For more on physical models related to the Tutte polynomial see [19, Chapter 4], [10] and [14]. For combinatorics associated with the Ising model see [11] and also the book [12]. Sokal's papers [15] and [16] give a lucid explanation of how combinatorial properties of the partition function of the general Potts model correspond to physical properties of a system. For more about exactly solved models in statistical physics see [2].

1 The Ising model

In the general Ising model on a graph $G = (V, E)$ each vertex i of G is assigned a *spin* σ_i, which is either $+1$ ("up") or -1 ("down"). An assignment of spins to all the vertices of G is called a *configuration* or *state* and denoted by σ.

Each edge $e = ij$ of G has an *interaction energy* J_{ij} which is constant on the edge, but may vary from edge to edge.

For each state σ the Hamiltonian $H(\sigma)$ is defined by

$$H(\sigma) = -\sum_{ij \in E} J_{ij}\sigma_i\sigma_j - \sum_{i \in V} M\sigma_i,$$

where M represents the energy of the external field.

The Hamiltonian $H(\sigma)$ measures the energy of the state σ. In a ferromagnet the J_{ij} are positive, which has the consequence that a configuration of spins in which adjacent vertices have parallel spins ($\sigma_i = \sigma_j$ for $ij \in E$) has a lower energy than a non-magnetized state in which spins are arbitrary. The external field has the effect of aligning spins with the direction of the field, thus again favouring states of low energy.

The *partition function* $Z = Z(G; \beta, J, M)$ is defined by

$$Z(G) = \sum_{\sigma} e^{-\beta H(\sigma)},$$

where the sum is over all $2^{|V|}$ possible spin configurations and $\beta = 1/kT$ is a parameter determined by the absolute temperature T and where k is Boltzmann's constant. The probability of finding the system in a state σ is given by

$$\Pr(\sigma) = e^{-\beta H(\sigma)}/Z(G).$$

This is the probability distribution on states σ which has maxiumum entropy for a given mean value $-\frac{\partial}{\partial\beta}\log Z(G)$ of the energy $H(\sigma)$. See [7] and [8] for more on information theory in statistical physics. A high temperature gives a low value of β and the probability distribution of states becomes more flat. On the other hand, a low temperature gives high β and correspondingly greater probability to low energy states.

The entropy of a finite probability distribution $(p_1, \ldots p_N)$ is defined by

$$h(p_1, \ldots, p_N) = -\sum_k p_k \log_2 p_k,$$

and is a measure of uncertainty in the system whose states follow the given distribution. The entropy of the Ising model system is

$$h(G; \beta) = -\sum_{\sigma} \Pr(\sigma) \log_2 \Pr(\sigma),$$

which gives

$$h(G;\beta) = \left[\frac{1}{Z(G)}(\beta \log_2 e)\sum_{\sigma} H(\sigma)e^{-\beta H(\sigma)}\right] + \log_2 Z(G).$$

Seeing that

$$\frac{\partial}{\partial\beta}\log Z(G) = \frac{1}{Z(G)}\frac{\partial Z(G)}{\partial\beta} = -\sum_{\sigma}\frac{H(\sigma)}{Z(G)}e^{-\beta H(\sigma)},$$

we have

$$h(G;\beta) = -(\beta \log_2 e)\frac{\partial}{\partial\beta}\log Z(G) + \log_2 Z(G).$$

The quantity $-\frac{\partial}{\partial\beta}\log Z(G)$ is called the *internal energy* and the quantity $\log Z(G) = \log_2 Z(G)/\log_2 e$ is the *free energy*.

Consider some countably infinite graph such as the two-dimensional square lattice (vertices $\mathbb{Z}^2$, with vertex (a,b) adjacent to vertices $(a\pm1,b)$ and $(a,b\pm1)$) and an increasing sequence of finite subgraphs $G_n = (V_n, E_n)$. Then, under reasonable hypotheses on the G_n, it can be shown that the *(limiting) free energy per lattice site*

$$\lim_{n\to\infty}\frac{\log Z(G_n)}{|V_n|}$$

exists for non-degenerate physical values of the parameters β, J, M of $Z(G)$.

