Contents

Contents i

List of Tables ii

List of Figures iii

6 Classical Interacting Systems 1

6.1 References . . . 1

6.2 The Ising Model . . . 2

6.2.1 Spontaneous symmetry breaking . . . 3

6.2.2 Solution in $d = 1$ dimension . . . 4

6.2.3 Domain walls . . . 7

6.2.4 Ising model in two dimensions : Peierls' argument . . . 9

6.2.5 High temperature expansion . . . 13

6.2.6 Kramers-Wannier duality . . . 16

6.3 Nonideal Classical Gases . . . 17

6.3.1 The configuration integral . . . 17

6.3.2 One-dimensional Tonks gas . . . 18

6.3.3 Mayer cluster expansion . . . 19

6.3.4 Examples . . . 27

6.4 Liquid State Physics . . . 33

6.4.1 The many-particle distribution function . . . 33

6.4.2 Averages over the distribution . . . 35
6.4.3 Virial equation of state . . . 39
6.4.4 Correlations and scattering . . . 40
6.4.5 Correlation and response . . . 43
6.4.6 BBGKY hierarchy . . . 45
6.4.7 Ornstein-Zernike theory . . . 47
6.4.8 Percus-Yevick and hypernetted chains approximations . . . 50
6.5 Coulomb Systems : Plasmas and the Electron Gas . . . 52
6.5.1 Electrostatic potential . . . 52
6.5.2 Debye-Hückel theory . . . 52
6.5.3 The electron gas : Thomas-Fermi screening . . . 56
6.6 Polymers . . . 59
6.6.1 Basic concepts . . . 59
6.6.2 Polymers as random walks . . . 60
6.6.3 The Rouse model . . . 63
6.6.4 Flory theory of self-avoiding walks . . . 66
6.6.5 Polymers and solvents . . . 68
6.7 Lee-Yang Theory . . . 70
6.7.1 Analytic properties of the partition function . . . 70
6.7.2 Electrostatic analogy . . . 71
6.7.3 Example . . . 72
6.8 Appendix: Potts Model in One Dimension . . . 74
6.8.1 Hamiltonian . . . 74
6.8.2 Transfer matrix . . . 74

List of Tables

6.1 Exact, Percus-Yevick, and hypernetted chains results for hard spheres 51

List of Figures

6.1 Domain walls in the $d = 2$ Ising model 8
6.2 Clusters and boundaries for the square lattice Ising model 10
6.3 A two-dimensional square lattice mapped onto a one-dimensional chain 12
6.4 High temperature expansion diagrams 14
6.5 High temperature expansion for the correlation function 15
6.6 The Lennard-Jones potential 20
6.7 Diagrams and the Mayer cluster expansion 21
6.8 Vertex labels in the configuration integral 22
6.9 Symmetry factors for cluster diagrams 23
6.10 Connected *versus* irreducible clusters 26
6.11 The overlap of hard sphere Mayer functions 27
6.12 Phase diagram of the $d = 3$ hard sphere gas 29
6.13 Mayer function for an attractive spherical well with a repulsive core 30
6.14 Density of hard spheres in the presence of a hard wall 32

6.15 Hard sphere pair distribution functions: simulation and experiment 37
6.16 Monte Carlo pair distribution functions for liquid water 38
6.17 Elastic and inelastic scattering 42
6.18 Static structure factor of the Lennard-Jones fluid 44
6.19 The Thomas-Fermi atom 58
6.20 Some examples of linear chain polymers 60
6.21 *trans* and *gauche* orientations in carbon chains 61
6.22 The polymer chain as a random coil 62
6.23 Radius of gyration R_{g} of polystyrene in a toluene and benzene solvent 68
6.24 Singularities of the partition function 71
6.25 Fugacity z and $pv_0/k_{\mathrm{B}}T$ *versus* dimensionless specific volume v/v_0 73

Chapter 6

Classical Interacting Systems

6.1 References

- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- L. E. Reichl, *A Modern Course in Statistical Physics* (2nd edition, Wiley, 1998)
  A comprehensive graduate level text with an emphasis on nonequilibrium phenomena.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.
- J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, 1990)
  An advanced, detailed discussion of liquid state physics.

6.2 The Ising Model

The simplest model of an interacting system consists of a lattice $\mathcal{L}$ of N sites, each of which hosts a spin σ_i which may be either up ($\sigma_i = +1$) or down ($\sigma_i = -1$). The Hamiltonian is

$$\hat{H} = -J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j - \mu_0 H \sum_i \sigma_i \quad . \tag{6.1}$$

The energy J accounts for an interaction between sites i and j, here taken to be nearest neighbors on $\mathcal{L}$. When $J = 0$, the model describes N noninteracting spins. The geometry is then irrelevant, and as we saw in §4.8.1, for each spin we have $\langle \sigma_i \rangle =$
$tanh(\mu_0 H / k_{\rm B} T)$.

