。如果子网络 g 的一条边的两个端点至少有一个在 g 中的连通度为1，我们就称其为子网络 g 的一条**摇摆边** (dangling edge)。现在来考虑任意一个包含摇摆边的子网络 g 对配分函数的修正贡献 L_g 。

假设子网络 g 有一条摇摆边 (i, a) ，且与之相连的变量节点 i 在子网络 g 中的连通度为1。这样一个子网络 g 对配分函数的修正贡献 L_g 为

$$L_g = \prod_{j \in g \setminus i} \left[\sum_{\sigma_j} \omega_j(\sigma_j) \right] \prod_{b \in g \setminus \partial b} \left[\sum_{\sigma_b} \omega_b(\sigma_b) \right] \prod_{(k,c) \in g \setminus (i,a)} \Delta_{(k,c)}(\sigma_k^c, \sigma_k) \\ \times \left[\frac{\hat{q}_{i \rightarrow a}(\sigma_i^a) \sum_{\sigma} q_{i \rightarrow a}(\sigma) p_{a \rightarrow i}(\sigma)}{q_{i \rightarrow a}(\sigma_i^a) \sum_{\sigma} \hat{q}_{i \rightarrow a}(\sigma) p_{a \rightarrow i}(\sigma)} - 1 \right], \quad (2.43)$$

其中，记号 $g \setminus i$ 表示除节点 i 外子网络 g 的所有其它变量节点的集合； $g \setminus (i, a)$ 表示除边 (i, a) 外子网络 g 的所有其它边的集合； $\hat{q}_{i \rightarrow a}(\sigma_i^a)$ 是一个概率分布函数，它的表达式为

$$\hat{q}_{i \rightarrow a}(\sigma) = \frac{\psi_i(\sigma) \prod_{b \in \partial i \setminus a} p_{b \rightarrow i}(\sigma)}{\sum_{\sigma'} \psi_i(\sigma') \prod_{b \in \partial i \setminus a} p_{b \rightarrow i}(\sigma')}, \quad (2.44)$$

其中 $\partial i \setminus a$ 表示集合 ∂i 除去元素 a 后的子集。注意到如果函数 $\hat{q}_{i \rightarrow a}(\sigma)$ 刚好等于函数 $q_{i \rightarrow a}(\sigma)$ ，那么由 (2.43) 有 $L_g = 0$ 。

继续考虑另一种情形，即子网络 g 有一条摇摆边 (i, a) ，该边所连的功能

节点 a 在 g 中的连通度为1。这样一个子网络对配分函数的修正贡献 L_g 为

$$L_g = \prod_{b \in g \setminus a} \left[\sum_{\underline{\sigma}_{\partial b}} \omega_b(\underline{\sigma}_{\partial b}) \right] \prod_{j \in g} \left[\sum_{\sigma_j} \omega_j(\sigma_j) \right] \prod_{(k,c) \in g \setminus (i,a)} \Delta_{(k,c)}(\sigma_k^c, \sigma_k) \\ \times \left[\frac{\hat{p}_{a \rightarrow i}(\sigma_i) \sum_{\sigma} q_{i \rightarrow a}(\sigma) p_{a \rightarrow i}(\sigma)}{p_{a \rightarrow i}(\sigma_i) \sum_{\sigma} q_{i \rightarrow a}(\sigma) \hat{p}_{a \rightarrow i}(\sigma)} - 1 \right]. \quad (2.45)$$

上式中 $\hat{p}_{a \rightarrow i}(\sigma)$ 为另外一个概率分布函数，它的表达式为

$$\hat{p}_{a \rightarrow i}(\sigma_i) = \frac{\sum_{\underline{\sigma}_{\partial a}} \delta(\sigma_i^a, \sigma_i) \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}(\sigma_j^a)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}(\sigma_j^a)}, \quad (2.46)$$

其中 $\partial a \setminus i$ 表示集合 ∂a 除去元素 i 后的子集。若函数 $\hat{p}_{a \rightarrow i}(\sigma)$ 刚好等于 $p_{a \rightarrow i}(\sigma)$ ，那么由(2.45)有 $L_g = 0$ 。

由上面的这些讨论可知，配分函数的展开表达式(2.39)中引入的辅助概率函数 $\{q_{i \rightarrow a}(\sigma), p_{a \rightarrow i}(\sigma)\}$ 应该选取为如下方程组的解（或称为不动点）：

$$q_{i \rightarrow a}(\sigma) = I_{i \rightarrow a}(p_{\partial i \setminus a}) \equiv \frac{1}{Z_{i \rightarrow a}} \psi_i(\sigma) \prod_{b \in \partial i \setminus a} p_{b \rightarrow i}(\sigma), \quad (2.47a)$$

$$p_{a \rightarrow i}(\sigma) = A_{a \rightarrow i}(q_{\partial a \setminus i}) \equiv \frac{1}{Z_{a \rightarrow i}} \sum_{\underline{\sigma}_{\partial a}} \delta(\sigma_i, \sigma) \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}(\sigma_j). \quad (2.47b)$$

上式中， $\underline{\sigma}_{\partial a}$ 表示功能节点 a 所连的变量节点集合 ∂a 的状态， $\underline{\sigma}_{\partial a} = \{\sigma_i | i \in \partial a\}$ ； $p_{\partial i \setminus a}$ 代表概率分布函数集合 $\{p_{b \rightarrow i}(\sigma_i) | b \in \partial i \setminus a\}$ ； $q_{\partial a \setminus i}$ 代表概率分布函数集合 $\{q_{j \rightarrow a}(\sigma_j) | j \in \partial a \setminus i\}$ ；而 $Z_{i \rightarrow a}$ 和 $Z_{a \rightarrow i}$ 是两个归一化常数，

$$Z_{i \rightarrow a} = \sum_{\sigma} \psi_i(\sigma) \prod_{b \in \partial i \setminus a} p_{b \rightarrow i}(\sigma), \quad Z_{a \rightarrow i} = \sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}(\sigma_j). \quad (2.48)$$

方程组(2.47)称为**信念传播方程** (belief propagation, 简称BP)，它首先是在不包含任何回路的树状因子网络上被推导出来的[31]。对任意一个包含回路的因子网络 G ，从配分函数展开的观点来看，为了使每一个包含摇摆边的子网络 g 对配分函数的修正贡献 $L_g = 0$ ，概率分布函数 $\{q_{i \rightarrow a}(\sigma), p_{a \rightarrow i}(\sigma)\}$ 必然被选择为BP方程的一个不动点。从这个意义上说，BP方程是概率分布函数需要满足的一组自洽方程，对任意由(2.1)所定义的因子网络模型都适用。注意到我们在推导BP方程组的过程中并没有要求变量节点状态必须为两分量自旋，所以(2.47)实际上对更一般的系统都成立（更详细的讨论见[42]）。

如果变量节点的状态为两分量自旋, 那么概率分布 $p_{a \rightarrow i}(\sigma)$ 和 $q_{i \rightarrow a}(\sigma)$ 分别可以用磁场 $u_{a \rightarrow i}$ 和 $h_{i \rightarrow a}$ 来完全表征, 即

$$p_{a \rightarrow i}(\sigma) = \frac{e^{\beta u_{a \rightarrow i} \sigma}}{2 \cosh(\beta u_{a \rightarrow i})}, \quad (2.49a)$$

$$q_{i \rightarrow a}(\sigma) = \frac{e^{\beta h_{i \rightarrow a} \sigma}}{2 \cosh(\beta h_{i \rightarrow a})}. \quad (2.49b)$$

由(2.47)可以推导出磁场 $h_{i \rightarrow a}$ 和 $u_{a \rightarrow i}$ 所满足的BP方程:

$$h_{i \rightarrow a} = h_i^0 + \sum_{b \in \partial i \setminus a} u_{b \rightarrow i}, \quad (2.50a)$$

$$u_{a \rightarrow i} = \frac{1}{2\beta} \ln \left[\frac{\sum_{\sigma_a} \delta(\sigma_i, +1) \psi_a(\sigma_a) \prod_{j \in \partial a \setminus i} e^{\beta h_{j \rightarrow a} \sigma_j}}{\sum_{\sigma_a} \delta(\sigma_i, -1) \psi_a(\sigma_a) \prod_{j \in \partial a \setminus i} e^{\beta h_{j \rightarrow a} \sigma_j}} \right]. \quad (2.50b)$$

在(2.50a)式中, h_i^0 代表作用于变量节点 i 的外场, 其表达式为

$$h_i^0 \equiv \frac{1}{2\beta} \ln \left[\frac{\psi_i(+1)}{\psi_i(-1)} \right] = -\frac{E_i(+1) - E_i(-1)}{2}. \quad (2.51)$$

如果给定相互作用能量的具体表达式, 方程(2.50b)有可能写成更简单的形式。下面的两个例子在统计物理学的研究文献中很常见:

习题 2.1 考虑只涉及变量节点 i 和 j 的相互作用 a , 能量 $E_a(\sigma_i, \sigma_j) = -J_{ij} \sigma_i \sigma_j$, 其中 J_{ij} 是耦合常数。验证方程(2.50b)可以写为

$$u_{a \rightarrow i} = \frac{1}{\beta} \operatorname{atanh} \left[\tanh(\beta J_{ij}) \tanh(\beta h_{j \rightarrow a}) \right]. \quad (2.52)$$

习题 2.2 考虑一个多体相互作用 a , 其能量函数为 $E_a = -J_a \prod_{j \in \partial a} \sigma_j$, 其中 J_a 是耦合常数。验证方程(2.50b)可以写为

$$u_{a \rightarrow i} = \frac{1}{\beta} \operatorname{atanh} \left[\tanh(\beta J_a) \prod_{j \in \partial a \setminus i} \tanh(\beta h_{j \rightarrow a}) \right]. \quad (2.53)$$

信念传播方程(2.47)是一组涉及很多概率分布函数的自洽方程, 寻找其不动点并非平庸的任务。我们可以将信念传播方程(2.47)和(2.50)看成是因子网络 G 上的信息传播和更新过程。例如对于示意图2.5所示的局部网络结构:

变量节点 i 收集从最近邻功能节点 b 和 c 传来的消息 $u_{b \rightarrow i}$ 和 $u_{c \rightarrow i}$, 通过方程(2.50a)得到 $h_{i \rightarrow a}$ 并将其作为输出消息传递给功能节点 a ; 功能节点 a 则收集

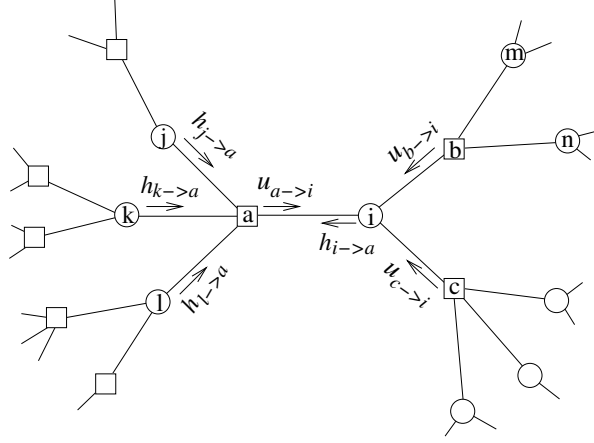


图 2.5: 信念传播示意图。在每条边上都有两个磁场消息，例如边 (i, a) 上的磁场 $h_{i \rightarrow a}$ 和 $u_{a \rightarrow i}$ 。这些磁场消息的值根据BP方程 (2.50) 不停地更新。

从所有其它最近邻变量节点 j, k, l 传递来的消息 $h_{j \rightarrow a}$ 、 $h_{k \rightarrow a}$ 和 $h_{l \rightarrow a}$ ，通过方程 (2.50b) 得到 $u_{a \rightarrow i}$ 并将其传递给变量节点 i 。这样在因子网络 G 的每一条边上都有两个磁场消息在不停地更新，直到所有的边上的这一对消息都不再发生变化为止。