Complex singularities of $\log Z(G_n)$ (i.e., zeroes of $Z(G_n)$) may approach the real axis in the limit $n\to\infty$, and in this case the points of physical *phase transitions* are precisely the real limit points of such complex zeroes. In the ferromagnetic Ising model (positive interaction energies J_{ij}), a cooling slab of iron becomes magnetized at the *critical temperature* that gives a phase transition.

In particular, the main problem of the Ising model on the two-dimensional lattice is to determine

$$\lim_{n\to\infty}\frac{\log Z(L_{n,n})}{n^2}$$

where $L_{n,n}$ is the $n\times n$ grid. (In practice, in order to facilitate anlysis $L_{n,n}$ is replaced by the $n\times n$ toroidal grid.)

1.1 Constant interaction energies, no external field

Assume that $M = 0$, so that there is no external field, and that $J_{ij} = J$ is constant over all edges of G.

The partition function is now

$$Z(G) = Z(G;\beta,J) = \sum_{\sigma} e^{-\beta H(\sigma)},$$

where

$$H(\sigma) = -\sum_{ij\in E} J\sigma_i\sigma_j.$$

Theorem 1.1. *The partition function for the Ising model on $G=(V,E)$ when there is constant edge interaction J and no external field is given by*

$$Z(G) = 2^{|V|}e^{-\beta Jn(G)}(\sinh\beta J)^{r(G)}T(G;\coth\beta J, e^{2\beta J}),$$

where $T(G)$ is the Tutte polynomial of G and $\beta = 1/kT$.

Theorem 1.2 (Van der Waerden, 1941)**.**

$$Z(G) = 2^{|V|}(\cosh\beta J)^{|E|}C(G;\tanh\beta J),$$

where

$$C(G;x) = \sum_{\substack{A\subseteq E\\(V,A)\text{ Eulerian}}} x^{|A|}.$$

The Eulerian subgraph expansion of the partition function of the Ising model of Theorem 1.2 is the starting point for reducing the Ising model problem for square lattices to a dimer (matching) problem, and thence via Pfaffian orientations to Onsager's solution in 1944 of the problem of finding $\lim_{n\to\infty}\frac{\log Z(L_{n,n})}{n^2}$. In particular it enabled the critical temperature T_c to be found for the two-dimensional lattice.

2 The q-state Potts model

The q-state Potts model on a graph $G = (V, E)$ is a generalization of the Ising model in which there are q possible states at a vertex rather than the two up/down states. In this model introduced by Askin and Teller (1943) and Potts (1952) the energy between two adjacent spins at vertices i and j is taken to be zero if the spins are the same and equal to a constant J_{ij} if they are different. For a state σ the Hamiltonian is defined by

$$H(\sigma) = \sum_{ij\in E} J_{ij}(1-\delta(\sigma_i,\sigma_j)),$$

where δ is the Kronecker delta function ($\delta(a,b) = 1$ if $a = b$ and $\delta(a,b) = 0$ if $a \neq b$). We shall assume there is no external magnetic field. The Hamiltonian $H(\sigma)$ represents the energy of the state σ. The partition function of the q-state Potts model is defined by

$$Z(G) = \sum_{\sigma} e^{-\beta H(\sigma)},$$

where the sum is over all $q^{|V|}$ possible states σ and β is the inverse temperature $\beta = \frac{1}{kT}$ as for the Ising model.

Just as for the Ising model, we have

$$\Pr(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z(G)},$$

the Boltzmann maximum entropy distribution on the state space subject to a given expected value of $H(\sigma)$. (This expected value is the internal energy of the system, which is constant when the system is isolated/ in equilibrium with its environment. This is the First Law of Thermodynamics, expressing the principle of conservation of energy.)

If we replace J_{ij} by $-2J_{ij}$ then the partition function of the 2-state Potts model is the same as that of the Ising model scaled by $e^{-\beta\sum_{ij\in E}J_{ij}}$.

