Problem sheet 7 issued: Tuesday 19th November 2013
Solutions available from: Wednesday 27th November 2013

Statistical Mechanics Answer Sheet 7

1. *The spin-spin correlation function and scaling relations.* (**RF Question**)

   (a) The spin-spin correlation function

$$\begin{aligned} g(\mathbf{r}_i, \mathbf{r}_j) &= \langle\, (s_i - \langle s_i \rangle)\, (s_j - \langle s_j \rangle)\, \rangle \\ &= \langle\, s_i s_j - \langle s_i \rangle s_j - s_i \langle s_j \rangle + \langle s_i \rangle \langle s_j \rangle\, \rangle \\ &= \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle - \langle s_i \rangle \langle s_j \rangle + \langle s_i \rangle \langle s_j \rangle \\ &= \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle, \end{aligned} \tag{1}$$

   where we use that the ensemble average operation $\langle \cdot \rangle$ is a linear operation and that the ensemble average of a constant is the constant itself.

   (b) Assuming that the system is translationally invariant, we substitute $m = \langle s_i \rangle = \langle s_j \rangle$ and find

$$\begin{aligned} g(\mathbf{r}_i, \mathbf{r}_j) &= \langle s_i s_j \rangle - m^2 \\ &= \langle s_j s_i \rangle - m^2 \\ &= g(\mathbf{r}_j, \mathbf{r}_i) \end{aligned} \tag{2}$$

   from which it follows that the correlation function is symmetric and thus a function of the relative distance between the spins at positions $\mathbf{r}_i$ and $\mathbf{r}_j$ only, that is,

$$g(\mathbf{r}_i, \mathbf{r}_j) = g(|\mathbf{r}_i - \mathbf{r}_j|). \tag{3}$$

   (c) (i) When $|\mathbf{r}_i - \mathbf{r}_j| \to \infty$, the spins become uncorrelated, assuming that we are not at the critical point that is! Thus

$$\begin{aligned} g(\mathbf{r}_i, \mathbf{r}_j) &= \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle \\ &\to \langle s_i \rangle \langle s_j \rangle - \langle s_i \rangle \langle s_j \rangle \quad \text{for } |\mathbf{r}_i - \mathbf{r}_j| \to \infty \\ &= 0. \end{aligned} \tag{4}$$

   (ii) By definition the spin-spin correlation function of spin i with itself

$$g(\mathbf{r}_i, \mathbf{r}_i) = \langle s_i s_i \rangle - \langle s_i \rangle \langle s_i \rangle = \langle s_i^2 \rangle - \langle s_i \rangle^2. \tag{5}$$

Because $s_i = \pm 1 \Leftrightarrow s_i^2 = 1$ we have $\langle s_i^2 \rangle = \langle 1 \rangle = 1$. Also $\langle s_i \rangle = m$, so

$$g(\mathbf{r}_i, \mathbf{r}_i) = 1 - m^2. \tag{6}$$

We assume the external magnetic field $H = 0$ so we can replace m with $m_0(T)$. If $T \geq T_c$, the magnetisation $m_0 = 0$ so that

$$g(\mathbf{r}_i, \mathbf{r}_i) = \begin{cases} 1 & \text{for } T \geq T_c \\ 1 - m_0^2(T) & \text{for } T < T_c. \end{cases} \tag{7}$$

The zero-field magnetisation per spin $m_0(T) \to \pm 1$ for $T \to 0$, implying

$$g(\mathbf{r}_i, \mathbf{r}_i) \to 0 \quad \text{for } T \to 0. \tag{8}$$

This result emphasises that the correlation function measures the fluctuations of the spins *away* from the average magnetisation as is clear from the original definition

$$g(\mathbf{r}_i, \mathbf{r}_i) = \langle\ (s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle)\ \rangle. \tag{9}$$

(iii) In the limit $J/(k_B T) \ll 1$ (high temperatures relative to the coupling constant), the spins will be orientated randomly, that is, there are no correlations between the spins, so we expect $g(\mathbf{r}_i, \mathbf{r}_j) \to 0$.

In the limit $J/(k_B T) \gg 1$ (low temperatures relative to the coupling constant), the spins will be aligned, that is, there are no fluctuations away from the average spin, so we expect $g(\mathbf{r}_i, \mathbf{r}_j) \to 0$.

(d) Because the susceptibility per spin diverges at the critical temperature in zero external field

$$\chi(T, 0) \propto |T - T_c|^{-\gamma} \quad \text{for } T \to T_c, H = 0 \tag{10}$$

the volume integral of the correlation function must also diverge at the critical temperature. Defining $r = |\mathbf{r}_i - \mathbf{r}_j|$, we have

$$\int_V g(\mathbf{r}_i, \mathbf{r}_j) d^d \mathbf{r}_j \propto \int_a^\infty g(r) r^{d-1}\, dr \to \infty \quad \text{for } T \to T_c, H = 0, \tag{11}$$

where a is a lower cutoff = lattice constant. This implies that $g(r)$ cannot decay exponentially with distance r at the critical point $(T, H) = (T_c, 0)$ since this would make the integral convergent in the upper limit. However, the divergence is consistent with an algebraic decay. Assuming

$$\begin{aligned} g(\mathbf{r}_i, \mathbf{r}_j) &\propto |\mathbf{r}_i - \mathbf{r}_j|^{-(d-2+\eta)} \\ &= r^{-(d-2+\eta)} \quad \text{for } T = T_c, H = 0, \text{ and all } r = |\mathbf{r}_i - \mathbf{r}_j| \end{aligned} \tag{12}$$

then

$$\begin{aligned}\int_V g(\mathbf{r}_i,\mathbf{r}_j)d^d\mathbf{r}_j &\propto \int_a^\infty g(r)r^{d-1}\,dr \\ &\propto \int_a^\infty r^{-(d-2+\eta)}r^{d-1}\,dr \\ &= \int_a^\infty r^{1-\eta}\,dr \\ &= \begin{cases}\left[\frac{1}{2-\eta}r^{2-\eta}\right]_a^\infty & \text{if } \eta\neq 2\\ \left[\ln(r)\right]_a^\infty & \text{if } \eta = 2\end{cases}\end{aligned}$$

that is, the integral will only diverge if the critical exponent $\eta \leq 2$. The divergence is logarithmic if $\eta = 2$ and algebraic if $\eta < 2$.

(e) (i) The correlation length diverges as $\xi(T,0) \propto |T_c - T|^{-\nu}$ for $T \to T_c, H = 0$. The critical exponent ν is independent of whether T_c is approached from below or above, however, the amplitude might differ, as indicated in Figure 1 below.

For $T > T_c$, the correlation length sets the upper linear distance over which spins are correlated. It is also identified as the linear size of the typical (characteristic) largest cluster of correlated spins and measures the typical largest fluctuation away from states with randomly oriented spins.

For $T < T_c$, the correlation length measures the fluctuations away from the fully ordered state, that is, the upper linear size of the holes in the cluster of aligned spins. There will be holes on all scales up to the correlation length.

(ii) When $T \neq T_c$ a finite correlation length ξ is introduced and

$$g(|\mathbf{r}_i - \mathbf{r}_j|) \propto |\mathbf{r}_i - \mathbf{r}_j|^{-(d-2+\eta)}\mathcal{G}_\pm(|\mathbf{r}_i - \mathbf{r}_j|/\xi) \quad \text{for } T \to T_c, \tag{13}$$

where

$$\xi(T,0) \propto |T_c - T|^{-\nu} \quad \text{for } T \to T_c, H = 0. \tag{14}$$

Consider the relation between the susceptibility per spin and the correlation function

$$k_B T\chi \propto \int_V g(\mathbf{r}_i,\mathbf{r}_j)d^d\mathbf{r}_j. \tag{15}$$

The left-hand side (LHS):

$$k_B T\chi(T,0) \propto |T - T_c|^{-\gamma} \quad \text{for } T \to T_c, H = 0. \tag{16}$$

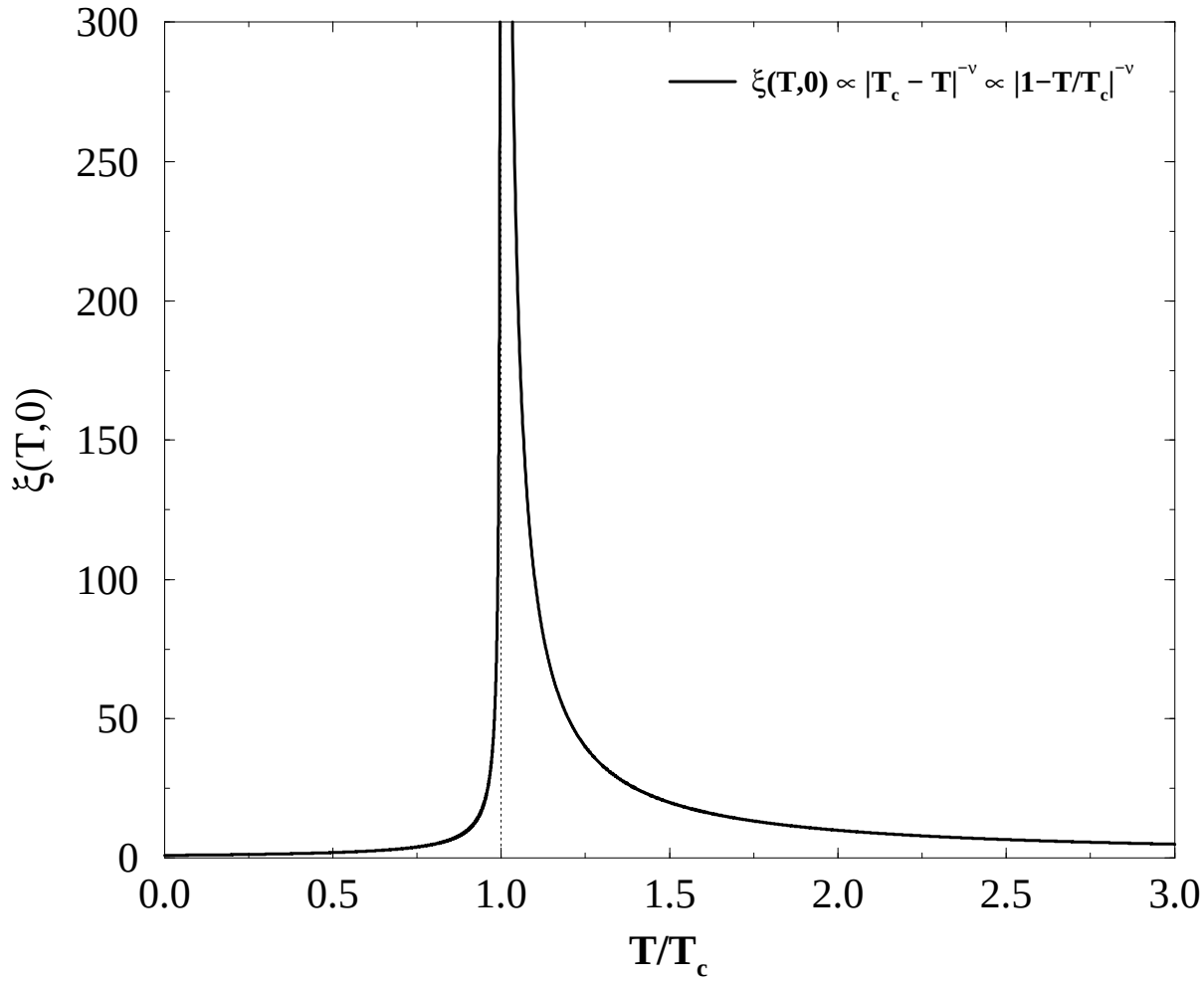


Figure 1: Sketch of the correlation length $\xi(T,0)$ as a function of the temperature T in units of the critical temperature T_c.

The right-hand side (RHS):

$$\begin{aligned}\int_V g(\mathbf{r}_i,\mathbf{r}_j)d^d\mathbf{r}_j &\propto \int_a^\infty r^{-(d-2+\eta)}\mathcal{G}_\pm(r/\xi)r^{d-1}\,dr \\ &= \int_a^\infty r^{1-\eta}\mathcal{G}_\pm(r/\xi)\,dr \\ &= \int_a^\infty (\tilde{r}\xi)^{1-\eta}\mathcal{G}_\pm(\tilde{r})\,d\tilde{r}\xi \quad \text{with } r=\tilde{r}\xi \\ &= \xi^{2-\eta}\int_a^\infty \tilde{r}^{1-\eta}\mathcal{G}_\pm(\tilde{r})\,d\tilde{r} \\ &= |T-T_c|^{-\nu(2-\eta)}\int_a^\infty \tilde{r}^{1-\eta}\mathcal{G}_\pm(\tilde{r})\,d\tilde{r} \quad \text{for } T\to T_c^\pm. \quad (17)\end{aligned}$$

The integral is just a number (which numerical value, however, depends on from which side T_c is approached due to the two different scaling functions $\mathcal{G}_\pm$), so we can conclude by comparing the LHS with the RHS that

$$\gamma = \nu(2-\eta). \quad (18)$$

(iii) We assume $T \leq T_c$ and consider the situation in zero external field $H = 0$ with m_0 replacing m. We define

$$\tilde{g}(r) = g(r) + m_0^2 = \langle s_i s_j\rangle. \quad (19)$$

For $T < T_c$, the correlation length $\xi < \infty$. As the correlation length sets the upper limit of the linear scale over which spins are correlated, the spins will be uncorrelated in the limit $r \to \infty$ as $r \gg \xi$. Thus

$$\tilde{g}(r) = \langle s_i s_j\rangle \to \langle s_i\rangle\langle s_j\rangle = m_0^2 \propto (T_c - T)^{2\beta} \quad \text{for } T \to T_c^-. \quad (20)$$

This is the reason for considering the function $\tilde{g}(r)$ and not $g(r)$ since the latter will approach zero for $r \gg \xi$.

At $T = T_c$ where the correlation length in infinite, the magnetisation is zero in zero external field, i.e., $m_0(T_c) = 0$. Thus

$$\tilde{g}(r) = g(r) \propto r^{-(d-2+\eta)} \quad \text{at } T = T_c. \quad (21)$$

One would thus expect, à la finite-size scaling in percolation theory, that

$$\tilde{g}(r) \propto \begin{cases} r^{-(d-2+\eta)} & \text{for } r \ll \xi \\ \xi^{-(d-2+\eta)} & \text{for } r \gg \xi. \end{cases} \quad (22)$$

Thus for $T < T_c$ where the correlation length is finite, we expect

$$\tilde{g}(r) \propto \xi^{-(d-2+\eta)} \propto |T-T_c|^{\nu(d-2+\eta)} \quad \text{for } r \gg \xi. \quad (23)$$

Comparing Eq.(23) and Eq.(20) we identify the scaling relation

$$2\beta = \nu(d-2+\eta) \Leftrightarrow d-2+\eta = 2\beta/\nu. \quad (24)$$

2. *Critical exponents inequality.*

   Given the thermodynamic relation

$$\chi \left(C_H - C_M\right) = T \left(\frac{\partial \langle M \rangle}{\partial T}\right)_H^2 \tag{25}$$

   As $C_M \geq 0$ and $\chi \geq 0$ it follows that

$$\chi \, C_H \geq T \left(\frac{\partial \langle M \rangle}{\partial T}\right)_H^2 . \tag{26}$$

   Using the scaling of the different quantities close to the critical point

$$\begin{aligned}
\chi(T,0) &\propto |T - T_c|^{-\gamma} && \text{for } T \to T_c, H = 0, \\
C_H &\propto |T - T_c|^{-\alpha} && \text{for } T \to T_c, H = 0, \\
\langle M \rangle &\propto (T_c - T)^{\beta} && \text{for } T \to T_c^-, H = 0 \text{ implying,} \\
\frac{\partial \langle M \rangle}{\partial T} &\propto -(T_c - T)^{\beta - 1} && \text{for } T \to T_c^-, H = 0
\end{aligned}$$

   so by substituting into Equation (26) we find

$$\begin{aligned}
(T_c - T)^{-\gamma} \;\; (T_c - T)^{-\alpha} &\geq T_c \;\left(-(T_c - T)^{\beta - 1}\right)^2 && \text{for } T \to T_c^- \\
(T_c - T)^{-\gamma - \alpha} &\geq T_c \, (T_c - T)^{2\beta - 2} && \text{for } T \to T_c^-
\end{aligned}$$

   from which we can conclude that

$$\begin{aligned}
-\gamma - \alpha &\leq 2\beta - 2 \Leftrightarrow \\
\gamma + \alpha &\geq 2 - 2\beta \Leftrightarrow \\
\alpha + 2\beta + \gamma &\geq 2.
\end{aligned} \tag{27}$$

   Notice that the inequality can be repalced by an *equality* for $d = 1, 2, 3$, and 4 and the mean-field exponents for the Ising Model.