在计算机上实现上述过程常常用随机排序迭代更新方法，附录??给出了具体实现的伪代码。

对有些问题，BP方程(2.47)或(2.50)的迭代过程会出现振荡现象。为了使迭代过程更快地收敛到不动点，可以在迭代过程中适当加入阻尼，例如将(2.50a) 改写成

$$h_{i \rightarrow a}(t+1) = \eta h_{i \rightarrow a}(t) + (1 - \eta) \left[h_i^0 + \sum_{b \in \partial i \setminus a} u_{b \rightarrow i}(t) \right], \quad (2.54)$$

其中 t 表示迭代时间， $0 \leq \eta < 1$ 是引入的一个阻尼系数。

当环境温度 T 足够低时，常常会出现这样一种情况，即上述的迭代算法总是不能收敛到一个不动点。这种不收敛通常意味着BP方程有多个不动点。我们将在下一章详细讨论如何处理这种情况。

2.4 Bethe-Peierls近似

从配分函数展开的观点来看，为了使所有带有摇摆边的子网络对配分函

数的修正贡献为零，辅助概率分布函数必须要满足BP方程(2.47)。为了进一步理解BP方程，我们现在以Bethe-Peierls近似[4, 32] 作为基础，从物理的角度“推导”出BP方程。

先回顾Bethe-Peierls近似的关键思想。考虑示意图2.5中功能节点 a 所代表的相互作用能量 $E_a(\underline{\sigma}_{\partial a})$ 。变量节点 i, j, k, l 参与到了该相互作用，即 $\underline{\sigma}_{\partial a} = (\sigma_i, \sigma_j, \sigma_k, \sigma_l)$ 。在系统处于热力学平衡时能量 E_a 的平均值为

$$\langle E_a(\underline{\sigma}_{\partial a}) \rangle = \frac{\sum_{\underline{\sigma}} e^{-\beta E(\underline{\sigma})} E_a(\underline{\sigma}_{\partial a})}{\sum_{\underline{\sigma}} e^{-\beta E(\underline{\sigma})}} = \frac{\sum_{\underline{\sigma}} e^{-\beta E_a^{(cv)}(\underline{\sigma})} \psi_a(\underline{\sigma}_{\partial a}) E_a(\underline{\sigma}_{\partial a})}{\sum_{\underline{\sigma}} e^{-\beta E_a^{(cv)}(\underline{\sigma})} \psi_a(\underline{\sigma}_{\partial a})},$$

其中

$$E_a^{(cv)}(\underline{\sigma}) \equiv E(\underline{\sigma}) - E_a(\underline{\sigma}_{\partial a})$$

是将功能节点 a 从系统中挖除后剩下的子系统的总能量（子系统称为一个空腔（cavity）系统）。经过简单的推导可将平均能量进一步写为

$$\langle E_a(\underline{\sigma}_{\partial a}) \rangle = \frac{\sum_{\underline{\sigma}_{\partial a}} Q_a^{(cv)}(\underline{\sigma}_{\partial a}) \psi_a(\underline{\sigma}_{\partial a}) E_a(\underline{\sigma}_{\partial a})}{\sum_{\underline{\sigma}_{\partial a}} Q_a^{(cv)}(\underline{\sigma}_{\partial a}) \psi_a(\underline{\sigma}_{\partial a})}. \quad (2.55)$$

上式中概率分布 $Q_a^{(cv)}$ 是在除去功能节点 a 的空腔系统中，集合 ∂a 中的变量节点（即图2.5中的 i, j, k, l ）自旋值的边缘概率分布：

$$Q_a^{(cv)}(\underline{\sigma}_{\partial a}) \equiv \frac{1}{\sum_{\underline{\sigma}} e^{-\beta E_a^{(cv)}(\underline{\sigma})}} \sum_{\underline{\sigma} \setminus \underline{\sigma}_{\partial a}} e^{-\beta E_a^{(cv)}(\underline{\sigma})}, \quad (2.56)$$

其中 $\sum_{\underline{\sigma} \setminus \underline{\sigma}_{\partial a}}$ 表示对除集合 ∂a 外所有其它变量节点的自旋态进行求和。

习题 2.3 验证表达式(2.55)的正确性并理解其统计物理含义。

在不包含功能节点 a 的空腔系统中（参见示意图2.6，左），集合 ∂a 中的变量节点不再因为相互作用能量 E_a 而相互关联。如果这些节点 i, j, k, l 在该空腔系统中是互相不连通的，那它们的自旋态必定完全独立，从而 $Q_a^{(cv)}(\underline{\sigma}_{\partial a})$ 将是这些节点的自旋边缘概率分布的乘积。但如果这些节点在空腔系统中仍然是通过其它功能节点而相互连通的，那它们之间就会因为空腔系统中的其它相互作用而互相关联。作为第一级近似，先不考虑 $\sigma_i, \sigma_j, \sigma_k, \sigma_l$ 在该空腔系统中的关联，而将 $Q_a^{(cv)}(\underline{\sigma}_{\partial a})$ 写成概率分布乘积的形式，即

$$Q_a^{(cv)}(\underline{\sigma}_{\partial a}) \approx \prod_{j \in \partial a} q_{j \rightarrow a}^{(cv)}(\sigma_j), \quad (2.57)$$

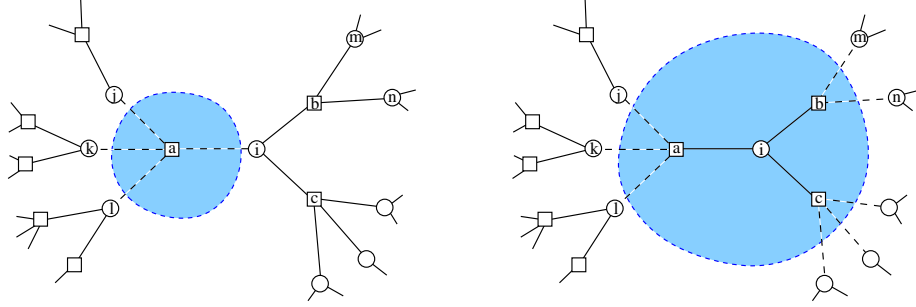


图 2.6: 空腔子网络示意图。(左) 将功能节点 a 从因子网络中挖掉; (右) 将变量节点 i 及其所有最近邻功能节点从因子网络中挖掉。

其中 $q_{j \rightarrow a}^{(cv)}(\sigma_j)$ 是变量节点 j 在空腔系统的自旋态边缘概率分布 (称为空腔概率分布):

$$q_{j \rightarrow a}^{(cv)}(\sigma_j) \equiv \frac{1}{\sum_{\underline{\sigma}} e^{-\beta E_a^{(cv)}(\underline{\sigma})}} \sum_{\underline{\sigma} \setminus \sigma_j} e^{-\beta E_a^{(cv)}(\underline{\sigma})}. \quad (2.58)$$

方程 (2.57) 是 Bethe-Peierls 近似的一种形式 [4, 32, 39, 2]。该表达式假定与功能节点 a 相连的变量节点在 a 被挖去后的子系统中不存在任何关联。将表达式 (2.57) 代入 (2.55) 就得到

$$\langle E_a(\underline{\sigma}_{\partial a}) \rangle = \frac{\sum_{\underline{\sigma}_{\partial a}} E_a(\underline{\sigma}_{\partial a}) \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j)}. \quad (2.59)$$

功能节点 a 的最近邻变量节点集合为 ∂a , 这些节点自旋态的联合概率分布记为 $p_{\partial a}(\underline{\sigma}_{\partial a})$:

$$p_{\partial a}(\underline{\sigma}_{\partial a}) \equiv \frac{\sum_{\underline{\sigma} \setminus \underline{\sigma}_{\partial a}} e^{-\beta E(\underline{\sigma})}}{\sum_{\underline{\sigma}} e^{-\beta E(\underline{\sigma})}}. \quad (2.60)$$

由平均能量表达式 (2.59) 可以看出, 在 Bethe-Peierls 近似下, 该联合概率分布的近似表达式为

$$p_{\partial a}(\underline{\sigma}_{\partial a}) \approx \frac{\psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j)}. \quad (2.61)$$

习题 2.4 应用 Bethe-Peierls 近似 (2.57) 从定义式 (2.60) 推导出联合概率分布近似表达式 (2.61)。理解 (2.61) 的统计物理含义。

再来考察变量节点自旋态的边缘概率分布。示意图 2.5 中变量节点 i 通过功能节点 a, b, c 与 $j, k, \dots$ 等 9 个变量节点发生相互作用, 故自旋 σ_i 受到 $\sigma_j, \sigma_k, \dots$ 等

其它9个自旋值的影响。类似于方程 (2.55)，自旋 σ_i 的平均值的表达式为

$$\langle \sigma_i \rangle = \frac{\sum_{\sigma_i} \sum_{\underline{\sigma}_{\partial a \setminus i}} \sum_{\underline{\sigma}_{\partial b \setminus i}} \sum_{\underline{\sigma}_{\partial c \setminus i}} Q_{i+\partial i}^{(cv)}(\underline{\sigma}_{\partial a \setminus i}, \underline{\sigma}_{\partial b \setminus i}, \underline{\sigma}_{\partial c \setminus i}) \psi_a \psi_b \psi_c \psi_i(\sigma_i) \sigma_i}{\sum_{\sigma_i} \sum_{\underline{\sigma}_{\partial a \setminus i}} \sum_{\underline{\sigma}_{\partial b \setminus i}} \sum_{\underline{\sigma}_{\partial c \setminus i}} Q_{i+\partial i}^{(cv)}(\underline{\sigma}_{\partial a \setminus i}, \underline{\sigma}_{\partial b \setminus i}, \underline{\sigma}_{\partial c \setminus i}) \psi_a \psi_b \psi_c \psi_i(\sigma_i)}. \quad (2.62)$$

上式中， $\underline{\sigma}_{\partial a \setminus i} \equiv \{\sigma_j : j \in \partial a \setminus i\}$ 是 i 最近邻功能节点 a 所连的其它变量节点，对于示意图2.5而言 $\underline{\sigma}_{\partial a \setminus i} = \{\sigma_j, \sigma_k, \sigma_l\}$ 。概率分布 $Q_{i+\partial i}^{(cv)}(\underline{\sigma}_{\partial a \setminus i}, \underline{\sigma}_{\partial b \setminus i}, \underline{\sigma}_{\partial c \setminus i})$ 表示在节点 i 及其所有最近邻节点被挖去后剩下的空腔子系统中， $j, k, \dots$ 等9个变量节点的自旋态联合概率分布，参见示意图2.6 (右)。如果我们假定在这一个空腔系统中变量节点 $j, k, \dots$ 等是彼此独立的，那么类似于表达式 (2.57) 可有如下的近似：

$$Q_{i+\partial i}^{(cv)}(\underline{\sigma}_{\partial a \setminus i}, \underline{\sigma}_{\partial b \setminus i}, \underline{\sigma}_{\partial c \setminus i}) \approx \prod_{a \in \partial i} \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}^{(cv)}(\sigma_j). \quad (2.63)$$

将 (2.63) 代入 (2.62) 就得到

$$\langle \sigma_i \rangle \approx \frac{\sum_{\sigma_i} \sigma_i \psi_i(\sigma_i) p_{a \rightarrow i}^{(cv)}(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}{\sum_{\sigma_i} \psi_i(\sigma_i) p_{a \rightarrow i}^{(cv)}(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}, \quad (2.64)$$

其中

$$p_{a \rightarrow i}^{(cv)}(\sigma_i) \equiv \frac{\sum_{\underline{\sigma}_{\partial a \setminus i}} \psi_a(\underline{\sigma}_a) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}^{(cv)}(\sigma_j)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_a) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}^{(cv)}(\sigma_j)}. \quad (2.65)$$

由 (2.64) 可知变量节点 i 的边缘概率分布 $q_i(\sigma_i)$ 在Bethe-Peierls近似下的表达式为

$$q_i(\sigma_i) \approx \frac{\psi_i(\sigma_i) p_{a \rightarrow i}^{(cv)}(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}{\sum_{\sigma_i} \psi_i(\sigma_i) p_{a \rightarrow i}^{(cv)}(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}. \quad (2.66)$$

