Diamagnetism

Substances with a negative magnetic susceptibility are called *diamagnetic*. The magnetic susceptibility per unit volume is defined as

$$\chi = M/H,$$

where M is the magnetic moment per unit volume, or the *magnetization*, and H is the magnetic field intensity.

Diamagnetism is associated with the tendency of electrical charges partially to shield the interior of a body from an applied magnetic field. In electromagnetism we are familiar with *Lenz's law*, which states that on attempting to change the flux through an electrical circuit an induced current is set up in such a direction as to oppose the flux change. In a resistanceless circuit, in a superconductor, or in an electron orbit within an atom, the induced current persists as long as the field is present, and the magnetic moment associated with the current is a diamagnetic moment.

DERIVATION OF THE LANGEVIN DIAMAGNETISM EQUATION

The usual derivation employs the Larmor theorem, which states that for an atom in a magnetic field the motion of the electrons is, to the first order in H, the same as a possible motion in the absence of H except for the superposition of a common precession of angular frequency

$$\omega_L = -eH/2mc,$$

or $f_L = 1.40$ mc/oersted. Furthermore, if the field is applied slowly, the motion in the rotating reference system will be the same as the original motion in the rest system before the application of the field. The precession of the electron distribution is equivalent to diamagnetic current

$$I = -(Ze)(eH/2mc)/2\pi c,$$

in electromagnetic units. As the magnetic moment μ of a current loop is given by the product of the current by the area of the loop, we have

$$\mu/H = -(Ze^2/4mc^2)\overline{\rho^2},$$

for Z electrons, where $\overline{\rho^2} = \overline{x^2} + \overline{y^2}$ is the average of the square of the perpendicular distance of the electron from the field axis. In terms of the mean square distance $\overline{r^2} = \overline{x^2} + \overline{y^2} + \overline{z^2}$ from the nucleus, we have

$$\overline{r^2} = \tfrac{3}{2}\overline{\rho^2}$$

for a distribution of charge which on the average is spherically symmetrical, so that $\overline{x^2} = \overline{y^2} = \overline{z^2}$. Then the diamagnetic susceptibility per unit volume is, if N is the number of atoms per unit volume,

$$\chi = -\frac{Ze^2N}{6mc^2}\overline{r^2},$$

which is the Langevin expression as corrected by Pauli.

The problem of calculating the diamagnetic susceptibility is thus reduced to the calculation of $\overline{r^2}$; this means that we must determine the electron charge distribution within the atom. The charge distribution can in principle be calculated by quantum mechanics, but exact solutions are available only for the hydrogen atom and isoelectronic ions. The quality of the approximate solutions which have been worked out deteriorates as the number of electrons increases. By and large, the best we can do is to use the charge distributions calculated by the "self-consistent field" method. An index of wave functions obtained by this method has been prepared by Hartree[1]; Stoner[2] utilized Hartree functions in early susceptibility calculations. Other approximate schemes have been devised by Slater,[3] Brindley,[4] Sommerfeld, and others. A comparison of experimental and theoretical results is shown in Tables 8.1 and 8.2.

QUANTUM THEORY OF DIAMAGNETISM OF MONONUCLEAR SYSTEMS

The magnetic vector potential $\mathbf{A}$ is defined by the relation $\mathbf{H} = \text{curl } \mathbf{A}$. In a magnetic field the generalized momentum $\mathbf{p}$ of a particle of charge e is

$$\mathbf{p} = \mathbf{p}_{\text{kin}} + \mathbf{p}_{\text{pot}} = m\dot{\mathbf{r}} + e\mathbf{A}/c,$$

so the kinetic energy is

$$T = \frac{1}{2}m\dot{r}^2 = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2$$

$$= \frac{1}{2m}p^2 - \frac{e}{mc}\mathbf{p}\cdot\mathbf{A} + \frac{e^2}{2mc^2}A^2.$$

In quantum mechanics in the Schrödinger coordinate representation the momentum $\mathbf{p}$ is the operator $-i\hbar\nabla$; therefore the effect of a magnetic field is to add to the Hamiltonian the terms

$$\mathcal{H}' = \frac{ie\hbar}{2mc}(\nabla\cdot\mathbf{A} + \mathbf{A}\cdot\nabla) + \frac{e^2}{2mc^2}A^2,$$

which may usually be treated as a small perturbation. If the magnetic field is uniform and in the z direction, we may write

$$\vec{A} = -1/2\,\vec{r}\times\vec{H} \qquad \vec{H} = H_z\hat{z} = H\hat{z}$$

$$\vec{A} = -1/2(yH_z - zH_y, zH_y - xH_z, zH_y - yH_x) = -1/2(yH_z, -xH_z, 0)$$

$$A_x = -\tfrac{1}{2}yH, \qquad A_y = \tfrac{1}{2}xH, \qquad A_z = 0,$$

$$\frac{ie\hbar}{2mc}\vec{r}\times\vec{H}.\vec{\nabla} = -\frac{ie\hbar}{2mc}\vec{H}.\vec{r}\times\vec{\nabla} = \frac{e}{2mc}\vec{H}.\vec{r}\times\frac{\hbar}{i}\vec{\nabla}$$

$\vec{p}$ $\qquad$ $\vec{L}$

$$\mathcal{H}' = \frac{ie\hbar H}{2mc}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) + \frac{e^2H^2}{8mc^2}(x^2 + y^2).$$

The first term on the right is proportional to the orbital angular momentum component L_z if $\mathbf{r}$ is measured from the nucleus, and in mononuclear systems gives rise only to paramagnetism. The

[4] J. Pirenne, Helv. Phys. Acta **22**, 479 (1949).

second term gives for a spherically symmetric system a diamagnetic contribution

$$W' = \frac{e^2H^2}{12mc^2}\overline{r^2}$$

to the perturbation energy, and the associated magnetic moment is

$$\mu = -\partial W'/\partial H = -(e^2\overline{r^2}/6mc^2)H,$$

in agreement with the classical result (8.5). For further details of the derivation the book by Van Vleck may be consulted.

Paramagnetism

Electronic paramagnetism is found in:

(a) All atoms and molecules possessing an odd number of electrons, as here the total spin of the system cannot be zero. Examples: free sodium atoms; gaseous nitric oxide (NO); organic free radicals such as triphenylmethyl, $C(C_6H_5)_3$.

(b) All free atoms and ions with an unfilled inner shell: transition elements; ions isoelectronic with transition elements; rare earth and actinide elements. Examples: Mn^{2+}, Gd^{3+}, U^{4+}. Paramagnetic properties are exhibited by many of these ions when incorporated into solids, and as ions in solution, but not invariably.

(c) A few miscellaneous compounds, including molecular oxygen and organic biradicals.

(d) Metals: the paramagnetism of conduction electrons is treated in Chapters 12 and 13.

3.9. The statistics of paramagnetism

Next, we study a system of N magnetic dipoles, each having a magnetic moment μ. In the presence of an external magnetic field $\boldsymbol{H}$, the dipoles will experience a torque tend[s] to align them in the direction of the field. If there were nothing else to check [i]s tendency, the dipoles would align themselves precisely in this direction and we would achieve a complete magnetization of the system. In reality, however, thermal agitation in the system offers [r]esistance to this tendency and, in equilibrium, we obtain only a *partial* magnetization. Clearly, as $T \to 0$ K, the thermal agitation becomes ineffective and the system exhibits a complete orientation of the dipole moments, whatever the strength of the applied field; at the other extreme, as $T \to \infty$, we approach a state of complete randomization of the dipole moments, which implies a vanishing magnetization. At intermediate temperatures, the situation is governed by the parameter $(\mu H/kT)$.

The model adopted for this study consists of N identical, localized (and, hence, distinguishable), practically static, mutually non-interacting and freely orientable dipoles. It is obvious that the only energy we need to consider here is the potential energy of the dipoles which arises from the presence of the external field $\boldsymbol{H}$ and is determined by the orientations of the dipoles with respect to the direction of the field:

$$E = \sum_{i=1}^{N} E_i = -\sum_{i=1}^{N} \boldsymbol{\mu}_i \cdot \boldsymbol{H} = -\mu H \sum_{i=1}^{N} \cos\theta_i.$$

The partition function of the system is then given by

$$Q_N(\beta) = [Q_1(\beta)]^N.$$

where

$$Q_1(\beta) = \sum_{\theta} \exp(\beta\mu H \cos\theta).$$

The mean magnetic moment $\boldsymbol{M}$ of the system will obviously be in the direction of the field $\boldsymbol{H}$; for its magnitude we shall have

$$M_z = N\langle\mu\cos\theta\rangle = N\frac{\sum_{\theta}\mu\cos\theta\exp(\beta\mu H\cos\theta)}{\sum_{\theta}\exp(\beta\mu H\cos\theta)}$$

$$= \frac{N}{\beta}\frac{\partial}{\partial H}\ln Q_1(\beta).$$

Thus, to determine the degree of magnetization in the system all we have to do is to evaluate the single-dipole partition function (3).

First, we do it classically (after Langevin, 1905a,b). Using $(\sin\theta\, d\theta\, d\phi)$ as the elemental solid angle representing a small range of orientations of the dipole, we get

$$Q_1(\beta) = \int_0^{2\pi}\int_0^{\pi} e^{\beta\mu H\cos\theta}\sin\theta\, d\theta\, d\phi = 4\pi\frac{\sinh(\beta\mu H)}{\beta\mu H}, \tag{5}$$

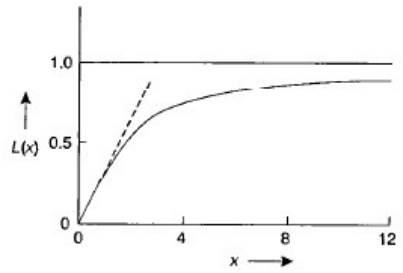
whence

$$\overline{\mu}_z \equiv \frac{M_z}{N} = \mu\left\{\coth(\beta\mu H) - \frac{1}{\beta\mu H}\right\} = \mu L(\beta\mu H), \tag{6}$$

where $L(x)$ is the so-called *Langevin function*

$$L(x) = \coth x - \frac{1}{x}; \tag{7}$$

we note th. ıe parameter $\beta\mu H$ denotes the strength of the (magnetic) potential energy μH against the (thermal) kinetic energy kT. A plot of the Langevin function is shown in Fig. 3.6.



If we have N_0 dipoles per unit volume in the system, then the magnetization of the system, viz. the mean magnetic moment per unit volume, is given by

$$M_{z0} = N_0\overline{\mu}_z = N_0\mu L(x) \quad (x = \beta\mu H).$$

For magnetic fields so strong (or temperatures so low) that the parameter $x \gg 1$, the function $L(x)$ is almost equal to 1; the system then acquires a state of magnetic saturation:

$$\overline{\mu}_z \simeq \mu \quad \text{and} \quad M_{z0} \simeq N_0\mu.$$

For temperatures so high (or magnetic fields so weak) that the parameter $x \ll 1$, the function $L(x)$ may be written as

$$\frac{x}{3} - \frac{x^3}{45} + \cdots$$

which, in the lowest approximation, gives

$$M_{z0} \simeq \frac{N_0\mu^2}{3kT}H.$$

The high-temperature susceptibility of the system is, therefore, given by

$$\chi = \operatorname*{Lim}_{H\to 0}\left(\frac{\partial M_{z0}}{\partial H}\right) \simeq \frac{N_0\mu^2}{3kT} = \frac{C}{T}, \text{ say.}$$

Equation (12) is the *Curie law* of paramagnetism, the parameter C being the *Curie constant* of the system. Figure 3.7 shows a plot of the susceptibility of a powdered sample of copper–potassium sulphate hexahydrate as a function of T^{-1}; the fact that the plot is linear and passes almost through the origin vindicates the Curie law for this particular salt.

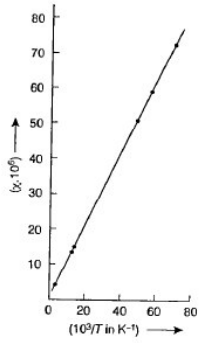


FIG. 3.7. χ vs. $1/T$ plot for a powdered sample of copper–potassium sulphate hexahydrate (after Hupse, 1942).