3. *Eigenvalues, eigenvectors and diagonalisation.*

   (a) Assume $\mathbf{x} \neq \mathbf{0}$ is an eigenvector for f with eigenvalue λ, that is

$$f(\mathbf{x}) = \lambda \mathbf{x}. \tag{28}$$

   Since f is linear,

$$f(\alpha \mathbf{x}) = \alpha f(\mathbf{x}) = \alpha \lambda \mathbf{x} = \lambda(\alpha \mathbf{x}) \tag{29}$$

   so $\alpha \mathbf{x}$ is also an eigenvector with the same eigenvalue λ when $\alpha \neq 0$ (ensuring $\alpha \mathbf{x} \neq \mathbf{0}$.

(b) Assume $\mathbf{x} \neq \mathbf{0}$ is an eigenvector for f with eigenvalue λ. If $\mathbf{A}$ is the associated matrix for the linear function f then

$$\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}, \tag{30}$$

where $\mathbf{I}$ is the identity matrix. If $\det(\mathbf{A} - \lambda\mathbf{I}) \neq 0$ the matrix $\mathbf{A} - \lambda\mathbf{I}$ would be invertible and the only solution to the Equation (30) would be the trivial solution $\mathbf{x} = \mathbf{0}$. Equation (30) can only have non-trivial solutions $\mathbf{x} \neq \mathbf{0}$ if the matrix $\mathbf{A} - \lambda\mathbf{I}$ is not invertible. Therefore, we have

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{31}$$

Equation (31) is called the characteristic equation or the secular equation for the matrix $\mathbf{A}$ and the solutions λ are the eigenvalues of $\mathbf{A}$ (or f).

(c) We need to show that $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$ assuming that

$$f(\mathbf{x}_1) = \lambda_1\mathbf{x}_1 \quad \text{and} \quad f(\mathbf{x}_2) = \lambda_2\mathbf{x}_2 \qquad \text{with } \lambda_1 \neq \lambda_2. \tag{32}$$

$$f(\mathbf{x}_1) \cdot \mathbf{x}_2 = \lambda_1\mathbf{x}_1 \cdot \mathbf{x}_2 \tag{33a}$$
$$\mathbf{x}_1 \cdot f(\mathbf{x}_2) = \lambda_2\mathbf{x}_1 \cdot \mathbf{x}_2. \tag{33b}$$

Since f is symmetric

$$\lambda_1\mathbf{x}_1 \cdot \mathbf{x}_2 = \lambda_2\mathbf{x}_1 \cdot \mathbf{x}_2. \tag{34}$$

However, $\lambda_1 \neq \lambda_2$ from which we conclude

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = 0. \tag{35}$$

(d) (i) Consider the real and symmetric matrix

$$\mathbf{T} = \begin{pmatrix} \exp(\beta J + \beta H) & \exp(-\beta J) \\ \exp(-\beta J) & \exp(\beta J - \beta H) \end{pmatrix}. \tag{36}$$

The eigenvalues $\lambda_\pm$ of $\mathbf{T}$ are the solutions to the characteristic equation

$$\det(\mathbf{T} - \lambda\mathbf{I}) = 0. \tag{37}$$

The determinant

$$\begin{aligned} \det(\mathbf{T} - \lambda\mathbf{I}) &= \begin{vmatrix} \exp(\beta J + \beta H) - \lambda & \exp(-\beta J) \\ \exp(-\beta J) & \exp(\beta J - \beta H) - \lambda \end{vmatrix} \\ &= \lambda^2 - [\exp(\beta J + \beta H) + \exp(\beta J - \beta H)]\,\lambda + \exp(2\beta J) - \exp(-2\beta J) \\ &= \lambda^2 - 2\exp(\beta J)\cosh(\beta H)\lambda + \exp(2\beta J) - \exp(-2\beta J), \end{aligned} \tag{38}$$

so the solutions to the characteristic Equation (37) are

$$\begin{aligned}\lambda_{\pm} &= \frac{2\exp(\beta J)\cosh(\beta H) \pm \sqrt{4\exp(2\beta J)\cosh^2(\beta H) - 4[\exp(2\beta J) - \exp(-2\beta J)]}}{2} \\ &= \exp(\beta J)\left(\cosh(\beta H) \pm \sqrt{\cosh^2(\beta H) - 1 + \exp(-4\beta J)}\right) \\ &= \exp(\beta J)\left(\cosh(\beta H) \pm \sqrt{\sinh^2(\beta H) + \exp(-4\beta J)}\right). \end{aligned} \tag{39}$$

(ii) Since $\lambda_+ > \lambda_-$, the associated eigenvectors must be orthogonal. To determine the eigenvectors for $\mathbf{T}$ we must solve the equations

$$\mathbf{T}\mathbf{x}_+ = \lambda_+\mathbf{x}_+ \tag{40a}$$
$$\mathbf{T}\mathbf{x}_- = \lambda_-\mathbf{x}_- \tag{40b}$$

or equivalently

$$(\mathbf{T} - \lambda_+\mathbf{I})\mathbf{x}_+ = \mathbf{0} \tag{40c}$$
$$(\mathbf{T} - \lambda_-\mathbf{I})\mathbf{x}_- = \mathbf{0} \tag{40d}$$

Then we construct the matrix of eigenvectors

$$\mathbf{U} = (\mathbf{x}_+\ \mathbf{x}_-) \tag{41}$$

that will satisfy

$$\mathbf{U}^{-1}\mathbf{T}\mathbf{U} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}. \tag{42}$$

Problem sheet 8 issued: Tuesday 26th November 2013
Solutions available from: Wednesday 4th December 2013

Statistical Mechanics Answer Sheet 8

1. *Second-order PT in a mass-spring system: Landau theory.* (**RF Question**)

   (a) The total energy of the mass-spring system

$$\begin{aligned} U(\theta) &= \text{elastic potential energy} + \text{gravitational potential energy} \\ &= \frac{1}{2}k(a\theta)^2 + mg(a\cos\theta - a) \\ &= \frac{1}{2}ka^2\theta^2 + mga(\cos\theta - 1). \end{aligned} \tag{1}$$

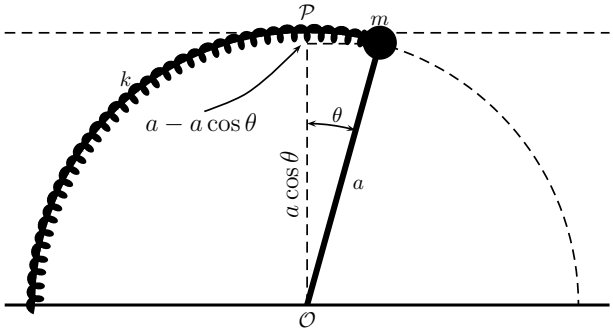


Figure 1: The projection of the rod of length a onto the vertical dashed line has length $a\cos\theta$ where the angle θ is measured (positive clockwise) from the vertical. Hence, the position of the center of mass of the variable mass m is $a - a\cos\theta = a(1 - \cos\theta)$ below the zeroth-level of the gravitational potential energy indicated by the horizontal dashed line.

   (b) (i) We expand the cosine to fourth order to find

$$\begin{aligned} U(\theta) &= \frac{1}{2}ka^2\theta^2 + mga(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots - 1) \\ &= \frac{a}{2}(ka - mg)\theta^2 + \frac{mga}{24}\theta^4 + \mathcal{O}(\theta^6), \end{aligned} \tag{2}$$

   where the coefficient of the fourth-order term is positive while the coefficient of the second-order term is zero for $ka = mg$ and changes sign from positive when $ka > mg$ to negative when $ka < mg$.

(ii) As the total energy $U(\theta)$ is an even function in θ (reflecting the symmetry of the problem), all the odd terms in the Taylor expansion around $\theta = 0$ are zero.

(iii) We denote the angle of equilibrium with θ_0. When $ka > mg$, the unique minimum is at $\theta_0 = 0$. When $ka = mg$, the unique minimum is at $\theta_0 = 0$. When $ka < mg$, there are two minima at $\pm\theta_0 \neq 0$.

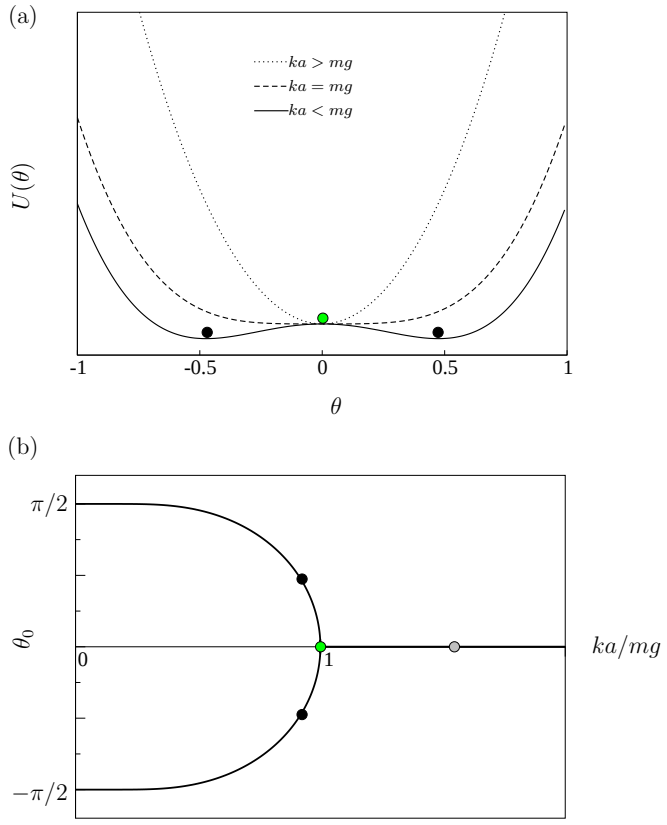


Figure 2: (a) The energy, $U(\theta)$, versus the angle θ. The solid circles show the position of the minima of the energy of the corresponding graph. For $ka > mg$, the minimal energy implies $\theta = 0$. For $ka = mg$, the trivial solution $\theta = 0$ is marginally stable However, for $ka < mg$, the minimal energy implies $\theta = \pm\theta_0 \neq 0$. (b) The angle of equilibrium, θ_0 as a function of the ratio ka/mg.

(iv) The system is in equilibrium when $dU/d\theta = 0$. Hence

$$\begin{aligned} \frac{dU}{d\theta} &= a(ka - mg)\theta + \frac{mga}{6}\theta^3 \\ &= mga\theta\left(\frac{ka}{mg} - 1 + \frac{1}{6}\theta^2\right) \\ &= 0 \end{aligned} \tag{3}$$

with solutions

$$\theta_0 = \begin{cases} 0 & \text{for } \frac{ka}{mg} \geq 1 \\ \pm\sqrt{6(1-ka/mg)} & \text{for } \frac{ka}{mg} < 1 \end{cases} = \begin{cases} 0 & \text{for } \frac{m_c}{m} \geq 1 \\ \pm\sqrt{6[(m-m_c)/m]} & \text{for } \frac{m_c}{m} < 1, \end{cases} \tag{4}$$

where $m_c = ka/g$.

(v) See Figure 4.

(vi) Landau suggested a simplistic general theory of second-order phase transitions based on expanding the free energy in powers of the order parameter. In the absence of a magnetic-like field, symmetry dictates that only even powers of the order parameter appear in the expansion. For example, in the Ising model

$$f - f_0 = a_2(T - T_c)m^2 + a_4 m^4 \quad \text{with } a_2, a_4 > 0,$$

where an expansion up to fourth order is sufficient to give a qualitative description of second-order phase transitions occurring at temperature T_c. The term f_0 is an unimportant constant, while $a_4 > 0$ in order for the free energy to be physically realistic, i.e. not minimised by extreme values of the order parameter. As written, the left-hand side is given by a quartic polynomial which always has one trivial solution, $m = 0$, and two non-trivial solutions, $m = \pm m_0(T)$, so long as $T < T_c$. As T passes through T_c from above, the trivial solution becomes unstable and two stable non-trivial solutions appear. Below T_c, therefore, the order parameter of the system is non-zero.

(vii) The order parameter of the mass-spring system is the equilibrium angle θ_0 which is zero for $m \leq m_c$ and non-zero for $m > m_c$. The critical value of the variable mass $m_c = ka/g$.

2. *Diluted Ising model.*

(a) A spin s_i is situated on each lattice site $\mathbf{r}_i$. However, the spin only interacts with with the nearest neighbours with probability p. Identifying a nonzero coupling constant $J_{ij} = J > 0$ as an occupied bond and $J_{ij} = 0$ as an empty bond, we have an exact mapping onto a bond percolation theory problem.

(b) (i) At $T = 0$, the total free energy $F = \langle E \rangle - TS = \langle E \rangle$. Because an equilibrium system will minimise the free energy, at $T = 0$ it will minimise its energy. In order to minimise the energy, all spins within a given cluster will point in the same direction. However, spins belonging to different clusters need not point in the same direction.

(ii) Within a cluster, $s_i = s_j$ so $s_i s_j = s_i^2 = 1$ implying $\langle s_i s_j \rangle = 1$ if the spins belong to the same cluster. If the spins i and j belong to different clusters, they are not correlated at all, that is, given, e.g., that $s_i = 1$ then $s_j = 1$ with probability 0.5 and $s_j = -1$ with probability 0.5 leaving $\langle s_i s_j \rangle = 0.5 \cdot 1 + 0.5 \cdot (-1) = 0$. Hence

$$\langle s_i s_j \rangle = \begin{cases} 1 & i, j \text{ in the same percolation cluster} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

(iii) For $p < p_c$ all clusters are finite. Since the clusters are not correlated, the average magnetisation must be zero.
For $p > p_c$, we can argue that all the finite clusters do not contribute to the magnetisation as their magnetisation would average out to zero. Hence, the magnetisation then becomes equal to $P_\infty(p)$, the density of the infinite cluster. In zero external field, the orientation of the spins in the infinite cluster is either up + or down −.
For $p = p_c$ the argument is the same as for $p > p_c$ with the additional information that the density of the infinite cluster $P_\infty(p)$ is zero at $p = p_c$ and hence there is no net magnetisation.
In summary

$$m_0(p) = \pm P_\infty(p) = \begin{cases} 0 & \text{for } p \leq p_c \\ \neq 0 & \text{for } p > p_c. \end{cases} \tag{6}$$

(c) (i) $P_\infty(p)$ is the probability for a spin to belong to the percolating infinite cluster. As $\tanh(sH/k_BT) \to 0$ for $H \to 0^\pm$, the last term will vanish and

$$m_0(p) = \lim_{H \to 0^\pm} m(p, H) = \pm P_\infty(p)$$

consistent with the result of (b)(iii).

(ii) The susceptibility per spin in zero external field

$$\chi(T, 0) = \left(\frac{\partial m}{\partial H} \right)_T \Bigg|_{H=0}.$$

Assuming $H \ll k_BT$ we use the Taylor expansion $\tanh(sH/k_BT) \approx sH/k_BT + \mathcal{O}\left((sH/k_BT)^3\right)$. Since $P_\infty(p)$ does not depend on the external field, we find,

$$\chi(T, 0) = \left(\frac{\partial m}{\partial H} \right)_T \Bigg|_{H=0} = \sum_{s=1}^{\infty} \frac{s^2 n(s, p)}{k_BT} = \beta \chi(p) \propto |p - p_c|^{-\gamma} \tag{7}$$

as the divergence of the second moment of the cluster number density $n(s, p)$ is characterized by the exponent γ when $p \to p_c$.

(d) When $p < p_c$, the magnetisation in zero external field $m_0(p) = 0$. We are at low temperature, so we may assume that within a cluster $\langle s_i s_j \rangle = 1$. In a cluster of size s there are a total of s^2 different pairs, so $\frac{1}{k_B T} \sum_i \sum_j \langle s_i s_j \rangle = \frac{1}{k_B T} s^2$. We can calculate the average susceptibility per spin by summing over all possible cluster sizes weighted by the cluster number density, that is,

$$\chi(T, 0) = \sum_{s=1}^{\infty} \left(\frac{1}{k_B T} \sum_i \sum_j \langle s_i s_j \rangle \right) n(s, p) = \frac{1}{k_B T} \sum_{s=1}^{\infty} s^2 n(s, p). \qquad (8)$$

3. (a) *Landau theory for the Ising model.*

(i) On each site $\mathbf{r}_i$, there is a spin variable $s_i = \pm 1$ that can take on only two values: spin up $(+1)$ or spin down (-1).

(ii) It is energetically favorable for neighbouring spins to be parallel. So, $J > 0$ so that a pair of parallel spins where $s_i s_j = +1$ has energy $-J$ and a pair of anti-parallel spins where $s_i s_j = -1$ has energy $+J$.

