

# Chapter 2

## Statistical Mechanics

### 2.1 Distributions

The formalism described in Chap. 1 can be regarded as a macroscopic perspective on a system with a microscopic structure consisting of a collection of atoms or molecules of one or more types. The numbers of particles of each type, or component, will be among the set of extensive variables  $Q_1, \dots, Q_{n_f}$ .

A configuration of the particles of the system is called a *microstate* and is denoted by  $\sigma$ . For each microstate there is a value  $E(\sigma)$  of the *statistical mechanical internal energy*  $E$  and values  $\widehat{Q}_i(\sigma)$ ,  $i = 1, \dots, \eta$ , of the dependent statistical mechanical extensive variables.<sup>1</sup> The Hamiltonian<sup>2</sup>

$$\widehat{H}_\eta(\sigma) := E(\sigma) - \sum_{i=1}^{\eta} \xi_i \widehat{Q}_i(\sigma), \quad (2.1.1)$$

is a crucial quantity in the statistical mechanics of the system with independent variables  $Q_{\eta+1}, \dots, Q_{n_f}, T, \xi_1, \dots, \xi_\eta$ . It is the microscopic equivalent of the thermodynamic enthalpy defined by (1.1.8).

As indicated above we shall be concerned with the case where the system is in equilibrium and then the probability function for the *distribution* of  $\sigma$  is

$$p_\eta(T; \sigma) := \frac{\exp[-\widehat{H}_\eta(\sigma)/T]}{Z_\eta(T)}, \quad (2.1.2)$$

<sup>1</sup> In cases where we wish to use the same letter for a thermodynamic variable and its classical statistical mechanical counterpart we make the distinction by adding a ‘hat’ to the statistical mechanical variable. In particular this relationship applies to the enthalpy  $H_\eta$  and the Hamiltonian  $\widehat{H}_\eta$  (Eq. 2.1.7).

<sup>2</sup> For a comment concerning the usage of the subscript  $\eta$ , specifying the number of independent fields (and hence the number of independent extensive variables, or densities) see the footnote on page 7. For the sake of brevity it is convenient in most cases to drop reference to the independent fields and extensive variables as arguments, retaining only  $T$ .

where

$$Z_\eta(T) := \sum_{\{\sigma\}} \exp[-\widehat{H}_\eta(\sigma)/T], \quad (2.1.3)$$

is the *partition function*.<sup>3</sup> This is usually referred to as the *Gibbs distribution*. In the case when the system is a perfect gas it is the *Maxwell-Boltzmann distribution*. The classic account which justifies the role of (2.1.2) as the equilibrium distribution in statistical mechanics is that of Khinchin (1949), who established the result for the case  $\eta = 0$ , called the *canonical distribution*. This situation can be pictured as a system in a heat bath which controls the temperature. Other cases, where, for example,  $\eta = 1$  and  $\xi_1$  is the chemical potential (called the *grand-canonical distribution*, where in addition to the heat bath the system has permeable walls with the passage of particles controlled by the chemical potential), or  $\eta = 1$  and  $\xi_1$  is the pressure (called the *constant-pressure distribution*, where for a fluid system the volume is controlled by a piston used to vary the pressure), can be established in a similar way.<sup>4</sup> Functions of the microstate of the system, like  $\widehat{C}(\sigma)$ , are random variables for which the expectation value

$$\langle \widehat{C}(\sigma) \rangle := \sum_{\{\sigma\}} p_\eta(T; \sigma) \widehat{C}(\sigma), \quad (2.1.4)$$

variance and covariance between pairs of such variables play a fundamental role in connecting the distribution to thermodynamics (Sect. 2.1.2).

### 2.1.1 Quantum Systems

This is a text on *classical* statistical mechanics. However, it is useful to give a brief indication of the corresponding development for quantum systems where the microstate  $\sigma$  will be some set of quantum operators. Then the Hamiltonian  $\widehat{H}_\eta(\sigma)$  is an operator and  $p_\eta(T; \sigma)$  is the *density operator*, which in terms of a representation of the system becomes the *density matrix*. Equations (2.1.3) and (2.1.4) are replaced by

$$Z_\eta(T) := \text{Trace}\{\exp[-\widehat{H}_\eta(\sigma)/T]\}, \quad (2.1.5)$$

---

<sup>3</sup> In this and all subsequent formulae, the summation is over all microstates  $\sigma$  compatible with the given values of the independent extensive variables.

<sup>4</sup> The different distributions, for different values of  $\eta$  are usually called *ensembles*. This term is associated with the (relative-frequency flavoured) picture of a collection (or ensemble) of systems whose phase points have the density (2.1.2) in the space of microstates (Gibbs 1902).

and

$$\langle \widehat{C}(\sigma) \rangle := \text{Trace}\{\rho_\eta(T; \sigma) \widehat{C}(\sigma)\}. \quad (2.1.6)$$

A *complete solution* to this problem, in both the quantum and statistical mechanical sense, would be a determination of all the eigenvalues and eigenvectors of the Hamiltonian. The complicating feature of a quantum system, as can be seen from the brief account of the Heisenberg model in Sect. 3.2, is that  $\sigma$  may contain non-commuting operators.