We shall now treat the problem of paramagnetism quantum-mechanically. The major modification here arises from the fact that the magnetic dipole moment $\boldsymbol{\mu}$ and its component μ_z in the direction of the applied field cannot have *arbitrary* values.

$$Q_1(\beta) = \sum_{m=-J}^{J} \exp(\beta g\mu_B mH). \tag{18}$$

Introducing a parameter x, defined by

$$x = \beta(g\mu_B J)H, \tag{19}$$

eqn. (18) becomes

$$\begin{aligned} Q_1(\beta) = \sum_{m=-J}^{J} e^{mx/J} &= \frac{e^{-x}\{e^{(2J+1)x/J} - 1\}}{e^{x/J} - 1} \\ &= \frac{e^{(2J+1)x/2J} - e^{-(2J+1)x/2J}}{e^{x/2J} - e^{-x/2J}} \\ &= \sinh\left\{\left(1 + \frac{1}{2J}\right)x\right\} \Big/ \sinh\left\{\frac{1}{2J}x\right\}. \end{aligned} \tag{20}$$

The mean magnetic moment of the system is then given by, see (4),

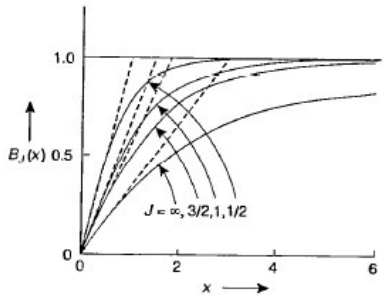
$$\begin{aligned} M_z = N\overline{\mu}_z &= \frac{N}{\beta}\frac{\partial}{\partial H}\ln Q_1(\beta) \\ &= N(g\mu_B J)\left[\left(1 + \frac{1}{2J}\right)\coth\left\{\left(1 + \frac{1}{2J}\right)x\right\} - \frac{1}{2J}\coth\left\{\frac{1}{2J}x\right\}\right]. \end{aligned} \tag{21}$$

Thus

$$\overline{\mu}_z = (g\mu_B J)B_J(x), \tag{22}$$

where $B_J(x)$ is the *Brillouin function* (of order J):

$$B_J(x) = \left(1 + \frac{1}{2J}\right)\coth\left\{\left(1 + \frac{1}{2J}\right)x\right\} - \frac{1}{2J}\coth\left\{\frac{1}{2J}x\right\}. \tag{23}$$



We shall now consider a few special cases. First of all, we note that for strong fields and low temperatures ($x \gg 1$), the function $B_J(x) \simeq 1$ *for all* J, which corresponds to a state of magnetic saturation. On the other hand, for high temperatures and weak fields ($x \ll 1$), the function $B_J(x)$ may be written as

$$\tfrac{1}{3}(1 + 1/J)x + \tag{24}$$

so that

$$\overline{\mu}_z \simeq \frac{(g\mu_B J)^2}{3kT}\left(1 + \frac{1}{J}\right)H = \frac{g^2\mu_B^2 J(J+1)}{3kT}H. \tag{25}$$

The Curie law, $\chi \propto 1/T$, is again obeyed; however, the Curie constant is now given by

$$C_J = \frac{N_0 g^2 \mu_B^2 J(J+1)}{3k} \tag{26}$$

$$C_{\frac{1}{2}} = \frac{N_0(2)^2\mu_B^2\frac{1}{2}(\frac{1}{2}+1)}{3k} = \frac{N_0\mu_B^2}{k} = 3C_{\text{Classic}}$$

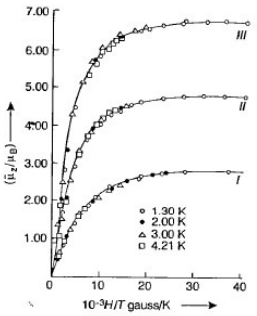


FIG. 3.9. Plots of $\overline{\mu}_z/\mu_B$ as a function of H/T. The solid curves represent the theoretical results, while the points mark the experimental findings of Henry (1952). Curve I is for potassium chromium alum ($J = \frac{3}{2}, g = 2$), curve II for iron ammonia alum ($J = \frac{5}{2}, g = 2$) and curve III for gadolinium sulphate octahydrate ($J = \frac{7}{2}, g = 2$).

Ferromagnetism and Antiferromagnetism

CURIE POINT AND THE EXCHANGE INTEGRAL

We call a substance ferromagnetic if it possesses a spontaneous magnetic moment, that is, a magnetic moment even in the absence of an applied magnetic field. The saturation magnetization M_s is defined as the spontaneous magnetic moment per unit volume. In technical literature the saturation flux density $B_s = 4\pi M_s$ is often used. The Curie point T_c is the temperature above which the spontaneous moment vanishes.

If we could add to a paramagnetic substance an interaction tending to make the ionic and atomic magnetic moments line up the same way, we would have a ferromagnetic substance. Let us postulate such an interaction and call it the Weiss field.[1] The orienting effect of the Weiss field is opposed by the motion of thermal agitation of the elementary moments. We consider the Weiss field the equivalent of an effective magnetic field H_E acting on the electron spins. The interaction energy of a spin with the Weiss field must be of the order of magnitude of the thermal energy of a spin at the Curie point.

If we could add to a paramagnetic substance an interaction tending to make the ionic and atomic magnetic moments line up the same way, we would have a ferromagnetic substance. Let us postulate such an interaction and call it the Weiss field.[1] The orienting effect of the Weiss field is opposed by the motion of thermal agitation of the elementary moments. We consider the Weiss field the equivalent of an effective magnetic field H_E acting on the electron spins. The interaction energy of a spin with the Weiss field must be of the order of magnitude of the thermal energy of a spin at the Curie point. Hence

$$gS\mu_B H_E \approx kT_c, \tag{10.1}$$

or

$$H_E \approx kT_c/gS\mu_B. \tag{10.2}$$

[1] Also called the molecular field or the exchange field; Weiss was the first to imagine such a field.

For iron we have $T_c \approx 1000°\text{K}$, $g \approx 2$, $S \approx 1$; therefore $H_E \approx 10^{-13}/2 \times 10^{-20} = 5 \times 10^6$ oersteds. This field is much stronger than that produced by the magnetic moments of the other ions in the crystal, as the magnetic interaction is only $\sim\mu_B/a^3 \sim 10^3$ oersteds, where a is the lattice constant.

Pierre Weiss (1907), inventor of this concept, showed that it will account for several important attributes of ferromagnetism provided that one assumes that the Weiss field H_E is proportional to the magnetization:

$$H_E = \lambda M, \tag{10.3}$$

where λ stands for a constant called the Weiss field constant. The susceptibility above the Curie point is deduced by postulating that the Curie law (9.4) holds if we take as the magnetic field the sum of the applied field H and the Weiss field H_E. Then

$$\frac{M}{H + \lambda M} = \frac{C}{T}, \tag{10.4}$$

or

$$\chi = \frac{M}{H} = \frac{C}{T - C\lambda}. \tag{10.5}$$

This gives a non-zero magnetization for zero applied field at the Curie point expressed by

$$T_c = C\lambda, \tag{10.6}$$

so that

$$\chi = \frac{C}{T - T_c}. \tag{10.7}$$

This expression, known as the Curie-Weiss law, describes quite well the observed susceptibility variation in the paramagnetic region above the Curie point.[2] From (10.6) and the definition (9.4) and (9.12) of the Curie constant C we may determine the value of the Weiss field constant:

$$\lambda^{-1} = C/T_c = Ng^2S(S+1)\mu_B{}^2/3kT_c, \tag{10.8}$$

so that for iron, $\lambda \approx (4 \times 10^{-13})/(8 \times 10^{-17}) \approx 5000$, in agreement with the earlier estimate of H_E.

The physical origin of the Weiss field is in the quantum-mechanical exchange integral, as pointed out by Heisenberg (1928). On certain assumptions it can be shown[3] that the energy of interaction of atoms i, j bearing spins S_i, S_j contains a term

$$W_{ex} = -2J\mathbf{S}_i \cdot \mathbf{S}_j, \tag{10.9}$$

where J is the exchange integral and is related to the overlap of the charge distributions i, j. The exchange energy has no classical analogue. It expresses the difference in Coulomb interaction energy of the systems when the spins are parallel or antiparallel. It is a consequence of the Pauli exclusion princple that in quantum mechanics one cannot usually change the relative direction of two spins without making changes in the spatial charge distribution in the overlap region. The resulting changes in the coulomb energy of the system may conveniently be written in the form[4] (10.9), so that it appears *as if* there were a direct coupling between the spins $\mathbf{S}_i$, $\mathbf{S}_j$.

We establish an approximate connection between the exchange integral J and the Weiss field constant λ. Suppose that the atom under consideration has z nearest neighbors, each connected with the central atom by the interaction J; for more distant neighbors we take J as zero. Then the interaction energy may be written, neglecting components of $\mathbf{S}$ perpendicular to the average magnetization,

$$W_{ex} \cong -2Jz\bar{S}^2 = -g\bar{S}\mu_B H_E = -g\bar{S}\mu_B\lambda(g\bar{S}\mu_B\Omega^{-1}),$$

where the term in parentheses is equal to M_s; here Ω is the atomic volume. Then

$$J = \lambda g^2\mu_B{}^2/2z\Omega, \tag{10.10}$$

and, using (10.8) and recalling that $N = 1/\Omega$,

$$J = \frac{3kT_c}{2zS(S+1)}, \tag{10.11}$$

This is the connection, as given by the Weiss field theory, between the exchange integral and the Curie point. More exact quantum statistics give somewhat different results. For a simple cubic lattice ($z = 6$)

ANTIFERROMAGNETISM

The antiferromagnetic state is characterized by an ordered antiparallel arrangement of electron spins. When the exchange integral J

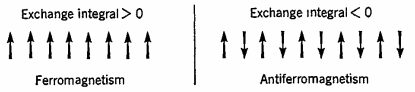


Fig. 10.19. Comparison of spin ordering in the ferromagnetic and antiferromagnetic states.

in (10.9) is positive, we have ferromagnetism; when J is negative, we have antiferromagnetism. On passing below the Curie point of an antiferromagnetic the spins lock in (Fig. 10.19) with antiparallel orientations, and at the Curie point the susceptibility attains its maximum value, as shown in Figs. 10.20 and 10.21. We recognize antiferromagnetism by a well-defined kink in the susceptibility vs. temperature curve. The transition is also marked by anomalies in the heat capacity and thermal expansion coefficient.

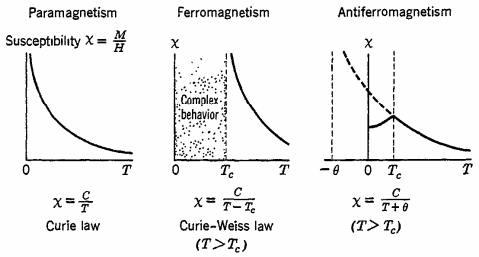


Fig. 10.20. Distinguishing features of the temperature dependence of the magnetic susceptibility in paramagnetism, ferromagnetism, and antiferromagnetism.

Antiferromagnetism was first discovered in 1938 by Bizetté, Squire, and Tsai,[23] while working with manganese oxide, which has a Curie temperature of 116° to 120°K. Néel[24] and Bitter[25] had presented theoretical discussions of the antiferromagnetic state earlier, and Van Vleck[26] gave the first detailed treatment.

[23] Bizette, Squire, and Tsai, Compt. rend. **207,** 449 (1938).
[24] L. Néel, Ann. phys. **18,** 5 (1932); **5,** 232 (1936).
[25] F. Bitter, Phys. Rev. **54,** 79 (1937).

The simplest situation in antiferromagnetism arises when the lattice of paramagnetic ions can be divided into two interpenetrating sublattices A, B such that all nearest neighbors of an ion on sublattice A lie on sublattice B. This condition is, for example, satisfied by the sc and bcc lattices, but not by the fcc lattice. If the only interactions are antiferromagnetic interactions between nearest neighbors, we may write for the magnetization above the Curie point on the Weiss field theory:

$$TM_A = C'(H - \lambda M_B); \tag{10.53}$$
$$TM_B = C'(H - \lambda M_A).$$

Here C' is the Curie constant for one sublattice, and the effective field on sublattice A is written as $H - \lambda M_B$, which for positive λ corresponds to antiferromagnetic interactions between A and B. Adding,

$$TM = T(M_A + M_B) = 2C'H - C'\lambda M,$$

$$\chi = \frac{2C'}{T + C'\lambda},$$

$$\chi = \frac{C}{T + \theta}$$

$$C = 2C'; \qquad \theta = C'\lambda$$

The transition temperature is that below which each sublattice A and B possesses a magnetic moment even without a field. Below the Curie point it is not legitimate to treat the moment as a linear function of the effective field, but, as saturation is not important close to the Curie point, linearity may be assumed in the equations for the Curie point.

The transition temperature T_c is then the temperature at which equations (10.53) have a non-trivial solution for $H = 0$. The condition for this is that the determinant of the coefficients of the unknowns M_A, M_B should be zero:

$$\begin{vmatrix} T & \theta \\ \theta & T \end{vmatrix} = 0$$

$$T_c = \theta.$$

METHOD OF MEASUREMENT OF SUSCEPTIBILITIES

The usual methods of measuring diamagnetic and paramagnetic susceptibilities depend in one way or another on the force exerted on a specimen by a non-uniform magnetic field. The force F is given by minus the gradient of the magnetic energy, so that,

$$\mathbf{F} = \tfrac{1}{2} \operatorname{grad} \int \chi H^2 \, dV,$$

where now χ is the volume susceptibility.

If we suppose that the specimen is in the form of a thin rod, as in the Gouy method shown in Fig. 8.1, we may write for the downward pull

$$F_z = \frac{1}{2} \chi A \int \frac{d}{dz} H^2 \, dz = \frac{1}{2} \chi A (H_1^2 - H_2^2),$$

where A is the sectional area of the specimen. Other methods of measurement are discussed, for example, in the book by Stoner cited in the references. A high sensitivity method has been described by Lewis, Calvin, and Kasha.[9]

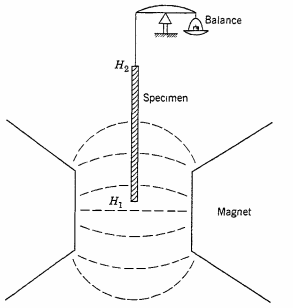


Fig. 8.1. Gouy method for measuring susceptibilities.