In zero external field (*i.e.* $H = 0$), energetic considerations dictate that the interaction term on each link $\langle ij \rangle$ prefers neighboring spins to be *ferromagnetically* aligned in a $|\uparrow\uparrow\rangle$ or $|\downarrow\downarrow\rangle$ configuration (*i.e.* $\sigma_i \, \sigma_j = +1$) when $J > 0$, and *antiferromagnetically* aligned in a $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$ configuration (*i.e.* $\sigma_i \, \sigma_j = -1$) when $J < 0$.

This model is not exactly solvable in general. In one dimension, the solution is quite straightforward. In two dimensions, Onsager's solution of the model (with $H = 0$) is among the most celebrated results in statistical physics. In higher dimensions the system has been studied by numerical simulations (the Monte Carlo method), by series expansions, and by field theoretic calculations (renormalization group), but no exact solutions exist.

One important aspect of interacting systems is the emergence of *correlations*. We define the *correlation matrix* of the Ising model $C_{ij}(T, H) \equiv \langle \sigma_i \, \sigma_j \rangle$, which is the average of the product $\sigma_i \, \sigma_j$[1]. When $J = 0$ the system is noninteracting, and the average of the product is the product of the averages for all $i \neq j$, *i.e.* $C_{ij}(T, H) = m_i \, m_j$, where $m_i \equiv \langle \sigma_i \rangle$. Thus the *connected correlator*,

$$\widetilde{C}_{ij}(T, H) \equiv \langle \sigma_i \, \sigma_j \rangle - \langle \sigma_i \rangle \, \langle \sigma_j \rangle \quad , \tag{6.2}$$

which is the average of the product *minus* the product of the averages, vanishes identically for $J = 0$ whenever $i \neq j$.

The general expression for the *thermal density matrix* at temperature T is

$$\varrho(T, N, H) = \frac{\sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \, |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|}{\sum_{\boldsymbol{\sigma}'} e^{-E(\boldsymbol{\sigma}')/k_{\rm B}T}} = \frac{1}{Z(T, N, H)} \sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \, |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}| \quad . \tag{6.3}$$

where

$$Z(T, N, H) = \mathsf{Tr} \; e^{-\hat{H}/k_{\rm B}T} = \sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \tag{6.4}$$

is the partition function. Here N is the number of sites, which is the volume $|\mathcal{L}|$ of the lattice (*i.e.* the total number of sites). Thermodynamic averages of any quantity $\mathcal{O}(\boldsymbol{\sigma})$ are computed via

$$\langle \mathcal{O} \rangle = \mathsf{Tr} \, (\mathcal{O} \varrho) = \frac{1}{Z(T, N, H)} \sum_{\boldsymbol{\sigma}} \mathcal{O}(\boldsymbol{\sigma}) \, e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \quad . \tag{6.5}$$

[1]The correlation matrix $C_{ij}(T, H)$ is defined for arbitrary sites i and j. In the thermodynamic limit, it is independent of the system size N.

When $J \neq 0$, there are nontrivial correlations between different sites. Consider, for example, the case $H = 0$. In this case the model has what we call a $\mathbb{Z}_2$ symmetry, which means that for any spin configuration $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$, the energy $E(\boldsymbol{\sigma})$ is invariant under reversing the direction of all the spins[2] . *I.e.* if $\varepsilon\boldsymbol{\sigma} \equiv \{-\sigma_1, \ldots, -\sigma_N\}$, we have $E(\varepsilon\boldsymbol{\sigma}) = E(\boldsymbol{\sigma})$ for each of the 2^N possible spin configurations $\boldsymbol{\sigma}$. This means that $\langle\sigma_i\rangle = 0$ for all i . At $T = 0$ there are two ground states, $|\Uparrow\rangle \equiv |\uparrow\uparrow\uparrow\cdots\rangle$ and $|\Downarrow\rangle \equiv |\downarrow\downarrow\downarrow\cdots\rangle$, each with energy $E_0 = -N_{\rm nn}J$, where $N_{\rm nn}$ is the number of nearest neighbor links[3] on $\mathcal{L}$. The thermal density matrix as $T \to 0$ is given by $\varrho(T = 0, H = 0) = \frac{1}{2}|\Uparrow\rangle\langle\Uparrow| + \frac{1}{2}|\Downarrow\rangle\langle\Downarrow|$, from which we obtain $\langle\sigma_i\rangle = 0$ and $\widetilde{C}_{ij}(T = 0, H = 0) = 1$. If we apply an external field $H > 0$, then the two ground states are no longer degenerate in energy, and $E_\Uparrow = -N_{\rm nn}J - N\mu_0 H$ while $E_\Downarrow = -N_{\rm nn}J + N\mu_0 H$. Thus, as $T \to 0$, even for very small $H > 0$, in the thermodynamic limit the $|\Uparrow\rangle$ state is selected. and we obtain a very different density matrix: $\varrho(T = 0, H > 0) = |\Uparrow\rangle\langle\Uparrow|$. Using this density matrix, we obtain $\langle\sigma_i\rangle = +1$ and $\langle\sigma_i\sigma_j\rangle = +1$ $(i \neq j)$, hence $\widetilde{C}_{ij}(T = 0, H > 0) = 0$.