Returning to the q-state Potts model, if $J_{ij} = J$ is constant over all edges and we write $K = \beta J$ then the partition function can be written in terms of the Tutte polynomial as follows:

$$Z(G) = q^{c(G)}(e^K-1)^{r(G)}e^{-K|E|}T(G;\frac{e^K-1+q}{e^K-1},e^K).$$

The point $(\frac{e^K-1+q}{e^K-1}, e^K)$ lies on the hyperbola $(x-1)(y-1)=q$.

Here is a summary of correspondences between the Potts model and the Tutte plane (taken from [18]):

Potts model on G	Tutte polynomial $T(G;x,y)$
Ferromagnetism	Positive $(x,y>1)$ branch of $(x-1)(y-1)=q$
Antiferromagnetism	Negative $(x<0)$ branch of $(x-1)(y-1)=q$ with $y>0$
High temperature	Asymptote of $(x-1)(y-1)=q$ to $y=1$
Low temp. ferromagnetic	Positive branch of $(x-1)(y-1)=q$ asymptotic to $x=1$
Zero temp. antiferromagnetic	Proper vertex q-colourings, $x=1-q, y=0$.

2.1 The Fortuin-Kasteleyn random cluster model

The random cluster model on a connected graph $G=(V,E)$ with parameters p and q is a probability space on all spanning subgraphs of G. The probability measure of a subgraph $A \subseteq E$ is

$$\mu(A) = \frac{1}{Z(G)} p^{|A|}(1-p)^{|E\setminus A|} q^{c(A)},$$

where as usual $c(A)$ denotes the number of connected components of the subgraph (V,A), and $Z(G)$ is the normalizing constant

$$Z(G) = \sum_{A\subseteq E} p^{|A|}(1-p)^{|E\setminus A|} q^{c(A)}.$$

When $q=1$ this is the *bond percolation* model on G, where an edge is open with probability p and otherwise closed. This model is used for such processes as molecules penetrating a porous solid, diffusion, and the spread of infection through a community (passage/contagion is possible along open edges).

When q is a positive integer the random cluster model is equivalent to the q-state Potts model with $p=1-e^{-K}$.

Proposition 2.1. *The partition function of the random cluster model on a connected graph $G=(V,E)$ with parameters $0\le p\le 1$ and $q>0$ is given by*

$$Z(G) = q(1-p)^{|E|-|V|+1} p^{|V|-1} T(G; 1+\frac{(1-p)q}{p}, \frac{1}{1-p}),$$

where $T(G)$ is the Tutte polynomial of G, and the probability measure of the subgraph A is given by

$$\mu(A) = \frac{\left(\frac{p}{1-p}\right)^{|A|} q^{c(A)-1}}{\left(\frac{p}{(1-p)}\right)^{|V|-1} T(G; \frac{p+q-pq}{p}, \frac{1}{1-p})}.$$

When $q>1$ there is a bias towards edges joining vertices in an existing component than edges uniting two old components, since a larger number of components are favoured. More precisely, given $B\subseteq E$ and $e\in E\setminus B$, under the probability distribution μ we have

$$\begin{aligned}\Pr(e\in A \,|\, A\setminus\{e\}=B) &= \frac{\Pr(A=B\cup\{e\})}{\Pr(A-\{e\}=B)} = \frac{\mu(B\cup\{e\})}{\mu(B\cup\{e\})+\mu(B)} \\ &= \begin{cases} p & \text{if } c(B\cup\{e\})=c(B), \\ \frac{p}{p+q(1-p)} & \text{if } c(B\cup\{e\})=c(B)-1, \end{cases}\end{aligned}$$

where, for $0 < p < 1$,

$$\frac{p}{p+q(1-p)} \begin{cases} < p & \text{if } q > 1 \\ > p & \text{if } 0 < q < 1. \end{cases}$$

Percolation in the random cluster model (the existence of an infinite component of open edges) is intimately related to *two-point correlation* (long-distance correlation between vertex colours) in the q-state Potts model. Given fixed vertices i and j, in the Ising model the two-point correlation between i and j is defined to be the expected value of $\sigma_i\sigma_j$ over all states σ. For the Potts model the two-point correlation is the expected value of $\delta(\sigma_i, \sigma_j)$, i.e., the probability that σ_i equals σ_j.