E. C. Stoner, *Magnetism and matter*, Methuen and Co., Ltd., London, 1934.

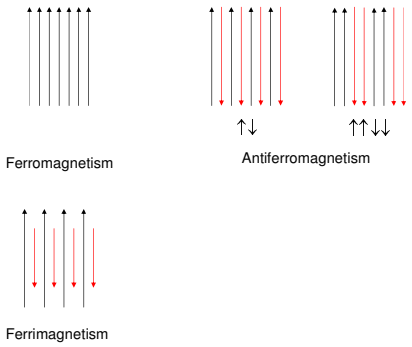
One can classify magnetic materials into two following types:

1) Collinear systems

2) Non-collinear systems

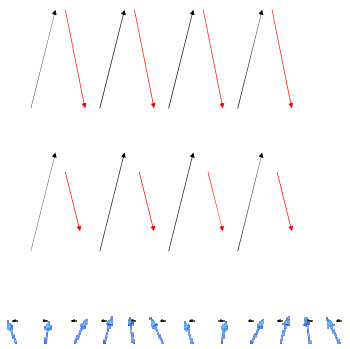
Collinear magnetic systems can be listed as follows:

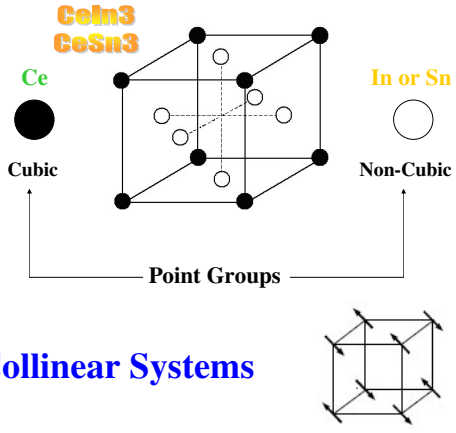
1) Ferromagnetism
2) Antiferromagnetism
3)Ferrimagnetism
4)Diamagnetism



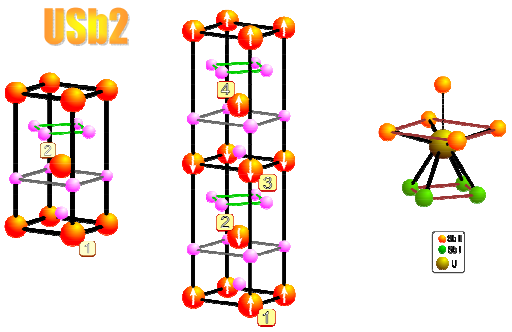
Non-collinear magnetic systems can be listed as follows:

1) Canted Antiferromagnetic materials
2) Canted Ferrimagnetic materials
3) Helical spin array



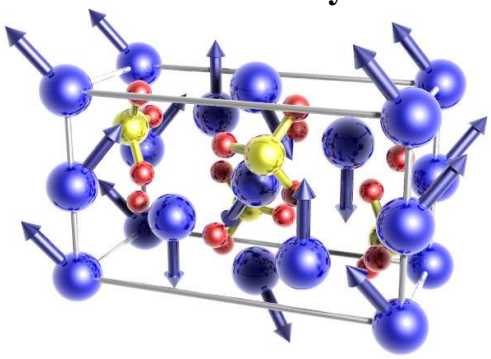


Collinear Systems

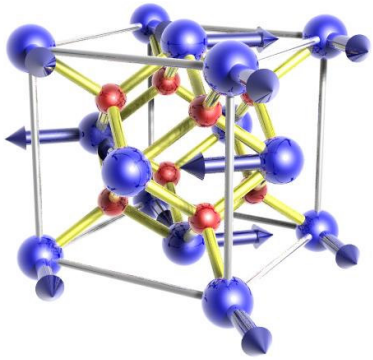


Collinear Systems

Non-Collinear Systems

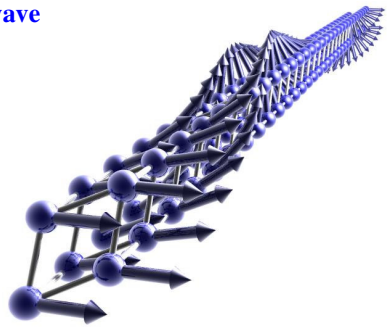


Non-Collinear Systems

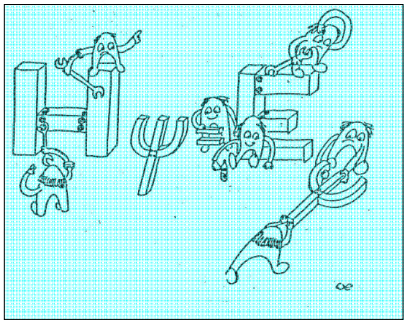


Non-Collinear Systems

Spin wave



What is a difference between them?



Why?

$$\begin{pmatrix} H_{11} & 0 \\ 0 & H_{11} \end{pmatrix}$$

Non-Magnetic systems

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \qquad \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix}$$

Magnetic systems

$$\rho(\vec{r}) = \frac{n(\vec{r})}{2} I + \frac{\vec{m}(\vec{r})}{2}.\vec{\sigma}$$

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V_{eff} + \mu_B \vec{\sigma}.\vec{B}_{eff}$$

$$V_{eff} = V_{ext} + V_H + V_{xc}$$
$$\vec{B}_{eff} = \vec{B}_{ext} + \vec{B}_{xc}$$

$$V_{xc} = \frac{\partial E_{xc}(n,\vec{m})}{\partial n}$$
$$\vec{B}_{xc} = \frac{\partial E_{xc}(n,\vec{m})}{\partial \vec{m}}$$

LDA:

$$E_{xc}(n,\vec{m}) = \int n\varepsilon_{xc}(n,m)d^3r$$

$$V_{xc} = \varepsilon_{xc}(n,m) + n\frac{\partial \varepsilon_{xc}(n,m)}{\partial n}$$
$$\vec{B}_{xc} = \frac{\partial \varepsilon_{xc}(n,m)}{\partial m}\hat{m}$$

$$\vec{B}_{xc} = \frac{\partial \varepsilon_{xc}(n,\vec{m})}{\partial \vec{m}}\hat{m}$$

LDA

$\vec{B} \parallel \vec{m}$

In a collinear case moments are aligned along an arbitrary direction. The spin quantization axis can be chosen along m direction. Thus x and y components of the exchange field will vanish resulting in:

$$\vec{\sigma}.\vec{B}_{eff} = \sigma_z B_{eff}$$

Since: $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Potential matrix is diagonal in spin space:

$$\hat{V} = V_{eff} I + \mu_B B_{eff} \sigma_z = \begin{bmatrix} V_{eff} + \mu_B B_{eff} & 0 \\ 0 & V_{eff} - \mu_B B_{eff} \end{bmatrix}$$

Up and down panels of the Hamiltonian are decupled. Each part is set up and diagonalized separatly.

Note that spin orbit coupling (SOC) couples these panels. Usually the off-diagonal term of density matrix, resulting from this coupling is ignored.

In a non-collinear case moments vary in space and all three components of $\vec{\sigma}.\vec{B}_{eff}$ the scalar product must be considered.

Since: $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Potential matrix is not diagonal:

$$\vec{B}_{eff}.\vec{\sigma} = \mu_B\left(B_x\sigma_x + B_y\sigma_y + B_z\sigma_z\right)$$

$$\hat{V} = V_{eff} I + \mu_B \vec{B}_{eff}.\vec{\sigma} = \begin{bmatrix} V_{eff} + \mu_B B_z & \mu_B\left(B_x - iB_y\right) \\ \mu_B\left(B_x + iB_y\right) & V_{eff} - \mu_B B_z \end{bmatrix}$$

WIENNCM

ORDER–DISORDER TRANSITIONS

An important class of systems that can be studied using methods of equilibrium statistical mechanics are those which exhibit a transition from a disordered to an ordered state. One of the simplest examples of such a system consists of a lattice of N spin-$\frac{1}{2}$ objects with magnetic moment μ. Each lattice site can interact with the magnetic fields of their nearest neighbors and with any external applied magnetic fields that might be present. The Hamiltonian for such a system can be written

$$H = \sum_{\{ij\}} \epsilon_{ij} s_i s_j - \mu B \sum_{i=1}^{N} s_i,$$

where $\sum_{ij}$ denotes the sum over nearest-neighbor pairs ij (one must be careful to count each pair only once), ϵ_{ij} is the magnetic interaction energy between nearest neighbors i and j, s_i is the z component of spin at the ith lattice site, and B is the external magnetic field. For spin-$\frac{1}{2}$ objects, $s_i = +1(-1)$ if the spin of site i is oriented in the positive (negative) z direction. There is no kinetic energy in Eq. (7.75). *The Hamiltonian only contains information about spin orientation and the spatial distribution of lattice sites.* If $\epsilon_{ij} < 0$, then for $B = 0$ the lattice will have its lowest energy when all the lattice sites have spin up or all the lattice sites have spin down, both cases being equally probable. If $B \neq 0$, then the configuration in which all lattice sites are oriented with spin up will be energetically favored. Similarly, if $\epsilon_{ij} > 0$, then for $B = 0$ the configuration in which neighboring spins are oriented opposite to one another will be favored.

The partition function for this spin lattice can be written

$$Z_N(T) = \sum_{\text{all config.}} \exp\left[-\beta \sum_{\{ij\}} \epsilon_{ij} s_i s_j + \beta \mu B \sum_{i=1}^{N} s_i\right]$$

where $\sum_{\text{allconfig.}}$ denotes the sum over all 2^N possible different configurations of spin on the lattice. The partition function introduces an additional influence, that of thermal energy, $k_B T$. While the magnetic interaction energy, H, will cause the spins on the lattice to become ordered, the thermal energy, $k_B T$, will tend to randomize the spins on the lattice. It is these two competing influences which lead to an *order–disorder phase transition* on the spin lattice. At low temperature, the lattice will be ordered. As we raise the temperature, at some point the order dissappears and the spins become randomly oriented. The system, described by the partition function, Eq. (7.76), is called the *Ising model* and was originally used by Ising [10] as a model for ferromagnetism. However, the model also has been used to describe the behavior of lattice gases, binary alloys "melting" of DNA, and so on.

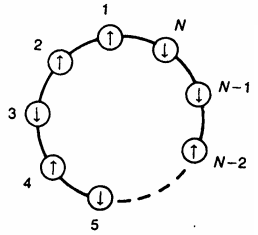
The Ising model can be solved exactly, and analytic expressions for its thermodynamic properties can be found in one and two dimensions, but no one has ever succeeded in solving it analytically in three dimensions. In one dimension it does not exhibit a phase transition at finite temperature, but in two dimensions it does. In one dimension the lattice does not have enough nearest neighbors for the ordering effects of the interaction energy to compete effectively with the disordering thermal energy. However, for two or more spatial dimensions it does. The Ising model was first solved in two dimensions by Onsager [11–13]. It is one of the few exactly soluble models which exhibit a phase transition.

In this section we shall consider two simple versions of the Ising model. We will first obtain exact expressions for the thermodynamic properties of a one-dimensional Ising lattice. The one-dimensional lattice does not exhibit a phase transition at finite temperature. However, the method of solution contains some ideas which are used to solve the much more difficult case of a two-dimensional lattice. We will solve the two-dimensional case in Chapter 8. In this section we will also compute the thermodynamic properties of an Ising lattice with d dimensions in the mean field approximation. The mean field approximation of the d-dimensional lattice does have a phase transition at finite temperature.