6.2.1 Spontaneous symmetry breaking

In quantum mechanics, the eigenstates of a Hamiltonian $\hat{H}_0$ which commutes with all the *generators* of a symmetry group G may be classified according to the *representations* of that group. Typically this entails the appearance of degeneracies in the eigenspectrum, with degenerate states transforming into each other under the group operations, as we saw above. Adding a perturbation V to the Hamiltonian which breaks G down to a subgroup H will accordingly split these degeneracies, and the new multiplets of $\hat{H} = \hat{H}_0 + \hat{V}$ are characterized by representations of the lower symmetry group H.

In quantum field theory, or in the thermodynamic limit of a classical system, as a consequence of the infinite number of degrees of freedom, symmetries may be *spontaneously broken*. This means that even if the Hamiltonian H (or action S) for the field theory is invariant under a group G of symmetry transformations, the ground state or thermodynamic density matrix may not be invariant under the full symmetry group G. The presence or absence of spontaneous symmetry breaking (SSB), and its detailed manifestations, will in general depend on the couplings, or the temperature in the case of quantum statistical mechanics. SSB is usually associated with the presence of a local *order parameter* which transforms nontrivially under some group operations, and whose quantum statistical average vanishes in a fully symmetric phase, but takes nonzero values in symmetry-broken phase[4]. The parade example is the Ising model, $\hat{H} = -\sum_{i<j} J_{ij}\,\sigma_i\,\sigma_j$, where each $\sigma_i = \pm 1$, the subscript i indexes a physical location in space, such as a site $\boldsymbol{R}_i$ on a particular lattice. The model is explicitly $\mathbb{Z}_2$ symmetric under $\sigma_i \to \varepsilon\sigma_i$ for all i, where $\varepsilon \in \{+1, -1\}$, yet if the interaction matrix $J_{ij} = J(\boldsymbol{R}_i - \boldsymbol{R}_j)$ is short-ranged and the space dimension d is greater than one, there is a *critical temperature* $T_{\rm c}$ below which SSB sets in, and the system develops a spontaneous magnetization $m = \langle\sigma_i\rangle$. You know how in quantum mechanics, the eigenstates of a particle moving in one-dimensional double-well potential $V(x) = V(-x)$ can be classified by their parity eigenvalues P, and the lowest two energy states are respectively symmetric (P $= +1$) and

[2]The group $\mathbb{Z}_2$ contains two elements, $\{1, \varepsilon\}$, and the multiplication table is $1 \times 1 = 1$, $1 \times \varepsilon = \varepsilon \times 1 = \varepsilon$, and $\varepsilon \times \varepsilon = 1$.

[3]On a lattice with coordination number z and with periodic boundary conditions, $N_{\rm nn} = \frac{1}{2}zN$.

[4]While SSB is generally associated with the existence of a phase transition, not all phase transitions involve SSB. Exceptions include the Kosterlitz-Thouless transition, and also those topological phases which have no local order parameter.

antisymmetric ($\mathsf{P} = -1$), and are delocalized among both wells. For a quantum field theory, however, with (Euclidean) Lagrangian density $\mathcal{L}_{\rm E} = \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi)$, for $d > 1$ and $T < T_{\rm c}$, the system actually picks the left or the right well, so that $\langle\phi(\boldsymbol{r})\rangle \neq 0$. Another example is the spontaneously broken $\mathsf{O}(2)$ invariance of superfluids, where the boson annihilation operator $\psi(\boldsymbol{r})$ develops a spontaneous average $\big\langle\psi(\boldsymbol{r})\big\rangle = \sqrt{n_0}\, e^{i\theta}$, where n_0 is the condensate density and θ the condensate phase.

But this is an obvious swindle, because, as we have seen, $\langle\sigma_i\rangle = 0$ in zero external field, due to the $\mathbb{Z}_2$ symmetry of the model. Rather, we may understand the phenomenon of spontaneous symmetry breaking in either of the following ways:

- First, rather than defining the order parameter of the Ising model, for example, to be the expected value $m = \langle\sigma_i\rangle$ of the local spin[5], consider instead the behavior of the correlator $C_{ij} = \langle\sigma_i\,\sigma_j\rangle$ in the limit $d_{ij} = |\boldsymbol{R}_i - \boldsymbol{R}_j| \to \infty$. In a disordered phase, there is no correlation between infinitely far separated spins, hence $\lim_{d_{ij}\to\infty} C_{ij} = 0$. In the ordered phase, this is no longer true, and we define the *spontaneous magnetization* m from the long distance correlator: $m^2 \equiv \lim_{d_{ij}\to\infty}\langle\sigma_i\,\sigma_j\rangle$. In this formulation, SSB is associated with the emergence of *long-ranged order* in the correlators of operators which transform nontrivially under the symmetry group.
- Second, we could impose an external field which *explicitly* breaks the symmetry, such as a Zeeman term $\hat{V} = -\mu_0 H\sum_i \sigma_i$ in the Ising model. We compute the magnetization per site $m(T,N,H) = \langle\sigma_i\rangle$ as a function of temperature T, the external field H, *and* the system volume (*i.e.* number of sites) N. The order parameter $m(T)$ in zero field is then defined as