A key result of Fortuin and Kasteleyn (1969) is the following (see e.g. [5, Theorem 2.1]):

Theorem 2.2. *For any pair of vertices i and j and positive integer q, the probability that σ_i equals σ_j in the q-state Potts model is given by*

$$\frac{1}{q} + (1 - \frac{1}{q})\mu\{i \rightsquigarrow j\},$$

where μ is the random cluster probability measure on G obtained by taking $p = 1 - e^{-K}$ and $\{i \rightsquigarrow j\}$ is the event that there is an open path from i to j, i.e.,

$$\{i \rightsquigarrow j\} = \bigcup \{A \subseteq E : i \text{ and } j \text{ belong to the same component of } (V, A)\}.$$

The expression on the right-hand side in Theorem 2.2 can be regarded as being made up of two parts. The first term $1/q$ is the probability that under a uniformly random colouring of the vertices of G the vertices i and j have the same colour. The second term measures the probability of long-range interaction. So Theorem 2.2 expresses an equivalence between long-range spin correlations and percolatory behaviour.

Phase transition (in the infinite system) occurs at the onset of an infinite cluster (connected component) in the random cluster model and corresponds to spins on the vertices of the Potts model having long-range two-point correlation.

See [19, Chapter 4] for further discussion of percolation in the random cluster model, as well as the detailed account of [6] from the point of view of probability theory.

2.2 Monte Carlo methods

Computation of the probability of a state σ in the q-state Potts model is a difficult problem (#P-hard) so that approximation is required. One way to do this is to run a Markov chain whose stationary distribution is the desired Potts model distribution, and to sample from this Markov chain (when it has run a sufficiently long time, and if more than one sample is required, choosing samples sufficiently far apart).

Suppose we wish to sample from a probability distribution π on a finite set Ω (e.g. $\Omega = [q]^V$ for the Potts model on $G = (V, E)$, with $\pi(\sigma) = e^{-\beta H(\sigma)}/Z(G)$ for $\sigma \in [q]^V$). The idea is to construct an irreducible[1] Markov chain with transition matrix P and state space Ω and which has π as stationary distribution.[2]

[1] Strongly connected transition graph.

[2] If each state is positive recurrent (finite expected return time) the matrix P has a unique stationary distribution π, i.e., such that $\pi P = \pi$, given by $\pi(\sigma) = 1/\mathbb{E}(T_\sigma)$, where T_σ is the random variable giving the return time to σ (when starting from σ).

A sufficient condition that the irreducible stochastic matrix P has π as its stationary distribution is that P satisfies the *detailed balance condition* for π, namely

$$\pi(\sigma)P(\sigma,\tau) = \pi(\tau)P(\tau,\sigma).$$

This is to say the Markov chain is *reversible*: the expected number of accepting moves from σ to τ is equal to the expected number of moves from τ to σ. We have

$$\sum_\sigma \pi(\sigma)P(\sigma,\tau) = \sum_\sigma \pi(\tau)P(\tau,\sigma) = \pi(\tau),$$

i.e. $\pi P = \pi$ and π is the stationary distibution for P.

To construct P we proceed as follows. Start with any irreducible stochastic matrix A with state space Ω. Let $G : (0,\infty) \to [0,1]$ be any function satisfying $G(x) = xG(x^{-1})$, such as $G(x) = \min(x,1)$ or $G(x) = \frac{x}{1+x}$.