Exact Solution for a One-Dimensional Lattice

Let us consider a one-dimensional periodic lattice that consists of N lattice sites evenly spaced, as shown in Fig. 7.5. We will assume that all nearest neighbors have the same interaction energy, $\epsilon_{ij} = -\epsilon$, so that the configuration with lowest energy is one in which the spins are totally aligned. The periodicity of the lattice is imposed by assuming that $s_{i+N} = s_i$. The total energy for a given configuration, $\{s_i\}$, is

$$E\{s_i\} = -\epsilon \sum_{i=1}^{N} s_i s_{i+1} - \mu B \sum_{i=1}^{N} s_i$$



A one-dimensional periodic lattice with N lattice sites

The partition function can be written

$$Z_N(T,B) = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \exp\left[\beta \sum_{i=1}^{N} \left(\epsilon s_i s_{i+1} + \frac{1}{2}\mu B(s_i + s_{i+1})\right)\right]$$

where we have used the fact that $\sum_{i=1}^{N} s_i = \frac{1}{2}\sum_{i=1}^{N}(s_i + s_{i+1})$ for a periodic lattice. It is now convenient to introduce a 2×2 matrix,

$$\bar{\mathbf{P}} = \begin{pmatrix} e^{\beta(\epsilon+\mu B)} & e^{-\beta\epsilon} \\ e^{-\beta\epsilon} & e^{\beta(\epsilon-\mu B)} \end{pmatrix},$$

whose matrix elements are defined as

$$\langle s_i | \bar{\mathbf{P}} | s_{i+1} \rangle = e^{\beta(\epsilon s_i s_{i+1} + \frac{1}{2}\mu B(s_i + s_{i+1}))}.$$

The partition function may then be written

$$Z_N(T,B) = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \langle s_1|\bar{\mathbf{P}}|s_2\rangle\langle s_2|\bar{\mathbf{P}}|s_3\rangle \cdots \langle s_N|\bar{\mathbf{P}}|s_1\rangle$$
$$= \sum_{s_1=\pm 1} \langle s_1|\bar{\mathbf{P}}^N|s_1\rangle = \mathrm{Tr}[\bar{\mathbf{P}}^N] = \lambda_+^N + \lambda_-^N = \lambda_+^N\left[1 + \left(\frac{\lambda_-}{\lambda_+}\right)^N\right],$$

where $\lambda_\pm$ are the eigenvalues of the matrix $\bar{\mathbf{P}}$. We shall use the convention $\lambda_+ > \lambda_-$. The eigenvalues of $\bar{\mathbf{P}}$ are easily found to be

$$\lambda_\pm = e^{\beta\epsilon}\left[\cosh(\beta\mu B) \pm \sqrt{\cosh^2(\beta\mu B) - 2e^{-2\beta\epsilon}\sinh(2\beta\epsilon)}\right].$$

In the limit $N \to \infty$, only the largest eigenvalue, λ_+, contributes to the thermodynamic quantities. This is easily seen if we note that the Gibbs free energy per site is

$$g(T,B) = \lim_{N\to\infty}\frac{1}{N}G_N(T,B) = -k_BT\lim_{N\to\infty}\frac{1}{N}\ln[Z_N(T,B)]$$
$$= -k_BT\ln[\lambda_+].$$

In Eq. (7.83), we have used the fact that $\lim_{N\to\infty}(\lambda_-/\lambda_+)^N = 0$. Thus the Gibbs free energy per site is

$$g(T,B) = -\epsilon - k_B\ln\left[\cosh(\beta\mu B) + \sqrt{\cosh^2(\beta\mu B) - 2e^{-2\beta\epsilon}\sinh(2\beta\epsilon)}\right].$$

The order parameter is given by

$$\langle s\rangle = -\left(\frac{\partial g}{\partial \mu B}\right)_T = \frac{\sinh(\beta\mu B)}{\sqrt{\cosh^2(\beta\mu B) - 2e^{-2\beta\epsilon}\sinh h(2\beta\epsilon)}}.$$

we see that the one-dimensional Ising model cannot exhibit a phase transition because when $B \to 0$ the order parameter also goes to zero. Hence, no spontaneous nonzero value of the order parameter is possible. The exact solution to the two-dimensional Ising model is givin in Chapter 8, and it does allow a phase transition.

Mean Field Theory for a *d*-Dimensional Lattice

We can obtain approximate analytic expressions for the thermodynamic properties of a d-dimensional Ising lattice using the mean field approximation first introduced by Weiss [14]. In the mean field approximation, the Hamiltonian of a d-dimensional spin lattice with N lattice sites can be written

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nu\epsilon\langle s\rangle s_i - \mu B\sum_{i=1}^{N}s_i = -\sum_{i=1}^{N}E(\epsilon,B)s_i,$$

where ϵ is the coupling constant, ν is the number of nearest-neighbor spins, and $E(\epsilon,B) = \frac{1}{2}\nu\epsilon\langle s\rangle + \mu B$. The factor of $\frac{1}{2}$ ensures that we don't count the same pair of spins twice. The quantity $\langle s\rangle \equiv \langle s_i\rangle$ is the average spin per site. The quantity $\nu\epsilon\langle s\rangle s_i$ is an average magnetic interaction energy between site i and its nearest neighbours, assuming that the neigbors all have spin $\langle s\rangle$. As we shall show below, the average spin per site, $\langle s\rangle$, must be determined in a self-consistent manner.

The partition function can be written

$$Z_N = \left(\sum_{s_i=\pm 1} e^{\beta E s_i}\right)^N = (2\cosh(\beta E))^N.$$

The Gibbs free energy per site is

$$g(\epsilon,B) = -k_BT\lim_{N\to\infty}\left(\frac{1}{N}\ln(Z_N)\right) = -k_BT\ln[2\cosh(\beta E)].$$

The probability $P(s_i)$ that site i has spin s_i is

$$P(s_i) = \frac{e^{\beta E s_i}}{\sum_{s_i=\pm 1} e^{\beta E s_i}} = \frac{e^{\beta E s_i}}{2\cosh(\beta E)}.$$

Note that the probability $P(s_i)$has an especially simple form in the mean field approximation because coupling to nearest neighbours is taken into account in such a simple manner.

The average magnetization of the lattice is given by

$$\langle M\rangle = N\mu\langle s\rangle,$$

where

$$\langle s\rangle = \frac{\sum_{s_i=\pm 1} s_i e^{\beta E s_i}}{\sum_{s_j=\pm 1} e^{\beta E s_j}} = \tanh[\beta E] = \tanh\left[\beta\left(\frac{1}{2}\nu\epsilon\langle s\rangle + \mu B\right)\right].$$

The magnetization is the order parameter for the spin lattice. If $B = 0$, the magnetization will be zero for the high-temperature paramagnetic phase of the lattice (randomly ordered spins) and it will be nonzero at lower temperatures where the spins have spontaneously aligned.

We can determine the critical temperature, T_c, at which the lattice starts to become ordered as temperature is lowered (the Curie point) from the expression for the average spin per site, $\langle s\rangle$. Let us write $\langle s\rangle$ for the case $B = 0$;

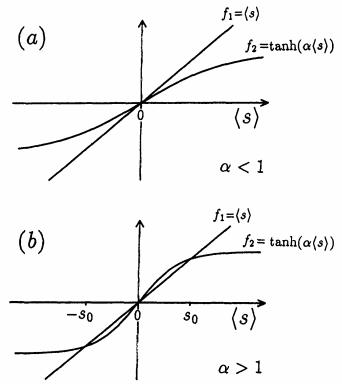
$$\langle s\rangle = \tanh\left[\frac{1}{2}\beta\nu\epsilon\langle s\rangle\right] = \tanh\left[\frac{\nu\epsilon\langle s\rangle}{2k_BT}\right].$$

We must solve Eq. (7.92) for $\langle s\rangle$. This can be done graphically by plotting $f_1 \equiv \langle s\rangle$ versus $\langle s\rangle$ and $f_2 \equiv \tanh(\alpha\langle s\rangle)$ versus $\langle s\rangle$ on the same graph. The solution to Eq. (7.92) is given by those points where the two curves cross—that is, where $f_1 = f_2$. In Figs. 7.6a and 7.6b, we plot f_1 and f_2 versus $\langle s\rangle$ for $\alpha < 1$ and $\alpha > 1$, respectively. For the case $\alpha < 1$, there is only one crossing point and it is at $\langle s\rangle = 0$. For $\alpha > 1$, there are three crossing points, at $\langle s\rangle = 0$ and at $\langle s\rangle = \pm s_0$. The free energy per site for these various cases is

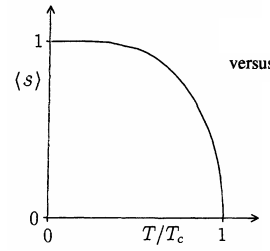
$$g(\epsilon, 0) = \begin{cases} -k_BT\ln(2) & \text{if } \langle s\rangle = 0, \\ -k_BT\ln\left[2\cosh(\frac{1}{2}\beta\nu\epsilon s_0)\right] & \text{if } \langle s\rangle = \pm s_0. \end{cases}$$

Thus, the values, $\langle s\rangle = \pm s_0$ [when they are solutions to Eq. (7.92)] describe possible states of thermodynamic equilibrium since they minimize the free energy. The transition point (critical point) occurs at $\alpha = 1$ in Fig. 7.6 and therefore when $\nu\epsilon/2k_BT = 1$. Thus the critical temperature in the mean field approximation is $T = T_c = \nu\epsilon/2k_B$. In Fig. 7.7, we plot the order parameter $\langle s\rangle$, versus T/T_c.

We see that mean field theory predicts a phase transition at a finite temperature for a d-dimensional lattice. This does not agree with our exact



Plots of $f_1 \equiv \langle s\rangle$ versus $\langle s\rangle$ and $f_2 \equiv \tanh(\alpha\langle s\rangle)$ versus $\langle s\rangle$. (a) $\alpha < 1$. (b) $\alpha > 1$.



A plot of the order parameter, $\langle s\rangle$, versus the reduced temperature, T/T_c.

result for the case $d = 1$ (cf. Section 7.F.1) where we found no phase transition at finite temperature. As we shall see in Chapter 8, mean field theory gives too high an estimate of the critical temperature for spatial dimensions, $d \leq 3$. It gives good estimates for $d \geq 4$ which is not of physical interest but is of mathematical interest.

Let us next examine the behavior of the heat capacity in the neighborhood of the critical temperature. The internal energy for $B = 0$ is

$$U = -\frac{1}{Z_N}\frac{\partial Z_N}{\partial\beta} = -\frac{1}{2}N\nu\epsilon\langle s\rangle^2.$$

The heat capacity is

$$C_N = \left(\frac{\partial U}{\partial T}\right)_N = -k_B\beta^2\left(\frac{\partial U}{\partial\beta}\right)_N = Nk_B\nu\epsilon\beta^2\langle s\rangle\left(\frac{\partial\langle s\rangle}{\partial\beta}\right)_N.$$

But

$$\left(\frac{\partial\langle s\rangle}{\partial\beta}\right)_N = \operatorname{sech}^2\left(\beta\frac{\nu\epsilon}{2}\langle s\rangle\right)\left[\frac{\nu\epsilon}{2}\langle s\rangle + \beta\frac{\nu\epsilon}{2}\left(\frac{\partial\langle s\rangle}{\partial\beta}\right)_N\right]$$

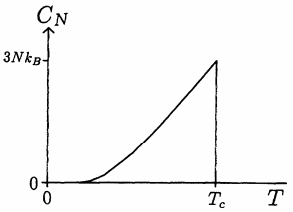
so

$$\left(\frac{\partial\langle s\rangle}{\partial\beta}\right)_N = \frac{\nu\epsilon\langle s\rangle}{3\cosh^2\left(\beta\frac{\nu\epsilon}{2}\langle s\rangle\right) - \beta\nu\epsilon}.$$

The heat capacity finally becomes

$$C_N = \frac{Nk_B\beta^2\nu^2\epsilon^2\langle s\rangle^2}{2\cosh^2\left(\beta\frac{\nu\epsilon}{2}\langle s\rangle\right) - \beta\nu\epsilon} = \frac{2Nk_B\langle s\rangle^2(T_c/T)^2}{[\cosh^2(\langle s\rangle T_c/T) - T_c/T]}.$$

we plot the heat capacity as a function of temperature. We see that the heat capacity has a finite jump at the critical temperature. It reaches a maximum value of $3Nk_B$ at $T = T_c$.



plot of the heat capacity, C_N, as a function of temperature in the neighborhood of the critical point.

The final quantity we wish to compute is the magnetic susceptibility, $\chi_{T,N}(B)$. The magnetic susceptibility is defined as

$$\chi_{T,N}(B) = \left(\frac{\partial\langle M\rangle}{\partial B}\right)_{T,N} = N\mu\left(\frac{\partial\langle s\rangle}{\partial B}\right)_{T,N}.$$

From Eq. (7.91) we can write

$$\left(\frac{\partial\langle s\rangle}{\partial B}\right)_{T,N} = \operatorname{sech}^2\left[\beta\frac{\nu\epsilon}{2}\langle s\rangle + \beta\mu B\right]\left[\beta\frac{\nu\epsilon}{2}\left(\frac{\partial\langle s\rangle}{\partial B}\right)_{T,N} + \beta\mu\right],$$

or

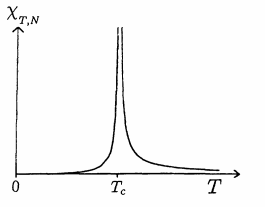
$$\left(\frac{\partial\langle s\rangle}{\partial B}\right)_{T,N} = \frac{2\beta\mu}{2\cosh^2\left[\beta\frac{\nu\epsilon}{2}\langle s\rangle + \beta\mu B\right] - \beta\nu\epsilon}.$$

The magnetic susceptibility, $\chi_{T,N}(B)$, is then given by

$$\chi_{T,N}(B) = \frac{2\beta N\mu^2}{2\cosh^2\left[\beta\frac{\nu\epsilon}{2}\langle s\rangle + \beta\mu B\right] - \beta\nu\epsilon}.$$

The magnetic susceptibility in the limit $B = 0$ is

$$\chi_{T,N}(0) = \frac{2\beta N\mu^2}{2\cosh^2\left[\beta\frac{\nu\epsilon}{2}\langle s\rangle\right] - \beta\nu\epsilon} = \frac{2N\mu^2}{\nu\epsilon}\frac{(T_c/T)}{\cosh^2(\langle s\rangle T_c/T) - T_c/T},$$



A plot of the magnetic susceptibility, $\chi(0)_{T,N}$, as a function of temperature in the neighborhood of the critical point.

with $\langle s\rangle$ given by Eq. (7.92). Examination of Eq. (7.103) shows that the magnetic susceptibility has an infinite jump at the critical point. In Fig. 7.9, we plot $\chi_{T,N}(0)$ as a function of temperature.