$$m(T) = \lim_{H\to 0}\lim_{N\to\infty} m(T,N,H) \quad . \tag{6.6}$$

  The order of limits here is crucially important. The thermodynamic limit $N \to \infty$ is taken first, which means that the energy difference between $|\Uparrow\rangle$ and $|\Downarrow\rangle$ diverges, being proportional to N, thus infinitely suppressing the $|\Downarrow\rangle$ state if $H > 0$ (and the $|\Uparrow\rangle$ state if $H < 0$). The magnitude of the order parameter will be independent of the way in which we take $H \to 0$, but its sign will depend on whether $H \to 0^+$ or $H \to 0^-$, with $\mathsf{sgn}(m) = \mathsf{sgn}(H)$. Physically, the direction in which a system orders can be decided by the presence of small stray fields or impurities.

Note that in both formulations, SSB is necessarily associated with the existence of a local operator $\mathcal{O}_i$ which is identified as the order parameter field. In the first scheme, the correlations $\big\langle\mathcal{O}_i\,\mathcal{O}_j\big\rangle$ exhibit long-ranged order in the symmetry-broken phase. In the second scheme, $\mathcal{O}_i$ is the operator to which the local external field H_i couples. We will discuss the mean field theory of phase transitions and spontaneous symmetry breaking in chapter 7.

6.2.2 Solution in $d = 1$ dimension

Consider a one-dimensional ring of L sites. The ordinary canonical partition function is then

$$Z_{\mathsf{PBC}}(T,L,H) = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_{\{\sigma_n\}}\prod_{n=1}^{L} e^{\beta J\sigma_n\sigma_{n+1}}\, e^{\beta\mu_0 H\sigma_n} \quad , \tag{6.7}$$

[5]We assume translational invariance, which means $\langle\sigma_i\rangle$ is independent of the site index i.

where $\sigma_{L+1} \equiv \sigma_1$ owing to periodic (ring) boundary conditions (PBC). We can replace the factor $e^{\beta\mu_0 H\sigma_n}$ in the above expression with $e^{\beta\mu_0 H(\sigma_n+\sigma_{n+1})/2}$, since the product over n yields the same result. We then obtain $Z = \mathsf{Tr}\left(R^L\right)$, where R is a 2×2 matrix with entries

$$R_{\sigma\sigma'} = e^{\beta J\sigma\sigma'}\, e^{\beta\mu_0 H(\sigma+\sigma')/2} = \begin{pmatrix} e^{\beta J}\, e^{\beta\mu_0 H} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J}\, e^{-\beta\mu_0 H} \end{pmatrix}_{\sigma\sigma'} \quad , \tag{6.8}$$

called the *transfer matrix*. Expressed in terms of the Pauli matrices, we have[6]

$$R = e^{\beta J}\cosh(\beta\mu_0 H)\, I + e^{-\beta J} X + e^{\beta J}\sinh(\beta\mu_0 H)\, Z \quad , \tag{6.9}$$

where I, X, and Z are the 2×2 identity, Pauli X, and Pauli Z matrices.

Since the trace of a matrix is invariant under a similarity transformation, we have

$$Z(T, L, H) = \lambda_+^L + \lambda_-^L \quad , \tag{6.10}$$

where $\lambda_\pm$ are the eigenvalues of R, *viz.*

$$\lambda_\pm(T, H) = e^{\beta J}\cosh(\beta\mu_0 H) \pm \sqrt{e^{2\beta J}\sinh^2(\beta\mu_0 H) + e^{-2\beta J}} \quad . \tag{6.11}$$

When $H = 0$, we have $\lambda_+ = 2\cosh(\beta J)$ and $\lambda_- = 2\sinh(\beta J)$. In the thermodynamic limit, $L \to \infty$, and the larger λ_+^L term dominates exponentially. We then have

$$F(T, L, H) = -Lk_{\rm B}T \log\lambda_+(T, H) \quad . \tag{6.12}$$

From the free energy, we can compute the magnetization,

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,L} = \frac{L\mu_0\, \sinh(\beta\mu_0 H)}{\sqrt{\sinh^2(\beta\mu_0 H) + e^{-4\beta J}}} \tag{6.13}$$

and the zero field isothermal susceptibility,

$$\chi(T) = \frac{1}{L}\,\frac{\partial M}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T}\, e^{2J/k_{\rm B}T} \quad . \tag{6.14}$$

Note that in the noninteracting limit $J \to 0$ we recover the familiar result for a free spin. The effect of the interactions at low temperature is to vastly increase the susceptibility. Rather than a set of independent single spins, the system effectively behaves as if it were composed of large blocks of spins, where the block size ξ is the *correlation length*, to be derived below.