(iii) The first sum runs over distinct nearest neighbour pairs (i.e., we assume that the spin-spin interaction J_{ij} fall off so rapidly that only nn interactions are present). If z denotes the coordination number, then

$$\sum_{\langle ij \rangle} 1 = \frac{z}{2} \sum_{i=1}^{N} 1 = \frac{1}{2} N z, \qquad (9)$$

where the factor of $1/2$ ensures that we are counting *distinct* nearest-neighbour pairs only.

(b) See Figure 3.

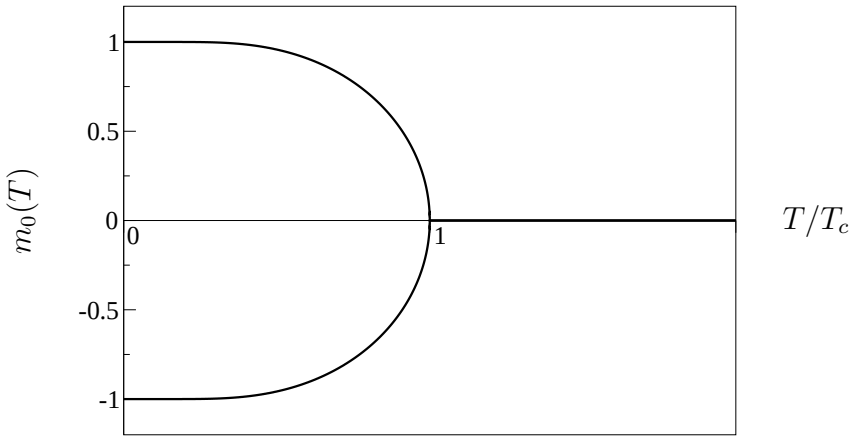


Figure 3: When $T \to 0$, then $m_0(T) \to \pm 1$. When $T \to T_c^-$, the magnetisation decreases sharply but continuously to zero at $T = T_c$. For all $T \geq T_c$, the magnetisation per spin in zero external field $m_0(T) = 0$.

(c) See Figure 4.

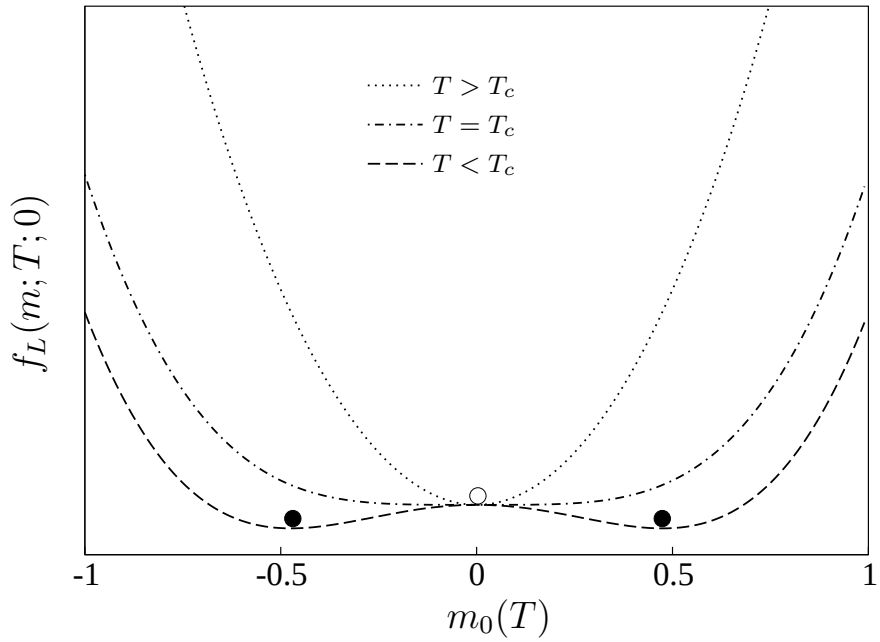


Figure 4: The Landau free energy per spin $f_L(m;T;0)$ vs. the average magnetisation per spin in zero external field $m_0(T)$. The solid circles show the position of the minima of the free energy of the corresponding graph. For $T \geq T_c$, unique minimum at $m_0(T) = 0$. For $T < T_c$, double minima at $m_0(T) = \pm m_0 \neq 0$.

Minimise f_L with respect to m:

$$\left(\frac{\partial f}{\partial m}\right)_{T,H} = 0 \Leftrightarrow 2a_2(T - T_c)m + 4a_4 m^3 - H = 0. \tag{10}$$

In zero external field, we have

$$2m_0(T)\left[a_2(T - T_c) + 2a_4 m_0^2(T)\right] = 0. \tag{11}$$

For $T \geq T_c$ this implies $m_0(T) = 0$ because $a_4 > 0$.
For $T < T_c$ this implies $m_0(T) = \pm\left[a_2(T_c - T)/2a_4\right]^{1/2} \propto \pm\,(T_c - T)^{1/2}$.
In summary

$$\begin{aligned} m_0(T) &= \begin{cases} 0 & \text{for } T \geq T_c \\ \pm\left[a_2(T_c - T)/2a_4\right]^{1/2} & \text{for } T \to T_c^- \end{cases} \\ &\propto \begin{cases} 0 & \text{for } T \geq T_c \\ \pm(T_c - T)^{1/2} & \text{for } T \to T_c^-. \end{cases} \end{aligned} \tag{12}$$

(d) See Figure 5.

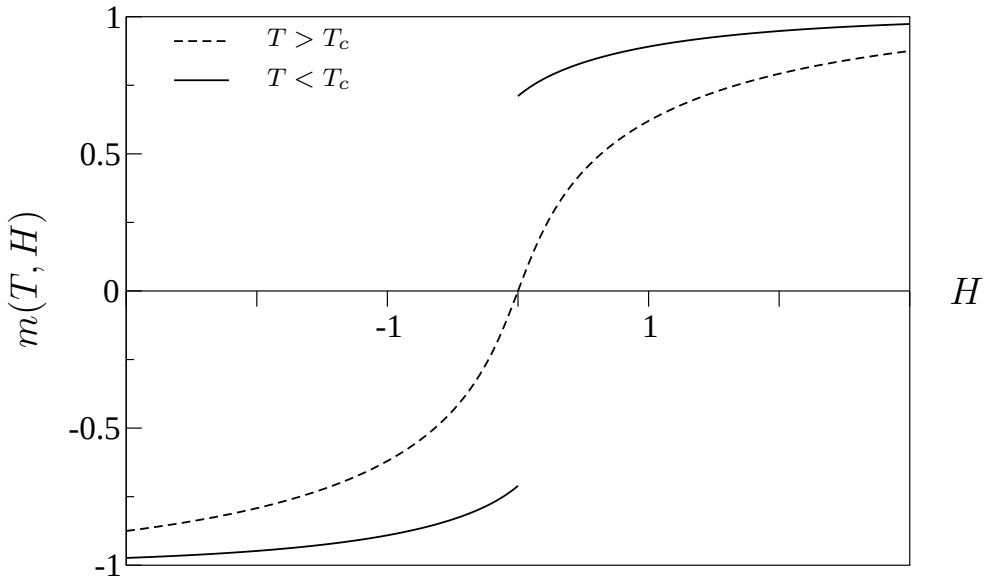


Figure 5: Sketch of the magnetisation per spin $m(T, H)$ versus the external field H for two different temperatures $T > T_c$ and $T < T_c$. For large external fields, the magnetisation saturates to $m = \pm 1$ for both graphs. When $H = 0$: For $T > T_c$ graph is continuous and it crosses the point $(0, 0)$ because $m(T, 0) = 0$. For $T < T_c$, the graph have a discontinuous jump at $H = 0$ because $\lim_{H \to 0^{\pm}} m(T, H) = \pm m(T, 0) \neq 0$.

(e) (i) I have been quite 'naughty' posing you this question as it is tempting you to make wrong conclusions in order to reach an almost (a factor of 2 will be missing) correct answer. Of course, I would never do that in an exam situation. However, it might be very instructive because it exposes two types of wrong-doing (that are intimately linked) that are found quite frequently in literature on the Ising model.

Case 1 – wrong mean-field theory:
First, we derive the result for non-interacting spins in an external field H (Sec. 2.2 in notes). The total energy is

$$E_{\{s_i\}} = -H \sum_{i=1}^{N} s_i. \tag{13}$$

The partition function is

$$Z(T,H) = \sum_{\{s_i\}} \exp\left(-\beta E_{\{s_i\}}\right) = \sum_{\{s_i\}} \exp\left(\beta H \sum_{i=1}^{N} s_i\right)$$
$$= \sum_{\{s_i\}} \prod_{i=1}^{N} \exp\left(\beta H s_i\right) = \left(2\cosh\beta H\right)^N . \tag{14}$$

The free energy per spin

$$f(T,H) = -k_B T \ln\left(2\cosh\beta H\right), \tag{15}$$

and hence, the magnetisation per spin is given by

$$m(T,H) = -\left(\frac{\partial f}{\partial H}\right)_T = k_B T \frac{2\sinh\beta H}{2\cosh\beta H}\,\beta = \tanh\beta H. \tag{16}$$

The total energy of the Ising model in this version of the mean-field model:

$$\begin{aligned}
E_{\{s_i\}} &= -J\sum_{\langle ij\rangle} s_i s_j - H\sum_{i=1}^{N} s_i \\
&\approx -J\sum_{\langle ij\rangle} s_i m - H\sum_{i=1}^{N} s_i \\
&= -Jm\frac{z}{2}\sum_{i=1}^{N} s_i - H\sum_{i=1}^{N} s_i \quad \text{each site } i \text{ has } \frac{z}{2} \text{ distinct nn} \\
&= -\left(Jm\frac{6}{2} + H\right)\sum_{i=1}^{N} s_i \quad z = 6 \text{ in } d = 3 \text{ cubic lattice} \\
&= -\left(3Jm + H\right)\sum_{i=1}^{N} s_i \\
&= -H_{\text{eff}}\sum_{i=1}^{N} s_i,
\end{aligned} \tag{17}$$

where we have introduced an effective external field

$$H_{\text{eff}} = 3Jm + H. \tag{18}$$

So far nothing illegal has taken place. However, now the argument (wrongly) states that using Eq. (16), we find that the magnetisation must satisfy the equation

$$m(T,H) = \tanh\beta H_{\text{eff}} = \tanh\left(\beta 3Jm + \beta H\right). \tag{19}$$

To see that this is indeed a conclusion that **cannot** be drawn, we need to go through the second case that is often presented in the literature.

Case 2 - wrong mean-field theory:
We start with the mean-field energy

$$E_{\{s_i\}} \approx -\left(3Jm + H\right) \sum_{i=1}^{N} s_i. \tag{20}$$

The associated partition function is

$$\begin{aligned} Z &= \sum_{\{s_i\}} \exp\left(\left(\beta 3Jm + \beta H\right) \sum_{i=1}^{N} s_i \right) \\ &= \sum_{\{s_i\}} \prod_{i=1}^{N} \exp\left[\left(\beta 3Jm + \beta H\right) s_i \right] \\ &= \left[2\cosh\left(\beta 3Jm + \beta H\right) \right]^N . \end{aligned} \tag{21}$$

The free energy per spin is

$$\begin{aligned} f &= -\frac{1}{N} k_B T \ln\left[\left[2\cosh\left(\beta 3Jm + \beta H\right) \right]^N \right] \\ &= -k_B T \ln\left[2\cosh\left(\beta 3Jm + \beta H\right) \right]. \end{aligned} \tag{22}$$

Hence, the magnetisation per spin is

$$\begin{aligned} m &= -\left(\frac{\partial f}{\partial H} \right)_T \\ &= k_B T \frac{2\sinh(\beta 3Jm + \beta H)}{2\cosh(\beta 3Jm + \beta H)} \left(\beta 3J \left(\frac{\partial m}{\partial H} \right)_T + \beta \right) \\ &= \tanh(\beta 3Jm + \beta H) \left(3J \left(\frac{\partial m}{\partial H} \right)_T + 1 \right). \end{aligned} \tag{23}$$

However it is common to use a dirty trick and (wrongly) briefly assume that m is independent of H. In doing so, we arrive at the following equation for determining m in the mean-field picture

$$\begin{aligned} m &= -\left(\frac{\partial f}{\partial H} \right)_T \\ &= k_B T \frac{2\sinh(\beta 3Jm + \beta H)}{2\cosh(\beta 3Jm + \beta H)} \\ &= \tanh(\beta 3Jm + \beta H). \end{aligned} \tag{24}$$

This is in effect what is tacitly done in case 1 also. Indeed, Eq.(16) states

$$m(T,H) = k_B T \frac{2\sinh\beta H}{2\cosh\beta H}\left(\frac{\partial \beta H}{\partial H}\right)_T \tag{25}$$

and substituting in this equation $H_{\text{eff}} = 3Jm + \beta H$ we arrive at the same result.

Case 3 - correct mean-field theory:
Following the derivation in the notes (Sec. 2.5), we find that the correct equation for determining m in the mean-field picture is, in fact

$$m = \tanh(\beta 6Jm + \beta H), \tag{26}$$

that is, there is an extra factor of 2 in the contribution to the 'internal field'.

Also, cases 1 and 2 would also fail to yield an equation for the magnetisation m using the equation

$$\left(\frac{\partial f}{\partial m}\right)_{T,H} = 0 \tag{27}$$

(ii) Mean field theory result:

$$m = \tanh(\beta 6Jm + \beta H) = \tanh\left[\left(6Jm + H\right)/(k_B T)\right]. \tag{28}$$

Uncorrelated spins in an effective field $H_{\text{eff}} = H + 6Jm$. Each spin feels the external field H. Also, each spin feels the average magnetisation m of each of its 6 ($d = 3$ cubic lattice) neighbouring spins. In the exchange interaction: $-Js_i s_j$ this corresponds to an effective internal field of $6Jm$.

As you can see from the discussion above, this interpretation is strictly speaking, not scientifically sound but nevertheless frequently used in the literature.

Problem sheet 9 issued: Tuesday 3rd December 2013
Solutions available from: Wednesday 11th December 2013

Statistical Mechanics Answer Sheet 9

1. *Scaling ansatz of free energy per spin and scaling relations.* (**RF Question**)

   (a) Consider the Ising model on a d-dimensional lattice in an external field H.

   (i) The total energy for a system of N spins $s_i = \pm 1$ with constant nearest-neighbour interactions $J > 0$ placed in a uniform external field H is

   $$E_{\{s_i\}} = -J \sum_{\langle ij \rangle} s_i s_j - H \sum_{i=1}^{N} s_i, \tag{1}$$

   where the notation $\langle ij \rangle$ restricts the sum to run over all distinct nearest-neighbour pairs.

   (ii) Let $M_{\{s_i\}} = \sum_{i=1}^{N} s_i$ denote the total magnetisation and $\langle M \rangle$ the average total magnetisation. The order parameter for the Ising model is defined as the magnetisation per spin

   $$m(T, H) = \lim_{N \to \infty} \frac{\langle M \rangle}{N}. \tag{2}$$

   Consider the free energy $F = \langle E \rangle - TS$. The ratio of the average total energy, $\langle E \rangle$, to the temperature times entropy, TS, defines a dimensionless scale $J/(k_B T)$. A competition exists between the tendency to randomise the orientation of spins for $J \ll k_B T$, and a tendency to align spins for $J \gg k_B T$. In the former case, the free energy is minimised by maximising the entropic term: the magnetisation is zero because the spins point up and down randomly. In the latter case, the free energy is minimised by minimising the total energy: the magnetisation is non-zero because the spins tend to align. Since the entropy in the free energy is multiplied by temperature, for sufficiently low temperatures, the minimisation of the free energy is dominated by the minimisation of the total energy. Therefore, at least qualitatively, there is a possibility of a phase transition from a phase with zero magnetisation at relatively high temperatures, to a phase with non-zero magnetisation at relatively low temperatures.

We assume that the singular part of free energy per spin is a generalised homogeneous function,

$$f(t,h) = b^{-d} f(b^{y_t} t, b^{y_h} h) \quad \text{for } t \to 0^{\pm}, h \to 0, \forall\ b > 0. \tag{3}$$

(b) Below, we use the notation $\frac{\partial f}{\partial t} = f'_t$ and $\frac{\partial^2 f}{\partial t^2} = f''_{tt}$ and similar for partial derivatives w.r.t. h.