在Bethe-Peierls近似中，非常重要的一个概率分布是变量节点 i 的空腔概率分布，例如 $q_{i \rightarrow a}^{(cv)}(\sigma_i)$ 等等。根据定义式 (2.58)，概率分布 $q_{i \rightarrow a}^{(cv)}(\sigma_i)$ 是当节点 i 不受到相互作用 a 的影响的情况下的边缘概率分布。在Bethe-Peierls近似的情况下这一空腔概率分布的表达式为

$$q_{i \rightarrow a}^{(cv)}(\sigma_i) \approx \frac{\psi_i(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}{\sum_{\sigma_i} \psi_i(\sigma_i) p_{b \rightarrow i}^{(cv)}(\sigma_i) p_{c \rightarrow i}^{(cv)}(\sigma_i)}. \quad (2.67)$$

习题 2.5 按照推导表达式 (2.66) 的同一方法推导出表达式 (2.67)。

从表达式 (2.66) 和 (2.67) 可以看出, 概率分布 $p_{a \rightarrow i}^{(cv)}(\sigma_i)$ 的统计物理含义是变量节点 i 的自旋态在只受到功能节点 a 的影响下的边缘概率分布。为了使这一点更清楚, 我们可以思考如下的问题: 对于图 2.6 (右) 所示的空腔系统, 现在将功能节点 a 添上, 这样节点 i 就会与变量节点 j, k, l 的状态发生关联, 那么自旋 σ_i 的边缘概率是多少? 重复本节前面的推导过程可以得出该边缘概率在 Bethe-Peierls 近似下即为 $p_{a \rightarrow i}^{(cv)}(\sigma_i)$, 由表达式 (2.65) 计算出来。

将表达式 (2.65) 和 (2.67) 与 BP 方程 (2.47) 相比较, 可以发现如果对每个概率分布都作如下的替换:

$$p_{a \rightarrow i}(\sigma_i) \Rightarrow p_{a \rightarrow i}^{(cv)}(\sigma_i), \quad q_{i \rightarrow a}(\sigma_i) \Rightarrow q_{i \rightarrow a}^{(cv)}(\sigma_i) \quad (2.68)$$

那么 BP 方程 (2.47) 与就变成方程 (2.65) 和 (2.67)。

通过这一节的讨论, 配分函数展开表达式 (2.39) 中辅助概率分布函数 $\{p_{a \rightarrow i}(\sigma_i), q_{i \rightarrow a}(\sigma_i)\}$ 的物理意义就清楚了: $p_{a \rightarrow i}(\sigma_i)$ 是变量节点 i 只参与相互作用 a 的话, 它的微观态 σ_i 的概率分布; 而 $q_{i \rightarrow a}(\sigma_i)$ 是变量节点 i 不参与相互作用 a 的话, 它的微观态 σ_i 的概率分布。

2.5 复本对称空腔理论

由节 2.3 的分析可知, 如果我们将满足信念传播方程 (2.47) 的一组概率分布 $\{q_{i \rightarrow a}(\sigma_i^a), p_{a \rightarrow i}(\sigma_i)\}$ 作为展开辅助函数, 配分函数 Z 的展开表达式为

$$Z = Z_0 \times \left(1 + \sum_{\mathbb{G} \subseteq G} L_{\mathbb{G}}\right). \quad (2.69)$$

在上式中, $\mathbb{G}$ 代表因子网络 G 的一个不包含任何摇摆边的子网络, 它的每个节点的连通度至少为 2。这样的子网络称为**圈图**。方程 (2.69) 称为配分函数的**圈图展开式**。相应的, 系统自由能 F 的表达式为

$$F = -k_B T \ln Z = F_0 + \Delta F \quad (2.70)$$

其中 $F_0 = -k_B T \ln Z_0$; 而圈图修正项 ΔF 为

$$\Delta F = -k_B T \ln \left[1 + \sum_{\mathbb{G} \subseteq G} L_{\mathbb{G}}\right]. \quad (2.71)$$

如果我们忽略 ΔF ，则得到系统自由能的近似表达式 $F \approx F_0$ ，

$$F_0 = \sum_{a \in G} f_a + \sum_{i \in G} f_i - \sum_{(i,a) \in G} f_{(i,a)}, \quad (2.72)$$

其中 f_a 、 f_i 和 $f_{(i,a)}$ 的表达式分别由(2.27)、(2.33)和(2.35)给出。 F_0 在文献中常常称为**Bethe-Peierls自由能**，简称BP自由能。 F_0 是所有功能节点的自由能贡献 f_a 与所有变量节点的贡献 f_i 之和和再从中减去所有边 (i,a) 的贡献 $f_{(i,a)}$ 后所得到的结果。对于要将每一条边的自由能贡献从总自由能中减去的原因，我们有一个直观的理解。在计算功能节点 a 的自由能贡献时，与之相连的边 (i,a) 的自由能贡献已经被考虑进去了；在计算变量节点 i 的自由能贡献时，边 (i,a) 的效应又被考虑进去了；因此为了抵消这种重复考虑导致的偏差，需要在总自由能的结果中将边 (i,a) 的自由能贡献 $f_{(i,a)}$ 减去一次。

BP自由能也可以看成是概率分布函数 $\{q_{i \rightarrow a}, p_{a \rightarrow i}\}$ 的泛函：

$$\begin{aligned} \beta F_0 = & - \sum_a \ln \left[\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i) \right] - \sum_i \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i) \right] \\ & + \sum_{(i,a)} \ln \left[\sum_{\sigma_i} q_{i \rightarrow a}(\sigma_i) p_{a \rightarrow i}(\sigma_i) \right]. \end{aligned} \quad (2.73)$$

该泛函对函数 $q_{i \rightarrow a}$ 和 $p_{a \rightarrow i}$ 的变分分别为

$$\frac{\delta \beta F_0}{\delta q_{i \rightarrow a}(\sigma_i)} = \frac{p_{a \rightarrow i}(\sigma_i)}{\sum_{\sigma} q_{i \rightarrow a}(\sigma) p_{a \rightarrow i}(\sigma)} - \frac{\sum_{\underline{\sigma}_{\partial a \setminus i}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a \setminus i} q_{j \rightarrow a}(\sigma_j)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j)}, \quad (2.74a)$$

$$\frac{\delta \beta F_0}{\delta p_{a \rightarrow i}(\sigma_i)} = \frac{q_{i \rightarrow a}(\sigma_i)}{\sum_{\sigma} q_{i \rightarrow a}(\sigma) p_{a \rightarrow i}(\sigma)} - \frac{\psi_i(\sigma_i) \prod_{b \in \partial i \setminus a} p_{b \rightarrow i}(\sigma_i)}{\sum_{\sigma} \psi_i(\sigma) \prod_{b \in \partial i} p_{b \rightarrow i}(\sigma)}. \quad (2.74b)$$

在BP方程(2.47)的不动点处自由能 F_0 作为泛函的一阶变分为零：

$$\left. \frac{\delta F_0}{\delta q_{i \rightarrow a}(\sigma_i)} \right|_{BP} = 0, \quad \left. \frac{\delta F_0}{\delta p_{a \rightarrow i}(\sigma_i)} \right|_{BP} = 0, \quad \forall (i,a) \in G. \quad (2.75)$$

这是BP自由能(2.73)一个非常重要的性质。

习题 2.6 对于所有自旋变量为 $\sigma_i = \pm 1$ 的情形，BP自由能(2.73)可以表达成磁场 $\{h_{i \rightarrow a}\}$ 和磁场 $\{u_{a \rightarrow i}\}$ 的函数。验证这一表达式为

$$\begin{aligned} \beta F_0 = & - \sum_a \ln \left[\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) e^{\sum_{i \in \partial a} \beta h_{i \rightarrow a} \sigma_i} \right] - \sum_i \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) e^{\sum_{a \in \partial i} \beta u_{a \rightarrow i} \sigma_i} \right] \\ & + \sum_{(i,a)} \ln \left[2 \cosh(\beta(u_{a \rightarrow i} + h_{i \rightarrow a})) \right]. \end{aligned} \quad (2.76)$$

Bethe-Peierls自由能 (2.73) 及信念传播方程 (2.47) 构成系统 (2.1) 的复本对称空腔理论 (replica-symmetric (RS) cavity theory) [23, 21]。该理论将平衡自由能的所有圈图修正贡献都忽略, 假设系统的平衡自由能等于BP自由能 F_0 。我们将以BP自由能 F_0 为出发点计算系统平均能量, 熵, 变量节点自旋平均值及自旋之间的关联函数。在此之前, 先对自由能修正贡献 ΔF 稍微进行一点讨论。

如果模型所对应的因子网络 G 不存在任何回路, 即因子网络是树状的, 那么在BP方程不动点处Bethe-Peierls自由能就严格等于系统的平衡自由能。对于这种情况, $\Delta F = 0$ 。而且可以证明BP方程 (2.47) 一定只存在一个不动点, 该不动点可以很容易通过局部迭代而得到 (从树状因子网络的边缘开始向内部进行BP迭代)。树状系统的平衡自由能计算是简单的。

对于带有回路的一般因子网络, 如果 $\Delta F \leq 0$, 那么就有 $F \leq F_0$, 即自由能的BP近似(2.72)给出系统自由能 F 的一个上限; 反之如果 $\Delta F \geq 0$, 则 F_0 是自由能 F 的一个下限。判断 ΔF 的正负并非容易的事情, 只在非常简单的系统能够做到, 例如如下的习题:

习题 2.7 考虑处在包含 N 个节点及 N 条边的一维环上的伊辛模型。自旋之间有最近邻相互作用, 能量函数为 $E(\sigma_1, \sigma_2, \dots, \sigma_N) = -J \sum_{i=1}^N \sigma_i \sigma_{i+1}$, 其中 $\sigma_i = \pm 1$, $\sigma_{N+1} \equiv \sigma_1$ 。该模型系统只有一个不含摇摆边的子网络。(1) 验证BP方程(2.50)的解为 $h_{i \rightarrow a} = u_{a \rightarrow i} = 0$ 。(2) 验证BP自由能为 $F_0 = -Nk_B T \ln[2 \cosh(\beta J)]$ 。(3) 验证自由能的修正贡献为

$$\Delta F = -k_B T \ln(1 + \tanh^N(\beta J)). \quad (2.77)$$

在这一习题中, 如果相互作用为铁磁型($J > 0$), 则有 $\Delta F < 0$ 。如果相互作用为反铁磁型($J < 0$), 则当 N 为偶数时 $\Delta F < 0$, 而当 N 为奇数时 $\Delta F > 0$ 。

对反铁磁相互作用伊辛环, 当相互作用数 N 为奇时, 环中存在阻挫, 即不存在一个自旋微观构型 $(\sigma_1, \sigma_2, \dots, \sigma_N)$ 能够使所有 N 个反铁磁相互作用的能量同时为最低($= -|J|$)。对任意自旋构型, 回路中至少有一个反铁磁相互作用能量为 $|J|$ 。上面的习题显示, 对这样一个存在阻挫的回路, 它的自由能修正贡献为正, 即BP自由能 F_0 比系统真正的自由能 F 要小。

阻挫广泛地存在于自旋玻璃系统中, 当一个圈图中回路的数目很多时, 阻挫的情况很复杂, 这导致圈图对配分函数的修正贡献的正负符号可能是多种因素相互竞争的结果。近年来在Sherrington-Kirkpatrick自旋玻璃模型[37]上展开的一些数学工作[40, 13, 15, 14], 倾向于支持这样的猜测: 对一个包含

很多阻挫的完全连通自旋玻璃系统，其自由能的BP近似 F_0 是系统真正自由能 F 的下限。这方面进一步的理论工作当然还很艰巨。

当温度 $T \rightarrow 0$ 时，系统的自由能趋向于系统的基态能量。对许多非平庸自旋玻璃模型系统而言，圈图对系统基态能量的修正贡献常常是不可忽略的。在文献[45, 46]中讨论了阻挫效应对一些随机自旋玻璃系统基态能量的影响。随机系统中的回路长度平均而言随变量节点数目 N 的增加而对数式增长。文献[45, 46]将这一类系统中的阻挫效应看成是一种浸润现象，并提出了一种长程阻挫平均场理论。该理论对系统的基态能量密度能给出较好的估计。