The physical properties of the system are often elucidated by evaluation of various correlation functions. Accordingly, we define $C(n) \equiv \langle\sigma_1\,\sigma_{n+1}\rangle$, where

$$\big\langle\sigma_1\,\sigma_{n+1}\big\rangle = \frac{\mathsf{Tr}\left(\sigma_1\, R_{\sigma_1\sigma_2}\cdots R_{\sigma_n\sigma_{n+1}}\,\sigma_{n+1}\, R_{\sigma_{n+1}\sigma_{n+2}}\cdots R_{\sigma_L\sigma_1}\right)}{\mathsf{Tr}\,\left(R^L\right)} = \frac{\mathsf{Tr}\left(Z\, R^n\, Z\, R^{L-n}\right)}{\mathsf{Tr}\,\left(R^L\right)} \quad , \tag{6.15}$$

[6]Take care not to confuse Pauli Z with the partition function!

with $0 < n < L$, and where Z is the Pauli matrix. To compute this ratio, we decompose R in terms of its eigenvectors, writing $R = \lambda_+ \, |+\rangle\langle+| + \lambda_- \, |-\rangle\langle-|$. Then

$$C(n) = \frac{\lambda_+^L \, Z_{++}Z_{++} + \lambda_-^L \, Z_{--}Z_{--} + \left(\lambda_+^{L-n} \, \lambda_-^n + \lambda_+^n \, \lambda_-^{L-n}\right) Z_{+-}Z_{-+}}{\lambda_+^L + \lambda_-^L} \quad , \tag{6.16}$$

with $Z_{\mu\mu'} = \langle \, \mu \, | \, Z \, | \, \mu' \, \rangle$ being the matrix elements of Z in the eigenbasis of R.

Zero external field

Consider $H = 0$, where $R = e^{\beta J} + e^{-\beta J} \, X$. Then $|\pm\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\rangle \pm |\downarrow\rangle\big)$, *i.e.* the eigenvectors of R are

$$\psi_\pm = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \quad , \tag{6.17}$$

which entails $Z_{++} = Z_{--} = 0$, while $Z_{+-} = Z_{-+} = 1$. The corresponding eigenvalues are given by $\lambda_+ = 2\cosh(\beta J)$ and $\lambda_- = 2\sinh(\beta J)$. The correlation function is then found to be

$$\begin{aligned} C(n) \equiv \big\langle \sigma_1 \, \sigma_{n+1} \big\rangle &= \frac{\lambda_+^{L-|n|} \, \lambda_-^{|n|} + \lambda_+^{|n|} \, \lambda_-^{L-|n|}}{\lambda_+^L + \lambda_-^L} \\ &= \frac{\tanh^{|n|}(\beta J) + \tanh^{L-|n|}(\beta J)}{1 + \tanh^L(\beta J)} \approx \tanh^{|n|}(\beta J) \quad \text{for } L \to \infty \quad . \end{aligned} \tag{6.18}$$

This result is also valid for $n < 0$, provided $|n| \le L$. We see that we may write $C(n) = e^{-|n|/\xi(T)}$, where the *correlation length* is

$$\xi(T) = \frac{1}{\log \operatorname{ctnh}(J/k_{\rm B}T)} \quad . \tag{6.19}$$

Note that $\xi(T)$ grows as $T \to 0$ as $\xi \approx \frac{1}{2} \, e^{2J/k_{\rm B}T}$.

Chain with free ends

When the chain has free ends, *i.e.* open boundary conditions (OBC), there are $(L-1)$ links and the partition function is

$$Z_{\rm OBC}(T, L, H) = \sum_{\sigma,\sigma'} \big(R^{L-1}\big)_{\sigma\sigma'} = \sum_{\sigma,\sigma'} \Big\{ \lambda_+^{L-1} \, \psi_+(\sigma) \, \psi_+(\sigma') + \lambda_-^{L-1} \, \psi_-(\sigma) \, \psi_-(\sigma') \Big\} \quad , \tag{6.20}$$

where $\psi_\pm(\sigma) = \langle \, \sigma \, | \, \pm \, \rangle$. When $H = 0$, we make use of eqn. 6.17 to obtain

$$R^{L-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \big(2\cosh \beta J\big)^{L-1} + \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \big(2\sinh \beta J\big)^{L-1} \quad , \tag{6.21}$$

and therefore $Z_{\text{OBC}} = 2^L \cosh^{L-1}(\beta J)$. There's a simple trick to obtain the zero field partition function which amounts to a change of variables. We define $\nu_n \equiv \sigma_n \sigma_{n+1}$ for $1 \le n < L$. Thus, $\nu_1 = \sigma_1\sigma_2$, $\nu_2 = \sigma_2\sigma_3$, *etc.* Note that each ν_j takes the values ± 1. The Hamiltonian for the chain is

$$H_{\text{OBC}} = -J \sum_{n=1}^{L-1} \sigma_n \, \sigma_{n+1} = -J \sum_{n=1}^{L-1} \nu_n \quad . \tag{6.22}$$