The results we have obtained in this section for the thermodynamic properties of a spin lattice in the neighborhood of the phase transition are qualitatively similar to the results we obtained in Chapter 3 using mean field theories. Using statistical mechanics to obtain these results allows us to compute various constants that appear in the thermodynamic expressions in terms of microscopic interaction energies and magnetic moments. A generalization of mean field methods has been used by Suzuki to study critical phenomena [15]. As we shall see in Chapter 8, where we give some exact results for the two-dimensional spin lattice and use renormalization theory, mean field theory gives a rough qualitative picture of the phase transition, but it is not quantitatively correct.

▶ S8.B.EXACT SOLUTION OF THE TWO-DIMENSIONAL ISING MODEL [14, 15]

14. L. Onsager, *Phys. Rev.* **65**, 117 (1944).
15. H. S. Robertson, *Statistical Thermophysics* (Prentice-Hall, Englewood Cliffs, NJ, 1993). (A really excellent discussion of the one- and two-dimensional Ising model and its history.)

The two-dimensional Ising model is one of the simplest systems that exhibit a phase transition and one of the few that can be solved exactly. The phase transition in this system is an order–disorder transition. It has served as the paradigm system for describing order–disorder transitions in a variety of contexts. In physics it has been used to model the phase transition from the paramagnetic to the ferromagnetic state (the Curie point) in magnetic crystals. However, it has also had many applications outside of physics. It has been used to model learning [16] and information storage [17] in neural networks. In molecular biology, it as been used to model cooperative conformational changes due to ligand binding in macromolecules [18] and heat denaturation of DNA [19]. It has also been used in sociology to model cultural isolation [20]. In this section we shall use it in its physics context to model order–disorder in a spin system.

An analytic expression for the partition function of an Ising system can be obtained in one and two dimensions, but no one has succeeded in solving it analytically in three dimensions. In one dimension it does not have a phase transition. In two dimensions, it does have a phase transition. In Section 7.F we considered the one-dimensional Ising model and the mean field approximation to the Ising model. In this section we will consider the case of a planar Ising spin lattice and show how an analytic expression for the partition function may be obtained and how the phase transition appears [21].

▶ S8.B.1. Partition Function

Let us consider a planar lattice of N spin-$\frac{1}{2}$ objects with magnetic moment, μ. A planar lattice is one which can be laid out flat in a two-dimensional plane without any bonds intersecting. For example, a lattice with periodic boundary conditions is not planar. Each lattice site can interact with the magnetic fields of its nearest neighbors. Some examples of planar lattices are shown in Fig. 8.5.

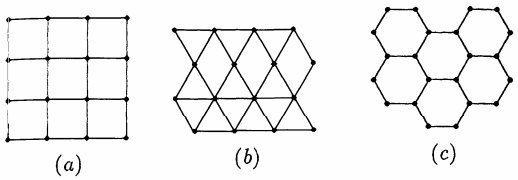


Fig. 8.5. Three types of planar lattice. (a) Square lattice. (b) Triangular lattice. (c) Hexagonal lattice.

The Hamiltonian for a two-dimensional planar lattice can be written

$$H = -\sum_{\{ij\}} J(s_i s_j - 1), \tag{8.161}$$

where $\Sigma_{\{ij\}}$ denotes the sum over nearest neighbors, i and j, J is the magnetic interaction energy between nearest-neighbor pairs, and s_i is the z component of spin at the ith lattice site. The factor, -1, in Eq. (8.161) merely shifts the zero point of energy and will prove useful in subsequent calculations. For spin$-\frac{1}{2}$ objects, $s_i = +1(-1)$ if the spin of site i is oriented in the positive (negative) z direction. If $J > 0$, the lattice will have its lowest energy when all the lattice sites have spin-oriented in like manner. If $J < 0$, the configuration in which neighboring spins are oriented opposite to one another will be favored.

The partition function for this planar lattice can be written

$$\begin{aligned} Z_N(T) &= \sum_{a.c.} \exp\left(\sum_{(ij)} K(s_i s_j - 1)\right) \\ &= e^{-N_{nn}K} \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \prod_{(ij)} e^{K s_i s_j}, \end{aligned} \tag{8.162}$$

where $K = \beta J$, $\Sigma_{a.c.}$ denotes the sum over all 2^N possible different configurations of spin on the lattice, and N_{nn} is the number of nearest neighbor pairs. The sum, $\Sigma_{(ij)}$ and the product, $\prod_{(ij)}$, are taken only over nearest-neighbor pairs. Let us now note that since $s_i = \pm 1$, we can write

$$e^{K s_i s_j} = \cosh(K) + s_i s_j \sinh(K) = \cosh(K)[1 + \tanh(K) s_i s_j]. \tag{8.163}$$

Therefore, the partition function can be written in the form

$$Z_N(T) = e^{-N_{nn}K}(\cosh(K))^{N_{nn}} \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \prod_{(ij)} [1 + \tanh(K) s_i s_j]. \tag{8.164}$$

To evaluate the partition function, we must expand the product and then sum over each spin variable, s_i. There are several properties of the spin that should be noticed and that simplify the expression for the partition function. First note that

$$\sum_{i=\pm 1} s_i^{2n} = 2 \quad \text{and} \quad \sum_{i=\pm 1} s_i^{2n+1} = 0, \tag{8.165}$$

where n is an integer (including zero). Thus, only even powers of the spin variable can appear in the partition function. Also, only nearest-neighbor pairs can appear and a given nearest-neighbor pair can appear only once. Each summation over a spin variable yields a factor of 2. Expansion of the product in Eq. (8.164) yields a sum of terms depending on varying powers of $\tanh(K)$. After summation over spins is performed, each term in the sum can be represented by a picture which we call a *closed graph*. We then can write the partition function in the form

$$Z_N(T) = 2^N e^{-N_{nn}K}(\cosh(K))^{N_{nn}} \sum(\text{all closed graphs}) \tag{8.166}$$

The closed graphs are obtained as follows:

(1) Lay out the grid of lattice sites as they appear on the original planer lattice (cf. Fig. 8.5) but do not yet draw connections between lattice sites.

(2) Find all different ways to draw nearest-neighbor bonds between lattice sites so they form closed loops (any number of closed loops including zero). We only require that no nearest-neighbor bond appears more than once. There may be more than one closed loop on a graph, and they may intersect, as long as they have *no bonds in common* (cf. Fig. 8.6) for an example of an allowed and forbidden graph).

(3) A given graph contributes a factor, $\tanh^r(K)$, where r is the number of bonds in the graph.

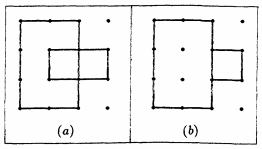
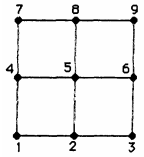


Fig. 8.6. Some closed graphs on a 3 × 3 square lattice. (a) An allowed graph. (b) A forbidden graph because two closed loops share a bond.

An example of the use of these rules is given in Exercise 8.2

In writing Eq. (8.166), we have made progress toward obtaining an analytic expression for the partition function, but we still have a ways to go. In the next section, we will change the subject slightly and show how to represent an antisymmetric matrix in terms of a *directed graph*. We will also show how to represent the determinant of an antisymmetric matrix (actually its square root) in terms of *dimer graphs*. Then in Section S8.B.3 we will make a connection between the *closed graphs* described above and *dimer graphs*.

■ **EXERCISE 8.2** Consider the 3 × 3 spin lattice shown in the figure. Assign a coupling constant, $J_x(J_y)$ to horizontal (vertical) bonds. Obtain an analytic expression for the partition function.



Answer: First note that the lattice has $N = 9$ lattice sites and $N_{nn} = 12$ nearest-neighbor pairs. The partition function can be written

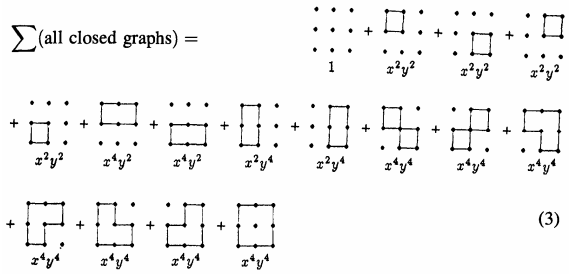
$$\begin{aligned} Z_9 = e^{-6K_x} e^{-6K_y} &\sum_{s_1=\pm 1} \cdots \sum_{s_9=\pm 1} e^{K_x s_1 s_2} e^{K_y s_1 s_4} e^{K_x s_2 s_3} e^{K_y s_2 s_5} e^{K_y s_3 s_6} \\ &\times e^{K_y s_4 s_7} e^{K_x s_4 s_5} e^{K_x s_5 s_6} e^{K_y s_5 s_8} e^{K_y s_6 s_9} e^{K_x s_7 s_8} e^{K_x s_8 s_9} \end{aligned}, \tag{1}$$

where $K_x = \beta J_x$ and $K_y = \beta J_y$. We may write this as

$$Z_9 = 2^9 e^{-6K_x} e^{-6K_y} \cosh^6(K_x) \cosh^6(K_y) \sum(\text{all closed graphs}). \tag{2}$$

Next we note that

$\sum$(all closed graphs) =



(3)

We therefore obtain the following analytic expression for Z_9:

$$Z_9 = 2^9 e^{-6K_x} e^{-6K_y} \cosh^6(K_x)\cosh^6(K_y) \times [1 + 4x^2y^2 + 2x^4y^2 + 2x^2y^4 + 7x^4y^4]$$

where $x \equiv \tanh(K_x)$ and $y \equiv \tanh(K_y)$.

▶ S8.B.2. Antisymmetric Matrices and Dimer Graphs [22]

Let us now return to the planar lattice whose partition function we wish to compute. We again draw a picture of the planar lattice, but now associate a direction to each bond. The resulting picture we call a *directed graph*. We can associate an antisymmetric matrix, A with a *directed graph* by assigning to the directed bond between lattice sites, i and j, a matrix element, a_{ij}. Only nearest-neighbor lattice sites give rise to nonzero matrix elements. We can associate the following matrix elements with the directed graph:

$$a_{ij} = \begin{cases} \alpha & \text{for bond directed } i \to j, \\ -\alpha & \text{for bond directed } i \leftarrow j, \\ = 0 & \text{if } i \text{ and } j \text{ have no bond} \end{cases} \qquad (8.167)$$

Here $\alpha = \tanh(K)$. It is possible to generalize Eq. (8.167) and assign different weights, α, to different bonds. For the case of a square directed graph, for example, for the horizontal bonds we let $\alpha = x = \tanh(K_x)$ and for the vertical bonds we let $\alpha = y = \tanh(K_y)$ (such a case is considered in Exercise 8.3)

If we want all the terms in the determinant of the antisymmetric matrix, A, to be positive, then the choice of directions given the nearest-neighbor bonds must be carefully made. Kasteleyn [22] showed that *for graphs with an even number of lattice sites, directions can be chosen in such a way that all terms in the determinant of the antisymmetric matrix, A, are positive*. This is important when we later connect the antisymmetric matrix generated by a directed graph to the partition function. The direction is chosen so that in going around any *even loop* on the directed graph in a clockwise manner, one encounters only an odd number of clockwise directed bonds. An *even loop* is a closed path containing an even number of lattice sites and enclosing an even number (including zero) of lattice sites. An even loop never passes through the same lattice site twice. Correct directions for some directed graphs are shown in Fig. 8.7. An example an antisymmetric graph obtained in this manner is given in Exercise 8.3.

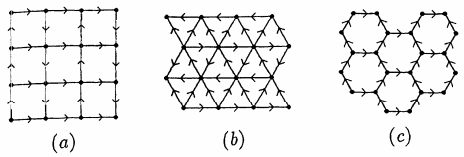
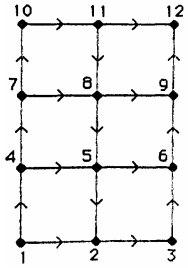


Fig. 8.7. Three types of directed planar lattice. (a) Directed square lattice. (b) Directed triangular lattice. (c) Directed hexagonal lattice

■ **EXERCISE 8.3.** Consider the directed planar lattice shown in the accompanying figure. Write the 12×12 antisymmetric matrix, A, associated with this lattice, and compute the determinant of the matrix, A.