2.5.1 Bethe-Peierls自由能的其它两种形式

在BP方程的不动点处，我们可以将BP自由能表达式(2.73)中的所有概率分布函数 $p_{a \rightarrow i}(\sigma)$ 用方程(2.47b)右侧的表达式来替代，这样BP自由能就写成了如下的形式，即

$$\begin{aligned} \beta F_0 = & - \sum_i \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{b \in \partial i} \sum_{\sigma_{\partial b}} \delta(\sigma_i^b, \sigma_i) \psi_b(\sigma_{\partial b}) \prod_{k \in \partial b \setminus i} q_{k \rightarrow b}(\sigma_k) \right] \\ & + \sum_a (|\partial a| - 1) \ln \left[\sum_{\sigma_{\partial a}} \psi_a(\sigma_{\partial a}) \prod_{j \in \partial a} q_{j \rightarrow a}(\sigma_j) \right]. \end{aligned} \quad (2.78)$$

BP自由能(2.73)也可以表达成概率分布 $\{p_{a \rightarrow i}(\sigma)\}$ 的函数，即

$$\begin{aligned} \beta F_0 = & - \sum_a \ln \left[\sum_{\sigma_{\partial a}} \psi_a(\sigma_{\partial a}) \prod_{j \in \partial a} \psi_j(\sigma_j) \prod_{b \in \partial j \setminus a} p_{b \rightarrow j}(\sigma_j) \right] \\ & + \sum_i (|\partial i| - 1) \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{b \in \partial i} p_{b \rightarrow i}(\sigma_i) \right]. \end{aligned} \quad (2.79)$$

自由能表达式(2.78)和(2.79)在文献中用得很多，它们相比于表达式(2.73)有一些数值计算上的优越性。

2.5.2 平均能量和熵

系统的平均能量与自由能的关系由(2.15)给出。利用自由能的圈图展开表达式(2.70)可以得到

$$\langle E \rangle = \langle E \rangle_0 + \langle \Delta E \rangle. \quad (2.80)$$

上式的 $\langle E \rangle_0$ 称为系统平均能量的Bethe-Peierls近似值。利用(2.73)以及(2.75)，

我们可以导出

$$\begin{aligned} \langle E \rangle_0 \equiv \frac{\partial[\beta F_0]}{\partial \beta} = & \sum_a \left[\frac{\sum_{\underline{\sigma}_{\partial a}} E_a(\underline{\sigma}_{\partial a}) \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i)}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i)} \right] \\ & + \sum_i \left[\frac{\sum_{\sigma_i} E_i(\sigma_i) \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i)}{\sum_{\sigma} \psi_i(\sigma) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma)} \right], \quad (2.81) \end{aligned}$$

即 $\langle E \rangle_0$ 是所有功能节点 a 与变量节点 i 的平均能量之和。平均能量的修正项 $\langle \Delta E \rangle$ 为

$$\langle \Delta E \rangle \equiv \frac{\partial[\beta \Delta F]}{\partial \beta} = \frac{-1}{1 + \sum_{\mathbb{G}} L_{\mathbb{G}}} \sum_{\mathbb{G}} \frac{\partial L_{\mathbb{G}}}{\partial \beta}. \quad (2.82)$$

由(2.14)可知, 在逆温度 β 处系统的熵 S 为

$$S = \frac{1}{T} (\langle E \rangle - F) = S_0 + \Delta S, \quad (2.83)$$

其中 S_0 称为熵的 Bethe-Peierls 近似,

$$S_0 = \frac{\langle E \rangle_0 - F_0}{T}, \quad (2.84)$$

而熵的修正项为

$$\Delta S = \frac{\langle \Delta E \rangle - \Delta F}{T}. \quad (2.85)$$

2.5.3 边缘概率分布及其相容性

由 (2.81) 可以看出, 在复本对称空腔理论中变量节点 i 的自旋态边缘概率分布的表达式为

$$q_i(\sigma_i) = \frac{1}{\sum_{\sigma} \psi_i(\sigma) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma)} \psi_i(\sigma_i) \prod_{a \in \partial i} p_{a \rightarrow i}(\sigma_i); \quad (2.86)$$

而任一功能节点 a 所连的变量节点自旋态的联合概率分布的表达式则为

$$p_a(\underline{\sigma}_{\partial a}) = \frac{1}{\sum_{\underline{\sigma}_{\partial a}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i)} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in \partial a} q_{i \rightarrow a}(\sigma_i). \quad (2.87)$$

这两个表达式也可以根据 Bethe-Peierls 近似推导出来, 参见 (2.66) 和 (2.61)。

由功能节点 a 的边缘概率 $p_a(\underline{\sigma}_{\partial a})$ 出发可以求得每个参与该相互作用的变量节点 i 的自旋边缘概率分布。例如, 对于示意图 2.5 而言, 由功能节点 a 推断

出变量节点 i 的自旋态边缘概率分布等于 $\sum_{\sigma_j, \sigma_k, \sigma_l} p_a(\sigma_i, \sigma_j, \sigma_k, \sigma_l)$ 。通过这种方式得到的自旋态边缘概率分布是否与方程(2.86)所得到的结果相同? 答案是肯定的。在BP方程(2.47)的不动点处边缘概率分布有如下**相容性**:

$$q_i(\sigma_i) = \sum_{\underline{\sigma}_{\partial a \setminus i}} p_a(\underline{\sigma}_{\partial a}), \quad \forall (i, a) \in G. \quad (2.88)$$

上式中, $\sum_{\underline{\sigma}_{\partial a \setminus i}}$ 表示对功能节点 a 的所有最近邻变量节点的自旋态进行求和, 但不包括节点 i 的自旋 σ_i 。上述关系的证明很容易, 作为练习题留给读者:

习题 2.8 证明边缘概率分布 $q_i(\sigma_i)$ 和 $p_a(\underline{\sigma}_{\partial a})$ 在信念传播方程(2.47)的不动点处满足相容关系(2.88)。

相容性关系(2.88)有重要的意义, 它表明复本对称空腔理论对于单个变量节点状态的描述至少是自洽的。由这一关系可以确保当一个变量节点 i 参与到多个相互作用(例如图2.5中的 a, b, c)时, 不同的概率分布函数 p_a, p_b, p_c 对于节点 i 状态的统计描述都相同。

如果因子网络 G 不含任何回路, 那么方程(2.86)和(2.87)所给出的边缘概率分布是完全精确的。对于这种树状系统, 系统的平衡玻耳兹曼分布(2.4)实际上可以写为如下的乘积形式:

$$P_B(\sigma_1, \sigma_2, \dots, \sigma_N) = \prod_{a \in G} p_a(\underline{\sigma}_{\partial a}) \prod_{i \in G} [q_i(\sigma_i)]^{1-k_i}, \quad (2.89)$$

其中 $k_i \equiv |\partial i|$ 是变量节点 i 的连通度。

习题 2.9 证明表达式(2.89)对于定义在任何不包含回路的因子网络上的自旋玻璃模型都成立, 其中概率分布函数 $q_i(\sigma_i)$ 及 $p_a(\underline{\sigma}_{\partial a})$ 由BP方程(2.47)的不动点分别通过(2.86)和(2.87)得到。

若因子网络中包含回路, 表达式(2.89)将不再成立, 但该式右侧可能仍然是平衡分布 $P_B(\underline{\sigma})$ 一个很好的近似(但它并不一定是归一化的)。

2.5.4 自旋关联函数

在这一小节, 为讨论方便起见, 我们局限于节点状态为 ± 1 伊辛自旋的情形。在平衡时, 变量节点 i 的自旋平均值为

$$\langle \sigma_i \rangle \equiv \sum_{\underline{\sigma}} \sigma_i P_B(\underline{\sigma}). \quad (2.90)$$

这一平均值可以由自由能 F 对外场 h_i^0 求偏导数而得到, 即

$$\langle \sigma_i \rangle \equiv -\frac{\partial F}{\partial h_i^0} = \langle \sigma_i \rangle_0 + \langle \Delta \sigma_i \rangle, \quad (2.91)$$

其中 $\langle \sigma_i \rangle_0$ 是自旋平均值的Bethe-Peierls近似值

$$\langle \sigma_i \rangle_0 = \sum_{\sigma_i} \sigma_i q_i(\sigma_i) = \tanh \left[\beta \left(h_i^0 + \sum_{a \in \partial i} u_{a \rightarrow i} \right) \right], \quad (2.92)$$

而平均自旋的圈图修正贡献项 $\langle \Delta \sigma_i \rangle$ 则为

$$\Delta \langle \sigma_i \rangle = -\frac{\partial \Delta F}{\partial h_i^0}. \quad (2.93)$$

两个变量节点 i 和 j 的自旋态之间的关联可由**连接关联** (connected correlation) $\langle \sigma_i \sigma_j \rangle_c$ 来表征:

$$\langle \sigma_i \sigma_j \rangle_c \equiv \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle = -\frac{1}{\beta} \frac{\partial^2 F}{\partial h_i^0 \partial h_j^0} = \frac{1}{\beta} \frac{\partial \langle \sigma_i \rangle}{\partial h_j^0}. \quad (2.94)$$

如果忽略自由能圈图修正贡献 ΔF , 那么两点关联的近似表达式为

$$\langle \sigma_i \sigma_j \rangle_c \approx \frac{1}{\beta} \frac{\partial \langle \sigma_i \rangle_0}{\partial h_j^0} = \left(1 - \langle \sigma_i \rangle_0^2 \right) \left(\delta_i^j + \sum_{a \in \partial i} \frac{\partial u_{a \rightarrow i}}{\partial h_j^0} \right). \quad (2.95)$$

为了计算自旋 σ_j 和系统中的所有其它自旋 σ_i 之间的关联, 就需要先计算出所有空腔场 $u_{a \rightarrow i}$ (它们由方程 (2.49a) 所定义) 对外场 h_j^0 的偏导数。

若外场 h_j^0 作微小改变时 $u_{a \rightarrow i}$ 的改变也很小, 由BP方程(2.50)就可以得到

$$\frac{\partial h_{i \rightarrow a}}{\partial h_j^0} = \delta_j^i + \sum_{b \in \partial i \setminus a} \frac{\partial u_{b \rightarrow i}}{\partial h_j^0}, \quad (2.96a)$$

$$\begin{aligned} \frac{\partial u_{a \rightarrow i}}{\partial h_j^0} = \frac{1}{2} \sum_{k \in \partial a \setminus i} \frac{\partial h_{k \rightarrow a}}{\partial h_j^0} & \left[\frac{\sum_{\underline{\sigma}_{\partial a}} \sigma_k \delta(\sigma_i, +1) \psi_a(\underline{\sigma}_{\partial a}) \prod_{l \in \partial a \setminus i} e^{\beta h_{l \rightarrow a} \sigma_l}}{\sum_{\underline{\sigma}_{\partial a}} \delta(\sigma_i, +1) \psi_a(\underline{\sigma}_{\partial a}) \prod_{l \in \partial a \setminus i} e^{\beta h_{l \rightarrow a} \sigma_l}} \right. \\ & \left. - \frac{\sum_{\underline{\sigma}_{\partial a}} \sigma_k \delta(\sigma_i, -1) \psi_a(\underline{\sigma}_{\partial a}) \prod_{l \in \partial a \setminus i} e^{\beta h_{l \rightarrow a} \sigma_l}}{\sum_{\underline{\sigma}_{\partial a}} \delta(\sigma_i, -1) \psi_a(\underline{\sigma}_{\partial a}) \prod_{l \in \partial a \setminus i} e^{\beta h_{l \rightarrow a} \sigma_l}} \right]. \end{aligned} \quad (2.96b)$$

容易注意到 (2.96) 是与BP方程 (2.50) 很类似的迭代方程, 这组方程的不动点也可以通过消息传递算法来获得。具体的计算步骤如下:

- (1) 通过迭代获得BP方程 (2.50) 的一个不动点 $\{u_{a \rightarrow i}, h_{i \rightarrow a}\}$ 。
- (2) 在因子网络所有的边 (i, a) 上赋初值: $\partial u_{a \rightarrow i} / \partial h_j^0 \leftarrow 0$, $\partial h_{i \rightarrow a} / \partial h_j^0 \leftarrow \delta_j^i$ 。
- (3) 按照方程(2.96)进行迭代, 更新 $\partial h_{i \rightarrow a} / \partial h_j^0$ 及 $\partial u_{a \rightarrow i} / \partial h_j^0$, 直到达到一个不动点。
- (4) 利用 (2.95) 计算自旋两点关联。