The state of the system is defined by the L Ising variables $\{\sigma_1 \, , \, \nu_1 \, , \, \ldots \, , \, \nu_{L-1}\}$. Note that σ_1 doesn't appear in the Hamiltonian. Thus, the interacting model is recast as $L-1$ noninteracting Ising spins, and the partition function is

$$\begin{aligned} Z_{\text{OBC}}(T, L, H) = \operatorname{Tr} e^{-\beta H_{\text{OBC}}} &= \sum_{\sigma_1} \sum_{\nu_1} \cdots \sum_{\nu_{L-1}} e^{\beta J \nu_1} e^{\beta J \nu_2} \cdots e^{\beta J \nu_{L-1}} \\ &= \sum_{\sigma_1} \left(\sum_{\nu} e^{\beta J \nu} \right)^{L-1} = 2^L \cosh^{L-1}(\beta J) \quad . \end{aligned} \tag{6.23}$$

6.2.3 Domain walls

We have just seen how in one dimension, the Ising model never achieves long-ranged spin order. That is, the spin-spin correlation function decays asymptotically as an exponential function of the distance with a correlation length $\xi(T)$ which is finite for all $T > 0$. Only for $T = 0$ does the correlation length diverge. At $T = 0$, there are two ground states, $|\Uparrow\rangle$ and $|\Downarrow\rangle$. To choose between these ground states, we can specify a boundary condition at the ends of our one-dimensional chain, where we demand that the spins are up. Equivalently, we can apply a magnetic field H of order $1/L$, which vanishes in the thermodynamic limit, but which at zero temperature will select the 'all up' ground state. At finite temperature, there is always a finite probability for any consecutive pair of sites $(n, n+1)$ to be in a high energy state, *i.e.* either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$. Such a configuration is called a *domain wall*, and in one-dimensional systems domain walls live on individual links. Relative to the configurations $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, a domain wall costs energy $2J$. For a system with $M = xL$ domain walls, the free energy in the thermodynamic limit is

$$F = E_0 + 2MJ - k_{\text{B}} T \log \binom{L}{M} = L \cdot \Big\{ (2x-1)J + k_{\text{B}} T \Big[x \log x + (1-x) \log(1-x) \Big] \Big\} \quad , \tag{6.24}$$

where $E_0 = -LJ$ is the ground state energy. Minimizing the free energy with respect to x, one finds $x = 1/\big(e^{2J/k_{\text{B}}T} + 1\big)$, so the equilibrium concentration of domain walls is finite, meaning there can be no long-ranged spin order. In one dimension, entropy wins and there is always a thermodynamically large number of domain walls in equilibrium. And since the correlation length for $T > 0$ is finite, any boundary conditions imposed at spatial infinity will have no thermodynamic consequences since they will only be 'felt' over a finite range.

As we shall discuss in the following chapter, this consideration is true for any system with sufficiently short-ranged interactions and a discrete global symmetry. Another example is the q-state Potts model,

$$H = -J \sum_{\langle ij \rangle} \delta_{\sigma_i, \sigma_j} - h \sum_i \delta_{\sigma_i, 1} \quad . \tag{6.25}$$

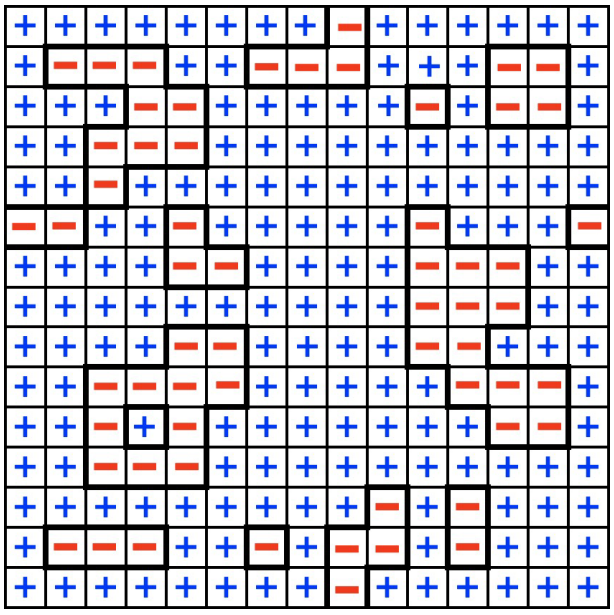


Figure 6.1: Domain walls (thick black lines) in the $d = 2$ Ising model.

Here, the spin variables σ_i take values in the set $\{1, 2, \ldots, q\}$ on each site. The equivalent of an external magnetic field in the Ising case is a field h which prefers a particular value of σ ($\sigma = 1$ in the above Hamiltonian). See the appendix in §6.8 for a transfer matrix solution of the one-dimensional Potts model.