Answer: The antisymmetric matrix associated with the accompanying graph is

$$A = \begin{pmatrix} 0 & x & 0 & y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -x & 0 & x & 0 & -y & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -x & 0 & 0 & 0 & y & 0 & 0 & 0 & 0 & 0 & 0 \\ -y & 0 & 0 & 0 & x & 0 & y & 0 & 0 & 0 & 0 & 0 \\ 0 & y & 0 & -x & 0 & x & 0 & -y & 0 & 0 & 0 & 0 \\ 0 & 0 & -y & 0 & -x & 0 & 0 & 0 & y & 0 & 0 & 0 \\ 0 & 0 & 0 & -y & 0 & 0 & 0 & x & 0 & y & 0 & 0 \\ 0 & 0 & 0 & 0 & y & 0 & -x & 0 & x & 0 & -y & 0 \\ 0 & 0 & 0 & 0 & 0 & -y & 0 & -x & 0 & 0 & 0 & y \\ 0 & 0 & 0 & 0 & 0 & 0 & -y & 0 & 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & y & 0 & -x & 0 & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -y & 0 & -x & 0 \end{pmatrix}, \qquad (1)$$

where $x = \tanh(K_x)$ and $y = \tanh(K_y)$. The determinant of this matrix is

$$\det(A) = 16x^8y^4 + 48x^6y^6 + 44x^4y^8 + 12x^2y^{10} + y^{12} = (4x^4y^2 + 6x^2y^4 + y^6)^2. \qquad (2)$$

The determinant, det(**A**), of an antisymmetric matrix, **A**, can always be written as the square of a polynomial containing the weighting factors, α (or whatever weighting factors are assigned to the graph). Furthermore, it is possible to express $\sqrt{\det(\mathbf{A})}$ (**A** obtained from a directed graph) in terms of a sum of graphs which we shall call *dimer graphs*. A dimer graph is obtained in the following way:

(1) Lay out the grid of lattice sites (but do not yet draw bonds) as they appear in the original *directed graph*.
(2) Connect all pairs of lattice sites by bonds so that each lattice site is attached to only one bond and all lattice sites are members of a bond. The lattice consists of a set of unconnected bonds (we call them dimer bonds).

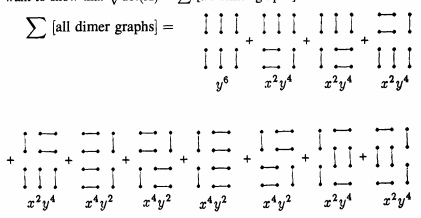
The quantity, $\sqrt{\det(\mathbf{A})}$, can be represented by the set of all different dimer graphs which may be obtained from the original planar lattice. Thus,

$$\sqrt{\det(A)} = \sum(\text{all different dimer graphs}). \tag{8.168}$$

Once a dimer graph is drawn, an analytic expression can be associated with it. If a graph has s dimer bonds, then we can assign a factor, $\tanh^s(K)$, to the graph. As we show in Exercise 8.4, we can also generalize this to give different weights to different bonds. For example, let us consider a square dimer graph with different bond strengths, K_x, for horizontal and, K_y, verticle bonds. If the graph has r horizontal dimer bonds and s verticle dimer bonds, we associate a factor, $\tanh^r(K_x)\tanh^s(K_y)$ with the graph. In Exercise 8.4, we construct the dimer graphs that represent $\sqrt{\det(\mathbf{A})}$ for the antisymmetric matrix, **A**, obtained in Exercise 8.3.

■ **EXERCISE 8.4.** Consider the directed lattice shown in Exercise 8.3 and the antisymmetric matrix, **A**, generated by that directed lattice. Draw all dimer graphs that represent $\sqrt{\det(\mathbf{A})}$ and show that they do indeed reproduce the analytic expression obtained in Exercise 8.3.
Answer: For the antisymmetric matrix, A, generated in Exercise 8.3, we want to show that $\sqrt{\det(\mathbf{A})} = \sum$ [all dimer graphs]. Note that

$\sum$ [all dimer graphs] =



(2)

Thus, if we add the contributions from the various graphs, we find

$$\sum[\text{all dimer graph}] = 4x^4y^2 + 6x^2y^4 + y^6,$$

in agreement with Exercise 8.3.

▶ S8.B.3. Closed Graphs and Mixed Dimer Graphs

Kasteleyn [22] and Fisher [23] have found a way to write the partition function in terms of dimer graphs and thereby have found a way to obtain an expression for the partition function in terms of the determinant of an antisymmetric matrix. The method is based on the following observation. Consider a lattice with two different coupling constants, J and J^*, for example. A bond with coupling constant, J^*, will contribute factor, $2e^{-K^*}\cosh(K^*)$ and $\tanh(K^*)$, where $K^* = \beta J^*$, to the partition function. [Note that the factor -1 in Eq. (8.161)] gives rise to this property.] However, note that

$$\lim_{K^*\to\infty} 2e^{-K^*}\cosh(K^*) = 1 \quad \text{and} \quad \lim_{K^*\to\infty} \tanh(K^*) = 1. \tag{8.169}$$

The next thing to notice is that the partition function of a given planar lattice depends on a sum of terms depending on different powers of the weighting factor, $\tanh(K)$. However, the dimer graphs which are generated from a directed graph all depend on the same power of the weighting factor, $\tanh(K)$. Therefore we need to construct a directed graph which can generate dimer graphs depending on different powers of the weighting factors. The limiting behavior in Eq. (8.169) enables us to do this in a unique and unambiguous way for planar graphs.

The expansion of the partition function in terms of dimer graphs requires two important pieces of information. On the one hand, we use the theorem of Kasteleyn which says that it is possible to draw a directed graph which generates an antisymmetric matrix whose determinant has all positive terms. On the other hand, we use a method of expanding the vertices of a planar lattice so that a directed graph can be associated with it (we call it a *mixed directed graph*) which produces dimer graphs (we call them *mixed dimer graphs*) whose analytic expressions are in one-to-one correspondence with the closed graphs for that planar lattice. Kasteleyn [21] and Fisher [22] had different ways to accomplish this. In this section, we use Fisher's approach [23].

We can construct mixed directed graphs according to the following procedure.

(1) Draw the planar lattice of interest. (Figure 8.5 gives some examples of planar lattices.) The planar lattice will have N vertices and N_{nn} nearest-neighbor bonds. On the planar lattice we shall call the bonds *primary bonds*. Represent primary bonds by solid lines.
(2) Replace each vertex of degree $q(q \geq 4)$ by a cluster of $q-2$ vertices connected by $q-3$ *internal* bonds. (A vertex of degree, q, is one at which q bonds meet.) Represent internal bonds by dashed lines. (See Fig. 8.8 for an example of this replacement.) With this replacement, all vertices are now either of degree $q = 2$ or 3.
(3) Replace each vertex of degree, $q = 2$, by two vertices connected by one internal bond. Replace each vertex of degree, $q = 3$, by a cluster of three vertices connected by three internal bonds. (See Fig. 8.9 for examples.)
(4) Assign directions to all the bonds (primary and internal) so that counterclockwise circulation around any *even loop* yields an odd number of clockwise directed bonds. (According to the theorem of Kasteleyn, this can always be done.)

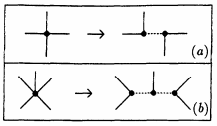


Fig. 8.8. Replace vertices of degree, $q(q \geq 4)$, by a cluster of $q-2$ vertices connected by $q-3$ internal bonds. (a) $q=4$. (b) $q=5$.

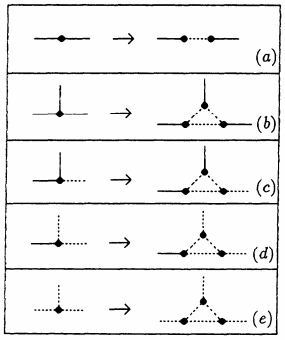


Fig. 8.9. (a) Replace vertex of degree 2 with two vertices and an internal bond. (b)–(e) Replace vertex of degree 3 with three vertices and three internal bonds.

The *mixed directed graph* which results from steps 1–4 is a collection of directed solid and dashed lines connecting vertices of degree 3 or smaller.

Once a *mixed directed graph* has been associated with a planar lattice, we can associate an *antisymmetric matrix*, A_m, with the directed mixed graph in the following manner. The antisymmetric matrix, A_m, is constructed by assigning to matrix elements, a_{ij}, the following values:

$$a_{ij} = \begin{cases} 1/\alpha & \text{for a solid bond directed } i \to j, \\ -1/\alpha & \text{for a solid bond directed } i \leftarrow j, \\ 1 & \text{for dashed bond directed } i \to j, \\ -1 & \text{for a dashed bond directed } i \leftarrow j, \\ 0 & \text{if } i \text{ and } j \text{ have no bond.} \end{cases} \tag{8.170}$$

As we found in Section S8.B.2, we can associate a collection of dimer graphs (we shall call them *mixed dimer graphs*) with each directed mixed graph. Dimers can now be solid or dashed depending on whether they occur at the position of primary or internal bonds, respectively, on the mixed directed graph which generated them. No more than one dimer bond attaches to any given vertex. To each dimer bond we assign a factor $1/\alpha$, if the dimer bond is a solid line and assign a factor 1 if the dimer bond is a dashed line. A given dimer graph has a factor $(1/\alpha)^m(1)^n$ associated with it if it has m primary (solid) dimer bonds and n internal (dashed) dimer bonds.

We can now relate $\sqrt{\det(\mathbf{A}_m)}$ to the sum over all mixed dimer graphs that can be derived from the mixed directed graph which generates the antisymmetric matrix, $\mathbf{A}_m$. We write

$$\sqrt{\det(\mathbf{A}_m)} = \sum(\text{all mixed dimer graphs}). \tag{8.171}$$

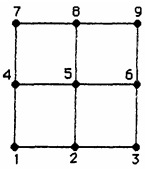
The mixed dimer graphs that result from a given planar lattice are in one-to-one correspondence with the closed graphs which are generated by that same planar lattice. The correspondence is "inverted," however. The primary (solid) dimer bonds on the mixed dimer graph correspond to the *lack* of bonds on the closed graph.

We can now relate $\sqrt{\det(\mathbf{A}_m)}$ to the partition function associated with the planar lattice. We can write

$$\begin{aligned} Z_N(T) &= 2^N e^{-N_{nn}K}(\cosh(K))^{N_{nn}} \alpha^s \sqrt{\det(\mathbf{A}_m)} \\ &= 2^N e^{-N_{nn}K}(\cosh(K))^{N_{nn}} \alpha^s \sum(\text{all mixed dimer graphs}), \end{aligned} \tag{8.172}$$

where s is the maximum number of primary (solid) dimer bonds that can appear on any of the mixed dimer graphs. With this definition, the dimer graph in which all dimers are primary (solid) bonds actually corresponds to the unit graph. In Exercise 8.5 we show these steps for the planar lattice considered in Exercise 8.2.

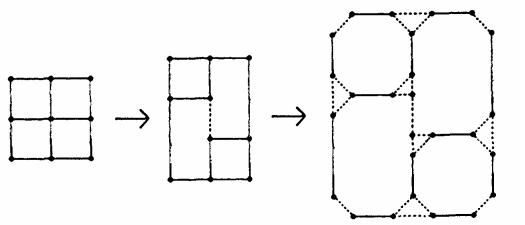
■ **EXERCISE 8.5.** Consider the planar lattice shown in the accompanying figure. (a) Construct a mixed directed graph for this lattice and construct the associated antisymmetric matrix, $\mathbf{A}_m$. (b) Compute the partition function from this antisymmetric matrix, $\mathbf{A}_m$. (c) Draw the mixed dimer graphs for this lattice and write an analytic expression for the partition function.



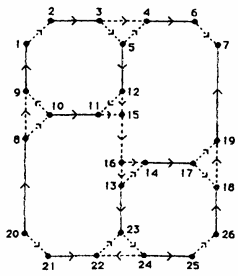
Answer:

(a) We first obtain the directed mixed lattice. We replace the vertex of degree 4 by two vertices of degree 3 and an internal line. We then replace each vertex of degree 2 by two vertices of degree 2 connected by an internal line. We replace each vertex of degree 3 by two vertices of degree 3 connected by three internal bonds. The results is a graph with 26 vertices.

This procedure is shown in the graphs below.



We next give directions to all the lines in the mixed graph so that for any *even loop*, clockwise circulation encounters an odd number of clockwise directed bonds. Such a choice is shown below. We have also numbered the vertices.