(i) The critical exponent α associated with the specific heat in zero external field characterises its divergence as $t \to 0$ and is defined by

$$c(t,0) \propto |t|^{-\alpha} \quad \text{for } t \to 0. \tag{4}$$

The specific heat is related to the free energy per spin:

$$c(t,h) \propto \left(\frac{\partial^2 f}{\partial t^2}\right) \propto b^{2y_t - d} f''_{tt}(b^{y_t} t, b^{y_h} h). \tag{5}$$

Choosing $b = |t|^{-1/y_t}$ and setting $h = 0$ we find

$$c(t,0) \propto |t|^{-\frac{2y_t - d}{y_t}} f''_{tt}(\pm 1, 0) \quad \text{for } t \to 0^{\pm}, \tag{6}$$

and since $f''_{tt}(\pm 1, 0)$ are just numbers, we identify

$$\alpha = \frac{2y_t - d}{y_t}. \tag{7}$$

(ii) The critical exponent β associated with the order parameter (magnetisation per spin) in zero external field characterises the pick up of the order parameter as $t \to 0^-$ and is defined by

$$m(t,0) \propto |t|^{\beta} \quad \text{for } t \to 0^-. \tag{8}$$

The magnetisation per spin is related to the free energy per spin:

$$m(t,h) \propto -\left(\frac{\partial f}{\partial h}\right) \propto b^{y_h - d} f'_h(b^{y_t} t, b^{y_h} h). \tag{9}$$

Choosing $b = |t|^{-1/y_t}$ and setting $h = 0$ we find

$$m(t,0) \propto |t|^{\frac{d - y_h}{y_t}} f'_h(-1, 0) \quad \text{for } t \to 0^-, \tag{10}$$

and since $f'_h(-1,0)$ is just a number, we identify

$$\beta = \frac{d - y_h}{y_t}. \tag{11}$$

(iii) The critical exponent γ associated with the susceptibility in zero external field characterises its divergence when $t \to 0$ and is defined by

$$\chi(t,0) \propto |t|^{-\gamma} \quad \text{for } t \to 0. \tag{12}$$

The susceptibility is related to the free energy per spin:

$$\chi(t,h) \propto -\left(\frac{\partial^2 f}{\partial h^2}\right) \propto b^{2y_h-d} f''_{hh}(b^{y_t}t, b^{y_h}h). \tag{13}$$

Choosing $b = |t|^{-1/y_t}$ and setting $h = 0$ we find

$$\chi(t,0) \propto |t|^{-\frac{2y_h-d}{y_t}} f''_{hh}(\pm 1, 0) \quad \text{for } t \to 0 \tag{14}$$

and since $f''_{hh}(\pm 1, 0)$ are just numbers, we identify

$$\gamma = \frac{2y_h - d}{y_t}. \tag{15}$$

(iv) The critical exponent δ associated with the order parameter at the critical temperature characterises how the magnetisation per spin vanishes for small external fields and is defined by

$$m(0,h) \propto \text{sign}(h)|h|^{1/\delta} \quad \text{for } h \to 0^{\pm}. \tag{16}$$

The magnetisation per spin is related to the free energy per spin:

$$m(t,h) \propto -\left(\frac{\partial f}{\partial h}\right) \propto b^{y_h-d} f'_h(b^{y_t}t, b^{y_h}h). \tag{17}$$

Choosing $b = |h|^{-1/y_h}$ and setting $t = 0$ we find

$$m(0,h) \propto |h|^{\frac{d-y_h}{y_h}} f'_h(0, \pm 1) \quad \text{for } h \to 0 \tag{18}$$

and since $f'_h(0, \pm 1)$ are just numbers, we identify

$$\delta = \frac{y_h}{d - y_h}. \tag{19}$$

(v) We find

$$\begin{aligned} \alpha + 2\beta + \gamma &= \frac{2y_t - d + 2d - 2y_h + 2y_h - d}{y_t} \\ &= 2 \end{aligned}$$

and

$$\begin{aligned}\beta(\delta-1) &= \frac{d-y_h}{y_t}\left(\frac{y_h}{d-y_h}-1\right)\\ &= \frac{d-y_h}{y_t}\left(\frac{2y_h-d}{d-y_h}\right)\\ &= \frac{2y_h-d}{y_t}\\ &= \gamma.\end{aligned}$$

2. *One-dimensional Ising model with periodic boundary conditions* (Exam 2007)

   (a) The total energy for a system of N spins $s_i = \pm 1$ with constant nearest-neighbour interactions $J > 0$ placed in a uniform external field H is

$$\begin{aligned}E_{\{s_i\}} &= -J\sum_{\langle ij\rangle} s_i s_j - H\sum_{i=1}^{N} s_i,\\ &= -J\sum_{i=1}^{N} s_i s_{i+1} - H\sum_{i=1}^{N} s_i. \qquad (20)\end{aligned}$$

   The sum over all distinct nearest-neighbour pairs $\langle ij\rangle$ reduces to the sum over all spins in $d = 1$ with $s_{N+1} = s_1$.

   (b) (i) At $T = 0$ all spins are aligned. Hence there are 2 microstates with all spins pointing up or all spins pointing down.

   (ii) When $T \to \infty$ all spins are pointing up and down at random without any correlations. Hence there are a total of 2^N microstates.

   (iii) The total energy $E_{\{s_i\}} = -J\sum_{i=1}^{N} s_i s_{i+1}$ in zero external field. At $T = 0$ all spins are aligned. At $T = \infty$ spins are pointing up and down at random. Hence, the energy per spin

$$\epsilon(T,0) = \frac{\langle E\rangle}{N} = \begin{cases} -J & \text{at } T = 0,\\ 0 & \text{for } T \to \infty. \end{cases} \qquad (21)$$

   (iv) The magnetisation per spin

$$m(T,0) = \frac{\langle M\rangle}{N} = \begin{cases} \pm 1 & \text{at } T = 0,\\ 0 & \text{for } T \to \infty \end{cases} \qquad (22)$$

   since at $T = 0$ all spins are aligned while at $T = \infty$ spins are pointing up and down at random.

(v) The entropy $S = k_B \ln \Omega$ where Ω is the number of microstates. Hence

$$S(T,0) = \begin{cases} k_B \ln 2 & \text{for } T = 0, \\ N k_B \ln 2 & \text{for } T \to \infty. \end{cases} \tag{23}$$

You may also arrive at the same result using

$$S(T,0) = -k_B \sum_{\{s_i\}} p_{\{s_i\}} \ln p_{\{s_i\}}, \tag{24}$$

where $p_{\{s_i\}}$ is given by the Boltzmann distribution

$$p_{\{s_i\}} = \frac{\exp\left(-\beta E_{\{s_i\}}\right)}{\sum_{\{s_i\}} \exp\left(-\beta E_{\{s_i\}}\right)}, \tag{25}$$

with $\beta = 1/(k_B T)$ the 'inverse temperature'.

For $T \to 0$, only the two ground states will have a non-zero probability and $p_{\{s_i=+1 \forall i\}} = p_{\{s_i=-1 \forall i\}} = 1/2$. Hence

$$S(0,0) = -k_B \sum_{\{s_i\}} p_{\{s_i\}} \ln p_{\{s_i\}} = -k_B \left(\frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{1}{2} \right) = k_B \ln 2. \tag{26}$$

For $T \to \infty$, $\beta \to 0$ and all 2^N microstates have equal probability with $p_{\{s_i\}} = 2^{-N}$. Hence

$$S(\infty,0) = -k_B \sum_{\{s_i\}} p_{\{s_i\}} \ln p_{\{s_i\}} = -k_B 2^N 2^{-N} \ln 2^{-N} = N k_B \ln 2. \tag{27}$$

(vi) The total free energy $\langle F \rangle = \langle E \rangle - TS$. Hence, using the results of (iii) and (v) we find that the free energy per spin

$$f(T,0) = \frac{F}{N} = \begin{cases} -J - \frac{1}{N} k_B T \ln 2 & \text{for } T = 0, \\ -k_B T \ln 2 & \text{for } T \to \infty. \end{cases} \tag{28}$$

(c) The magnetisation per spin

$$m(T,H) = \frac{\sinh \beta H}{\sqrt{\sinh^2 \beta H + \exp(-4\beta J)}}, \tag{29}$$

where $\beta = 1/(k_B T)$ and $J > 0$ the coupling constant. We note that $\sinh \beta H \to 0$ for $H \to 0^{\pm}$.

When $T > 0$, the term $\exp(-4\beta J)$ is finite and hence $\lim_{H \to 0} m(T,H) = 0$.

When $T = 0$, the term $\exp(-4\beta J)$ is zero and hence $\lim_{H \to 0^{\pm}} m(0,H) = \pm 1$.

Hence there is no phase-transition in the Ising model in zero external field at any finite temperature.

3. *One-dimensional Ising model with periodic boundary conditions* (Exam 2010)

(a) The partition function for the $d = 1$ Ising model:

$$\begin{aligned} Z_{\text{ring}} &= \sum_{\{s_i\}} e^{-\beta E_{\{s_i\}}} \\ &= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} e^{\beta J s_1 s_2} e^{\beta J s_2 s_3} \cdots e^{\beta J s_{N-1} s_N} e^{\beta J s_N s_1} \end{aligned} \tag{30}$$

where $\beta = 1/(k_B T)$ with k_B the Boltzmann constant and T the temperature.

(b) Using the notation of the transfer matrix, we find

$$\begin{aligned} Z_{\text{ring}} &= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} e^{\beta J s_1 s_2} e^{\beta J s_2 s_3} \cdots e^{\beta J s_{N-1} s_N} e^{\beta J s_N s_1} \\ &= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} T_{s_1 s_2} T_{s_2 s_3} \cdots T_{s_{N-1} s_N} T_{s_N s_1} \\ &= \sum_{s_1=\pm 1} \sum_{s_3=\pm 1} \cdots \sum_{s_{N-1}=\pm 1} \left(\sum_{s_2=\pm 1} T_{s_1 s_2} T_{s_2 s_3} \right) \cdots \left(\sum_{s_N=\pm 1} T_{s_{N-1} s_N} T_{s_N s_1} \right) \\ &= \sum_{s_1=\pm 1} \sum_{s_3=\pm 1} \cdots \sum_{s_{N-1}=\pm 1} T^2_{s_1 s_3} T^2_{s_3 s_5} \cdots T^2_{s_{N-3} s_{N-1}} T^2_{s_{N-1} s_1} \\ &= \sum_{s_1=\pm 1} \sum_{s_5=\pm 1} \cdots \sum_{s_{N-3}=\pm 1} T^4_{s_1 s_5} T^4_{s_5 s_9} \cdots T^4_{s_{N-3} s_1} \\ &= \sum_{s_1=\pm 1} T^N_{s_1 s_1} \\ &= \text{Tr}\left(\mathbf{T}^N\right), \end{aligned} \tag{31}$$

where we use the (general) fact of matrix multiplication

$$\sum_{s_k} T_{s_i s_k} T_{s_k s_j} = T^2_{s_i s_j}. \tag{32}$$

(c) The transfer matrix in zero external field ($H = 0$) is

$$\mathbf{T} = \begin{pmatrix} T_{+1+1} & T_{+1-1} \\ T_{-1+1} & T_{-1-1} \end{pmatrix} = \begin{pmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{pmatrix}. \tag{33}$$

The eigenvalues $\lambda_\pm$ of $\mathbf{T}$ are the solutions to the characteristic equation

$$\det(\mathbf{T} - \lambda \mathbf{I}) = 0. \tag{34}$$

The determinant

$$\begin{aligned} \det(\mathbf{T} - \lambda \mathbf{I}) &= \begin{vmatrix} e^{\beta J} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} - \lambda \end{vmatrix} \\ &= \lambda^2 - 2e^{\beta J}\lambda + e^{2\beta J} - e^{-2\beta J} \end{aligned} \tag{35}$$

so the solutions to the characteristic Equation (34) are

$$\begin{aligned}\lambda_{\pm} &= \frac{2e^{\beta J} \pm \sqrt{4e^{2\beta J} - 4\left[e^{2\beta J} - e^{-2\beta J}\right]}}{2} \\ &= e^{\beta J} \pm e^{-\beta J} \\ &= \begin{cases} 2\cosh \beta J, \\ 2\sinh \beta J. \end{cases}\end{aligned} \tag{36}$$

Hence, the partition function

$$\begin{aligned}Z_{\text{ring}} &= \lambda_+^N + \lambda_-^N \\ &= (2\cosh \beta J)^N + (2\sinh \beta J)^N \\ &= (2\cosh \beta J)^N \left[1 + \tanh^N \beta J\right].\end{aligned} \tag{37}$$

(d) In the high-temperature limit

$$\beta J \ll 1. \tag{38}$$

Taylor expansion to first order yields $\cosh \beta J \approx 1$ and $\tanh \beta J \approx \beta J$ so that the partition function

$$\begin{aligned}Z_{\text{ring}} &= (2\cosh \beta J)^N \left[1 + \tanh^N \beta J\right] \\ &\approx 2^N \left[1 + (\beta J)^N\right] \\ &\approx 2^N\end{aligned} \tag{39}$$

because $(\beta J)^N \ll 1$. Hence, the total free energy of the system

$$\begin{aligned}F_{\text{ring}} &= -k_B T \ln Z_{\text{ring}} \\ &\approx -k_B T \ln 2^N \\ &= -TNk_B \ln 2 \\ &= -TS,\end{aligned} \tag{40}$$

where the entropy is given by

$$S = k_B \ln 2^N, \tag{41}$$

and also given by

$$S = -k_B \sum_{\{s_i\}} p_{\{s_i\}} \ln p_{\{s_i\}}, \tag{42}$$

with

$$\lim_{T\to\infty} p_{\{s_i\}} = \lim_{T\to\infty} \frac{e^{-\beta E_{\{s_i\}}}}{\sum_{\{s_i\}} e^{-\beta E_{\{s_i\}}}} = \frac{1}{2^N}. \tag{43}$$

The N spins are effectively free spins because the thermal energy k_BT is much larger than the energy $2J$ it costs to one flip from a $\uparrow\uparrow$ to a $\uparrow\downarrow$ local configuration. Therefore the free energy is just entropic with $-Tk_B \ln 2$ per spin.

(e) (i) Assume extremely low temperature with $\beta J \gg 1$. Using that

$$2\cosh x = e^x + e^{-x} \approx e^x \quad \text{for } x \gg 1, \tag{44a}$$

$$2\sinh x = e^x - e^{-x} \approx e^x \quad \text{for } x \gg 1, \tag{44b}$$

we find that the partition function

$$\begin{aligned} Z_{\text{ring}} &= (2\cosh \beta J)^N + (2\sinh \beta J)^N \\ &\approx e^{N\beta J} + e^{N\beta J} \\ &= 2e^{N\beta J}. \end{aligned} \tag{45}$$

Hence the total free energy

$$\begin{aligned} F_{\text{ring}} &= -k_BT \ln Z_{\text{ring}} \\ &\approx -k_BTN\beta J - k_BT \ln 2 \\ &= -NJ - k_BT \ln 2. \end{aligned} \tag{46}$$

(ii) Recall that $F = \langle E \rangle - TS$. Hence, we identify the first term, $-NJ$ as the energy of the ground state where all spins are aligned at $T = 0$. The second term, $-k_BT \ln 2$, tells us that the entropy is $k_B \ln 2$ at low temperatures. This is because there are two degenerate ground states, all spins pointing up or all spins pointing down.

4. *Ising model in $d > 1$* (Exam 2006)

(a) (i) The total energy for a system of N spins $s_i = \pm 1$ with constant nearest-neighbour interactions $J > 0$ placed in a uniform external field H is

$$E_{\{s_i\}} = -J \sum_{\langle ij \rangle} s_i s_j - H \sum_{i=1}^{N} s_i, \tag{47}$$

where the notation $\langle ij \rangle$ restricts the sum to run over all distinct nearest-neighbour pairs.

(ii) Spins interact only with their nearest neighbours. The interaction strength is assumed to be a constant. The spins can only take one of two values $s_i = \pm 1$. Finally, the external field H is constant.

(b) (i) The free energy per spin

$$f(T, H) = -\frac{1}{N} k_BT \ln Z. \tag{48}$$

(ii) The average magnetisation per spin

$$m(T,H) = \left\langle \frac{1}{N}\sum_{i=1}^{N} s_i \right\rangle . \tag{49}$$

The statistical mechanical definition of the free energy yields

$$\begin{aligned} -\left(\frac{\partial f}{\partial H}\right)_T &= \frac{1}{N}k_B T \frac{\partial}{\partial H}\ln Z \\ &= \frac{1}{N}k_B T \frac{1}{Z}\frac{\partial}{\partial H} Z \\ &= \frac{1}{N}k_B T \frac{1}{Z}\frac{\partial}{\partial H}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) \\ &= \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}})\frac{1}{N}\sum_{i=1}^{N} s_i \\ &= \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) m_{\{s_i\}} \end{aligned} \tag{50}$$

which is indeed the average magnetisation per spin.