In dimensions $d > 1$ the Ising model exhibits spontaneous symmetry breaking below a critical temperature T_c. Consider the two-dimensional Ising model on a square lattice. We'll impose periodic boundary conditions on a $L \times L$ region and later imagine taking the thermodynamic limit. Each spin configuration $|\,\boldsymbol{\sigma}\,\rangle$ can be represented as an L^2-digit binary number $\boldsymbol{n}$, with $n_{\boldsymbol{r}} = 0$ if $\sigma_{\boldsymbol{r}} = -1$ and $n_{\boldsymbol{r}} = 1$ if $\sigma_{\boldsymbol{r}} = +1$. Thus $n_{\boldsymbol{r}} = \frac{1}{2}(1 + \sigma_{\boldsymbol{r}})$ where $\boldsymbol{r}$ labels a plaquette. We can also label each spin configuration by a configuration of *closed loops* in the following way. Let Γ represent a set of closed loops, represented by the thick black lines in fig. 6.1. Start with each $\sigma_{\boldsymbol{r}} = +1$ and in each plaquette $\boldsymbol{r}$ flip the spin to $\sigma_{\boldsymbol{r}} = -1$ if $\boldsymbol{r}$ is enclosed by an odd number of loops. The energy of this configuration $|\,\Gamma\,\rangle$ is given by $E_\Gamma = 2JL_\Gamma$, where L_Γ is the total perimeter of all the individual loops comprising Γ. This is because spins lying on opposite sides of a loop contribute an energy $-J\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}'} = +J$ instead of $-J$, which is their energy in the fiducial configuration with all spins up[7]. Thus the partition sum may be written as

$$Z = \mathrm{Tr}\; e^{-\beta\hat{H}} = 2\sum_\Gamma e^{-2\beta JL_\Gamma} = 2\sum_{P=0}^{\infty} N_P\, e^{-2\beta JP} \quad . \tag{6.26}$$

Here we have subtracted off the ground state energy from the start, setting $E_0 \equiv 0$. The second sum is over the total perimeter P, with N_P the number of loop configurations with total perimeter P. But what is N_P? For a very crude approximation, we might take $N_P \sim (z-1)^P$, where $z = 4$ is the coordination number of the square lattice[8]. This is because at each step of an individual loop, if we rule out retracing

[7] For each Γ there are in fact *two* corresponding spin configurations, the second obtained by flipping all the spins of the first.

[8] A better approximation for the asymptotic behavior of N_P is $N_P = CP^{-theta}(z-1)^P$, *i.e.* including a power

our previous step, the number of possible directions in which the domain wall can be continued is $z-1$. If we approximate $N_P \sim (z-1)^P$, then we have

$$N_P\, e^{-2\beta JP} \sim \exp\Big(\big[\log(z-1) - 2\beta J\big]P\Big) \tag{6.27}$$

and the argument of the exponential starts to increase with P when $T > T_{\rm c}$, where $k_{\rm B}T_{\rm c} = 2J/\log(z-1)$. Thus, below $T_{\rm c}$ the loops are suppressed, while above $T_{\rm c}$ the loops proliferate. This provides us a very rough picture of the phase transition and an estimate of where it occurs. For the square lattice, where $z = 4$, we obtain an estimate of $k_{\rm B}T_{\rm c}/J \approx 2/\log 3 = 1.8205$. The exact location of the phase transition, as we will derive below, is given by $k_{\rm B}T_{\rm c}/J = 2/\log(1+\sqrt{2}) = 2.2692$.

6.2.4 Ising model in two dimensions : Peierls' argument

A more careful version of our previous argument, due to Rudolf Peierls, shows rigorously that there will be a finite temperature phase transition for the Ising model on the square lattice[9]. Consider the Ising model, in zero magnetic field, on a $N = L_x \times L_y$ square lattice, with $L_{x,y} \to \infty$ in the thermodynamic limit. Along the perimeter of the system we impose the boundary condition $\sigma_i = +1$. Any configuration of the spins may then be represented uniquely in the following manner. Start with a configuration in which all spins are up. Next, draw a set of closed loops on the lattice, as we did previously. By definition, the loops cannot share any links along their boundaries, *i.e.* each link on the lattice is associated with at most one such loop. Now flip all the spins inside each loop from up to down. Identify each such loop configuration with a label Γ. The partition function is

$$Z = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_{\Gamma} e^{-2\beta J L_\Gamma} \quad , \tag{6.28}$$

where L_Γ is the total perimeter of the loop configuration Γ. The domain walls are now loops, rather than individual links, but as in the one-dimensional case, each link of each domain wall contributes an energy $+2J$ relative to the ground state.