### 2.1.2 The Connection to Thermodynamics

This is achieved by the identifications

$$H_\eta := \langle \widehat{H}_\eta \rangle = -\frac{\partial \ln(Z_\eta)}{\partial (1/T)}, \quad (2.1.7)$$

$$Q_i := \langle \widehat{Q}_i \rangle = T \frac{\partial \ln(Z_\eta)}{\partial \xi_i}, \quad i = 1, \dots, \eta, \quad (2.1.8)$$

$$\xi_i := \left\langle \frac{\partial \widehat{H}_\eta}{\partial Q_i} \right\rangle = -T \frac{\partial \ln(Z_\eta)}{\partial Q_i}, \quad i = \eta + 1, \dots, n_f. \quad (2.1.9)$$

The ‘averaging’ relations (2.1.7)–(2.1.9) become identical to (1.1.12) if we set

$$F_\eta(T) := -T \ln[Z_\eta(T)]. \quad (2.1.10)$$

In the special case of the canonical distribution,  $\eta = 0$ ,  $\widehat{H}_0(\sigma)$  is the energy  $E(\sigma)$ ,  $H_0$  is the internal energy  $U$  and  $F_0$  is the Helmholtz free energy, usually denoted by  $A$ . The second derivatives of  $F_\eta$  with respect to the independent variables  $T, \xi_1, \dots, \xi_\eta$  are the *response functions*. It is not difficult to show that the response function with respect to the variables  $\xi_i, \xi_j$  is related to the covariance of the random variables  $\widehat{Q}_i$  and  $\widehat{Q}_j$  by

$$\frac{\partial^2 F_\eta}{\partial \xi_i \partial \xi_j} = -\frac{\partial Q_j}{\partial \xi_i} = -\frac{\partial Q_i}{\partial \xi_j} = -\frac{1}{T} \text{Cov}[\widehat{Q}_i, \widehat{Q}_j]. \quad (2.1.11)$$

This is a *fluctuation-response function relation*.

It follows from (2.1.3) and (2.1.10) that

$$F_\eta(T) = -T \ln \left\{ \sum_{\{Q_\eta\}} \exp[\{\xi_\eta Q_\eta - F_{\eta-1}(T)\}/T] \right\}. \quad (2.1.12)$$

Differentiating with respect to  $T$  and using the expression for  $S$  given by (1.1.10) gives

$$F_\eta(T) = \langle F_{\eta-1}(T) \rangle - \xi_\eta \langle \mathcal{Q}_\eta \rangle, \quad (2.1.13)$$

where the expectations are taken with respect to the  $\eta$  distribution. This formula is the statistical mechanical analogue of (1.1.13).

## 2.2 Variations of the Probability Function

For the Hamiltonian (2.1.1), let

$$\mathcal{F}_\eta(\Pi) := \sum_{\{\sigma\}} \Pi(\sigma) \hat{H}_\eta(\sigma) + T \sum_{\{\sigma\}} \Pi(\sigma) \ln[\Pi(\sigma)], \quad (2.2.1)$$

where  $\Pi(\sigma) \geq 0$  satisfies the normalization condition

$$\sum_{\{\sigma\}} \Pi(\sigma) = 1. \quad (2.2.2)$$

It is easy to verify, from (2.1.2) and (2.1.10), that

$$\mathcal{F}(p_\eta) = F_\eta \quad (2.2.3)$$

and thus, from (1.1.10) and (2.1.7), that

$$S = - \sum_{\{\sigma\}} p_\eta(\sigma) \ln[p_\eta(\sigma)]. \quad (2.2.4)$$

Now consider variations of  $\Pi(\sigma)$  in (2.2.1), subject to the normalization condition (2.2.2), for which we use the undetermined multiplier  $\lambda$ . Then

$$\begin{aligned} \delta \mathcal{F}_\eta(\Pi) = \sum_{\{\sigma\}} \left\{ [\hat{H}_\eta(\sigma) + T \ln \Pi(\sigma) + T + \lambda] \delta \Pi(\sigma) \right. \\ \left. + T \frac{(\delta \Pi(\sigma))^2}{2 \Pi(\sigma)} + O((\delta \Pi(\sigma))^3) \right\}, \end{aligned} \quad (2.2.5)$$

yielding a minimum given by

$$\hat{H}_\eta(\sigma) + T \ln \Pi(\sigma) + T + \lambda = 0, \quad (2.2.6)$$

which is the unique solution  $\Pi(\sigma) = p_\eta(\sigma)$ .

## 2.3 Coupling Representations

In general the statistical mechanical internal energy is a linear function of a set of energy parameters  $\varepsilon_\ell$  of the form

$$E(\varepsilon_\ell; \sigma) := \sum_{\ell=1}^{n_e} E_\ell(\sigma) \varepsilon_\ell. \quad (2.3.1)$$

These parameters can be regarded as *internal fields* and it will be seen that the formulae of Sect. 2.1 can be simplified if the temperature is absorbed into the ratios

$$K_\ell := \varepsilon_\ell / T, \quad \ell = 1, \dots, n_e, \quad L_i := \xi_i / T, \quad i = 1, \dots, n_f, \quad (2.3.2)$$

called *internal* and *external couplings* respectively. This then gives

$$\hat{H}_\eta(K_\ell, L_i; \sigma) = \sum_{\ell=1}^{n_e} E_\ell(\sigma) K_\ell - \sum_{i=1}^{\eta} \hat{Q}_i(\sigma) L_i. \quad (2.3.3)$$

The discussion of scaling in Chap. 4 shows that the distinction between the internal couplings  $K_\ell$  and the external couplings  $L_i$  is physical rather than mathematical.<sup>5</sup> The latter arise from interactions between the system and external fields, like the magnetic field  $\mathcal{H}$  or the chemical potential  $\mu$ , whereas the former arise from the energies of interaction between the microsystems. The external couplings are, in general, open to change by the experimenter, either by individual variations of the fields  $\xi_i$  or together by variations of  $T$ . The internal couplings can be changed only by means of a change of  $T$ , so that

$$dK_\ell = -K_\ell dT/T, \quad \ell = 1, \dots, n_e. \quad (2.3.4)$$

From (2.1.3),<sup>6</sup>

$$Z_\eta(K_\ell, L_i) := \sum_{\{\sigma\}} \exp[-\hat{H}_\eta(K_\ell, L_i; \sigma)]. \quad (2.3.5)$$

Then, from (2.1.8)–(2.1.10),

$$\Phi_\eta(K_\ell, L_i) := F_\eta(T, \varepsilon_\ell, \xi_i) / T = -\ln[Z_\eta(K_\ell, L_i)] \quad (2.3.6)$$

---

<sup>5</sup> Although we do need to be careful about the consequences of the difference of sign in the two terms on the right of (2.3.10).