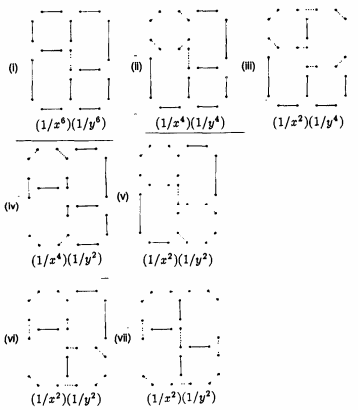
(b) Given the numbered mixed directed graph above, we can use it to construct an antisymmetric matrix, $\mathbf{A}_m$. Choose the matrix elements according to the directions of the bonds and their type. Assign a value, $\pm 1/x$, to horizontal bonds and assign a factor, $\pm\frac{1}{y}$ to vertical bonds. The nonzero matrix elements are $a_{1,2} = +1$, $a_{2,1} = -1$, $a_{2,3} = +1/x$, $a_{3,2} = -1/x$, $a_{3,5} = +1$, $a_{5,3} = -1$, $a_{5,12} = +1/y$, $a_{12,5} = -1/y$, $a_{1,3} = a_{3,1} = 0$, and so on. The matrix, A_L is too big to write here, but it is easy to construct given the directed graph above. Once it has been constructed we can find its determinant. It is

$$\det(A_L) = \frac{1}{x^{12}y^{12}}(1 + 4x^2y^2 + 2x^4y^2 + 2x^2y^4 + 7x^4y^4)^2. \qquad (1)$$

The partition function is

$$\begin{aligned} Z_N(T) &= 2^9 e^{-6K_x} e^{-6K_y} \cosh^6(K_x)\cosh^6(K_y) x^6 y^6 \\ &\quad \times \left[\frac{1}{x^6y^6}(1 + 4x^2y^2 + 2x^4y^2 + 2x^2y^4 + 7x^4y^4)\right] \\ &= 2^9 e^{-6K_x} e^{-6K_y} \cosh^6(K_x)\cosh^6(K_y) \\ &\quad \times (1 + 4x^2y^2 + 2x^4y^2 + 2x^2y^4 + 7x^4y^4). \end{aligned} \qquad (2)$$

Let us now draw the dimer graphs. There is one dimer graph for each closed graph in Eq. (3) of Exercise 8.2. We shall draw some of them here. We have



The analytic expression associated with each mixed dimer graph is shown below it. Graph (i) above corresponds to the first closed graph in Exercise 8.2, Eq. (3). Four mixed dimer graphs of the type (ii) above correspond to closed graphs, 2–5 in Exercise 8.2, Eq. (3). The two mixed dimer graphs of the type (iii) above correspond to closed graphs 6 and 7, and the two mixed dimer graphs of the type (iv) above correspond to closed graphs 8 and 9 in Exercise 8.2, Eq. (3). Two mixed dimer graphs of the type (v) above correspond to closed graphs 10 and 11 in Exercise 8.2, Eq. (3). Four mixed dimer graphs of the type (vi) above correspond to closed graphs 12–15 in Exercise 8.2, Eq. (3). And finally, graph (vii) corresponds to the last graph in Exercise 8.2, Eq. (3). Thus there are 16 possible mixed dimer graphs. They each consist of 26 vertices and a mixture of solid and dashed dimer bonds. The partition function is

$$\begin{aligned} Z_N(t) &= 2^9 e^{-K_x} e^{-6K_y} \cosh^6(K_x)\cosh^6(K_y) \\ &\quad \times x^6 y^6 \sum(\text{all mixed dimer graphs}). \end{aligned} \qquad (3)$$

▶ S8.B.4. Partition Function for the Infinite Planar Lattice

We now have enough information to obtain an analytic expression for the partition function of an infinite planar lattice. For simplicity we will consider a square planar lattice. The procedure is to extract an $n \times n$ segment from the infinite lattice, and then take the limit, $n \to \infty$. In order to demonstrate the procedure, we will first extract a 3×3 segment from the infinite planer lattice and describe its properties in detail. Then we will go on to write partition function for an $n \times n$ segment and take the limit, $n \to \infty$.

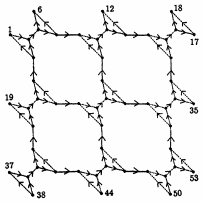


Fig. 8.10. A 3×3 segment of an infinite square directed graph corresponds to a 54×54 mixed directed graph.

$$\Gamma \equiv \Gamma_6 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 & 0 \end{pmatrix},$$

$$X \equiv X_6 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/x & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad (8.174)$$

$$Y \equiv Y_6 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/y & 0 & 0 & 0 & 0 \end{pmatrix},$$

The subscript 6 is added to show that Γ_6, X_6, and Y_6 are 6×6 matrices. This notation will be useful later. In constructing the mixed directed graph, Fig. 8.10, we have left out its connections to the rest of the infinite mixed directed graph, and the antisymmetric matrix, $\mathbf{A}_{54}$, does not include these bonds.

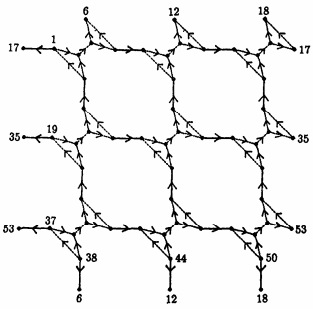


Fig. 8.11. The 54×54 directed lattice with bonds connecting opposite sides added.

▶ *S8.B.4.1. A 3×3 Segment of the Infinite Square Planar Lattice*

Let us consider a 3×3 segment of an infinite square planar lattice. We first construct a 54×54 mixed directed graph as shown in Fig 8.10 using the rules of Section S8.B.3. We can then assign an antisymmetric matrix, $\mathbf{A}_{54}$, to the directed mixed graph. We obtain

$$A_{54} = \begin{pmatrix} \Gamma & X & 0 & -Y^T & 0 & 0 & 0 & 0 & 0 \\ -X^T & \Gamma & X & 0 & -Y^T & 0 & 0 & 0 & 0 \\ 0 & -X^T & \Gamma & 0 & 0 & -Y^T & 0 & 0 & 0 \\ Y & 0 & 0 & \Gamma & X & 0 & -Y^T & 0 & 0 \\ 0 & Y & 0 & -X^T & \Gamma & X & 0 & -Y^T & 0 \\ 0 & 0 & Y & 0 & -X^T & \Gamma & 0 & 0 & -Y^T \\ 0 & 0 & 0 & Y & 0 & 0 & \Gamma & X & 0 \\ 0 & 0 & 0 & 0 & Y & 0 & -X^T & \Gamma & X \\ 0 & 0 & 0 & 0 & 0 & Y & 0 & -X^T & \Gamma \end{pmatrix}, \qquad (8.173)$$

where "0" in the matrix (8.173) represents a 6×6 matrix with all elements zero:

Let us now take our 54×54 directed lattice and change it by giving it periodic boundary conditions. We will connect opposite sides to one another by adding bonds and then add the contributions from these new bonds to the antisymmetric matrix. It turns out that as we consider larger and larger segments of the infinite planar lattice, these additional terms contribute less and less to the determinant of the antisymmetric matrix, and in the limit of an infinite lattice they do not affect the result. In Fig. 8.11 we redraw the 54×54 directed lattice so that bonds on opposite sides are connected. We then no longer have a planar lattice. Note the directions of these added bonds are opposite to normal connections to the infinite lattice. The antisymmetric matrix, A_{54}, that results from adding these new bonds to A_{54} is given by

$$A_{54} = \begin{pmatrix} \Gamma & X & X^T & -Y^T & 0 & 0 & -Y & 0 & 0 \\ -X^T & \Gamma & X & 0 & -Y^T & 0 & 0 & -Y & 0 \\ -X & -X^T & \Gamma & 0 & 0 & -Y^T & 0 & 0 & -Y \\ Y & 0 & 0 & \Gamma & X & X^T & -Y^T & 0 & 0 \\ 0 & Y & 0 & -X^T & \Gamma & X & 0 & -Y^T & 0 \\ 0 & 0 & Y & -X & -X^T & \Gamma & 0 & 0 & -Y^T \\ Y^T & 0 & 0 & Y & 0 & 0 & \Gamma & X & X^T \\ 0 & Y^T & 0 & 0 & Y & 0 & -X^T & \Gamma & X \\ 0 & 0 & Y^T & 0 & 0 & Y & -X & -X^T & \Gamma \end{pmatrix}, \qquad (8.175)$$

The antisymmetric matrix, $\tilde{A}_{54}$, is now in a form that is easy to deal with. First note that we can write $\tilde{A}_{54}$ as the tensor product of smaller matrices (tensor products are defined in the subsection below). We can write

$$\begin{aligned}\tilde{A}_{54} &= \begin{pmatrix} D_{18} & -I_3 \otimes Y_6^T & -I_3 \otimes Y_6 \\ I_3 \otimes Y_6 & D_{18} & -I_3 \otimes Y_6^T \\ I_3 \otimes Y_6^T & I_3 \otimes Y_6 & D_{18} \end{pmatrix} \\ &= I_3 \otimes D_{18} + E_3 \otimes (I_3 \otimes Y_6^T) - E_3^{-1} \otimes (I_3 \otimes Y_6), \end{aligned} \tag{8.176}$$

where

$$D_{18} = \begin{pmatrix} \Gamma_6 & X_6 & X_6^T \\ -X_6^T & \Gamma_6 & X_6 \\ -X_6 & -X_6^T & \Gamma_6 \end{pmatrix}, \tag{8.177}$$

$$E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{8.178}$$

■ **Tensor Product of Matrices.** We introduce the tensor product of two matrices, $\bar{A}$ and $\bar{B}$, as follows. We write $\bar{\bar{C}} = \bar{A} \otimes \bar{B}$, where $\bar{\bar{C}}$ is a fourth-order tensor whose $(ijkl)$th element is given by

$$(\bar{\bar{C}})_{ijkl} = (\bar{A} \otimes \bar{B})_{ijkl} = A_{ij}B_{kl}. \tag{1}$$

where A_{ij} and B_{kl} are the (ij)th and (kl)th elements of matrices $\bar{A}$ and $\bar{B}$, respectively. The product of two fourth-order tensors, $\bar{\bar{\mathbb{Q}}}$ and $\bar{\bar{\mathbb{R}}}$, is given by

$$(\bar{\bar{\mathbb{Q}}}\bar{\bar{\mathbb{R}}})_{ijkl} = \sum_{m,n} \mathbb{Q}_{imkn}\mathbb{R}_{mjnl}, \tag{2}$$

where $\mathbb{Q}_{imkn}$ and $\mathbb{R}_{mjnl}$ are the $(imkn)$th and $(mjnl)$th elements of the tensors $\bar{\bar{\mathbb{Q}}}$ and $\bar{\bar{\mathbb{R}}}$, respectively. The tensor product of a matrix $\bar{A}$ with the unit matrix $\bar{I}$ is

$$(\bar{A} \otimes \bar{I})_{ijkl} = A_{ij}\delta_{k,l}, \tag{3}$$

where $\delta_{k,l}$ is the Kronecker delta function. *If $\bar{B}$ and $\bar{C}$ commute*, then we can write

$$(\bar{A} \otimes \bar{B})(\bar{C} \otimes \bar{D}) = (\bar{A}\bar{C}) \otimes (\bar{B}\bar{D}). \tag{4}$$

The matrix D_{18} can also be written as a tensor product:

$$D_{18} = I_3 \otimes \Gamma_6 - E_3 \otimes X_6 + E_3^{-1} \otimes X_6^T. \tag{8.179}$$

Therefore, the antisymmetric matrix $(\tilde{A})_{54}$ takes the form

$$\begin{aligned}\tilde{A}_{54} = {} & I_3 \otimes (I_3 \otimes \Gamma_6) - I_3 \otimes (E_3 \otimes X_6) + I_3 \otimes (E_3^{-1} \otimes X_6^T) + E_3 \otimes (I_3 \otimes Y_6^T) \\ & - E_3^{-1} \otimes (I_3 \otimes Y_6), \end{aligned} \tag{8.180}$$

where we have used the fact that $E_3^2 = E_3^{-1}$. Let us next introduce the matrix P_3, which diagonalizes E_3. That is,

$$P_3^{-1}E_3P_3 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{8.181}$$

where λ_k is the kth eigenvalue ($k = 1, 2, 3$) of the matrix E_3 and is given by $\lambda_k = \exp(2\pi ik/3)$. Then we find that

$$\begin{aligned}\mathbf{A}'_{54} &\equiv (P_3^{-1} \otimes (P_3^{-1} \otimes I_6))\tilde{A}_{54}(P_3 \otimes (P_3 \otimes I_6)) \\ &= I_3 \otimes I_3 \otimes \Gamma_6 - I_3 \otimes (P_3^{-1}E_3P_3) \otimes X_6 \\ &\quad + I_3 \otimes (P_3^{-1}E_3^{-1}P_3) \otimes X_6^T \\ &\quad + (P_3^{-1}E_3P_3) \otimes I_3 \otimes Y_6^T - (P_3^{-1}E_3^{-1}P_3) \otimes I_3 \otimes Y_6. \end{aligned} \tag{8.182}$$

We next note, for example, that

$$I_3 \otimes \Gamma_6 = \begin{pmatrix} \Gamma_6 & 0 & 0 \\ 0 & \Gamma_6 & 0 \\ 0 & 0 & \Gamma_6 \end{pmatrix} \tag{8.183}$$

and

$$(P_3^{-1}E_3P_3) \otimes X_6 = \begin{pmatrix} \lambda_1 X_6 & 0 & 0 \\ 0 & \lambda_2 X_6 & 0 \\ 0 & 0 & \lambda_3 X_6 \end{pmatrix}. \tag{8.184}$$