消息迭代公式 (2.96) 对于一些简单类型的自旋相互作用可以写成明显的形式:

习题 2.10 (1) 如果功能节点 a 代表两个变量节点 i 和 k 之间的相互作用, 且能量函数为 $E_a = -J_{ik} \sigma_i \sigma_k$, 那么 (2.96b) 可以简化为

$$\frac{\partial u_{a \rightarrow i}}{\partial h_j^0} = \frac{\partial h_{k \rightarrow a}}{\partial h_j^0} \frac{\tanh(\beta J_{ik}) [1 - \tanh^2(\beta h_{k \rightarrow a})]}{1 - \tanh^2(\beta J_{ik}) \tanh^2(\beta h_{k \rightarrow a})}. \quad (2.97)$$

(2) 如果功能节点 a 代表多个变量节点之间的自旋耦合, 其能量函数为 $E_a = -J_a \prod_{k \in \partial a} \sigma_k$, 那么 (2.96b) 可以简化为

$$\frac{\partial u_{a \rightarrow i}}{\partial h_j^0} = \sum_{k \in \partial a \setminus i} \frac{\partial h_{k \rightarrow a}}{\partial h_j^0} \frac{\tanh(\beta J_a) [1 - \tanh^2(\beta h_{k \rightarrow a})] \prod_{l \in \partial a \setminus i, k} \tanh(\beta h_{l \rightarrow a})}{1 - \tanh^2(\beta J_a) \prod_{l \in \partial a \setminus i} \tanh^2(\beta h_{l \rightarrow a})}. \quad (2.98)$$

方程(2.97)和(2.98)也可以分别由(2.52)和(2.53)出发推导出来。

两个变量节点 i, j 之间的距离 $d(i, j)$ 定义为因子网络 G 中连接 i 和 j 的最短路径上功能节点的数目 (见示意图2.7)。例如, 如果 i 和 j 都属于某个功能节点 a 的最近邻节点集合 ∂a , 则 $d(i, j) = 1$ 。如果 i 和 j 分属于 G 中两个互不相连的连通单元, 则可定义 $d(i, j) = \infty$ 。自旋两点关联 $\langle \sigma_i \sigma_j \rangle_c$ 通常随距离 $d(i, j)$ 指数式衰减, 即 $\langle \sigma_i \sigma_j \rangle_c \sim \exp[-d(i, j)/\xi]$, 其中 ξ 是特征关联长度。系统也可能存在某个临界温度 T_c , 当环境温度 T 很接近 T_c 时, 自旋两点关联 $\langle \sigma_i \sigma_j \rangle_c$ 趋向于以幂次形式随距离而衰减, 即 $\langle \sigma_i \sigma_j \rangle_c \sim [d(i, j)]^{-\lambda}$, 其中 λ 称为一个**临界指数**。后一种情况意味着系统的自旋两点关联特征长度趋向于无穷大, 对应于系统将在 T_c 处发生一个连续相变, 系统的宏观性质有定性的改变。

但是另一方面, 系统宏观性质出现定性改变并不一定意味着自旋两点关联特征距离一定会趋向无穷大。例如, 如果系统在温度 T_c 处发生的是一个非连续相变, 那么在相变温度两侧的自旋两点关联特征距离都是有限的。定量

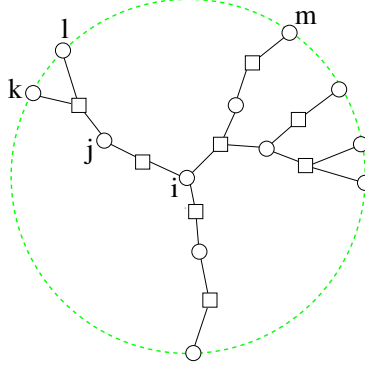


图 2.7: 变量节点自旋态之间的点-点关联及点-集合关联示意图。节点*i*和*j*的距离为1。而处于虚线圆圈上的节点*k, l, m*等与*i*的距离为2, 它们构成集合 $\partial_2 i$ 。

描述系统中的关联性质的另一个重要物理学量是**点和集合的关联** (point-to-set correlations) [20, 24]。考虑任一变量节点*i*并构造一个集合 $\partial_d i$, 集合 $\partial_d i$ 包含*G*中所有与*i*的距离为*d*的变量节点, 即 $\partial_d i \equiv \{j | d(i, j) = d\}$ 。为了考察变量节点*i* 的自旋状态 σ_i 与集合 $\partial_d i$ 的自旋状态 $\underline{\sigma}_{\partial_d i} \equiv \{\sigma_j | d(i, j) = d\}$ 的关联, 我们可以选择 σ_i 的某个合适函数, $f(\sigma_i)$, 以及自旋态 $\underline{\sigma}_{\partial_d i}$ 的某个合适函数, $g(\underline{\sigma}_{\partial_d i})$, 并研究这两个函数值之间的连接关联:

$$\langle f(\sigma_i)g(\underline{\sigma}_{\partial_d i}) \rangle_c \equiv \langle f(\sigma_i)g(\underline{\sigma}_{\partial_d i}) \rangle - \langle f(\sigma_i) \rangle \langle g(\underline{\sigma}_{\partial_d i}) \rangle. \quad (2.99)$$

如果 $\langle f(\sigma_i)g(\underline{\sigma}_{\partial_d i}) \rangle_c$ 随着距离*d*而指数式衰减, 即 $\langle f(\sigma_i)g(\underline{\sigma}_{\partial_d i}) \rangle_c \sim e^{-d/\xi}$, 则说明一个处于系统内部的变量节点*i* 的自旋状态不受系统边界条件的影响。如果 $\langle f(\sigma_i)g(\underline{\sigma}_{\partial_d i}) \rangle_c \sim d^{-\lambda}$ (幂次衰减), 则意味着系统内部的状态强烈的依赖于边界状态。

计算点和集合之间自旋态的关联有重要的意义, 但却是一项很困难的任务。文献[20, 24] 对于随机网络上的一些简单模型系统做了一些理论上的探讨, 但如何通过消息传递的迭代算法计算点和集合之间的关联仍然需要理论上的进一步发展。

2.6 复本对称种群动力学过程

上一节的复本对称空腔理论适用于研究一个给定的自旋玻璃样本。同一个自旋玻璃模型对应于许多不同样本, 这些样本有不同的相互作用参数或者

它们的网络连接方式不同。信念传播方程也可以用来研究由不同样本所构成的系综的平均统计物理性质,例如样本自由能的系综平均值,样本平均能量的系综平均值等等。具体的计算方案是复本对称种群动力学 (population dynamics) 迭代过程[22]。我们在此介绍该计算方案的要点。为简单起见假定系统的自旋都只有两个分量,相应的BP方程为方程 (2.50)。

首先定义一个长度为 $\mathcal{N}$ 的实数集合,记为 $P[h_{i \rightarrow a}]$,并按照某种分布产生 $\mathcal{N} \gg 1$ 个随机实数 $h_{i \rightarrow a}$ 作为这一集合的初始元素。一种常用的初始化方法是令集合 $P[h_{i \rightarrow a}]$ 的每一个元素都为0。集合 $P[h_{i \rightarrow a}]$ 可以看成是由因子网络所有的边上的空腔场 $h_{i \rightarrow a}$ 构成的一个种群。

然后我们对种群 $P[h_{i \rightarrow a}]$ 的元素不断进行更新,直到种群在概率的意义上达到稳态。每一次基本的更新包含如下几个步骤:

- (1) 根据给定自旋玻璃系统的性质,产生一个变量节点 i 的外场 h_i^0 ,并产生一个随机非负整数 k_i 作为节点 i 在因子网络中的连通度;
- (2) 根据所研究的自旋玻璃系统的统计性质,产生一个功能节点 b 的玻耳兹曼因子 ψ_b ,它是 k_b 个自旋 $\sigma_i, \sigma_j, \sigma_k, \dots$ 的函数;然后从种群中 $P[h_{i \rightarrow a}]$ 以相互独立且完全随机的方式选取 k_b 个空腔场 $h_{i \rightarrow b}, h_{j \rightarrow b}, h_{k \rightarrow b}, \dots$,并由方程 (2.50b) 得到一个空腔场 $u_{b \rightarrow i}$;功能节点 b 的平均能量及自由能贡献可由这 k_b 个空腔场计算出来,例如它的自由能贡献由表达式 (2.78) 可知为

$$f_b = -(k_b - 1) \frac{1}{\beta} \ln \left[\sum_{\underline{\sigma}_{\partial b}} \psi_b(\underline{\sigma}_{\partial b}) \prod_{j \in \partial b} q_{j \rightarrow b}(\sigma_j) \right];$$

- (3) 重复第 (2) 步直到一共产生了 k_i 个空腔场 $u_{a \rightarrow i}, u_{b \rightarrow i}, \dots$,然后通过方程 (2.50a) 得到 k_i 个新的空腔场 $h_{i \rightarrow a}, h_{i \rightarrow b}, \dots$ 并用它们来替代种群 $P[h_{i \rightarrow a}]$ 中随机选取的 k_i 个元素;变量节点 i 的平均能量及自由能贡献可以相应地计算出来,例如它的自由能贡献由表达式 (2.78) 可知为

$$f_i = -\frac{1}{\beta} \ln \left[\sum_{\sigma_i} \psi_i(\sigma_i) \prod_{b \in \partial i} \sum_{\underline{\sigma}_{\partial b}} \delta(\sigma_i^b, \sigma_i) \psi_b(\underline{\sigma}_{\partial b}) \prod_{k \in \partial b \setminus i} q_{k \rightarrow b}(\sigma_k) \right].$$

通过上述的种群动力学迭代过程,我们将获得一个稳态种群 $P[h_{i \rightarrow a}]$,并将得到热力学量的系综平均值。以自由能密度系综平均值 $\bar{f}$ 为例,其表达式根据 (2.78) 可知为

$$\bar{f} = \bar{f}_i - \alpha \bar{f}_b, \quad (2.100)$$

其中 $\alpha \equiv \frac{M}{N}$ 是功能节点数目与变量节点数目的比值，而 $\overline{f_i}$ 和 $\overline{f_b}$ 分别是变量节点和功能节点的自由能贡献的系综平均值。

2.7 随机规则网络模型上的应用

这一节通过随机规则网络上的铁磁模型和自旋玻璃模型来演示复本对称空腔理论的应用及其局限性。为讨论简单起见，只考虑两体相互作用，且系统中每个变量节点 i 的自旋态 $\sigma_i = \pm 1$ ，没有外部磁场 ($h_i^0 = 0$)。

2.7.1 铁磁系统

在一个随机规则网络中，每一个节点都与 K 个完全随机选取的其它节点相连，见第1.1.3节的介绍。随机规则网络上的铁磁伊辛模型的能量函数由能量表达式 (2.2) 定义，其中网络的每一条边 (i, j) 上的自旋耦合常数都为同一值 $J > 0$ (在本小节以后的讨论中，我们将系统的单位能量设为 J)。这个系统的信念传播方程为(2.50a) 和(2.52)。由于系统中每个节点都与 K 个其它节点发生铁磁相互作用，导致空腔磁场 $h_{i \rightarrow a}$ 和 $u_{a \rightarrow i}$ 都不依赖于节点 i 和相互作用 a ，即 $h_{i \rightarrow a} \equiv h$ ， $u_{a \rightarrow i} \equiv u$ ，而 h 与 u 为如下耦合方程的解：

$$h = (K - 1)u, \quad u = \frac{1}{\beta} \operatorname{atanh} [\tanh(\beta) \tanh(\beta h)] . \quad (2.101)$$

系统的平均磁化强度的BP近似值为

$$m = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle_0 = \tanh [\beta K u] . \quad (2.102)$$