(c) For $T \geq T_c$, the spins are equally likely to be pointing up and down on average so the magnetisation per spin is zero. The magnetisation picks up abruptly at $T = T_c$ and for $T < T_c$ a finite fraction of the spins are aligned. At $T = 0$, all spins point in the same direction. Hence, $m(0,0) = \pm 1$.

(d) (i) The average magnetisation per spin is

$$m(T,H) = -\left(\frac{\partial f}{\partial H}\right)_T . \tag{51}$$

Hence, for fixed temperature, T, the magnetisation per spin is the negative slope of the free energy per spin as a function of the external field H.

(ii) Clearly $T > T_c$ has zero slope at $H = 0$. For $T < T_c$ the slope is finite and take the same numerical value for $H \to 0^{\pm}$ but with different sign. At $T = T_c$, the slope is also zero but the second derivative (susceptibility) diverges, that is, the rate of change in the slope is infinite.

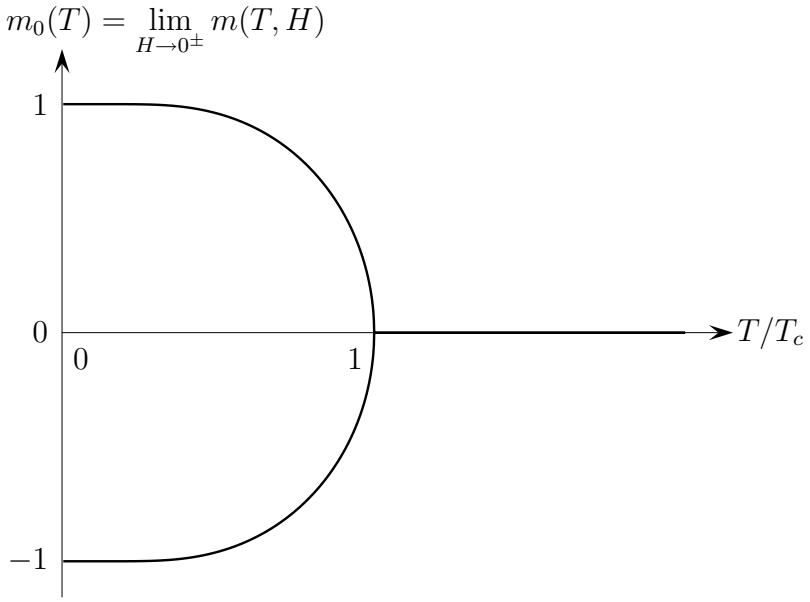


Figure 1: A sketch of the magnetisation per spin $m_0(T) = \lim_{H \to 0^{\pm}} m(T, H)$ versus the relative temperature T/T_c for the Ising model.

Problem sheet issued: Tuesday 8th October 2013
Solutions available from: Wednesday 16th October 2013

Statistical Mechanics Answer Sheet 1

Regular Fractals

1. Sierpinski carpet
   We need to deduce how the area A scales with the system size. Each square of side L contains 8 smaller squares of side $L/3 \Rightarrow A(L) = 8A(L/3) = 8^n A(L/3^n)$. Set $b = 3^n$ so that $n = \ln b / \ln 3$. Substitute into 8^n and rearrange to obtain

   $$A(L) = b^{\ln 8/\ln 3} A(L/b)\,.$$

   This is an implicit equation for A. To obtain an explicit equation, set $L/b = a$ for some small length scale a. Then, we can see that $A(L) = (L/a)^{d_f} A(a)$. In other words, the area obeys a power law: $A(L) \sim L^{d_f}$ with $d_f = \ln 8/\ln 3 = \log_3 8$.

   Alternatively, you can suggest the power law $A \sim L^{d_f}$ as an ansatz and substitute into the implicit scaling form above to give: $L^{d_f} = b^{\ln 8/\ln 3}(L/b)^{d_f} \Rightarrow d_f = \ln 8/\ln 3$.

2. Koch curve

   (a) The length of the perimeter P is just like the area in the previous problem. Here, at each iteration, the length of a straight segment is reduced by a factor of 3, but the number of segments has increased by a factor of 4. Following the same procedure as above,

   $$P(L) = 4^n P(L/3^n) \Rightarrow P(L) = b^{\ln 4/\ln 3} P(L/b),$$

   we can identify $d_f = 2\ln 2/\ln 3 \simeq 1.26$.

   (b) Qualitatively, the perimeter is very large because the curve wiggles at all length scales.
   Suppose we start with a triangle of side L and stop constructing the curve after n steps, when the sides are of length a. Then, setting $L/b = a$ gives $P(L) = (L/a)^{d_f} P(a)$. But the length of the portion of side a is simply a if we stopped the iteration at that scale. So

   $$P(L) = (L/a)^{d_f} a = L^{d_f} a^{1-d_f}.$$

Note that $d_f > 1$ and so $P(L) \propto 1/a^{0.26}$ scales as a *negative* power of a. This means that the perimeter diverges if we continue our construction to smaller and smaller scales ($n \to \infty$, $a \to 0$).

(c) Examine first the first iteration $n = 0 \to 1$. Each kink inserted adds two edges of length 1/3 of the original edge. This means that the extra enclosed area is a triangle with 1/9 of the area of the original. There are 3 such edges and so: $\Sigma_1 = \Sigma_0 + 3\Sigma_0/9 = \Sigma_0 + \Sigma_0/3$. Note that the resulting object has 4 times as many edges as the original.
Now consider the n'th iteration, which generates Σ_n from Σ_{n-1}. Each kink adds a triangle with 1/9 the area of the previous generation of added triangles, *i.e.*, $1/9^n$ of the original Σ_0. The number of kinks to add is the number of edges at the $(n-1)$-th generation. Since each generation multiplies the number of edges by 4, the number of edges is $3 \times 4^{n-1}$. Therefore, the extra area enclosed is $(\Sigma_0/9^n) \times (3 \times 4^{n-1})$, *i.e.*,

$$\Sigma_n = \Sigma_{n-1} + \left(\frac{4}{9}\right)^{n-1} \frac{\Sigma_0}{3}.$$

This is a recurrence relation which starts with a given $\Sigma_0 = \sqrt{3}L^2/4$ and generates the successive enclosed areas Σ_n at later generations. Each successive contribution to the area is a factor of 4/9 smaller than the previous one, giving a geometric series that we can sum analytically:

$$\begin{aligned} \Sigma_N &= \Sigma_0 + \frac{\Sigma_0}{3}\left[1 + \frac{4}{9} + \left(\frac{4}{9}\right)^2 + \ldots \left(\frac{4}{9}\right)^{N-1}\right] \\ \Rightarrow \lim_{N\to\infty} \Sigma_N &= \Sigma_0 \left[1 + \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{4}{9}\right)^n\right] = \Sigma_0\left[1 + \frac{1}{3}\cdot\frac{1}{1-(4/9)}\right] = \frac{8}{5}\Sigma_0. \end{aligned}$$

Random Fractals

3. 1D random walk

(a) The distribution is even in x. The walker is equally likely to go left or right at any step, so $\langle X \rangle = 0$.

(b) The RMS displacment is defined as $(\langle X^2 \rangle - \langle X \rangle^2)^{1/2}$. In this case, it is just $\langle X^2 \rangle^{1/2}$. For uncorrelated steps, the variances add up. In this case, all the steps obey the same distribution and so have the same variance:

$$\langle X^2 \rangle = \sum_{i=1}^{N} \langle x_i^2 \rangle = N \langle x_i^2 \rangle \,.$$

We now have to work out $\langle x_i^2 \rangle = \int_{-\infty}^{\infty} x^2 p(x)\, dx$. The distribution is proportional to the exponential and must be normalised, so it can be written in the

form:

$$p(x) = \frac{1}{A} e^{-|x|/a} \quad \text{with} \quad A = \int_{-\infty}^{\infty} e^{-|x|/a}\, dx = 2\int_{0}^{\infty} e^{-x/a}\, dx = 2a.$$

This gives

$$\langle x^2 \rangle = \frac{1}{A}\int_{-\infty}^{\infty} x^2 e^{-|x|/a}\, dx = \frac{1}{2a} \cdot 4a^3 = 2a^2\,.$$

Hence, since $\langle X^2 \rangle = N \times 2a^2$, the RMS displacement is $\sqrt{2N}a$.

(c) The central limit theorem states that the sum X of N independent and identically distributed random variables x_i obeys a probability distribution that converges, in the limit as $N \to \infty$, to a Normal distribution with mean $N\mu$ and variance $N\sigma^2$, where μ and σ^2 are the mean and variance of the individual terms x_i.
The Normal distribution for a variable z with mean μ and variance σ^2 is

$$\mathcal{N}(z-\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/2\sigma^2}.$$

Using the mean and variance we know from the previous parts of this question, we replace μ by zero and σ^2 by $N \cdot 2a^2$ to obtain

$$P(X) = \frac{1}{\sqrt{4\pi N a^2}} e^{-X^2/4Na^2}\,.$$

4. Ideal polymer

(a) (i) The distribution $p(x,y,z)$ can be factorized into independent distributions for x, y and z separately: $p(x,y,z) = f(x)f(y)f(z)$ with $f(x) = e^{-|x|/a}/2a$. It follows that each component performs an independent one-dimensional random walk as in the previous question. The 1D normalisation factor $1/2a$ was also derived in Q3. The normalised 3D probability distribution is

$$p(x,y,z) = A^{-1} e^{-(|x|+|y|+|z|)/a},$$

where $A = (2a)^3$.
Borrowing the 1D results, we find

$$\begin{aligned} \langle \boldsymbol{R} \rangle &= (\langle X \rangle, \langle Y \rangle, \langle Z \rangle) = (0,0,0), \\ \langle \boldsymbol{R}^2 \rangle &= \langle X^2 + Y^2 + Z^2 \rangle = \langle X^2 \rangle + \langle Y^2 \rangle + \langle Z^2 \rangle = 3 \times (2Na^2) = 6Na^2. \end{aligned}$$

Note that this argument *fails* if the three components are not independent. It would not work, for instance, for a spherically symmetric distribution of the form $p(x,y,z) \propto \exp(-\sqrt{x^2+y^2+z^2}/a) = \exp(-|\boldsymbol{r}|/a)$, which does not factorise.

(ii) Using the central limit theorem (as in the previous question):

$$\begin{aligned} P(\boldsymbol{R}) &= P(X)\,P(Y)\,P(Z) = \frac{1}{(4\pi Na^2)^{3/2}}\, e^{-X^2/4Na^2} e^{-Z^2/4Na^2} e^{-Z^2/4Na^2} \\ &= \frac{1}{(4\pi Na^2)^{3/2}}\, \exp\left(-\frac{\boldsymbol{R}^2}{4Na^2}\right). \end{aligned}$$

(iii) Normalisation requires $\int P(\boldsymbol{R})\, d^3R = 1$, so $P(\boldsymbol{R})$ must have dimensions of $1/(\text{length})^3$. The normalisation factor in $P(\boldsymbol{R})$ is indeed $\propto 1/a^3$. Also, the exponent in the Gaussian factor must be dimensionless, *i.e.*, $\boldsymbol{R}^2$ appears with the only length scale in the problem as $\boldsymbol{R}^2/a^2$.

(b) (i) The number of ways $W(\boldsymbol{R})$ to arrange the polymer so that its end-to-end displacement is $\boldsymbol{R}$ is proportional to the probability density $P(\boldsymbol{R})$: $W(\boldsymbol{R}) = AP(\boldsymbol{R})$ for some proportionality constant A. Therefore, $S(\boldsymbol{R}) = k_B \ln W(\boldsymbol{R}) = k_B(\ln P(\boldsymbol{R}) + \ln A)$. Using the probability density

$$P(\boldsymbol{R}) = \frac{1}{(4\pi Na^2)^{3/2}}\, \exp\left(-\frac{\boldsymbol{R}^2}{4Na^2}\right)$$

from the problem sheet, we obtain

$$S(\boldsymbol{R}) = -k_B\boldsymbol{R}^2/4Na^2 + k_B \ln A\,.$$

(ii) Configurations with different end-to-end distances now have different internal energies U. If the end experiences a force $\boldsymbol{f}$, this is equivalent to a potential energy for the end monomer of $U(\boldsymbol{R}) = -\boldsymbol{f}\cdot\boldsymbol{R}$ (+ irrelevant constant), *i.e.*, the end wants to roll down a potential hill with gradient $-\boldsymbol{f}$.
We must include the Boltzmann factor $e^{-U(\boldsymbol{R})/k_BT}$ to weight these different configurations, so

$$\begin{aligned} P(\boldsymbol{R}) &\propto \exp\left(-\frac{\boldsymbol{R}^2}{4Na^2}\right) \times e^{-U(\boldsymbol{R})/k_BT} \\ &\propto \exp\left(-\frac{\boldsymbol{R}^2}{4Na^2} + \frac{1}{k_BT}\boldsymbol{f}\cdot\boldsymbol{R}\right). \end{aligned}$$

(iii) I will give two equivalent derivations. The normalized probability distribution is

$$P(\boldsymbol{R}) = \frac{1}{Z}\exp\left(-\frac{\boldsymbol{R}^2}{4Na^2} + \frac{1}{k_BT}\boldsymbol{f}\cdot\boldsymbol{R}\right).$$

If we complete the square in the exponent,

$$\frac{\boldsymbol{R}^2}{4Na^2} - \frac{1}{k_BT}\boldsymbol{f}\cdot\boldsymbol{R} = \frac{1}{4Na^2}\left(\boldsymbol{R} - \frac{2Na^2}{k_BT}\boldsymbol{f}\right)^2 - \frac{Na^2}{k_B^2T^2}\boldsymbol{f}^2,$$

we see that

$$P(\boldsymbol{R}) \propto \exp\left[-\frac{1}{4Na^2}\left(\boldsymbol{R} - \frac{2Na^2}{k_BT}\boldsymbol{f}\right)^2\right] \times (\text{const indep of } \boldsymbol{R})$$

is again a Gaussian with the same variance but mean shifted to

$$\langle \boldsymbol{R} \rangle = \frac{2Na^2}{k_BT}\boldsymbol{f} = \frac{\langle \boldsymbol{R}^2 \rangle_0}{3k_BT}\boldsymbol{f},$$

where I have used $\langle \boldsymbol{R}^2 \rangle_0 = 6Na^2$ as computed above.

An alternative approach is to consider the free energy $F = U - TS$. At equilibrium this is minimised at constant temperature T, so $dF = dU - TdS = 0$ for any small change $d\boldsymbol{R}$ in the end-to-end displacement. Using the expression for the entropy from part (b)(i):

$$dF = -\boldsymbol{f} \cdot d\boldsymbol{R} - T \times (-2\boldsymbol{R} \cdot d\boldsymbol{R})/4Na^2 = (-\boldsymbol{f} + k_BT\boldsymbol{R}/2Na^2) \cdot d\boldsymbol{R},$$

where I have used $d(\boldsymbol{R}^2) = 2\boldsymbol{R} \cdot d\boldsymbol{R}$. (You can see this explicitly by writing out $\boldsymbol{R}^2$ in terms of the coordinates, X, Y and Z and working out how it changes due to dX, dY and dZ.)
At thermal equilibrium $dF = 0$, so the equilibrium value of $\langle \boldsymbol{R} \rangle$ is given by $\boldsymbol{f} = (k_BT/2Na^2)\langle \boldsymbol{R} \rangle$. This agrees with the result derived above.

(iv) The rubber band consists of cross-linked polymer chains. Each strand of polymer obeys Hooke's law as above. Note that the stretching of the chain is proportional to $1/T$ and so *decreases* with increasing temperature. Therefore, the rubber band contracts when heated and the weight rises. This should be contrasted with a metal rod which expands on heating.

The origin of this tension in the rubber band comes from the term TdS in dF. In other words, it is completely *entropic*. When we stretch the polymer chain, the random walk becomes constrained compared to the case with no applied force. Although the stretching lowers the mechanical energy of the system, it is costly in entropy. From the expression for dF, we see that the entropic contribution to the free energy is more important at higher temperatures. Therefore, the polymer chain is less willing to pay the entropic cost of stretching and the amount of stretching for a given applied force is reduced.

Problem sheet issued: Tuesday 15th October 2013
Solutions available from: Wednesday 23rd October 2013

Statistical Mechanics Answer Sheet 2

Percolation in One Dimension

1. Scaling form of order parameter

   (a) If $s \leq L-2$, an s-cluster must be bounded by two empty sites; if $s = L-1$, there is only one empty site in the system; if $s = L$, all sites are occupied. Thus

   $$n(s,p) = \begin{cases} p^s(1-p)^2 & \text{for } s \leq L-2, \\ p^{L-1}(1-p) & \text{for } s = L-1, \\ p^L & \text{for } s = L. \end{cases}$$

   (b) A cluster with size $s = L$ is percolating and becomes infinite in size as $L \to \infty$. Such a cluster ought not to be regarded as finite. Therefore, $\sum_{s=1}^{L-1} sn(s,p)$ represents the probability that a site belongs to a *finite* cluster.