Define the matrix $Q = Q(\sigma,\tau)$, $\sigma,\tau \in \Omega$, by

$$Q(\sigma,\tau) = G\left(\frac{\pi(\tau)A(\tau,\sigma)}{\pi(\sigma)A(\sigma,\tau)}\right).$$

Modify matrix A into P by setting

$$P(\sigma,\tau) = A(\sigma,\tau)Q(\sigma,\tau) \qquad \sigma \neq \tau,$$

$$P(\sigma,\sigma) = 1 - \sum_{\tau\neq\sigma} P(\sigma,\tau).$$

Then P satisfies, for $\sigma \neq \tau$,

$$\begin{aligned}
\pi(\sigma)P(\sigma,\tau) &= \pi(\sigma)A(\sigma,\tau)G(\frac{\pi(\tau)A(\tau,\sigma)}{\pi(\sigma)A(\sigma,\tau)} \\
&= \pi(\tau)A(\tau,\sigma)G(\frac{\pi(\sigma)A(\sigma,\tau)}{\pi(\tau)A(\tau,\sigma)} \\
&= \pi(\tau)A(\tau,\sigma)Q(\tau,\sigma) \\
&= \pi(\tau)P(\tau,\sigma)
\end{aligned}$$

In other words, the detailed balance condition is satisfied by P and it defines a reversible Markov chain.

Now let us apply this to the Potts model distibution $\pi(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z(G)}$.

Metropolis method

Choose A with $A(\sigma,\tau) = \frac{1}{q-1}$ when σ and τ differ at exactly one vertex, $A(\sigma,\tau) = 0$ otherwise, and $G(x) = \min(1,x)$. Then, for σ and τ which differ at just one vertex,

$$\begin{aligned}
Q(\sigma,\tau) &= \min\left(\frac{e^{-\beta H(\tau)}A(\tau,\sigma)}{e^{-\beta H(\sigma)A(\sigma,\tau)}}, 1\right) \\
&= \min(e^{-\beta(H(\tau)-H(\sigma))}, 1)
\end{aligned}$$

The Markov chain whose transition matrix P is defined by

$$P(\sigma,\tau) = \frac{1}{q-1}Q(\sigma,\tau) = \begin{cases} \frac{1}{q-1} & H(\tau) \leq H(\sigma) \\ \frac{1}{q-1}e^{-\beta(H(\tau)-H(\sigma))} & H(\tau) > H(\sigma) \end{cases},$$

$$P(\sigma,\sigma) = 1 - \sum_{\tau\neq\sigma} P(\sigma,\tau)$$

then has the desired Potts model stationary distribution.

A random walk on this Markov chain can be descibed thus. When at a current state $\sigma \in [q]^V$ (the chain is initiated by choosing an arbitrary state to begin with):

(i) choose a vertex $v \in V$ uniformly at random and colour $c \in [q]$ uniformly at random;

(ii) change the colour of v to c and thereby change state σ to a new state τ differing just at v *either* if $H(\tau) \leq H(\sigma)$ *or* if $H(\tau) > H(\sigma)$ then change from σ to τ with probability $e^{-\beta(H(\tau)-H(\sigma))}$.

Glauber dynamics

A variation on the Metroplis method is to choose $G(x) = \frac{x}{1+x}$, which leads to the following rule for follwing a random walk on states whose stationary distibution is the q-state Potts model probability distribution.

When at current state σ,

(i) choose a vertex $v \in V$ uniformly at random and colour $c \in [q]$ uniformly at random;

(ii) move from σ to the state τ differing (at most) from σ by having colour c at v with probability $\frac{1}{1+e^{\beta(H(\tau)-H(\sigma)}}$.

A drawback with the random walks on such Markov chains as defined by the Metropolis method and Glauber dynamics as defined above is that convergence to the stationary distribution is not fast, especially near the critical temperature. Successive states are not statistically independent (they differ in at most one vertex) and the autocorrelation is $O(|V|)$ (intuitively, and very roughly, there has to be a large separation of $O(|V|)$ steps in order to have uncorrelated samples from the same Markov chain).

What is required is a *rapidly mixing* Markov chain, where the distance (suitably defined) from the stationary distribution is sufficiently small after time polynomial in $|V|$, rather than the exponential time that the Metropolis method requires. See [19, Section 8.6] and [9] for one example of such a chain. There is much other research, surveyed e.g. in [13], on obtaining rapidly mixing Markov chains (of which I am ignorant). See e.g. talk slides at http://mae.ucdavis.edu/dsouza/Talks/msri-June06.pdf for coupled Markov chains and simulation of the Ising model.