通过一些分析可以发现自洽方程 (2.101) 存在一个临界温度 T_c ：

$$(K - 1) \tanh \left[\frac{1}{k_B T_c} \right] = 1 . \quad (2.103)$$

当温度 $T > T_c$ 时，方程 (2.101) 只有唯一解 $h = u = 0$ ，对应于顺磁态，系统平均磁化强度为零。当温度 $T < T_c$ 时，方程 (2.101) 出现两组非零稳定解，对应于系统出现自发的铁磁性，系统的平均磁化强度 $m = +m_0$ 或 $m = -m_0$ ，其中

$$m_0 = \tanh [\beta K |u|] \quad (u \neq 0) . \quad (2.104)$$

对于 $K = 4$ 的情况， $T_c \simeq 2.8854$ (对应于临界逆温度 $\beta_c \simeq 0.347$)；而对于 $K = 6$ 的情况， $T_c \simeq 4.9326$ (对应于 $\beta_c \simeq 0.203$)。图2.7.1比较了平均磁化

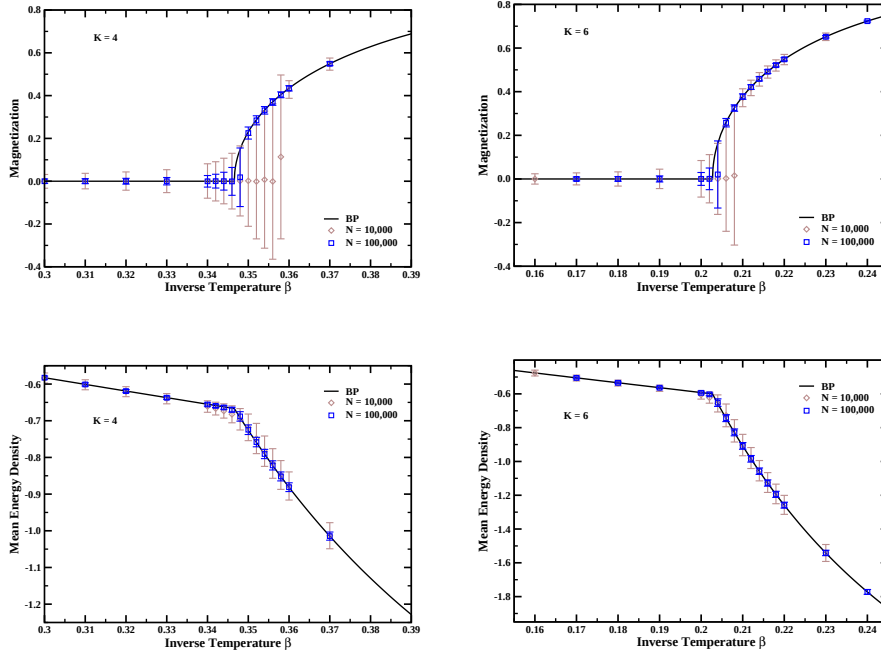


图 2.8: 随机规则网络铁磁伊辛模型的平均磁化强度（上）及平均能量密度（下）与逆温度 β 的关系。实线是信念传播方程预言的结果，而圆点是在单个随机网络（包含 $N = 10^4$ 个节点或 $N = 10^5$ 个节点）上进行Glauber动力学模拟[27]得到的结果。左图对应于 $K = 4$ ，而右图对应于 $K = 6$ 。在铁磁相变点 $T = T_c$ 附近有较大的有限尺度效应，所以只有当节点数 N 足够多时模拟结果才和理论曲线完全一致。

强度的理论预言与计算机模拟结果的比较，可以看到二者相互吻合得很好。这说明信念传播方程能够非常精确地描述随机规则网络铁磁伊辛模型的统计物理性质。在铁磁相变点 T_c 附近有较大的有限尺度效应，故系统包含的节点数 N 越多，复本对称空腔理论与计算机模拟的定量符合就越好。

具有周期性边界条件的正方晶格（维数 $D = 2$ ）、立方晶格（ $D = 3$ ）及超立方晶格（ $D \geq 4$ ）上的铁磁伊辛模型的信念传播方程的不动点也同样由方程（2.101）确定。但这些有限维系统的铁磁相变温度却不同于信念传播方程预言的结果。例如正方晶格体系的铁磁相变温度的精确值为 $T_c = 2/\ln(1 + \sqrt{2}) \simeq 2.2692$ [17, 18, 28]，低于信念传播方程的预言2.8554。造成复本对称空腔理论在有限维晶格体系铁磁相变附近不适用的原因是系统中存

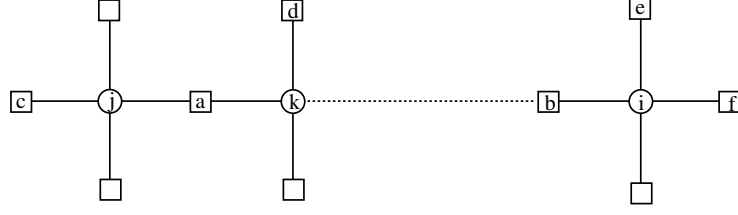


图 2.9: 随机规则网络 (变量节点连通度 $K = 4$) 铁磁伊辛模型中的节点状态关联。由于网络中回路的长度在热力学极限 ($N \rightarrow +\infty$) 时发散, 节点 j 和 i 之间的状态关联仅仅通过它们之间的最短路径 $j-a-k \cdots b-i$ 来传递。当节点 j 上加入一个外场 h_j^0 后, 所有指向这一最短路径的消息, 例如 $u_{c \rightarrow j}$ 和 $u_{d \rightarrow k}$ 都不会受到 h_j^0 的影响, 即 $\frac{\partial u_{c \rightarrow j}}{\partial h_j^0} = \frac{\partial u_{d \rightarrow k}}{\partial h_j^0} = 0$, 而 i 和 j 之间的最短路径上所有沿 j 指向 i 方向上的消息 (例如 $h_{j \rightarrow a}$, $u_{a \rightarrow k}$, $u_{b \rightarrow i}$ 等等) 都与 h_j^0 相关。

在许多短程回路。在接近临界温度 T_c 时, 系统的关联长度变得越来越长, 远远长于短程回路的长度, 导致圈图对自由能的修正贡献不可忽略不计。我们将在第六章讨论信念传播方程在有限维系统的改进。

我们现在来计算随机规则网络体系中两个最短距离为 d 的变量节点 i 与 j 之间的状态关联, 其中距离 d 等于 i 和 j 之间的最短路径上功能节点的数目。当 $N \rightarrow +\infty$ 时, i 和 j 之间的关联几乎完全通过它们之间的最短路径来传递。由于随机网络的局部结构是树状的, 当 d 为有限值时, i 和 j 之间的最短路径只有一条, 参见示意图 2.9。对于 $N \rightarrow \infty$ 的随机规则网络, 由于每一个节点的有限半径内的局域子网络都是树状结构, 当节点 j 上有一个非零外场 h_j^0 时只有少数一些边 (k, b) 上的 BP 消息 $u_{a \rightarrow k}$ 和 $h_{k \rightarrow b}$ 会受到外场 h_j^0 的影响。对于示意图 2.9 中的边 (i, b) 而言, 由方程 (2.97) 和 (2.96a) 可知

$$\frac{\partial u_{b \rightarrow i}}{\partial h_j^0} = \left(\frac{\tanh(\beta J)[1 - \tanh^2(\beta h)]}{1 - \tanh^2(\beta J) \tanh^2(\beta h)} \right)^d, \quad (2.105)$$

而对于节点 i 周围的其它功能节点 (例如节点 e 和 f) 传递给 i 的消息都不随 h_j^0 而改变。将这些计算结果代入表达式 (2.95) 就得到自旋 σ_i 和 σ_j 的关联随距离 d 的衰减关系

$$\langle \sigma_i \sigma_j \rangle_c = (1 - m_0^2) \left(\frac{\tanh(\beta J)[1 - \tanh^2(\beta h)]}{1 - \tanh^2(\beta J) \tanh^2(\beta h)} \right)^d \propto \exp\left(-\frac{d}{\xi_2}\right), \quad (2.106)$$

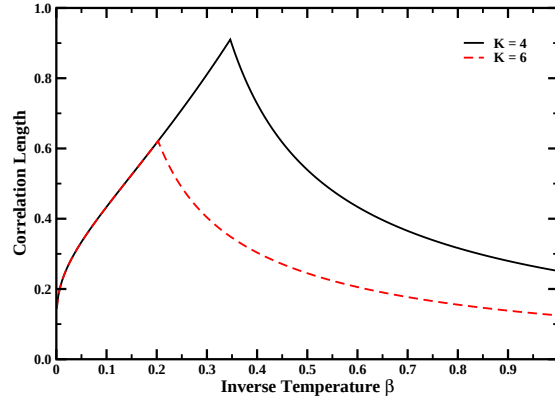


图 2.10: 在热力学极限 $N = \infty$ 下, 随机规则网络铁磁伊辛模型的两点关联长度 ξ_2 与逆温度 β 的关系。网络的连通度 $K = 4$ 或 $K = 6$ 。

其中 ξ_2 是自旋两点关联特征长度:

$$\xi_2 = \frac{1}{\ln \left(\frac{1 - \tanh^2(\beta J) \tanh^2(\beta h)}{\tanh(\beta J) [1 - \tanh^2(\beta h)]} \right)}. \quad (2.107)$$

图2.10显示了关联长度 ξ_2 随温度的变化曲线。在任意温度 T 处关联长度 ξ_2 都是有限的。 ξ_2 虽然在温度 $T = T_c$ 处达到极大值, 但它并未发散。这与晶格体系的铁磁伊辛模型的性质有显著的不同。对于后一类系统, 在铁磁相变处系统的两点关联长度趋向于无穷大[1, 2]。对于热力学极限 $N = \infty$ 下的随机网络, 如果我们随机选取两个(或有限多个)变量节点, 这些节点之间有趋向于1的概率相互之间不存在关联。由于这一原因, 导致信念传播方程(及其背后的Bethe-Peierls近似)能够在所有的温度下精确描述随机网络系统中铁磁伊辛模型的性质。

既然随机网络系统的两点关联函数在任何温度下都是有限的, 那为何系统仍然存在一个相变临界温度 T_c ? 实际上在 $T = T_c$ 处系统的确存在一个发散的关联长度, 即点与集合之间的特征关联长度[20, 24, 25]。参考示意图2.7。粗略地说, 在临界温度 T_c 处, 无穷大随机网络系统中的每一节点 i 和距离为任意大 d 的边界节点集合的集体状态 $\sigma_{\partial_d i}$ 都是相关的。虽然节点 i 的自旋 σ_i 在 T_c 处有同样的概率为 $+1$ 或 -1 , 但在给定边界自旋集体状态 $\sigma_{\partial_d i}$ 后, 自旋 σ_i 的条件概率不再是均匀分布。因此我们由边界构型 $\sigma_{\partial_d i}$ 能够获得中心节点 i 的状态信息, 而无论边界离中心节点有多远。

点与集合之间的关联相比于两点之间的关联而言，是一种更弱的层次上的关联。随机网络系统的特征是两点关联长度总是有限，但点与集合之间的关联长度在临界温度处发散。而规则晶格系统中的关联更强，它的两点关联长度在临界温度处也是发散的。

2.7.2 自旋玻璃系统

我们考虑随机规则网络上的自旋玻璃模型 (2.3)，其中每条边 (i, j) 上的耦合 J_{ij} 是一个随机常数，有相同的概率等于 $+J$ 和 $-J$ (在本小节以后的讨论中，我们将系统的单位能量设为 J)。我们在系统的一个给定样本上运行了BP迭代方程，发现BP迭代过程在逆温度 β 小于某个临界值 β_c 时总是收敛到唯一的不动点，且所有的BP空腔场 $h_{i \rightarrow a}$ 和 $u_{a \rightarrow i}$ 都等于零，即系统处于顺磁态。但如果 $\beta > \beta_c$ ，那么BP迭代过程将不能在单个样本上收敛，这时系统处于自旋玻璃态。对于 $K = 4$ 的随机规则网络， $\beta_c \approx 0.659$ ；而对于 $K = 6$ 的随机规则网络， $\beta_c \approx 0.481$ 。在BP迭代过程收敛的顺磁态 ($\beta < \beta_c$)，BP方程所预言的样本平均能量密度与计算机模拟的结果完全吻合。