<sup>6</sup> Here it is clearer if we explicitly display representative members  $K_\ell$  and  $L_i$  of the  $n_e$  internal couplings (all independent) and the  $\eta$  independent external couplings.

is the *dimensionless free energy*, with

$$Q_i = \langle \widehat{Q}_i(\boldsymbol{\sigma}) \rangle = -\frac{\partial \Phi_\eta}{\partial L_i}, \quad i = 1, \dots, \eta, \quad L_i = \frac{\partial \Phi_\eta}{\partial Q_i}, \quad i = \eta + 1, \dots, n_f, \quad (2.3.7)$$

$$\langle E_\ell(\boldsymbol{\sigma}) \rangle = \frac{\partial \Phi_\eta}{\partial K_\ell}, \quad \ell = 1, \dots, n_e. \quad (2.3.8)$$

Since  $\Phi_\eta(K_\ell, L_i)$  is a function of  $T$  only through the couplings this must also be the case for the expectations  $\langle \widehat{Q}_i(\boldsymbol{\sigma}) \rangle$  and  $\langle E_\ell(\boldsymbol{\sigma}) \rangle$  and thus, from (2.3.1) and (2.3.3),

$$U(K_\ell) := \sum_{\ell=1}^{n_e} \langle E_\ell(\boldsymbol{\sigma}) \rangle K_\ell, \quad (2.3.9)$$

$$H_\eta(K_\ell, L_i) := \sum_{\ell=1}^{n_e} \langle E_\ell(\boldsymbol{\sigma}) \rangle K_\ell - \sum_{i=1}^{\eta} \langle \widehat{Q}_i(\boldsymbol{\sigma}) \rangle L_i. \quad (2.3.10)$$

The *coupling-density representation*, like the field-density representation, is obtained by forming densities by dividing the independent extensive variables by a size parameter which we again take to be  $Q := Q_{n_f}$ . Displaying a representative independent density  $\rho_i$  and extensive variable  $Q_i := \rho_i Q$  explicitly, the *dimensionless free-energy density*<sup>7</sup>

$$\phi_\eta(\rho_i, K_\ell, L_i) := \frac{\Phi_\eta(Q, \rho_i Q, K_\ell, L_i)}{Q} \quad (2.3.11)$$

is a function of the independent couplings and densities, without explicit reference to the temperature, making them convenient tools for the formulation of scaling theory and renormalization group methods.

We can also, of course, defined densities  $\hat{\rho}_i := \widehat{Q}_i/Q$ , for the dependent extensive variables,  $i = 1, 2, \dots, \eta$ . These are random variables with  $\rho_i := \langle \hat{\rho}_i \rangle$  defining thermodynamic densities and, from (2.1.11),

$$\text{Var}[\hat{\rho}_i] = -\frac{1}{Q} \frac{\partial^2 \phi_\eta}{\partial L_i^2} \quad i = 1, 2, \dots, \eta. \quad (2.3.12)$$

In the thermodynamic limit  $Q \rightarrow \infty$ , unless the response function  $\partial^2 \phi_\eta / \partial L_i^2$  diverges faster than  $Q$ , fluctuations in  $\hat{\rho}_i$  will tend to zero and the distinction between dependent and independent densities, and hence between the different choices of

---

<sup>7</sup> In later work, when the context makes it unnecessary we shall omit the qualification ‘dimensionless’.

distribution (canonical, grand-canonical etc.), will disappear. We have a *thermodynamic equivalence of distributions* away from the exceptional points where the response functions diverge, which are associated with phase transitions (Sect. 4.1).

As in Sect. 1.3 we now omit explicit reference to  $Q$  unless we need to be specific about the system being finite. With

$$u_\ell := \langle E_\ell(\sigma) \rangle / Q, \quad v := u/T = \sum_{\ell=1}^{n_e} u_\ell K_\ell, \quad (2.3.13)$$

from (1.2.3)–(1.2.5),

$$v = s + L_{n_f} + \sum_{i=1}^{n_f-1} \rho_i L_i, \quad (2.3.14)$$

$$\phi_\eta = L_{n_f} + \sum_{i=\eta+1}^{n_f-1} \rho_i L_i \quad (2.3.15)$$

and, from (1.2.6)–(1.2.9) and (2.3.4),

$$ds = \sum_{\ell=1}^{n_e} K_\ell du_\ell - \sum_{i=1}^{n_f-1} L_i d\rho_i, \quad (2.3.16)$$

$$0 = \sum_{\ell=1}^{n_e} u_\ell dK_\ell - dL_{n_f} - \sum_{i=1}^{n_f-1} \rho_i dL_i, \quad (2.3.17)$$

$$d\phi_\eta = \sum_{\ell=1}^{n_e} u_\ell dK_\ell - \sum_{i=1}^{\eta} \rho_i dL_i + \sum_{i=\eta+1}^{n_f-1} L_i d\rho_i. \quad (2.3.18)$$