Then one can readily show we have obtained an expression for A'_{54} which is block diagonal. It can be written

$$\mathbf{A}'_{54} = \begin{pmatrix} D^{(3)}_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & D^{(3)}_{1,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & D^{(3)}_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & D^{(3)}_{2,1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D^{(3)}_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & D^{(3)}_{2,3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & D^{(3)}_{3,1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & D^{(3)}_{3,2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & D^{(3)}_{3,3} \end{pmatrix}, \tag{8.185}$$

where

$$\begin{aligned} D^{(3)}_{i,j} &= \Gamma_6 - \lambda_i X_6 + \frac{1}{\lambda_i}X_6^T + \lambda_j Y_6^T - \frac{1}{\lambda_j}Y_6 \\ &= \begin{pmatrix} 0 & 1 & 1 & 0 & x/\lambda_i & 0 \\ -1 & 0 & 1 & 0 & 0 & y\lambda_j \\ -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \\ -x\lambda_i & 0 & 0 & -1 & 0 & 1 \\ 0 & -y/\lambda_j & 0 & -1 & -1 & 0 \end{pmatrix}. \end{aligned} \tag{8.186}$$

The determinant of $D^{(3)}_{i,j}$ is given by

$$\begin{aligned}\det[D^{(3)}_{i,j}] = {} & 1 + \left(\frac{1}{x}\right)^2 + \left(\frac{1}{y}\right)^2 + \left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)^2 \\ & - 2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi i}{3}\right) - 2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi j}{3}\right) \\ & + 2\left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi j}{3}\right) + 2\left(\frac{1}{y}\right)^2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi i}{3}\right). \end{aligned} \tag{8.187}$$

Let us now note that due to the cyclic permeability of matrices under the determinant and the fact the determinant of a product of matrices is equal to the product of determinants of those matrices, we obtain

$$\begin{aligned}\det[\tilde{A}_{54}] = \det[A'_{54}] = \prod_{i=1}^{3}\prod_{j=1}^{3}\Bigg[& 1 + \left(\frac{1}{x}\right)^2 + \left(\frac{1}{y}\right)^2 + \left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)^2 - 2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi i}{3}\right) \\ & - 2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi j}{3}\right) + 2\left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi j}{3}\right) + 2\left(\frac{1}{y}\right)^2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi i}{3}\right)\Bigg]. \end{aligned} \tag{8.188}$$

These results are easily generalized to an $n \times n$ segment of the infinite planar lattice.

▶ ***S8.B.4.2. An $n \times n$ Segment of an Infinite Square Planar Lattice***

Let us now consider an $n \times n$ segment of the infinite square spin lattice. The mixed directed lattice that we construct will have $6n^2$ lattice sites, and therefore the antisymmetric matrix that we associate with the directed lattice will be $6n^2 \times 6n^2$-dimensional. We proceed exactly as before. We introduce the $n \times n$-dimensional matrix, E_n, which consists of all zeros except for (i) matrix elements -1 along the nearest upper off-diagonal and (ii) the matrix element 1 in the bottom left corner. We shall restrict outselves of odd values of n. Then the eigenvalues of E_n are given by $\lambda_k = \exp(2\pi ik/n)$, where $k = 1, 2, \ldots, n$. The determinant of the antisymmetric matrix, $\tilde{A}_{6n^2}$, obtained by connecting opposite bonds of the $6n^2 \times 6n^2$-dimensional directed lattice will be given by

$$\det[\tilde{A}_{6n^2}] = \prod_{j=1}^{n}\prod_{k=1}^{n}\left[1 + \left(\frac{1}{x}\right)^2 + \left(\frac{1}{y}\right)^2 + \left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)^2 - 2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi j}{n}\right) - 2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi k}{n}\right) + 2\left(\frac{1}{x}\right)^2\left(\frac{1}{y}\right)\cos\left(\frac{2\pi k}{n}\right) + 2\left(\frac{1}{y}\right)^2\left(\frac{1}{x}\right)\cos\left(\frac{2\pi j}{n}\right)\right] = \prod_{j=1}^{n}\prod_{k=1}^{n} D_{jk}^{(n)}. \tag{8.189}$$

The partition function can be written

$$Z_n(T) = 2^{n^2} e^{-K_x n(n-1)} e^{-K_y n(n-1)} \cosh^{n(n-1)}(K_x)\cosh^{n(n-1)}(K_y) \times x^{n(n-1)} y^{n(n-1)} \sqrt{\det[\tilde{\mathbf{A}}_{6n^2}]}. \tag{8.190}$$

If we remember that $x = \tanh(K_x)$ and $y = \tanh(K_y)$ and note that $e^{-K}\sinh(K) = \tanh(K)/[1 + \tanh(K)]$, then with the help of Eq. (8.189) we can write the partition function as

$$Z_n(T) = 2^{n^2}\left(\frac{x}{1+x}\right)^{n(n-1)}\left(\frac{y}{1+y}\right)^{n(n-1)}\sqrt{\prod_{j=1}^{n}\prod_{k=1}^{n} D_{jk}^{(n)}}. \tag{8.191}$$

The logarithm of the partition function becomes

$$\ln[Z_n(T)] = n^2\ln(2) + n(n-1)\ln\left(\frac{x}{1+x}\right) + n(n-1)\ln\left(\frac{y}{1+y}\right) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\ln(D_{ij}^{(n)}). \tag{8.192}$$

We now can take the thermodynamic limit. It is useful to note that $\lim_{n\to\infty}\sum_{i=1}^{n}(1/n)\cos(2\pi i/n) = (1/2\pi)\int_0^{2\pi} d\phi\cos(\phi)$. Then for the free energy density (free energy per lattice site), $a(T)$, we find

$$-\frac{a(T)}{k_B T} = \lim_{n\to\infty}\left[\frac{1}{n^2}\ln[Z_n(T)]\right] = \ln 2 + \ln\left(\frac{x}{1+x}\right) + \ln\left(\frac{y}{1+y}\right) + \frac{1}{2}\frac{1}{4\pi^2}\int_0^{2\pi} d\phi_1\int_0^{2\pi} d\phi_2 \ln\left[1 + \frac{1}{x^2} + \frac{1}{y^2} + \frac{1}{x^2y^2} - \frac{2}{x}\cos(\phi_1) - \frac{2}{y}\cos(\phi_2) + \frac{2}{x^2y}\cos(\phi_2) + \frac{2}{xy^2}\cos(\phi_1)\right]. \tag{8.193}$$

We can simplify subsequent calculations if we let $J_x = J_y$. Then $x = y = \tanh(K)$ and the free energy density takes the form

$$-\frac{a(T)}{k_B T} = \ln(2) + 2\ln\left(\frac{x}{1+x}\right) + \frac{1}{2}\frac{1}{4\pi^2}\int_0^{2\pi} d\phi_1\int_0^{2\pi} d\phi_2 \ln\left[1 + \frac{2}{x^2} + \frac{1}{x^4} - \left(\frac{2}{x} - \frac{2}{x^3}\right)(\cos(\phi_1) + \cos(\phi_2))\right] = \ln(2) - 2\ln(1+x) + \frac{1}{2}\frac{1}{4\pi^2}\int_0^{2\pi} d\phi_1 \times \int_0^{2\pi} d\phi_2 \ln\left[1 + 2x^2 + x^4 + 2x(1-x^2)(\cos(\phi_1) + \cos(\phi_2))\right].$$

We now can make use of the identities

$$2x(1-x^2) = \sinh(2K)/\cosh^4(K) \quad \text{and} \quad 1 + 2x^2 + x^4 = \cosh^2(2K)/\cosh^4(K). \tag{8.195}$$

Then the free energy density becomes

$$-\frac{a(T)}{k_B T} = \ln(2) - 2K + \frac{1}{2}\frac{1}{4\pi^2}\int_0^{2\pi} d\phi_1\int_0^{2\pi} d\phi_2 \ln[\cosh^2(2K) + \sinh(2K)(\cos(\phi_1) + \cos(\phi_2))].$$

Equation (8.196) can be simplified still further. Note that $\cos(\phi_1) + \cos(\phi_2) = 2\cos[\frac{1}{2}(\phi_1 + \phi_2)]\cos[\frac{1}{2}(\phi_1 - \phi_2)]$. If we make the change of variables, $\theta_1 = \frac{1}{2}(\phi_1 - \phi_2)$ and $\theta_2 = \frac{1}{2}(\phi_1 + \phi_2)$, and let $\kappa = 2\sinh(2K)/\cosh^2(2K)$, then the free energy density becomes

$$-\frac{a(T)}{k_B T} = \ln(2) - 2K + \ln[\cosh(2K)] + \frac{1}{4\pi^2}\int_{-\pi}^{\pi} d\theta_1\int_0^{2\pi} d\theta_2 \ln[1 + \kappa\cos(\theta_1)\cos(\theta_2)]. \tag{8.197}$$

We can integrate Eq. (8.197) if we note that [24]

$$\int_0^{2\pi} dx \ln(a + b\cos(x)) = 2\pi\ln[(a + \sqrt{a^2 - b^2})/2]. \tag{8.198}$$

Then we obtain

$$-\frac{a(T)}{k_B T} = \ln(2) - 2K + \ln[\cosh(2K)] + \frac{2}{\pi}\int_0^{\pi/2} d\theta \ln\left[\frac{1}{2}\left(1 + \sqrt{1 - \kappa^2\sin^2(\theta)}\right)\right]. \tag{8.199}$$

The internal energy density is easily obtained from Eq. (8.199). We find

$$u(T) = J\frac{\partial}{\partial K}\left(\frac{a(T)}{k_B T}\right) = 2J - J\coth(2K)[1 + (2\tanh^2(2K) - 1)\frac{2}{\pi}\mathrm{K}(\kappa)], \tag{8.200}$$

where $\mathrm{K}(\kappa)$ is the complete elliptic integral of the first kind [25] and is defined as

$$\mathrm{K}(\kappa) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - \kappa^2\sin^2(\theta)}}. \tag{8.201}$$

The specific heat is given by

$$c = \frac{\partial u}{\partial T} = -\frac{K}{T}\frac{\partial u}{\partial K} = \frac{2}{\pi}k_B K^2\coth^2(2K)\Big\{2\mathrm{K}(\kappa) - 2\mathrm{E}(\kappa) - 2\,\mathrm{sech}^2(2K)\Big[\frac{\pi}{2} + (2\tanh(2K) - 1)\mathrm{K}(\kappa)\Big]\Big\}.$$

The complete elliptic integral of the first kind, $\mathrm{K}(\kappa)$, has a singularity for $\kappa = 1$. The phase transition occurs at this point. The temperature at which the phase transition occurs is then given by the condition

$$\kappa_c = \frac{2\sinh(2K_c)}{\cosh^2(2K_c)} = 1. \tag{8.203}$$

This gives a critical temperature, $T_c = 2.269J/k_B$. In Fig 8.12 we plot the internal energy per site and the specific heat as a function of K, which is proportional to the inverse temperature.

Ising-like phase transitions have been measured in the two-dimensional Ising-like antiferromagnets, K_2CoF_4 and Rb_2CoF_4. These substances behave like two-dimensional spin systems because they consist of strongly coupled antiferromagnetic CoF_2 planes separated by weakly coupled planes containing the remaining molecules. In Fig 8.13 we show two measurements of the heat capacity of Rb_2CoF_4. We see the characteristic Ising-like singularity in the heat capacity.

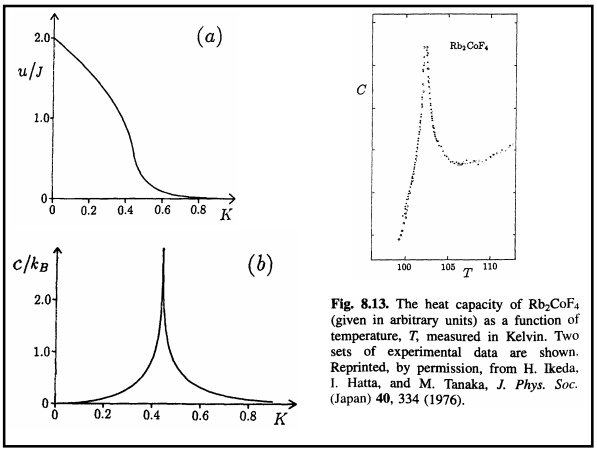


Fig. 8.13. The heat capacity of Rb_2CoF_4 (given in arbitrary units) as a function of temperature, T, measured in Kelvin. Two sets of experimental data are shown. Reprinted, by permission, from H. Ikeda, I. Hatta, and M. Tanaka, *J. Phys. Soc.* (Japan) **40**, 334 (1976).

In the section we have obtained analytic expressions for the free energy, the internal energy, and the heat capacity of a two-dimensional square planar Ising lattice. From Eq. (8.202) we can obtain the critical exponent, α. The type of singularity that occurs in the heat capacity is easy to find from properties of the complete elliptic integrals. Let us note that as $\kappa \to 1$, $E(\kappa) \to 1$ but $K(\kappa) \to \ln(4/\sqrt{1-\kappa^2})$. Thus, the singularity in the heat capacity is logarithmic and the critical exponent, α, for the two-dimensional Ising lattice is $\alpha = 0$.

It is interesting to compare Fig. 8.12, the exact heat capacity of the two-dimensional Ising lattice, to that of the mean field approximation given in Fig. 7.8. Mean field theory predicts a finite discontinuity, while the exact solution gives a logarithmic singularity. The two theories also give different critical temperatures.