Swendsen–Wang algorithm

An improvement on the Metroplis method both with respect to statistical independence (autocorrelation) and speed of convergence (although still not rapidly mixing, I think: see [4]) is the Swendsen–Wang algorithm [17] for the Potts model on graph $G = (V, E)$ with parameters β and J (interaction energy), with Hamiltonian $H(\sigma) = J\#\{ij \in E : \sigma_i \neq \sigma_j\})$.

When at current state σ move to state τ defined as follows:

(i) let

$$E(\sigma) = \{ij \in E : \sigma_i = \sigma_j\}$$

be the set of monochromatic edges of the state σ. Delete each edge of $E(\sigma)$ independently with probability $e^{-\beta J}$, giving a random subset A of $E(\sigma)$.

(ii) For each connected component of the graph (V, A), choose a colour c uniformly at random from $[q]$, and for all vertices i within that component set $\tau_i = c$. This defines the new state τ.

The correctness of this algorithm relies on the equivalence between the q-state Potts model and the Fortuin–Kasteleyn random cluster model alluded to in Section 2.1 above, with percolation probability $p = 1 - e^{-\beta J}$ (open edges between like-coloured vertices).

References

[1] M. Aigner, *A Course in Enumeration*, Springer-Verlag, Berlin, 2007.

[2] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, London, 1982. [Available online at http://tpsrv.anu.edu.au/Members/baxter/book/Exactly.pdf]

[3] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998

[4] C. Borgs, J.T. Chayes and P. Tetali, Tight bounds for mixing of the Swendsen-Wang Algorithm at the Potts transition point, *Probab. Theory Relat. Fields.*, to appear. arXiv:1011.3058v1 [math.PR]

[5] G. Grimmett, *The Random-Cluster Model*, arXiv:math/0205237v2, 2002.

[6] G. Grimmett, *The Random-Cluster Model*, Grundlehren der mathematischen Wissenschaften **333**, Springer, 2006. Corrected reprint 2009.

[7] E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* **106** (1957), 620–630.

[8] E.T. Jaynes, Information theory and statistical mechanics II, *Phys. Rev.* **108** (1957), 171–190.

[9] M.R. Jerrum and A. Sinclair, Polynomial-time approximation algorithms for the Ising model, *Proc. 17th ICALP, EATCS* (1990), 462–475.

[10] P.W. Kasteleyn, Graph theory and crystal physics, In: F. Harary (ed.), *Graph Theory and Theoretical Physics*, Academic Press, London, 1969, pp. 43-110.

[11] M. Loebl, Some discrete tools in statistical physics. In: J.P. Gazeau, J. Nešetřil, B. Rovan (eds.), *Physics and Theoretical Computer Science*, IOS Press, 2006.

[12] M. Loebl, *Discrete Mathematics in Statistical Physics*, Advanced Lectures in Mathematics, Vieweg and Teubner, 2009.

[13] F. Martinelli, Lectures on Glauber dynamics for discrete spin models, *Lectures on probability theory and statistics (Saint-Flour, 1997)*, Lecture Notes in Math., **1717**, 93–191, Springer, Berlin, 1999.

[14] J.K. Percus, *Combinatorial Methods*, Springer, New York, 1971.

[15] A.D. Sokal, Chromatic polynomials, Potts models and all that, *Physica***279** (2000), 324–332.

[16] A.D. Sokal, The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In: B. S. Webb (ed.), *Surveys in Combinatorics*, Cambridge Univ. Press, Cambridge, 2005, pp. 173–226.

[17] R. Swendsen and J-S.Wang, Non-universal critical dynamics in Monte-Carlo simulation, *Phys. Rev. Lett.*, **58** (1987), 86–88.

[18] D.J.A. Welsh and C. Merino, The Potts model and the Tutte polynomial, *J. Math. Phys.* **41**:3 (2000), 1127–1152.

[19] D.J.A. Welsh, *Complexity: Knots, Colourings and Counting*, London Math. Soc. Lecture Notes Ser. 186, 1993.