### 2.3.1 The Case $n_f = 2$

As in Sect. 1.4 we set  $\rho := \rho_1$  with the one independent external coupling  $L := L_1$ . In place of the temperature we now have  $K_\ell$ ,  $\ell = 1, \dots, n_e$  internal couplings. To reduce these to one variable which corresponds to thermal variation we set  $\varepsilon := \varepsilon_1$ ,  $K := \varepsilon/T$  with the parameters  $\kappa_\ell := \varepsilon_\ell/\varepsilon$  kept constant and

$$v = u K, \quad u := \sum_{\ell=1}^{n_e} u_\ell K_\ell. \quad (2.3.19)$$

Now there are two free-energy densities:  $\phi_0(K, \rho)$  and  $\phi_1(K, L)$  with

$$d\phi_0 = u dK + L d\rho, \quad (2.3.20)$$

$$d\phi_1 = u dK - \rho dL. \quad (2.3.21)$$

They are related by the Legendre transformation

$$\phi_1 = \phi_0 - L\rho, \quad (2.3.22)$$

where  $\rho = \rho(K, L)$  is the solution of

$$\frac{\partial \phi_0}{\partial \rho} = L. \quad (2.3.23)$$

Alternatively

$$\phi_0 = \phi_1 + \rho L, \quad (2.3.24)$$

where  $L = L(K, \rho)$  is the solution of

$$\frac{\partial \phi_1}{\partial L} = -\rho. \quad (2.3.25)$$

In Sects. 1.4.1 and 1.4.2 we considered two examples of  $n_f = 2$ , a magnetic system where the two extensive variables are the magnetization  $\mathcal{M}$  and the number of dipoles  $M$  and a fluid system where the two extensive variables are the volume  $V$  and the number of particles  $M$ . In the former case there is an obvious choice for scaling; taking the magnetization per dipole as the density, conjugate to the magnetic field. For the fluid the choice is less obvious: volume per particle, with conjugate field  $-P$ , or particles per unit volume, with conjugate field  $\mu$ . It is useful to list the relationship between these two choices. Denoting one using a prime,

$$\begin{aligned} \rho' &= 1/\rho, & u' &= \rho' u, & \phi'(K, \rho') &= \rho' \phi_0(K, \rho), \\ L' &= \phi_1(K, L), & L &= \phi'(K, L'). \end{aligned} \quad (2.3.26)$$

It is now easy to check that Eqs. (2.3.20)–(2.3.25) take exactly the same form in the primed quantities. An example of the interchange between different choices of scaling variable is given for the van der Waals equation in Sect. 5.2.

We shall, in Chap. 4, need forms for the response functions, defined in Sect. 1.4, in terms of the coupling-density representation. From (2.3.20) and (2.3.21),

$$u = \frac{\partial \phi_0}{\partial K} = \frac{\partial \phi_1}{\partial K} \quad (2.3.27)$$



and, from (2.3.15),

$$s = -\phi_0 + \mathbf{u}K = -\phi_1 + \mathbf{u}K - L\rho. \quad (2.3.28)$$

Then from (1.4.5)–(1.4.7), (2.3.20)–(2.3.22) and (2.3.25),

$$c_\rho = -k_B K \left( \frac{\partial s}{\partial K} \right)_\rho = -k_B K^2 \left( \frac{\partial \mathbf{u}}{\partial K} \right)_\rho = -k_B K^2 \frac{\partial^2 \phi_0}{\partial K^2}, \quad (2.3.29)$$

$$\begin{aligned} c_\xi &= -k_B \left\{ K \left( \frac{\partial s}{\partial K} \right)_L + L \left( \frac{\partial s}{\partial L} \right)_K \right\} \\ &= -k_B \left\{ K^2 \left( \frac{\partial \mathbf{u}}{\partial K} \right)_L - 2KL \left( \frac{\partial \rho}{\partial K} \right)_K - L^2 \left( \frac{\partial \rho}{\partial L} \right)_K \right\} \\ &= -k_B \left( K^2 \frac{\partial^2 \phi_1}{\partial K^2} + 2KL \frac{\partial^2 \phi_1}{\partial K \partial L} + L^2 \frac{\partial^2 \phi_1}{\partial L^2} \right), \end{aligned} \quad (2.3.30)$$

$$\varphi_T = \left( \frac{\partial \rho}{\partial L} \right)_K = -\frac{\partial^2 \phi_1}{\partial L^2}, \quad (2.3.31)$$

$$\alpha_\xi = -K \left( \frac{\partial \rho}{\partial K} \right)_L - L \left( \frac{\partial \rho}{\partial L} \right)_K = L \frac{\partial^2 \phi_1}{\partial L^2} + K \frac{\partial^2 \phi_1}{\partial L \partial K}. \quad (2.3.32)$$

Alternatively  $c_\rho$  can be obtained from (2.3.30)–(2.3.32) using (1.4.8).

## 2.4 Lattice Systems

In Sect. 16.2.1 we define what is meant by a lattice, give diagrams for the regular lattices and consider the question of finite lattices and periodic boundary conditions. Here we need simply to assume that the lattice, denote by  $\mathcal{N}$ , is one of the regular types with periodic boundary conditions and  $N := |\mathcal{N}|$  sites, which is taken to be the one independent extensive variable  $Q_{n_f}$ .<sup>8</sup>

Chapter 3 contains a survey of lattice models, some, but not all, of which are treated in this book. They are distinguished by the types of cooperative interaction which contribute to the Hamiltonian. These fall broadly into two classes, those for which the state variables are associated with lattice sites and those for which they are associated with lattice edges (connecting first-neighbour pairs of sites).

---

<sup>8</sup> Of course, all the extensive quantities in the system, including the partition function and Hamiltonian are functions of  $N$ . However, it, or the lattice signifier  $\mathcal{N}$ , will be displayed explicitly only when it seems useful to do so. In the case of densities like  $\phi$  it is also sometimes useful to indicate the underlying lattice by including the infinite lattice signifier  $\mathcal{L}$ .