当 $\beta > \beta_c$ 时模型 (2.3) 处于自旋玻璃态，我们将在下一章探讨描述自旋玻璃态的理论方法。虽然BP迭代过程在自旋玻璃态不能收敛到一个不动点，但我们仍然可以利用BP迭代方程对随机样本所构成的系综的平均性质进行估计，即对系综的统计性质进行复本对称种群动力学模拟。图2.11是种群动力学过程所预言的系综平均能量密度及单个样本的平均能量密度的比较，我们发现在 $\beta < \beta_c$ 的区域这两个能量密度值完全吻合；当 β 超过 β_c 后，单个样本的平均能量密度要高于种群动力学预言的系综平均能量密度，但二者的差距只有在 $\beta \gg \beta_c$ 的区域才变得显著。对于给定逆温度 $\beta > \beta_c$ ，通过复本对称种群动力学模拟我们能够获得模型 (2.3) 的系综平均能量密度的一个下限。

在自旋玻璃态系统一个重要的序参量是Edwards-Anderson交叠 q_{EA} [9]，即

$$q_{EA} \equiv \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle^2, \quad (2.108)$$

其中 $\langle \sigma_i \rangle$ 是节点 i 的自旋平均值。图2.11比较了种群动力学模拟预言的 q_{EA} 值与单个样本上进行计算机模拟所得到的 q_{EA} 值的差距。对单个样本进行计算机模拟时，我们采取单自旋Glauber动力学过程[27]，每隔100步对系统的微观构型进行一次采样；在每一Monte Carlo步中，样本的每个自旋都通过Glauber动力学规则进行了一次更新。

复本对称种群动力学模拟预言系统在 $\beta > \beta_c$ 时处于自旋玻璃态。变量节点的平均磁矩不等于零，故 $q_{EA} > 0$ ；但系统不表现出宏观的磁矩，因为不同变量节点的平均磁矩有正有负，相互抵消。计算机Glauber动力学模拟的结果也支持这一结论，见图2.11。

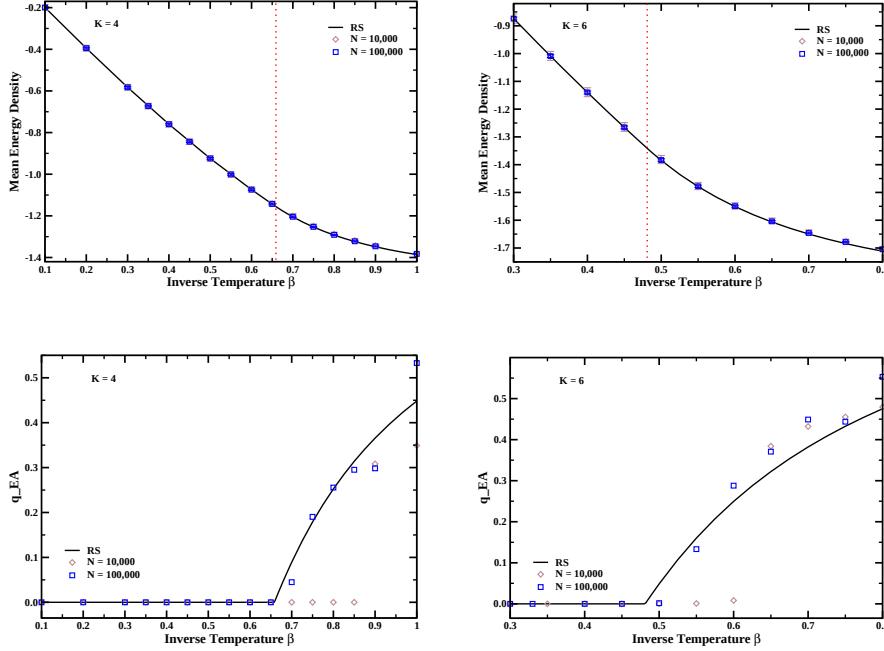


图 2.11: 随机规则网络Edwards-Anderson自旋玻璃模型的平均能量密度(上)及自旋玻璃序参量 q_{EA} (下)与逆温度 β 的关系。实线是对信念传播方程进行复本对称种群动力学迭代所得到的结果，而菱形和方形点是在单个随机网络(包含 $N = 10^4$ 或 $N = 10^5$ 个节点)上进行Glauber动力学模拟得到的结果。左图对应于 $K = 4$ ，而右图对应于 $K = 6$ 。图中的虚线标示出自旋玻璃相变的临界逆温度 β_c 。当 $\beta > \beta_c$ 时，信念传播迭代过程在单个随机网络样本上不能收敛。

我们也可以用BP方程来研究有限维Edwards-Anderson自旋玻璃模型。在二维正方晶格的计算机模拟结果[47]表明，BP方程的顺磁解在逆温度 $\beta > 0.371$ 时是不稳定的，失稳的原因是系统中的自旋两点关联长度在低温时超过晶格中的短程回路的长度。二维正方晶格上的Edwards-Anderson自旋玻璃系统可能并不存在有限温度的自旋玻璃态。

2.8 Kikuchi自由能泛函

Bethe-Peierls自由能 (2.73) 也可以表示成变量节点及功能节点自旋态边缘概率分布的泛函:

$$F_0 = \sum_a \left[\sum_{\underline{\sigma}_{\partial a}} p_a(\underline{\sigma}_{\partial a}) (E_a(\underline{\sigma}_{\partial a}) + \sum_{j \in \partial a} E_j(\sigma_j)) + \frac{1}{\beta} \sum_{\underline{\sigma}_{\partial a}} p_a(\underline{\sigma}_{\partial a}) \ln p_a(\underline{\sigma}_{\partial a}) \right] + \sum_i (1 - k_i) \left[\sum_{\sigma_i} q_i(\sigma_i) E_i(\sigma_i) + \frac{1}{\beta} \sum_{\sigma_i} q_i(\sigma_i) \ln q_i(\sigma_i) \right]. \quad (2.109)$$

这一表达式实际上是Kikuchi团簇变分自由能的一种最简单形式[16, 3, 26, 33, 47], 故我们将其称为**Kikuchi自由能泛函**。注意到这一泛函中的概率分布函数必须满足相容关系 (2.88)。

考虑由功能节点 a 及其最近邻变量节点集合 ∂a 所构成的子系统。当该子系统自旋态联合分布为 $p_a(\underline{\sigma}_{\partial a})$ 时, 由方程 (2.18) 知其自由能为

$$F_{a+\partial a}[p_a] \equiv \sum_{\underline{\sigma}_{\partial a}} p_a(\underline{\sigma}_{\partial a}) (E_a(\underline{\sigma}_{\partial a}) + \sum_{j \in \partial a} E_j(\sigma_j)) + \frac{1}{\beta} \sum_{\underline{\sigma}_{\partial a}} p_a(\underline{\sigma}_{\partial a}) \ln p_a(\underline{\sigma}_{\partial a}). \quad (2.110)$$

这一自由能就是Kikuchi自由能 (2.109) 的第一项。

再考虑由一个变量节点 i 所构成的另一子系统。当 i 的自旋态分布为 $q_i(\sigma_i)$ 时, 这一子系统的自由能为

$$F_i[q_i] \equiv \sum_{\sigma_i} q_i(\sigma_i) E_i(\sigma_i) + \frac{1}{\beta} \sum_{\sigma_i} q_i(\sigma_i) \ln q_i(\sigma_i). \quad (2.111)$$

注意在 (2.109) 中, 变量节点 i 对Kikuchi自由能的贡献要乘上一个因子 $(1 - k_i)$ 。对这一乘积因子直观地解释是: 每一个与 i 相连的功能节点 a 的自由能贡献 $F_{a+\partial a}[p_a]$ 中都包含了节点 i 的贡献, 导致节点 i 的贡献被重复考虑了 k_i 次, 因而需要将 i 的自由能贡献 $F_i[q_i]$ 减去 $(k_i - 1)$ 次。

习题 2.11 从Bethe-Peierls自由能 (2.73) 出发, 利用表达式 (2.86) 和 (2.87), 并考虑相容关系 (2.88) 及BP方程 (2.47a), 推导出Kikuchi自由能泛函表达式 (2.109)。

第六章将详细介绍团簇变分法及其推广, 我们将会讨论推导Kikuchi自由能泛函的另一种方式。

从配分函数出发, 我们推导出BP方程, 进而得到Kikuchi自由能表达式 (2.109)。由于Kikuchi自由能是概率分布 $\{q_i, p_a\}$ 的泛函, 我们关心它在相容

关系 (2.88) 的约束下的极小值, 这一极值问题对应于如下的泛函:

$$K[\{q_i, p_a\}] \equiv \sum_a F_{a+\partial a}[p_a] + \sum_i (1 - k_i) F_i[q_i] + \sum_a \lambda_a \sum_{\underline{\sigma}_{\partial a}} p_a(\underline{\sigma}_{\partial a}) \\ + \sum_{(i,a)} \sum_{\sigma_i} \lambda_{ia}(\sigma_i) \left[q_i(\sigma_i) - \sum_{\underline{\sigma}_{\partial a \setminus i}} p_a(\underline{\sigma}_{\partial a}) \right]. \quad (2.112)$$

上式中, 拉格朗日乘子 λ_a 对应于概率分布 $p_a(\underline{\sigma}_{\partial a})$ 的归一化约束; 而拉格朗日乘子 $\lambda_{ia}(\sigma_i)$ 则是自旋态 σ_i 的函数, 它对应于概率分布相容关系 (2.88) 及概率分布 $q_i(\sigma_i)$ 的归一化约束。将 (2.112) 求一阶变分得到

$$\frac{\partial K[\{q_i, p_a\}]}{\partial q_i(\sigma_i)} = \frac{1 - k_i}{\beta} \left[\ln q_i(\sigma_i) + 1 + \beta E_i(\sigma_i) \right] + \sum_{a \in \partial i} \lambda_{ia}(\sigma_i), \quad (2.113a)$$

$$\frac{\partial K[\{q_i, p_a\}]}{\partial p_a(\underline{\sigma}_{\partial a})} = \frac{1}{\beta} \left[\beta E_a(\underline{\sigma}_{\partial a}) + \sum_{j \in \partial a} \beta E_j(\sigma_j) + 1 + \ln p_a(\underline{\sigma}_{\partial a}) \right] \\ + \lambda_a - \sum_{j \in \partial a} \lambda_{ja}(\sigma_j). \quad (2.113b)$$

由极值条件 $\partial K / \partial q_i(\sigma_i) = 0$ 可推出如下结论: 如果变量节点 i 是孤立节点, $k_i = 0$, 那么 $q_i(\sigma_i) \propto e^{-\beta E_i(\sigma_i)}$; 如果 i 只参与一个相互作用 a , $k_i = 1$, 那么拉格朗日乘子 $\lambda_{ia}(\sigma_i) = 0$, 不依赖于自旋态 σ_i ; 而对于 $k_i \geq 2$ 的其它情况,

$$q_i(\sigma_i) \propto e^{-\beta E_i(\sigma_i)} \prod_{a \in \partial i} \exp \left[\frac{\beta \lambda_{ia}(\sigma_i)}{k_i - 1} \right], \quad k_i \geq 2. \quad (2.114)$$

由另一个极值条件 $\partial K / \partial p_a(\underline{\sigma}_{\partial a}) = 0$ 则得出:

$$p_a(\underline{\sigma}_{\partial a}) \propto e^{-\beta E_a(\underline{\sigma}_{\partial a})} \prod_{i \in \partial a} \exp [\beta \lambda_{ia}(\sigma_i) - \beta E_i(\sigma_i)]. \quad (2.115)$$

方程 (2.114) 和 (2.115) 中的拉格朗日乘子函数 $\lambda_{ia}(\sigma_i)$ 由边 (i, a) 上的相容关系

$$q_i(\sigma_i) = \sum_{\underline{\sigma}_{\partial a \setminus i}} p_a(\underline{\sigma}_{\partial a})$$

来确定。每一组满足相容关系的拉格朗日乘子函数集合 $\{\lambda_{ia}(\sigma_i) : (i, a) \in G\}$ 所对应的自由能极值²都可以由表达式 (2.109) 计算出来。如果有多组满足相容关系的解存在, 那么应该选取自由能最低的解。