   (c) It is obvious that the probability of an arbitrary site belonging to the spanning (infinite) cluster is $P_\infty(p,L) = p^L$, but we can double check this by finding the probability that the site does *not* belong to a finite cluster:

   $$\begin{aligned} P_\infty(p,L) &= p - \sum_{s=1}^{L-1} sn(s,p) \\ &= p - (L-1)p^{L-1}(1-p) - \sum_{s=1}^{L-2} sp^s(1-p)^2 \\ &= p - (L-1)p^{L-1}(1-p) - (1-p)^2 \left(p\frac{d}{dp}\right)\left(\sum_{s=1}^{L-2} p^s\right) \\ &= p - (L-1)p^{L-1}(1-p) - (1-p)^2 \left(p\frac{d}{dp}\right)\left(\frac{p-p^{L-1}}{1-p}\right) \\ &= p - (L-1)p^{L-1}(1-p) - (1-p)^2 p \frac{(1-p)(1-(L-1)p^{L-2}) + (p-p^{L-1})}{(1-p)^2} \\ &= p^L \,. \end{aligned}$$

(d) (i) We are given that $\xi(p) = -1/\ln p$, from which it follows that $\ln p = -1/\xi$ and hence that $p = \exp(-1/\xi)$. Since $P_\infty(p, L) = p^L$, we obtain

$$P_\infty(\xi, L) = p^L = \left[\exp\left(-\frac{1}{\xi}\right)\right]^L = \exp\left(-\frac{L}{\xi}\right) .$$

(ii) The expression above is already in scaling form,

$$P_\infty(\xi, L) = \xi^{-\beta/\nu}\mathcal{P}(L/\xi) ,$$

with $\beta/\nu = 0$ and $\mathcal{P}(x) = \exp(-x)$.
When $x \ll 1$, $\mathcal{P}(x) \approx 1$. When $x \gg 1$, $\mathcal{P}(x)$ decays rapidly to zero.

2. Moments and moment ratio of the cluster number density

(a)
$$\begin{aligned} M_k(p) &= \sum_{s=1}^{\infty} s^k n(s,p) \\ &= \sum_{s=1}^{\infty} s^k (1-p)^2 p^s \\ &= (1-p)^2 \sum_{s=1}^{\infty} s^k p^s \\ &= (1-p)^2 \sum_{s=1}^{\infty} s^k \exp(s \ln p) \\ &= (1-p)^2 \sum_{s=1}^{\infty} s^k \exp(-s/s_\xi) \qquad (\text{where } s_\xi = -1/\ln p) \\ &= (1-p)^2 s_\xi^k \sum_{s=1}^{\infty} (s/s_\xi)^k \exp(-s/s_\xi) . \end{aligned}$$

Since $s_\xi \to \infty$ as $p \to p_c^- = 1^-$, the values of s/s_ξ hardly change from term to term in the summation. Hence we can approximate the summation by an integral:

$$\begin{aligned} M_k(p) &= (1-p)^2 s_\xi^k \sum_{s=1}^{\infty} (s/s_\xi)^k \exp(-s/s_\xi)\, \Delta s \qquad (\text{where } \Delta s = 1) \\ &\approx (1-p)^2 s_\xi^k \int_1^{\infty} (s/s_\xi)^k \exp(-s/s_\xi) ds \\ &= (1-p)^2 s_\xi^{k+1} \int_{1/s_\xi}^{\infty} u^k \exp(-u) du \qquad (\text{where } u = s/s_\xi) . \end{aligned}$$

This is the required result.

(b) Re-expressing the result derived in (a) in terms of p gives

$$M_k(p) = (1-p)^2 \left(\frac{-1}{\ln p}\right)^{k+1} \int_{-\ln p}^{\infty} u^k \exp(-u) du.$$

As $p \to p_c^- = 1^-$, the lower limit tends to zero and the integral is easily evaluated by parts (or looked up in a table — it is the definition of the Gamma function). The result is

$$\begin{aligned} M_k(p) &= (1-p)^2 \left(\frac{-1}{\ln p}\right)^{k+1} k! \\ &\approx k!(1-p)^{1-k} \quad \text{as} \quad p \to p_c^- = 1^- , \end{aligned}$$

where I have used the approximation $\ln p = \ln[1-(1-p)] \approx -(1-p)$ valid as $p \to p_c = 1^-$. This is in the form required, so we identify

$$\Gamma_k = k! \quad \text{and} \quad \gamma_k = k - 1 .$$

(c) The result above tells us that $M_1 \to 1!(1-p)^0 = 1$ as $p \to p_c^- = 1^-$, which is consistent with the equation $\sum_{s=1}^{\infty} sn(s,p) = p$ discussed in the course notes. Hence $\Gamma_1 = 1$ and $\gamma_1 = 0$. The moment ratio becomes

$$\begin{aligned} g_k &= \frac{M_k M_1^{k-2}}{M_2^{k-1}} \qquad (k \geq 2) \\ &= \frac{\Gamma_k (1-p)^{1-k} \Gamma_1^{k-2} \left[(1-p)^0\right]^{k-2}}{\Gamma_2^{k-1} \left[(1-p)^{-1}\right]^{k-1}} \\ &= \frac{\Gamma_k \Gamma_1^{k-2}}{\Gamma_2^{k-1}} \\ &= \frac{k!}{(2!)^{k-1}} \quad \text{as} \quad p \to p_c^- = 1^- . \end{aligned}$$

3. Site-bond percolation in $d = 1$ **(Rapid Feedback question)**

(a) If there is a percolating (infinite) cluster, no sites or bonds can be empty. Hence $(p_c, q_c) = (1, 1)$.

(b) An s-cluster has s consecutive sites occupied, each with probability p, and $s-1$ consecutive bonds occupied, each with probability q. At each end of the s-cluster, either the site or the bond or both must be empty. Since pq is the probability for both the site *and* the bond at one end to be occupied, $(1-pq)^2$ is the probability that a cluster does not continue at either end. Therefore

$$n(s,p,q) = p^s q^{s-1} (1-pq)^2 .$$

(c) First let us confirm that $\sum_{s=1}^{\infty} sn(s,p,q) = p$:

$$\begin{aligned}
\sum_{s=1}^{\infty} sn(s,p,q) &= \sum_{s=1}^{\infty} sp^s q^{s-1}(1-pq)^2 \\
&= \frac{(1-pq)^2}{q} \sum_{s=1}^{\infty} s(pq)^s \\
&= \frac{(1-pq)^2}{q} \left(pq \frac{d}{d(pq)} \right) \left(\sum_{s=1}^{\infty} (pq)^s \right) \\
&= p(1-pq)^2 \frac{d}{d(pq)} \left(\frac{pq}{1-pq} \right) \\
&= p \,.
\end{aligned}$$

Similarly

$$\begin{aligned}
\sum_{s=1}^{\infty} s^2 n(s,p,q) &= \sum_{s=1}^{\infty} s^2 p^2 q^{s-1}(1-pq)^2 \\
&= \frac{1}{q}(1-pq)^2 \sum_{s=1}^{\infty} s^2 (pq)^s \\
&= \frac{1}{q}(1-pq)^2 \left(pq \frac{d}{d(pq)} \right)^2 \sum_{s=1}^{\infty} (pq)^s \\
&= \frac{1}{q}(1-pq)^2 \left(pq \frac{d}{d(pq)} \right)^2 \left(\frac{pq}{1-pq} \right) \\
&= p \frac{1+pq}{1-pq} \,.
\end{aligned}$$

Hence

$$\chi(p,q) = \frac{\sum_{s=1}^{\infty} s^2 n(s,p,q)}{\sum_{s=1}^{\infty} sn(s,p,q)} = \frac{1+pq}{1-pq}.$$

This result is identical to that for site percolation if we identify the occupation probability with pq. That is, the results for the two different models are equivalent if we if we consider an adjacent site-bond pair in the site-bond percolation problem as equivalent to a single site in the site percolation problem.

Problem sheet issued: Tuesday 22nd October 2013
Solutions available from: Wednesday 30th October 2013

Statistical Mechanics Answer Sheet 3

Site Percolation on a Bethe Lattice

1. Correlation function

   (a) Site i is already occupied by the definition of the correlation function. All the sites from i to j must also be occupied. There are l_{ij} such sites, excluding i but including j. Each site is occupied independently with probability p and so the joint probability that all these l_{ij} sites are occupied is $p^{l_{ij}}$.

   (b) The mean cluster size can be calculated by adding up the average number of occupied sites belonging to the same cluster as site i. (This uses the fact that the mean of a sum of independent random variables is the sum of the means of each variable.) But the average contribution from site j is the probability $g(i,j)$ that this site belongs to the same cluster as site i. So, the sum of the contributions from each site j on the lattice is the sum of $g(i,j)$ over all j. Notice that all sites are equivalent on the Bethe lattice, so the answer does not depend on the starting site i.

   (c) Using the expression for $g(i,j)$ gives the equation $\chi(p) = \sum_j p^{l_{ij}}$. To evaluate this summation, group together the sites at every distance $l = 0, 1, 2, 3, \ldots$ from site i. At $l = 0$ is the site i itself. We know that this is already occupied, so it contributes $g(0,0) = 1$ to the mean cluster size. There are z neighbours at distance $l = 1$, contributing $z \times p^1$. There are $z(z-1)$ neighbours at distance $l = 2$, contributing $z(z-1) \times p^2$. In general, for $l > 0$, there are $z(z-1)^{l-1}$ sites at distance l, contributing $z(z-1)^{l-1}p^l$. Therefore,

   $$\chi(p) = 1 + \sum_{l=1}^{\infty} z(z-1)^{l-1}p^l = 1 + zp\sum_{l=1}^{\infty}[p(z-1)]^{l-1} = 1 + zp\sum_{l=0}^{\infty}[p(z-1)]^l.$$

   The geometric series on the right-hand side of this equation converges provided $p(z-1) < 1$. We can then use the well-known formula for the sum of an infinite geometric series, $\sum_{l=0}^{\infty} x^l = (1-x)^{-1}$ for $|x| < 1$, to obtain

   $$\chi(p) = 1 + \frac{zp}{1-p(z-1)} = \frac{1+p}{1-p(z-1)}.$$

(d) $\chi(p)$ diverges as $p(z-1) \to 1^-$ (which is why the geometric series is restricted to $p(z-1) < 1$). Therefore, an infinite cluster stretching to infinity from site i can be found for $p > p_c = 1/(z-1)$.

2. Order parameter

(a) Choose an occupied site as the root of the Bethe lattice and trace out the cluster to which it belongs. Assume this cluster survives at least until level l, where it has t occupied "perimeter" or "surface" sites. Since every perimeter site has $z-1$ neighbours at level $l+1$, the expected number of occupied sites at level $l+1$ is $t(z-1)p$. If $t(z-1)p > t$, the expected number of occupied sites on the perimeter grows as l increases. Any cluster that survives until some high level l with a lot of occupied perimeter sites (and some clusters will always manage this, purely by chance) is then very likely to grow in strength as l increases and thus to percolate to infinity. The percolation threshold is given by $t(z-1)p = t$ or, equivalently, $p = 1/(z-1)$.

(b) For site A to be part of the infinite percolating cluster it must be occupied (which occurs with probability p) and at least one of the z branches emerging from it must reach to infinity. Since Q^z is the probability that *none* of the z branches is connected to infinity, the probability that at least one is connected to infinity is $1 - Q^z$. Thus

$$P_\infty(p) = p\left[1 - (Q(p))^z\right].$$

(c) We need each branch to be statistically independent. This means that each branch must contain a distinct set of sites from the other branches. This is only possible if two branches never meet again, *i.e.*, if there are no loops on the lattice. This is not the case for a square lattice.

(d) Consider a given occupied site, which we take to be the root of the Bethe lattice ($l = 0$), and a given neighbour at level $l = 1$. By definition, the probability that the neighbour is *not* the first site on a branch that links the $l = 0$ site to infinity is Q. There are two ways that the branch starting on a specific $l = 1$ neighbour can fail to link to infinity. Either the $l = 1$ site itself is unoccupied (probability $(1-p)$), or the $l = 1$ site is occupied (probability p) but *all* of its $(z-1)$ neighbours at level $l = 2$ fail to connect to infinity (probability Q^{z-1}). Hence

$$Q = (1-p) + pQ^{z-1}.$$

Rearranging gives $1 - Q = p(1 - Q^{z-1})$.

(e) $Q = 1$ is obviously a solution. Other solutions should satisfy $(1-Q^{z-1})/(1-Q) = 1/p$. As for any polynomial equation, we should try to factorize the polynomials. Since $Q = 1$ is a solution, we expect that the equation contains

a factor of $1-Q$. Recall that: $1-x^m = (1-x)(1+x+x^2+\ldots+x^{m-1})$. [Put another way, $(1-x^m)/(1-x)$ is the formula for the sum of a finite geometric series with first term 1 and ratio x.] Putting $m = z-1$ gives the necessary factorization in our case.

(f) For a continuous phase transition, the order parameter goes continuously to zero as the critical point is approached. So, we expect $P_\infty \ll 1$, and therefore $Q \simeq 1$, as p approaches the percolation threshold from above.

(g) We need to solve the equation $1 + Q + \ldots + Q^{z-1} = 1/p$. Since $Q \to 1^-$ as $p \to p_c$, we can write Q as $Q = 1 - \delta Q$ and perform a Taylor expansion around $Q = 1$, assuming that $\delta Q \ll 1$. Therefore,

$$\begin{aligned}
\frac{1}{p} &= \sum_{m=0}^{z-2} Q^m = \sum_{m=0}^{z-2} (1-\delta Q)^m \\
&\simeq \sum_{m=0}^{z-2} (1 - m\,\delta Q) \qquad \text{using the Taylor/binomial expansion} \\
&= \left(\sum_{m=0}^{z-2} 1\right) - \delta Q \left(\sum_{m=1}^{z-2} m\right) = (z-1) - \frac{1}{2}(z-1)(z-2)\delta Q.
\end{aligned}$$

To reach the last expression, I have used the arithmetic sum $1+2+\ldots+n = n(n+1)/2$ with $n = z-2$.
We now have a linear equation in δQ which can be solved easily to give

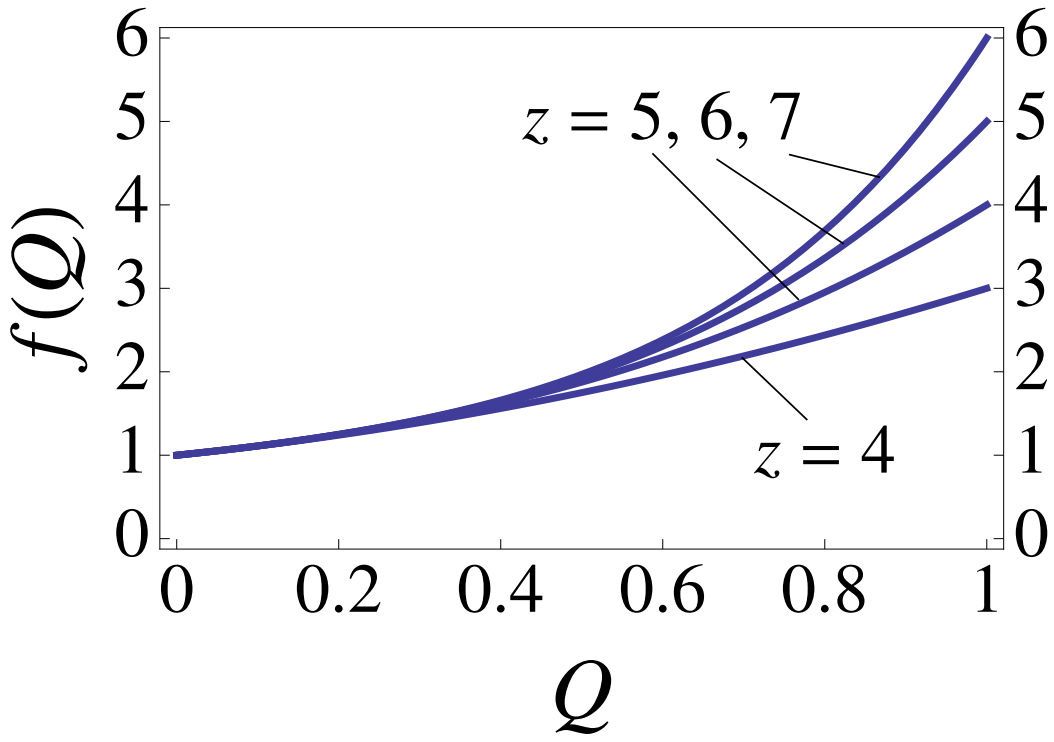
$$\delta Q = \frac{2}{z-2}\left(1 - \frac{1}{p(z-1)}\right) = \frac{2}{z-2}\left(1 - \frac{p_c}{p}\right).$$

We can now obtain P_∞ from: $P_\infty = p(1-Q^z)$. This gives

$$P_\infty = p\,[1-(1-\delta Q)^z] \simeq p\,[1-(1-z\delta Q)] = pz\,\delta Q = \frac{2z}{z-2}\,(p-p_c).$$

(h) The $z-2$ roots are solutions of the equation $f(Q) = 1+Q+\ldots+Q^{z-2} = 1/p$. Examine the function $f(Q)$ in the domain $0 \le Q \le 1$. Since all the coefficients in the polynomial $f(Q)$ are positive (all $+1$), this function increases monotonically (because $df/dQ > 0$ for $Q > 0$) for all positive Q, starting at $f(0) = 1$ and increasing indefinitely to infinity as $Q \to \infty$. See the figure overleaf. Therefore, $f(Q)$ will be equal to the value of $1/p$ (which is greater than 1) *once and once only* for some $Q > 0$.