### 2.4.1 Site-Variable Models

There is a microsystem at each site of the lattice and the state of the microsystem at site  $\mathbf{r}$  is specified by the variable  $\sigma(\mathbf{r})$ . The symbol  $\sigma$  represents the collection or vector specifying a set of values of the states of all the microsystems on the lattice. The simplest contributions to the Hamiltonian are single-site terms with the extensive variables  $\widehat{Q}_i$  expressed in the form

$$\widehat{Q}_i(\sigma) = \sum_{\{\mathbf{r}\}} q_i(\mathbf{r}; \sigma), \quad i = 1, \dots, n_f - 1, \quad (2.4.1)$$

where  $q_i(\mathbf{r}; \sigma)$  is a function of the one component  $\sigma(\mathbf{r})$  of  $\sigma$  at  $\mathbf{r}$ . Unless the states of the microsystems are very complicated there will be only a rather small number of independent extensive variables and different functions  $q_i$ .

The extensive internal energy variables  $E_\ell$  in the Hamiltonian involve interactions between groups of more than one site. They can be expressed in the form

$$E_\ell(\sigma) = \sum_{\{\mathbf{r}\}} \epsilon_\ell(\mathbf{r}; \sigma), \quad \ell = 1, \dots, n_e, \quad (2.4.2)$$

where  $\epsilon_\ell(\mathbf{r}; \sigma)$  is a function of  $\sigma(\mathbf{r})$  and the state variables of some set of sites around  $\mathbf{r}$ . In most cases of interest the functions  $\epsilon_\ell(\mathbf{r}; \sigma)$  are taken to be *local* in the sense that they involve only sites quite close to the site  $\mathbf{r}$ , although in Sects. 3.2 and 3.3 we consider cases where interactions, involving pairs of sites, are allowed to have an arbitrary range.

The simplest *pair-interaction models* are those where the terms are restricted to first-neighbour sites and this is the case for most site-variable models considered here. However, it is sometimes of interest to include second-neighbour interactions, particularly when there is ‘conflict’ between the types of ordering encouraged by the two different interactions (Sect. 3.5.1.1).

Another type of site-variable model which has attracted interest is the *face model* where the interaction term involves, in an indecomposable way, the variables  $\sigma(\mathbf{r})$  for all sites  $\mathbf{r}$  around a face of the lattice.<sup>9</sup> Such a face model is considered in Sect. 3.7.2.

### 2.4.2 Edge-Variable Models

In this case a variable is attached to each edge of the lattice. The edge-decorated Ising models discussed in Chap. 10 are obvious examples of this, although in that case there are also state variables at the sites of the lattice and the model can equally

---

<sup>9</sup> It is obvious that the Hamiltonian of a pair-interaction model can be represented (in a decomposable way) as a sum over lattice faces and this is often useful in cases where the Hamiltonian also includes an indecomposable face interaction.

well be thought of as a site-variable model on a lattice modified by the addition of a site on each lattice edge. The two types of true edge-variable models discussed in this book are the vertex models introduced in Sect. 3.8 and analyzed in Sect. 9.3 and Chap. 12, and the dimer model, introduced in Sect. 3.9 and analyzed in Chap. 13. In each model the edge variable has two states, which for the vertex model corresponds to the two directions of an arrow along the edge and for the dimer model indicates the presence or absence of a dimer on the edge.

Although vertex models can be formed on any regular two-dimensional lattice<sup>10</sup> we shall restrict attention to the two two-dimensional regular lattices with coordination number four; that is the square and kagomé lattices (Sect. 16.2.1). There are then sixteen arrow arrangements around a lattice site. However, the cases which have been solved are the six-vertex model (Sect. 3.8.2) where the arrow configuration is constrained so that there is the same number of arrows pointing towards and pointing away from the lattice site and the eight-vertex model (Sect. 3.8.1) where there is an even number of arrows pointing towards the lattice site.

The dimer configuration is also constrained by a condition at each lattice site. In this case there can be at most (and in the case of a complete covering of the lattice with dimers, exactly) one dimer end at each lattice site (Sect. 3.9).

## 2.5 Correlation Functions and Symmetry Properties

### 2.5.1 A General Hamiltonian

The coupling-density representation gives a particularly elegant formulation of thermodynamics in the special case  $\eta = n_f - 1$ , when there are no independent densities. The function  $\phi_{n_f-1}$  is dependent only on couplings and, from (2.3.15), is itself equal to the coupling  $L_{n_f}$ . We shall concentrate on this case, omitting subscripts indicating the number of independent external couplings.

The values of the two sets of quantities  $q_i(\mathbf{r}; \sigma)$ ,  $i = 1, \dots, n_f - 1$  and  $e_\ell(\mathbf{r}; \sigma)$ ,  $\ell = 1, \dots, n_e$  are determined by the microstate  $\sigma$  of the system. We shall therefore refer to them respectively as *external* and *internal state operators*. It is clear from the above discussion that the sets  $\{Q_i, L_i, \hat{Q}_i(\sigma), q_i(\mathbf{r}; \sigma)\}$  and  $\{U_i, K_i, E_i(\sigma), e_i(\mathbf{r}; \sigma)\}$  play the same mathematical role in the development of the theory.<sup>11</sup> We shall, therefore, use the generic set  $\{C_j, \zeta_j, \hat{C}_j(\sigma), c_j(\mathbf{r}; \sigma)\}$  to stand for members of either the external or internal sets with

$$\hat{C}_j(\sigma) = \sum_{\{\mathbf{r}\}} c_j(\mathbf{r}; \sigma). \quad (2.5.1)$$

<sup>10</sup> Or indeed on an arbitrary planar lattice (Baxter 1978) or a three-dimensional lattice.

<sup>11</sup> Although, of course, an internal coupling  $K_i$  is simply replaced by a  $\zeta_j$ , whereas the coupling corresponding to an external coupling  $L_i$  is a  $-\zeta_j$ .