²这一自由能极值可能是极小值, 也可能是极大值或鞍点, 这需要通过计算自由能泛函的二阶变分才能确定。

文献中对泛函 (2.112) 的极值问题讨论很多, 并提出了许多算法, 可参见[29, 41, 44, 43]。对于两体相互作用且自旋变量只有两个分量的一类问题, 相容关系 (2.88) 可以很容易地在概率分布函数的参数化过程中实现, 因而 (2.109) 转变为一组实变量的函数, 它的极值点可以通过信念优化 (belief optimization) 算法[41, 38]或其它算法[29, 44]获得。

现在让我们讨论Kikuchi自由能泛函极值点与BP方程不动点之间的关系。假定Kikuchi自由能极值点处的概率分布函数 (2.114) 和 (2.115) 确定了一组拉格朗日乘子函数 $\{\lambda_{ia}(\sigma_i) : (i, a) \in G\}$ 。在每条边 (i, a) 上定义一个新的函数 $\eta_{ia}(\sigma_i)$ 为

$$\eta_{ia}(\sigma_i) \equiv -\lambda_{ia}(\sigma_i) + \frac{1}{k_i - 1} \sum_{b \in \partial i} \lambda_{ib}(\sigma_i), \quad (k_i \geq 2); \quad (2.116)$$

如果变量节点 i 只连有一个功能节点 a , 即 $k_i = 1$, 那就定义 $\eta_{ia}(\sigma_i) = 0$ 。借助于这一组函数就可以将拉格朗日乘子函数 $\lambda_{ia}(\sigma_i)$ 写成如下的加和形式:

$$\lambda_{ia}(\sigma_i) = \sum_{b \in \partial i \setminus a} \eta_{ib}(\sigma_i). \quad (2.117)$$

将表达式 (2.116) 代人 (2.114) 和 (2.115) 就得到

$$q_i(\sigma_i) \propto e^{-\beta E_i(\sigma_i)} \prod_{a \in \partial i} e^{\beta \eta_{ia}(\sigma_i)}, \quad (2.118a)$$

$$p_a(\sigma_a) \propto e^{-\beta E_a(\sigma_a)} \prod_{i \in \partial a} [e^{-\beta E_i(\sigma_i)} \prod_{b \in \partial i \setminus a} e^{\beta \eta_{ib}(\sigma_i)}]. \quad (2.118b)$$

将这一结果与方程 (2.86) 和 (2.87) 相比较, 我们就得到了函数 $\eta_{ia}(\sigma_i)$ 与BP方程中的概率分布函数 $p_{a \rightarrow i}(\sigma_i)$ 之间的对应关系:

$$p_{a \rightarrow i}(\sigma_i) \propto e^{\beta \eta_{ia}(\sigma_i)}, \quad (2.119)$$

而另一概率分布函数 $q_{i \rightarrow a}(\sigma_i)$ 的表达式也可以类似地得到:

$$q_{i \rightarrow a}(\sigma_i) \propto e^{-\beta E_i(\sigma_i)} \prod_{b \in \partial i \setminus a} e^{\beta \eta_{ib}(\sigma_i)}. \quad (2.120)$$

上面的分析表明, Kirkuchi自由能泛函 (2.109) 的极值点与BP方程 (2.47) 的不动点有一一对应的关系[43, 44]: 给定描述Kikuchi 自由能泛函一个极值点的拉格朗日乘子函数 $\{\lambda_{ia}(\sigma_i)\}$, 由 (2.116) 就可构造BP方程 (2.47) 的一个不动点; 反过来, 给定BP 方程的一个不动点, 由 (2.119) 和 (2.117) 就

可构造Kikuchi自由能的一组极值点拉格朗日乘子函数。利用迭代的方法求解BP方程不能达到收敛的时候，我们可以利用这种对应关系，通过求泛函(2.109)极值的方法来获得BP方程的不动点[41, 44]。

本章小结

在本章，我们对定义在因子网络 G 上的自旋模型(2.1)推导了配分函数的圈图展开表达式(2.69)，并由配分函数圈图展开导出了信念传播方程(2.47)。系统总的自由能是Bethe-Peierls自由能(2.73)与圈图修正贡献项(2.71)二者之和。信念传播方程和Bethe-Peierls自由能构成自旋玻璃复本对称平均场空腔理论[22]，这一理论忽略了自由能圈图修正贡献项。

我们还从Kikuchi自由能泛函(2.109)的角度讨论了信念传播方程，并指出Kikuchi自由能泛函的极值点对应于BP方程(2.47)的不动点。

我们以随机规则网络上的伊辛模型及Edwards-Anderson模型为例讨论了信念传播方程的简单应用。

参考文献

- [1] 李政道: 李政道讲义—统计力学. 上海科学技术出版社, 上海(2006)
- [2] 林宗涵: 热力学与统计物理学. 北京大学出版社, 北京(2007)
- [3] An, G.: A note on the cluster variation method. J. Stat. Phys. **52**, 727–734 (1988)
- [4] Bethe, H.A.: Statistical theory of superlattices. Proc. R. Soc. London A **150**, 552–575 (1935)
- [5] Brout, R.: Phase Transitions. Benjamin, New York (1965)
- [6] Chertkov, M., Chernyak, V.Y.: Loop calculus in statistical physics and information science. Phys. Rev. E **73**, 065,102(R) (2006)
- [7] Chertkov, M., Chernyak, V.Y.: Loop series for discrete statistical models on graphs. J. Stat. Mech.: Theor. Exp. p. P06009 (2006)
- [8] Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley, New York (1991)
- [9] Edwards, S.F., Anderson, P.W.: Theory of spin glasses. J. Phys. F: Met. Phys. **5**, 965–974 (1975)
- [10] Frey, B.J.: Graphical Models for Machine Learning and Digital Communication. MIT Press, Cambridge, MA (1998)
- [11] Friedberg, R.: Dual trees and resummation theorems. J. Math. Phys. **16**, 20–30 (1975)

-
- [12] Georges, A., Yedidia, J.S.: How to expand around mean-field theory using high-temperature expansions. *J. Phys. A: Math. Gen.* **24**, 2173–2192 (1991)
 - [13] Guerra, F.: Sum rules for the free energy in the mean field spin glass model. *Fields Institute Communications* **30**, 161–170 (2001)
 - [14] Guerra, F.: Broken replica symmetry bounds in the mean field spin glass model. *Commun. Math. Phys.* **233**, 1–12 (2003)
 - [15] Guerra, F., Toninelli, F.L.: The thermodynamic limit in mean field spin glass models. *Commun. Math. Phys.* **230**, 71–79 (2002)
 - [16] Kikuchi, R.: A theory of cooperative phenomena. *Phys. Rev.* **81**, 988–1003 (1951)
 - [17] Kramers, H.A., Wannier, G.H.: Statistics of the two-dimensional ferromagnet. part i. *Phys. Rev.* **60**, 252–262 (1941)
 - [18] Kramers, H.A., Wannier, G.H.: Statistics of the two-dimensional ferromagnet. part ii. *Phys. Rev.* **60**, 263–276 (1941)
 - [19] Kschischang, F.R., Frey, B.J., Loeliger, H.A.: Factor graphs and the sum-product algorithm. *IEEE Trans. Inf. Theory* **47**, 498–519 (2001)
 - [20] Mézard, M., Montanari, A.: Reconstruction on trees and spin glass transition. *J. Stat. Phys.* **124**, 1317–1350 (2006)
 - [21] Mézard, M., Montanari, A.: *Information, Physics, and Computation*. Oxford Univ. Press, New York (2009)
 - [22] Mézard, M., Parisi, G.: The bethe lattice spin glass revisited. *Eur. Phys. J. B* **20**, 217–233 (2001)
 - [23] Mézard, M., Parisi, G., Virasoro, M.A.: *Spin Glass Theory and Beyond*. World Scientific, Singapore (1987)
 - [24] Montanari, A., Semerjian, G.: On the dynamics of the glass transition on bethe lattices. *J. Stat. Phys.* **124**, 103–189 (2006)

-
- [25] Montanari, A., Semerjian, G.: Rigorous inequalities between length and time scales in glassy systems. *J. Stat. Phys.* **125**, 23–54 (2006)
- [26] Morita, T., Suzuki, M., Wada, K., Kaburagi, M. (eds.): Foundations and Applications of Cluster Variation Method and Path Probability Method, *Prog. Theor. Phys. Suppl.*, vol. 115. Physical Society of Japan (1994)
- [27] Newman, M.E.J., Barkema, G.T.: Monte Carlo Methods in Statistical Physics. Oxford University Press, New York (1999)
- [28] Onsager, L.: Crystal statistics i. a two-dimensional model with an order-disorder transition. *Phys. Rev.* **65**, 117–149 (1944)
- [29] Oppen, M., Saad, D. (eds.): Advanced Mean Field Methods: Theory and Practice. MIT Press, Cambridge, MA (2001)
- [30] Pathria, R.K.: Statistical Mechanics, second edn. Elsevier (Singapore) Pte Ltd. (2001)
- [31] Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Francisco, CA, USA (1988)
- [32] Peierls, R.: On ising’s model of ferromagnetism. *Proc. Camb. Phil. Soc.* **32**, 477–481 (1936)
- [33] Pelizzola, A.: Cluster variation method in statistical physics and probabilistic graphical models. *J. Phys. A: Meth. Gen.* **38**, R309–R339 (2005)
- [34] Plefka, T.: Convergence condition of the tap equation for the infinite-ranged ising spin glass model. *J. Phys. A: Math. Gen.* **15**, 1971–1978 (1982)
- [35] Sessak, V.: Problèmes inverses dans les modèles de spin. Ph.D. thesis, l’Université Pierre et Marie CURIE (Paris, France) (2010). ArXiv:1010.1960
- [36] Sessak, V., Monasson, R.: Small-correlation expansions for the inverse ising problem. *J. Phys. A: Math. Theor.* **42**, 055,001 (2009)

-
- [37] Sherrington, D., Kirkpatrick, S.: Solvable model of a spin-glass. *Phys. Rev. Lett.* **35**, 1792–1796 (1975)
- [38] Shin, J.: The complexity of approximating a bethe equilibrium. *arXiv:1109.1724v3* (2011)
- [39] Suzuki, M., Hu, X., Hatano, N., Katori, M., Minami, K., Lipowski, A., Nonomura, Y.: *Coherent Anomaly Method: Mean Field, Fluctuations and Systematics*. World Scientific, Singapore (1995)
- [40] Talagrand, M.: *Spin Glasses: A Challengt for Mathematicians*. Springer-Verlag, Berlin (2003)
- [41] Welling, M., Teh, Y.W.: Belief optimization for binary networks: A stable alternative to loopy belief propagation. In: J.S. Breese, D. Koller (eds.) *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence* (Seattle, Washington, USA), pp. 554–561. Morgan Kaufmann, San Fransisco, USA (2001)
- [42] Xiao, J.Q., Zhou, H.J.: Partition function loop series for a general graphical model: free-energy corrections and message-passing equations. *J. Phys. A: Math. Theor.* **44**, 425,001 (2011)
- [43] Yedidia, J.S., Freeman, W.T., Weiss, Y.: Constructing free-energy approximations and generalized belief-propagation algorithms. *IEEE Trans. Inf. Theory* **51**, 2282–2312 (2005)
- [44] Yuille, A.L.: CCCP algorithms to minimize the bethe and kikuchi free energies: Convergent alternatives to belief propagation. *Neural Comput.* **14**, 1691–1722 (2002)
- [45] Zhou, H.J.: Long-range frustration in a spin-glass model of the vertex-cover problem. *Phys. Rev. Lett.* **94**, 217,203 (2005)
- [46] Zhou, H.J.: Long-range frustration in finite connectivity spin glasses: a mean-field theory and its application to the random k -satisfiability problem. *New J. Phys.* **7**, 123 (2005)

- [47] Zhou, H.J., Wang, C.: Region graph partition function expansion and approximate free energy landscapes: Theory and some numerical results. *J. Stat. Phys.* **148**, 513–547 (2012)