Note that $f(Q)$ will reach $1/p$ in the range $0 \le Q < 1$ only if $p > 1/(z-1)$. Otherwise, for $p \le 1/(z-1)$, we must choose the other solution: $Q = 1$ corresponding to no percolation. In other words, this tells us that the percolation threshold is $p_c = 1/(z-1)$.



The polynomial $f(Q)$ for coordination numbers $z = 4, 5, 6, 7$ in the physically meaningful range $0 \leq Q \leq 1$. Since $1 \leq f(Q) \leq z - 1$, the equation $f(Q) = 1/p$ only has solutions for $p \geq 1/(z-1)$.

Problem sheet issued: Tuesday 29th October 2013
Solutions available from: Wednesday 6th November 2013

Statistical Mechanics Answer Sheet 4

Questions from 2011 Exam

1. (a) (i) As $p \to p_c$, the mean cluster size $\chi(p)$ is expected to diverge. Hence $p_c = 1/(z-1)$.

   (ii) Since $p_c = 1/(z-1)$,

   $$\chi(p) = \frac{1+p}{1-(z-1)p} = \frac{p_c(1+p)}{p_c - p}.$$

   Hence $\chi(p) \sim (p_c - p)^{-\gamma}$ with $\gamma = 1$.

   (b) (i) Start by noting that

   $$N(l,p) = \frac{z}{z-1}[(z-1)p]^l = \frac{z}{z-1}e^{l\ln[(z-1)p]}.$$

   Since $(z-1)p < 1$, the argument of the logarithm is less than one and the logarithm is negative. Hence

   $$N(l,p) = \frac{z}{z-1}e^{-l/l_\xi(p)} \sim e^{-l/l_\xi(p)} \quad \text{where} \quad l_\xi(p) = \frac{1}{|\ln[(z-1)p]|}.$$

   (ii) As $p \to p_c^-$, $\xi(p)$ is expected to diverge. From the formula derived in (b)(i) we see that $l_\xi(p)$ diverges when $(z-1)p$ tends to 1 and the logarithm in the denominator tends to zero. This tells us that $p_c = 1/(z-1)$ in agreement with the result of part (a)(i).

   (iii) Setting $p = p_c(1-\delta) = (1-\delta)/(z-1)$ gives

   $$l_\xi(p) = \frac{1}{|\ln[(z-1)p]|} = \frac{1}{|\ln[1-\delta]|}$$

   In the limit as $p \to p_c^-$, $\delta \to 0^+$, so it makes sense to expand the logarithm as a power series in the small quantity δ:

   $$\ln(1-\delta) = -\delta - \delta^2/2 - \delta^3/3 - \ldots.$$

Substituting this expansion into the formula for l_ξ gives

$$l_\xi(p) = \frac{1}{\delta + \delta^2/2 + \ldots} \sim \frac{1}{\delta} = \frac{p_c}{p_c - p}.$$

Thus $l_\xi(p) \sim 1/(p_c - p)^\nu$ with $\nu = 1$ as $p \to p_c^-$.

(c) (i) χ is the expected number of occupied sites in the same cluster as a randomly chosen occupied site, which we can always take as the $l = 0$ site at the root of the Bethe lattice. Hence

$$\begin{aligned}\chi(p) &= 1 \text{ (for the site at } l = 0) \\ &\quad + \text{ expected numbers of sites in the same cluster at } l = 1, 2, \ldots. \\ &= 1 + \sum_{l=1}^{\infty} N(l, p).\end{aligned}$$

(ii) Use the formula

$$\chi(p) = 1 + \sum_{l=1}^{\infty} N(l, p)$$

with $N(l, p) = \frac{z}{z-1}[(z-1)p]^l$. This gives

$$\begin{aligned}\chi(p) &= 1 + \frac{z}{z-1} \sum_{l=1}^{\infty} [(z-1)p]^l \\ &= 1 + \frac{z}{z-1} \times \frac{(z-1)p}{1-(z-1)p} \quad \text{(sum of infinite geometric series)} \\ &= 1 + \frac{zp}{1-(z-1)p} \\ &= \frac{1-(z-1)p+zp}{1-(z-1)p} \\ &= \frac{1+p}{1-(z-1)p} \qquad \text{as required.}\end{aligned}$$

2. (a) (i) s_ξ diverges at $p = p_c$. Since $\mathcal{G}(0)$ is a non-zero constant, we see that $n(s, p_c) \sim s^{-\tau}$. (For the sketch see Fig. 1 below.) This means that the system is **scale invariant** or **fractal** at p_c.

(ii) The quantity $s_\xi(p)$ is the **characteristic cluster size**. It is exponentially unlikely to find clusters much larger than s_ξ.

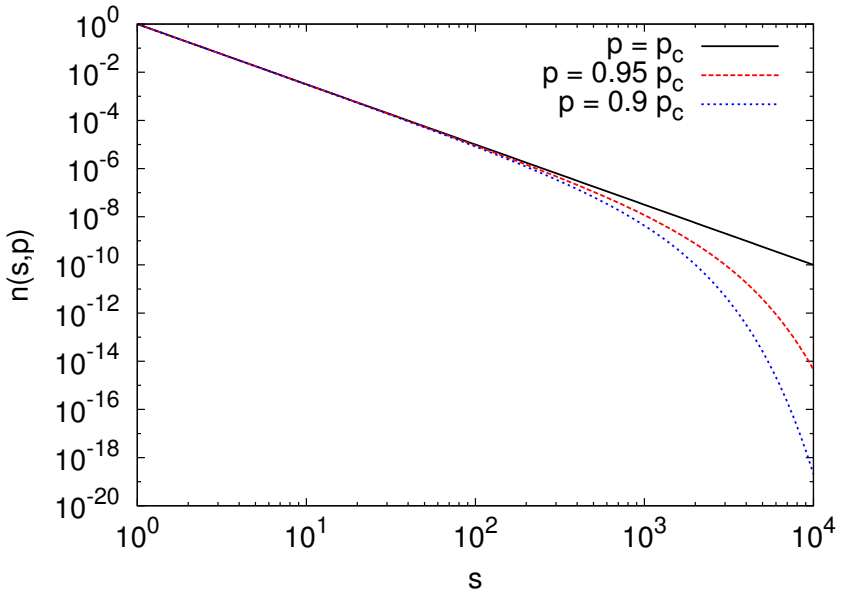


Figure 1: Log-log plot of $n(s, p)$ as a function of s for several values of p.

(b) $n(s, p)$ is by definition the density of clusters containing s sites.
Hence $sn(s, p)$ is the density of sites belonging to clusters containing s sites.
Hence $\sum_{s=1}^{\infty} sn(s, p)$ is the density of sites belonging to clusters of any size from $s = 1$ to $s = \infty$.
Since all sites belong to clusters of one of these sizes, we deduce that $\sum_{s=1}^{\infty} sn(s, p)$ is the total density of occupied sites, otherwise known as p:

$$\sum_{s=1}^{\infty} sn(s, p) = p.$$

(c) (i)
$$\begin{aligned}
\chi(p) &\sim \sum_{s=1}^{\infty} s^2 n(s, p) \\
&\sim \sum_{s=1}^{\infty} s^{2-\tau} \mathcal{G}(s/s_\xi) \qquad \text{(using scaling form valid near transition)} \\
&= s_\xi^{2-\tau} \sum_{s=1}^{\infty} \left(\frac{s}{s_\xi}\right)^{2-\tau} \mathcal{G}(s/s_\xi)\, \Delta s
\end{aligned}$$

where, by definition, $\Delta s \equiv 1$. In the limit as $p \to p_c^-$, $s_\xi \to \infty$ and the values of the summand change less and less as s changes by ± 1. We can therefore approximate the sum as an integral:

$$\begin{aligned}\chi(p) &\sim s_\xi^{2-\tau} \int_1^\infty \left(\frac{s}{s_\xi}\right)^{2-\tau} \mathcal{G}(s/s_\xi)\, ds \\ &= s_\xi^{3-\tau} \int_{1/s_\xi}^\infty u^{2-\tau} \mathcal{G}(u)\, du, \qquad \text{where } u = s/s_\xi.\end{aligned}$$

(ii) We know from the numerics that $\tau < 3$ and that $\mathcal{G}(0)$ is non-zero. This means that the integrand remains finite or diverges more slowly than $1/u$ as $u \to 0$. The integral therefore converges at the lower limit as $1/s_\xi \to 0$, allowing us to replace the lower limit by zero without affecting the result to leading order in $1/s_\xi$. The integral is therefore just some constant. This leads to the required result: $\chi(p) \sim s_\xi^{3-\tau}$ as $p \to p_c^-$.

(iii) Since $s_\xi \sim |p - p_c|^{-1/\sigma}$ as $p \to p_c^-$, we obtain

$$\chi(p) \sim s_\xi^{3-\tau} \sim 1/|p - p_c|^{(3-\tau)/\sigma} \quad \text{as} \quad p \to p_c^-.$$

Comparing this equation with the definition of γ,

$$\chi(p) \sim \frac{1}{|p - p_c|^\gamma} \quad \text{as} \quad p \to p_c^-,$$

we deduce that

$$\gamma = \frac{3 - \tau}{\sigma}$$

as required.

Since $\tau = 2.19$ and $\sigma = 0.45$, we decuce that $\gamma = 1.80$.

(d) When $p = p_c$, equation (2) from the question becomes

$$\sum_{s=1}^\infty s n(s, p_c) = p_c$$

Using the scaling form of $n(s, p_c)$ and the same technique as above, we deduce that

$$p_c \sim \int_1^\infty s^{1-\tau} \mathcal{G}(s/s_\xi) ds.$$

Since p_c is finite the integral must converge as $s_\xi \to \infty$. Noting that $\mathcal{G}(s/s_\xi) \to \mathcal{G}(0)$ as $s_\xi \to \infty$, we obtain

$$p_c \sim \mathcal{G}(0) \int_1^\infty s^{1-\tau} ds \quad \text{as} \quad s_\xi \to \infty.$$

This integral diverges at the upper limit if $\tau \leq 2$, so we deduce that $\tau > 2$.

Problem sheet issued: Tuesday 5th November 2013
Solutions available from: Wednesday 13th November 2013

Statistical Mechanics Answer Sheet 5

Finite-Size Scaling

1. Percolation probability

   (a) The probability that there is a percolating cluster in one dimension is the probability that all L sites are occupied:

   $$\Pi(p, L) = p^L = (e^{\ln p})^L = e^{L \ln p}\,.$$

   $\ln p < 0$ because $0 < p < 1$. Therefore, $\Pi(p, L)$ is a decaying exponential in the system size. We can write $\Pi(p, L) = g_{1\mathrm{D}}(L/\xi)$ with

   $$g_{1\mathrm{D}}(x) = e^{-x} \quad \text{and} \quad \xi(p) = \frac{1}{|\ln p|}\,.$$

   (b) (i) Since $\xi \sim |p - p_c|^{-\nu}$, we can substitute for $\xi(p)$ to obtain

   $$\Pi(p, L) = g_{\pm}(L|p - p_c|^{\nu}) = G((p - p_c)L^{1/\nu}),$$

   where

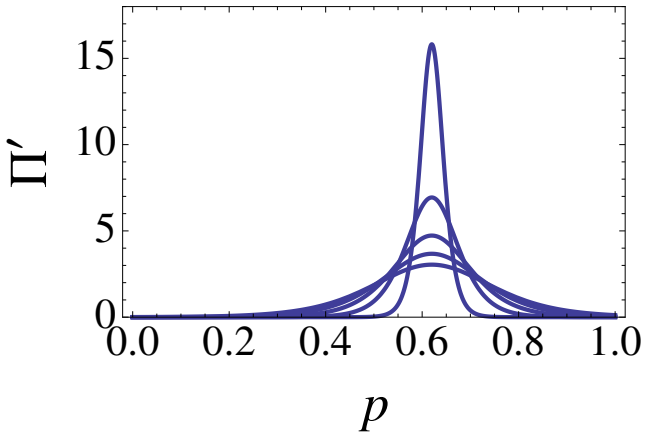
   $$G(u) = \begin{cases} G(u) = g_+(u^\nu), & u > 0, \\ G(u) = g_-(|u|^\nu), & u < 0. \end{cases}$$

   The function G and its derivative dG/du are continuous at $u = 0$ if Π and $d\Pi/dp$ are continuous at $p = p_c$.

   (ii) $d\Pi/dp = L^{1/\nu} G'((p - p_c)L^{1/\nu})$ has a sharp peak. The position of the peak is at $p = p_c$ and the peak height is $L^{1/\nu}G'(0)$. So, this height diverges as $L \to \infty$.
   We know that Π increases from 0 to 1 across the percolation threshold. So, the area under $d\Pi/dp$ in a plot against p remains constant at unity. So, as the peak height increases with system size, the peak must become narrower.

   The width of $d\Pi/dp$ can be determined by a criterion such as the point where $G'(u)$ reaches a fixed value, $G'(u) = 0.5G'(0)$, say. This gives two

values for $|p-p_c|L^{1/\nu} = C_\pm$ for either side of the transition. The width can be defined as the difference between these values: $\delta p = (C_+ - C_-)/L^{1/\nu} \sim L^{-1/\nu}$.
Alternatively, we can use as a measure of the width the inverse of the maximum slope $(d\Pi/dp)^{-1} = 1/[L^{1/\nu}G(0)]$.



The peak in Π' narrows and increases in height as the system size increases. The area under the peak remains constant at unity.

2. Mean cluster size

   (a) The scaling hypothesis postulates that the cluster length scale ξ is the only characteristic length scale in the problem (at lengths much greater than the lattice spacing). Finite-size effects therefore depends on the relative size of L compared to this sole length scale ξ. Therefore, we expect the function f to be a function of the ratio of ξ to L: $\chi(p, L) = \chi(p, \infty) f(\xi/L)$.
   Now, we know that $\chi(p, \infty) \sim |p-p_c|^{-\gamma}$ for the infinite system. In terms of the cluster length scale $\xi \sim |p - p_c|^{-\nu}$, we can write $\chi(p, \infty) \sim [\xi(p)]^{+\gamma/\nu}$. Therefore, we can express χ solely in terms of ξ and L without explicit dependence on p:

$$\chi(p, L) = \chi(p, \infty) f(\xi/L) \sim \xi^{\gamma/\nu} f(\xi/L)\,.$$

   (b) When $L \gg \xi$, the system is larger than any characteristic length scale in the system and L can be treated as effectively infinite. Therefore, $f(x \ll 1) \simeq 1$.
   Conversely if $L \ll \xi$, then the system does not see the characteristic length at all and $\chi(p, L \ll \xi)$ should become independent of ξ. This means that the small-x behaviour of $f(x)$ must be of a form that would cancel out the dependence on ξ: $f(x \gg 1) \sim x^{-\gamma/\nu}$, so that

$$\chi(p, L \ll \xi) \sim \xi^{+\gamma/\nu} f(\xi/L \gg 1) \sim \xi^{+\gamma/\nu} (\xi/L)^{-\gamma/\nu} \sim L^{\gamma/\nu}$$

   has no dependence on ξ.

(c) $\chi(p_c, L)$ belongs to the regime where $L \ll \xi$. The previous part tells us that $\chi(p_c, L) \sim L^{\gamma/\nu}$.
In other words, the mean cluster size does not diverge as p crosses the threshold. It simply has a peak whose height is large for large systems. The divergent and singular behaviour of χ is only found in the limit of an infinitely large system.