The state operators  $c_j(\mathbf{r}; \sigma)$ ,  $j = 1, \dots, n := n_f + n_e - 1$ , with  $\mathbf{r}$  ranging over all  $N$  sites of the lattice form a set of  $nN$  random variables distributed according to the probability function (2.1.2). In using (2.5.1) to express a general form for the Hamiltonian it is convenient, to include a configuration-independent term. Thus we set

$$\hat{H}(N, \zeta_j; \sigma) := N\zeta_0 + \sum_{j=1}^n \hat{C}_j(\sigma) \zeta_j. \quad (2.5.2)$$

This definition could, of course, be expressed in a more compact form by extending the summation to the range  $j = 0, 1, \dots, n$  and setting  $\hat{C}_0(\sigma) = N$ . It will also be observed that

$$\phi(\zeta_0, \zeta_1, \dots, \zeta_n) = \zeta_0 + \phi(0, \zeta_1, \dots, \zeta_n), \quad (2.5.3)$$

and that the calculation of an expectation value (2.1.4) using the probability formula (2.1.2) is unaffected by the presence of  $\zeta_0$ . For this reason it is often called the *trivial coupling*, although, as we shall see in Sects. 4.3.6 and 15.4, it plays a crucial role in scaling and renormalization group theory.

It is convenient to generalize this scheme further by supposing that the coupling is different at each site of the lattice with the corresponding generalized Hamiltonian

$$\hat{H}^{(G)}(N, \zeta_j(\mathbf{r}); \sigma) = \sum_{j=0}^n \sum_{\{\mathbf{r}\}} c_j(\mathbf{r}; \sigma) \zeta_j(\mathbf{r}), \quad (2.5.4)$$

where  $c_0(\mathbf{r}; \sigma) = 1$ , for all  $\mathbf{r}$ . With a similar notation for other quantities,

$$\langle c_j(\mathbf{r}; \sigma) \rangle = -\frac{1}{Z^{(G)}} \frac{\partial Z^{(G)}}{\partial \zeta_j(\mathbf{r})} = N \frac{\partial \phi^{(G)}}{\partial \zeta_j(\mathbf{r})}. \quad (2.5.5)$$

## 2.5.2 Correlation Functions

Let  $c_1(\mathbf{r}_1; \sigma), \dots, c_\tau(\mathbf{r}_\tau; \sigma)$  be some choice of a subset of these variables at  $\tau$  sites of the lattice. In terms of these a number of different  $\tau$ -point correlation functions can be defined:

- (i) The *total correlation function*  $\langle c_1(\mathbf{r}_1; \sigma) \cdots c_\tau(\mathbf{r}_\tau; \sigma) \rangle$ .
- (ii) The *fluctuation correlation function*  $\langle \delta c_1(\mathbf{r}_1; \sigma) \cdots \delta c_\tau(\mathbf{r}_\tau; \sigma) \rangle$  where

$$\delta c_j(\mathbf{r}; \sigma) := c_j(\mathbf{r}; \sigma) - \langle c_j(\mathbf{r}; \sigma) \rangle, \quad j = 1, 2, \dots, \tau, \quad (2.5.6)$$

is a measure of the fluctuation of  $c_j$  at site  $\mathbf{r}$  about its mean.

(iii) The *net* or *connected correlation function*, which is defined recursively by

$$\begin{aligned} \Gamma_\tau(c_1(\mathbf{r}_1) \cdots c_\tau(\mathbf{r}_\tau)) &:= \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \cdots c_\tau(\mathbf{r}_\tau; \boldsymbol{\sigma}) \rangle \\ &\quad - \sum \Gamma_{p_1}(c_{j_1}(\mathbf{r}_{j_1}), \dots, c_{p_1}(\mathbf{r}_{p_1})) \\ &\quad \cdots \Gamma_{p_k}(c_{j_k}(\mathbf{r}_{j_k}), \dots, c_{p_k}(\mathbf{r}_{p_k})), \end{aligned} \quad (2.5.7)$$

$$\Gamma_1(c_1(\mathbf{r}_1)) := \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle, \quad (2.5.8)$$

where the sum in (2.5.7) is over products of the connected correlation functions for all partitions of the  $\tau$  lattice sites.

For  $\tau = 1$  the total and connected correlation functions are the same and the fluctuation correlation function is zero. For  $\tau = 2$ ,

$$\begin{aligned} \Gamma_2(c_1(\mathbf{r}_1), c_2(\mathbf{r}_2)) &= \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle - \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle \langle c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle, \\ &= \langle [c_1(\mathbf{r}_1; \boldsymbol{\sigma}) - \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle] [c_2(\mathbf{r}_2; \boldsymbol{\sigma}) - \langle c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle] \rangle \end{aligned} \quad (2.5.9)$$

and, for  $\tau = 3$ ,

$$\begin{aligned} \Gamma_3(c_1(\mathbf{r}_1), c_2(\mathbf{r}_2), c_3(\mathbf{r}_3)) &= \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) c_2(\mathbf{r}_2; \boldsymbol{\sigma}) c_3(\mathbf{r}_3; \boldsymbol{\sigma}) \rangle \\ &\quad - \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle \Gamma_2(c_2(\mathbf{r}_2; \boldsymbol{\sigma}), c_3(\mathbf{r}_3; \boldsymbol{\sigma})) - \langle c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle \Gamma_2(c_3(\mathbf{r}_3; \boldsymbol{\sigma}), c_1(\mathbf{r}_1; \boldsymbol{\sigma})) \\ &\quad - \langle c_3(\mathbf{r}_3; \boldsymbol{\sigma}) \rangle \Gamma_2(c_1(\mathbf{r}_1; \boldsymbol{\sigma}), c_2(\mathbf{r}_2; \boldsymbol{\sigma})) - \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle \langle c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle \langle c_3(\mathbf{r}_3; \boldsymbol{\sigma}) \rangle \\ &= \langle [c_1(\mathbf{r}_1; \boldsymbol{\sigma}) - \langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \rangle] [c_2(\mathbf{r}_2; \boldsymbol{\sigma}) - \langle c_2(\mathbf{r}_2; \boldsymbol{\sigma}) \rangle] [c_3(\mathbf{r}_3; \boldsymbol{\sigma}) - \langle c_3(\mathbf{r}_3; \boldsymbol{\sigma}) \rangle] \rangle. \end{aligned} \quad (2.5.10)$$

The two- and three-point fluctuation and connected correlation functions are the same. That this identity does not persist for large values of  $\tau$  can be seen if we expand the general  $\tau$ -point fluctuation correlation function

$$\begin{aligned} \langle \delta c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \cdots \delta c_\tau(\mathbf{r}_\tau; \boldsymbol{\sigma}) \rangle &= \sum_{\{j_1, \dots, j_s\}} (-1)^{\tau-s} \langle c_{j_1}(\mathbf{r}_{j_1}; \boldsymbol{\sigma}) \cdots c_{j_s}(\mathbf{r}_{j_s}; \boldsymbol{\sigma}) \rangle \\ &\quad \times \langle c_{j_{s+1}}(\mathbf{r}_{j_{s+1}}; \boldsymbol{\sigma}) \rangle \cdots \langle c_{j_\tau}(\mathbf{r}_{j_\tau}; \boldsymbol{\sigma}) \rangle, \end{aligned} \quad (2.5.11)$$

where the sum is over all subsets  $\{j_1, \dots, j_s\}$  of the indices  $\{1, 2, \dots, \tau\}$ . Each term in the expansion is of the form of an  $s$ -point total correlation function multiplied by  $\tau - s$  one-point functions. As can be seen from (2.5.7), the expansion for  $\Gamma_\tau(c_1(\mathbf{r}_1) \cdots c_\tau(\mathbf{r}_\tau))$  in terms of total correlation functions will contain terms which are products of  $p$ - and  $q$ -point total correlation functions where *both*  $p$  and  $q$  are greater than unity when  $\tau > 3$ .

Using the site-specific couplings defined above,

$$\langle c_1(\mathbf{r}_1; \boldsymbol{\sigma}) \cdots c_\tau(\mathbf{r}_\tau; \boldsymbol{\sigma}) \rangle = \frac{(-1)^\tau}{Z^{(G)}} \frac{\partial^\tau Z^{(G)}}{\partial \zeta_1(\mathbf{r}_1) \cdots \partial \zeta_\tau(\mathbf{r}_\tau)}. \quad (2.5.12)$$

It can also be shown by induction (Binney et al. 1993) that

$$\Gamma_\tau(c_1(\mathbf{r}_1), \dots, c_\tau(\mathbf{r}_\tau)) = N(-1)^{\tau+1} \frac{\partial^\tau \phi^{(G)}}{\partial \zeta_1(\mathbf{r}_1) \cdots \partial \zeta_\tau(\mathbf{r}_\tau)}, \quad (2.5.13)$$

for connected correlation functions.

Returning now to the situation where the couplings  $\zeta_j$  are site independent

$$\begin{aligned} \frac{\partial^\tau \phi}{\partial \zeta_1 \cdots \partial \zeta_\tau} &= \sum_{\{\mathbf{r}_1, \dots, \mathbf{r}_\tau\}} \frac{\partial^\tau \phi^{(G)}}{\partial \zeta_1(\mathbf{r}_1) \cdots \partial \zeta_\tau(\mathbf{r}_\tau)} \\ &= \frac{(-1)^{\tau+1}}{N} \sum_{\{\mathbf{r}_1, \dots, \mathbf{r}_\tau\}} \Gamma_\tau(c_1(\mathbf{r}_1), \dots, c_\tau(\mathbf{r}_\tau)), \end{aligned} \quad (2.5.14)$$

where, in the second summation, the reduction to site-independent couplings is applied after differentiation. In the case of the connected two-point or *pair* correlation function this result can be derived directly without the use of (2.5.13). From (2.1.11), (2.5.1), and (2.5.9)

$$\frac{\partial^2 \phi}{\partial \zeta_1 \partial \zeta_2} = -\frac{1}{N} \text{Cov}[\widehat{C}_1, \widehat{C}_2] = -\frac{1}{N} \sum_{\{\mathbf{r}_1, \mathbf{r}_2\}} \Gamma_2(c_1(\mathbf{r}_1), c_2(\mathbf{r}_2)). \quad (2.5.15)$$

Like (2.1.11) this is a *fluctuation-response function relation*. Such relations play an important role in statistical mechanics, relating microscopic and macroscopic properties of a system. Of course, when the two variables are the same the covariance becomes the variance. Thus, for example, from (2.3.31), for the set  $\{Q, L, \widehat{Q}(\sigma), q(\mathbf{r}; \sigma)\}$ ,

$$\varphi_\tau = -\frac{\partial^2 \phi}{\partial L^2} = \frac{1}{N} \text{Var}[\widehat{Q}] = \frac{1}{N} \sum_{\{\mathbf{r}_1, \mathbf{r}_2\}} \Gamma_2(q(\mathbf{r}_1), q(\mathbf{r}_2)). \quad (2.5.16)$$