(d) This finite-size scaling of $\chi(p, L \ll \xi)$ is best obeyed when ξ is infinite, *i.e.*, at p_c. This allows us to extract the exponent γ/ν from $\chi(p_c, L) \sim L^{\gamma/\nu}$ using the slope of the plot of $\ln \chi(p_c, L)$ against $\ln L$.
We can also measure γ by itself by using data in the range $L \gg \xi(p)$. In other words, we need data sufficiently far away from the critical point that the cluster length scale $\xi(p)$ has not yet reached the system size L. Then, the typical clusters are much smaller than the system and are not cut off by the edges of the system. The value of the system size L should not make any difference to the mean cluster size. Therefore, $\chi(p, L \gg \xi) \simeq \chi(p, \infty) \sim |p - p_c|^{-\gamma}$ so that a log-log plot would give γ as the slope.
This gives us to a value for γ and we can deduce the value of ν using the measured value of γ/ν.

Scaling Relations

3. Cluster size distribution and mean cluster size

(a) (i) Use the same logic as the previous question for $\chi(p, L)$:

$$s_\xi(p, L) = s_\xi(p, \infty) g(\xi/L) \sim |p - p_c|^{-1/\sigma} g(\xi/L)\,.$$

with $g(0) = 1$ to match the $L \to \infty$ form.
As before, write the whole expression in terms of $\xi(p)$, *i.e.*, everything that is singular depends on p through $\xi(p) \sim |p - p_c|^{-\nu}$:

$$s_\xi(p, L) \sim \xi^D g(\xi/L) \quad \text{with} \quad D = 1/\sigma\nu\,.$$

The system should not see the scale ξ when its linear size L is much smaller than ξ. So, we need $s_\xi(p, L \ll \xi)$ to be independent of ξ and so $g(x)$ must be $\sim 1/x^D$ for large x. This gives:

$$s_\xi(p, L \ll \xi(p)) \sim \xi^D (L/\xi)^D \sim L^D = L^{1/\sigma\nu}\,.$$

(ii) At the threshold $p = p_c$, the cluster length scale $\xi(p)$ is infinite. So, our previous results says that $s_\xi(p_c, L) \sim L^D = L^{1/\sigma\nu}$ for *all* L. The scaling hypothesis says that $n(s, p, L)$ should be only controlled by this single

length scale. So, we can just substitute $s_\xi(p, L)$ for s_ξ into the form for $n(s, p, L \to \infty)$. Setting $p = p_c$ gives:

$$n(s, p_c, L) \sim s^{-\tau}\mathcal{G}(s/s_\xi(p_c, L)) \sim s^{-\tau}\mathcal{G}(s/L^D)$$

with $D = 1/\sigma\nu$.

(b) (i) The mean cluster size $\chi(p)$ is the expectation value of the cluster size to which an occupied site belongs. The density of clusters of size s is $n(s, p)$, so the density of occupied sites involved in clusters of size s is $sn(s, p)$. The normalised probability for finding a site belonging to cluster of size s is therefore $P(s) = sn(s, p)/\sum_s sn(s, p)$. To obtain the average, we need to sum over the cluster sizes s weighted by this probability distribution. Therefore,

$$\chi(p) = \sum_{s=1}^{\infty} sP(s) = \sum_{s=1}^{\infty} s \times sn(s, p) \Bigg/ \sum_{s=1}^{\infty} sn(s, p)\,.$$

But the total density occupied sites is p for a site occupation probability of p, *i.e.*, the denominator must be p.

$$\chi(p) = \frac{1}{p}\sum_{s=1}^{\infty} s^2 n(s, p)\,.$$

(ii) We know the finite-size behaviour for χ and n at criticality. We require that both sides of the relationship between n and χ scale in the same way with L:

$$\begin{aligned} L^{\gamma/\nu} &\sim \frac{1}{p_c}\sum_{s=1}^{\infty} s^2 \times s^{-\tau}\mathcal{G}(s/L^D) \\ &\sim \int_1^{\infty} ds\, s^{2-\tau}\mathcal{G}(s/L^D) \sim \int_{1/L^D}^{\infty} L^D dx\, (xL^D)^{2-\tau}\mathcal{G}(x) \\ &\sim L^{D(3-\tau)}\int_{1/L^D}^{\infty} dx\, x^{2-\tau}\mathcal{G}(x) \end{aligned}$$

with $D = 1/\sigma\nu$. Now, if $\mathcal{G}(0)$ is non-zero and $3 - \tau > 0$, the lower limit of the integral does not diverge if we take $1/L^D$ to zero. The error we make by moving the lower limit to zero will be small: of order $\mathcal{G}(0)\int_0^{1/L^D} x^{2-\tau}dx \sim (1/L^D)^{3-\tau}$, which vanishes as $L \to \infty$. Therefore,

$$L^{\gamma/\nu} \sim L^{D(3-\tau)}\int_0^{\infty} dx\, x^{2-\tau}\mathcal{G}(x)\,.$$

Assuming that $\mathcal{G}(x)$ decays exponentially at large x, the integral converges and is just a numerical factor with no dependence on physical parameters. We therefore deduce that $\gamma/\nu = D(3-\tau)$. Since $D = 1/\sigma\nu$, we arrive at:

$$\gamma = (3-\tau)/\sigma\,.$$

Problem sheet 6 issued: Tuesday 12th November 2013
Solutions available from: Wednesday 20th November 2013

Statistical Mechanics Answer Sheet 6

Real space renormalisation group transformation

1. *Site percolation on a one-dimensional lattice (Exam 2010).*

   (a) (1) Zoom out to scale b, that is, divide the lattice into blocks of linear scale b;
   (2) Perform coarse-graining to find new effective parameters $p \to p' = R_b(p)$ corresponding to occupation of the block site;
   (3) Change unit of length to new scale so that all lengths ℓ become ℓ/b.

   (b) Divide the chain into blocks of b sites. A block is part of a percolating section if all the sites are occupied. This has probability p^b. Therefore, the renormalised system should have occupation probability $R_b(p) = p^b$. Fixed points satisfy the equation $p^\star = R_b(p^\star) \Rightarrow p^\star = 0, 1$. $p^b < p$ for all $0 < p < 1$ for the power $b > 1$. Therefore, if we start off with any $p < 1$, we flow to $p^\star = 0$. In other words, $p^\star = 0$ is stable and $p^\star = 1$ is unstable.

   (c) As with all physical length scales, $\xi \to \xi/b$ under RSRG. Therefore, $\xi(p' = p^b) = \xi(p)/b \Rightarrow \left[\xi(p^b)\right]^{-1} = b\left[\xi(p)\right]^{-1}$. Alternatively, in one-dimensional percolation, $\xi(p) = -1/\ln p$. Hence, $\xi(p^b) = -1/\ln(p^b) = -1/(b \ln p) = \xi(p)/b$.

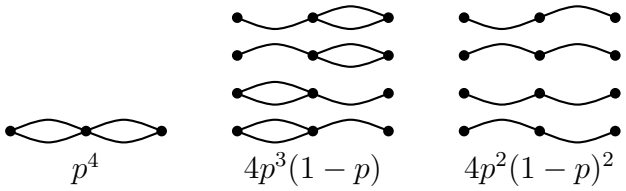
   (d) We identify the percolation threshold with the unstable fixed point: $p_c = 1$. $dR_b/dp = bp^{b-1} = b$ at $p = 1$ and so the formula gives $\nu = \ln b/\ln b = 1$ so that $\xi(p)$ diverges as $1/(p_c - p) = 1/(1 - p)$ as $p \to 1^-$.

   (e) To convert the given expression to a useful form for this problem, write $p = e^x$. Let $F(x) = F(\ln p) = f(p)$. Then, the condition becomes $F(bx) = bF(x) \Rightarrow F(b \ln p) = bF(\ln p) \Rightarrow f(p^b) = bf(p)$. Since the condition implies that $F(x) = \alpha x$, we have $f(p) = \alpha \ln p$. We can identify $f(p) = 1/\xi(p)$ so that $1/\xi(p) = \alpha \ln p$ as required.

   (f) Writing $p = 1 - \delta$, we see that $1/\ln p = 1/\ln(1 - \delta) \to -1/\delta = 1/(1 - p)$ for small δ. Therefore, both results for ξ agree as long as p is close to $p_c = 1$.

2. *Bond percolation on a square lattice in two dimensions.* (**RF Question)**

   (a) There are nine configurations that have a connected path from **A** to **B**:



Adding the probabilities for these configurations, we find

$$\begin{aligned} R_b(p) &= p^4 + 4p^3(1-p) + 4p^2(1-p)^2 \\ &= p^4 - 4p^3 + 4p^2. \end{aligned} \tag{1}$$

(b) (i) The fixed point Equation $R_b(p) = p$ is solved graphically by plotting the graph of $R_b(p)$ versus p and locating the intersections with the line $R_b(p) = p$.

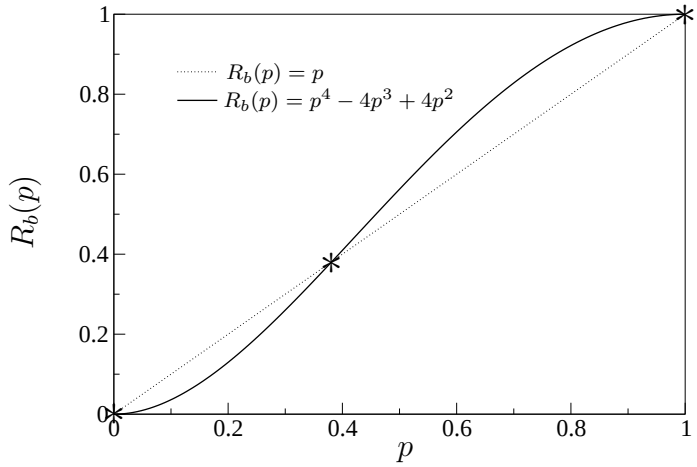


Figure 1: The fixed point Equation $R_b(p^\star) = p^\star$ are $p^\star = 0, 0.38, 1$.

tt inspection, we find the three fixed points

$$p^\star = \begin{cases} 0 & \text{trivial fixed point - empty lattice} \\ 0.38 & \text{non-trivial fixed point} \\ 1 & \text{trivial fixed point - fully occupied lattice.} \end{cases} \tag{2}$$

(ii) When performing the real-space renormalisation procedure, length scales are rescaled by the factor b.

If we start out with a finite correlation length, the rescaled correlation length $\xi' = \xi/b$ will decrease ($b > 1$) with an associated flow in p-space as indicated below. Starting out with $p < p^\star$, the flow will be toward $p^\star = 0$. If we started out with $p > p^\star$, the flow will be toward $p^\star = 1$.

(iii) The correlation length $\xi \to \xi/b$ only remains invariant if $\xi = 0$, associated with the trivial fixed points $p^\star = 0$ (empty lattice) or $p^\star = 1$ (fully

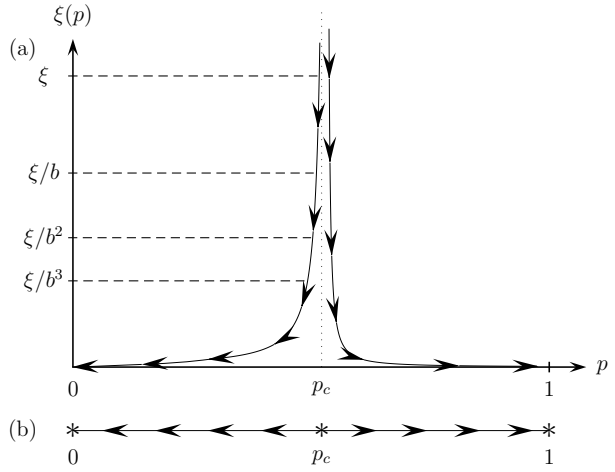


Figure 2: (a) A sketch of the correlation length as a function of occupation probability. The dotted line shows the position of p_c. (b) The corresponding flow in parameter space.

occupied lattice) or $\xi = \infty$, associated with the non-trivial fixed point $p^\star \approx 0.38$.

Since the correlation length is $\xi = 0$ or $\xi = \infty$ at the fixed point, there is no characteristic scale and scale invariance prevails.

(c) (i) Let A denote a constant. Then

$$\xi = A\,|p - p_c|^{-\nu} \tag{3a}$$

$$\xi' = A\,|R_b(p) - p_c|^{-\nu}. \tag{3b}$$

As $\xi' = \xi/b$ we find

$$|p - p_c|^{-\nu} = b|R_b(p) - p_c|^{-\nu} = b|R_b(p) - R(p_c)|^{-\nu}, \tag{4}$$

from which we find for $p \to p_c$

$$\nu = \frac{\log b}{\log\left(\frac{dR_b(p_c)}{dp}\right)}. \tag{5}$$

(ii) Now

$$\begin{aligned}\frac{dR_b}{dp}|_{p^\star} &= (4p^3 - 12p^2 + 8p)|_{p^\star=0.38} \\ &\approx 1.53\end{aligned} \tag{6}$$

and hence

$$\nu = \frac{\log 2}{\log 1.53} \approx 1.63. \tag{7}$$

The exact values in $d = 2$ are $p_c = 0.5$ and $\nu = 4/3$,

Ising Model.

3. *The entropy and the free energy of a system at equilibrium.* (**RF Question)**

   (a) According to the Boltzmann's distribution, the probability p_r to find an equilibrium system in a microstate r with energy E_r is given by

$$p_r = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} = \frac{1}{Z}\exp(-\beta E_r) \tag{8}$$

   where Z denotes the partition function and $\beta = 1/(k_B T)$. Therefore, the entropy

$$\begin{aligned} S &= -k_B \sum_r p_r \ln p_r \\ &= -k_B \sum_r \frac{1}{Z}\exp(-\beta E_r)\left[\ln\left(\exp(-\beta E_r)\right) - \ln Z\right] \\ &= k_B \frac{\ln Z}{Z}\sum_r \exp(-\beta E_r) - k_B \sum_r \frac{(-\beta E_r)\exp(-\beta E_r)}{Z} \\ &= k_B \ln Z + \frac{1}{T}\sum_r \frac{E_r \exp(-\beta E_r)}{Z} \\ &= k_B \ln Z + \frac{\langle E\rangle}{T}. \end{aligned} \tag{9}$$

   (b) From part (a) we find

$$\ln Z = \frac{1}{k_B}\left(S - \frac{\langle E\rangle}{T}\right), \tag{10}$$

   so the free energy

$$\begin{aligned} F &= -k_B T \ln Z \\ &= -T\left(S - \frac{\langle E\rangle}{T}\right) \\ &= \langle E\rangle - TS. \end{aligned} \tag{11}$$

4. *Fluctuation-dissipation theorem.*
   First we note that the average total energy

$$\langle E\rangle = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_H, \tag{12}$$

since

$$
\begin{aligned}
-\left(\frac{\partial \ln Z}{\partial \beta}\right)_H &= -\frac{1}{Z}\left(\frac{\partial Z}{\partial \beta}\right)_H \\
&= -\frac{1}{Z}\frac{\partial}{\partial \beta}\left(\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}})\right)_H \\
&= \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}}. \qquad (13)
\end{aligned}
$$

However, the *instantaneous* total energy will, of course, fluctuate around the average total energy. The magnitude of the fluctuations is determined by the standard deviation ΔE where

$$
(\Delta E)^2 = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 + \langle E \rangle^2 - 2E\langle E \rangle \rangle = \langle E^2 \rangle - \langle E \rangle^2. \qquad (14)
$$

Differentiating twice $\ln Z$ with respect to β we find

$$
(\Delta E)^2 = \left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_H \qquad (15)
$$

since

$$
\begin{aligned}
\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_H &= -\frac{\partial}{\partial \beta}\left(-\frac{\partial \ln Z}{\partial \beta}\right)_H \\
&= -\frac{\partial}{\partial \beta}\left(\frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}}\right)_H \\
&= \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}}^2 + \frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)_H \sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}} \\
&= \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}}^2 + \left(\frac{\partial \ln Z}{\partial \beta}\right)_H \frac{1}{Z}\sum_{\{s_i\}} \exp(-\beta E_{\{s_i\}}) E_{\{s_i\}} \\
&= \langle E^2 \rangle - \langle E \rangle^2. \qquad (16)
\end{aligned}
$$

However,

$$
\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_H = -\left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_H = -\left(\frac{\partial \langle E \rangle}{\partial T}\right)_H \frac{\partial T}{\partial \beta} = -C\frac{\partial (k_B \beta)^{-1}}{\partial \beta} = k_B T^2 C, \qquad (17)
$$

where C denotes the heat capacity at constant external parameter. Hence, with c denoting the heat capacity per spin, we have that

$$
k_B T^2 c = \frac{1}{N}\left(\langle E^2 \rangle - \langle E \rangle^2\right). \qquad (18)
$$