### 2.5.3 Symmetry Properties

Suppose that the lattice is  $d$ -dimensional hypercubic with periodic boundary conditions and lattice vectors given by (18.1.1). The  $\tau$ -point connected correlation function then has (Sect. 18.1.1) the Fourier transform

$$\begin{aligned} \Gamma_\tau^*(c_1, \dots, c_\tau; \mathbf{k}_1, \dots, \mathbf{k}_\tau) &= \sum_{\{\mathbf{r}_1, \dots, \mathbf{r}_\tau\}} \Gamma_2(c_1(\mathbf{r}_1), \dots, c_\tau(\mathbf{r}_\tau)) \\ &\quad \times \exp[-i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \cdots + \mathbf{k}_\tau \cdot \mathbf{r}_\tau)], \end{aligned} \quad (2.5.17)$$

with

$$\begin{aligned} \Gamma_\tau(c_1(\mathbf{r}_1), \dots, c_\tau(\mathbf{r}_\tau)) &= \frac{1}{N} \sum_{\{\mathbf{k}_1, \dots, \mathbf{k}_\tau\}} \Gamma_\tau^*(c_1, \dots, c_\tau; \mathbf{k}^{(1)}, \dots, \mathbf{k}^{(\tau)}) \\ &\quad \times \exp[i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \cdots + \mathbf{k}_\tau \cdot \mathbf{r}_\tau)], \end{aligned} \quad (2.5.18)$$

where the wave vectors are given by (18.1.6). It follows from (2.5.14) and (2.5.17) that

$$\frac{\partial^\tau \phi}{\partial \zeta_1 \cdots \partial \zeta_\tau} = \frac{(-1)^{\tau+1}}{N} \Gamma_\tau^*(\mathbf{c}_1, \dots, \mathbf{c}_\tau; \mathbf{0}, \dots, \mathbf{0}). \quad (2.5.19)$$

In many cases of interest the system has the property of invariance under *translations* of the form

$$\mathbf{r} \rightarrow \mathbf{r} + a[m^{(1)}\hat{\mathbf{r}}^{(1)} + m^{(2)}\hat{\mathbf{r}}^{(2)} + \cdots + m^{(d)}\hat{\mathbf{r}}^{(d)}], \quad (2.5.20)$$

for any integers  $m^{(1)}, \dots, m^{(d)}$ . The difference between the translation given by (18.1.2), (18.1.3) and (2.5.20), is that the former can be applied to each of the vectors  $\mathbf{r}_1, \dots, \mathbf{r}_\tau$  of  $\Gamma_\tau(\mathbf{c}_1(\mathbf{r}_1), \dots, \mathbf{c}_\tau(\mathbf{r}_\tau))$  independently, whereas (2.5.20) must be applied to them all simultaneously. One translation of the form (2.5.20) takes  $\mathbf{r}_1$  to the origin to give

$$\Gamma_\tau(\mathbf{c}_1(\mathbf{r}_1), \dots, \mathbf{c}_\tau(\mathbf{r}_\tau)) = \Gamma_\tau(\mathbf{c}_1(\mathbf{0}), \mathbf{c}_2(\bar{\mathbf{r}}_1), \dots, \mathbf{c}_\tau(\bar{\mathbf{r}}_{\tau-1})), \quad (2.5.21)$$

where  $\bar{\mathbf{r}}_j = \mathbf{r}_{j+1} - \mathbf{r}_1$ ,  $j = 1, \dots, \tau - 1$  are the relative lattice vectors. The one-point connected correlation function, given from (2.5.8) by  $\langle \mathbf{c}_1(\mathbf{r}_1) \rangle$  is independent of the site location  $\mathbf{r}_1$ . We also see, from (2.5.7) and (2.5.11), that the total and fluctuation correlation functions have a similar dependence on relative vectors. With translational invariance one of the summations in the final form of (2.5.14) can be performed to give

$$\frac{\partial^\tau \phi}{\partial \zeta_1 \cdots \partial \zeta_\tau} = (-1)^{\tau+1} \sum_{\{\bar{\mathbf{r}}_1, \dots, \bar{\mathbf{r}}_{\tau-1}\}} \Gamma_\tau(\mathbf{c}_1(\mathbf{0}), \mathbf{c}_2(\bar{\mathbf{r}}_1), \dots, \mathbf{c}_\tau(\bar{\mathbf{r}}_{\tau-1})), \quad (2.5.22)$$

and, if  $\bar{\Gamma}_\tau^*(\mathbf{c}_1, \dots, \mathbf{c}_\tau; \bar{\mathbf{k}}_1, \dots, \bar{\mathbf{k}}_{\tau-1})$  is the Fourier transform defined with respect to the relative vectors,

$$\frac{\partial^\tau \phi}{\partial \zeta_1 \cdots \partial \zeta_\tau} = (-1)^{\tau+1} \bar{\Gamma}_\tau^*(\mathbf{c}_1, \dots, \mathbf{c}_\tau; \mathbf{0}, \dots, \mathbf{0}). \quad (2.5.23)$$

Thus, for example, from (2.5.16),

$$\varphi_T = \bar{\Gamma}_2^*(\mathbf{q}; \mathbf{0}). \quad (2.5.24)$$

Finally we note that if, in addition to translational invariance, the system has  $N^{(1)} = N^{(2)} = \cdots = N^{(d)}$  and all its properties invariant under *rotation* between lattice directions the two-point correlation function is given by

$$\Gamma_2(\mathbf{c}_1(\mathbf{r}_1), \mathbf{c}_2(\mathbf{r}_2)) = \Gamma_2(\mathbf{c}_1(\mathbf{0}), \mathbf{c}_2(\mathbf{r}_2 - \mathbf{r}_1)) = \Gamma_2(\mathbf{c}_1, \mathbf{c}_2; |\mathbf{r}_2 - \mathbf{r}_1|). \quad (2.5.25)$$

Equilibrium Statistical Mechanics of Lattice Models

Lavis, D.

2015, XVII, 793 p. 101 illus., Hardcover

ISBN: 978-94-017-9429-9