Now we wish to compute the average magnetization of the central site (assume $L_{x,y}$ are both odd, so there is a unique central site). This is given by the difference $P_+(0) - P_-(0)$, where $P_\mu(0) = \big\langle \delta_{\sigma_0,\mu} \big\rangle$ is the probability that the central spin has spin polarization μ. If $P_+(0) > P_-(0)$, then the magnetization per site $m = P_+(0) - P_-(0)$ is finite in the thermodynamic limit, and the system is ordered. Clearly

$$P_+(0) = \frac{1}{Z} \sum_{\Gamma\in\Sigma_+} e^{-2\beta J L_\Gamma} \quad , \tag{6.29}$$

where the restriction on the sum indicates that only those configurations where the central spin is up ($\sigma_0 = +1$) are to be included. (see fig. 6.2a). Similarly,

$$P_-(0) = \frac{1}{Z} \sum_{\tilde{\Gamma}\in\Sigma_-} e^{-2\beta J L_{\tilde{\Gamma}}} \quad , \tag{6.30}$$

law prefactor in addition to the exponential.

[9] Here we modify slightly the discussion in chapter 5 of the book by L. Peliti.

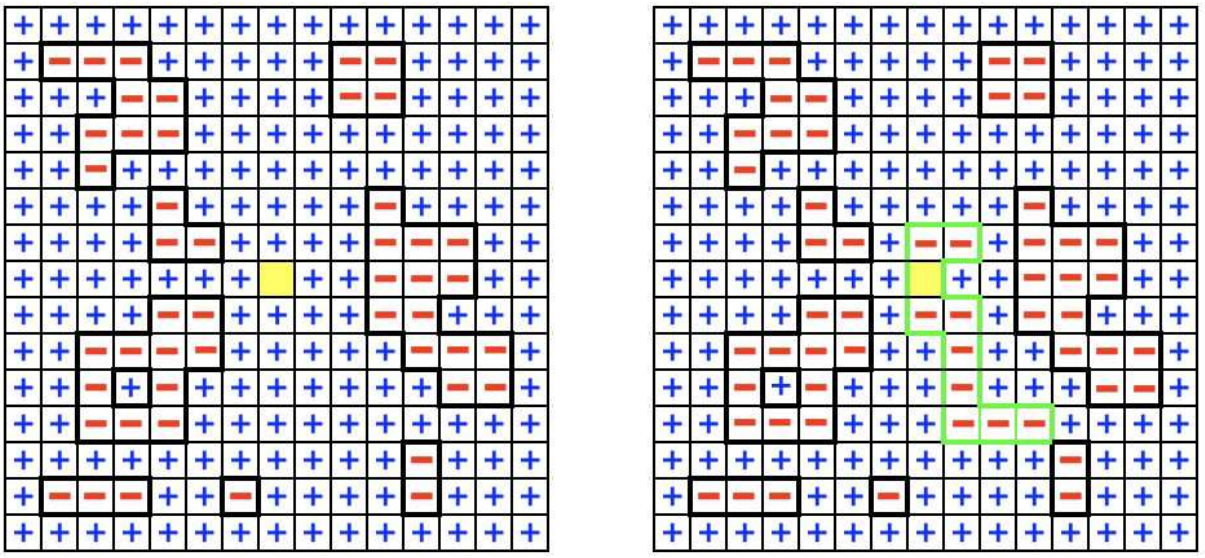


Figure 6.2: Clusters and boundaries for the square lattice Ising model. Left panel: a configuration Γ where the central spin is up. Right panel: a configuration $C_\gamma \circ \Gamma$ where the interior spins of a new loop γ containing the central spin have been flipped.

where only configurations in which $\sigma_0 = -1$ are included in the sum, where $\Sigma_\pm \equiv \Big\{ \Gamma \;\Big|\; \sigma_0 = \pm \Big\}$. That is, $\Sigma_+(\Sigma_-)$ is the set of configurations Γ in which the central spin is always up (down). Consider now the construction in fig. 6.2b. Any loop configuration $\widetilde{\Gamma} \in \Sigma_-$ may be associated with a unique loop configuration $\Gamma \in \Sigma_+$ by reversing all the spins within the loop of $\widetilde{\Gamma}$ which contains the origin. Note that the map from $\widetilde{\Gamma}$ to Γ is many-to-one. That is, we can write $\widetilde{\Gamma} = C_\gamma \circ \Gamma$, where C_γ overturns the spins within the loop γ, with the conditions that (i) γ contains the origin, and (ii) none of the links in the perimeter of γ coincide with any of the links from the constituent loops of Γ. Let us denote this set of loops as Υ_Γ:

$$\Upsilon_\Gamma = \Big\{ \gamma \;:\; 0 \in \mathrm{int}(\gamma) \text{ and } \gamma \cap \Gamma = \emptyset \Big\} \quad . \tag{6.31}$$

Then

$$m = P_+(0) - P_-(0) = \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} \left(1 - \sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} \right) \quad . \tag{6.32}$$

If we can prove that $\sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < 1$, then we will have established that $m > 0$. Let us ask: how many loops γ are there in Υ_Γ with perimeter L? We cannot answer this question exactly, but we can derive a rigorous upper bound for this number, which, following Peliti, we call $g(L)$. We claim that

$$g(L) < \frac{2}{3L} \cdot 3^L \cdot \left(\frac{L}{4} \right)^2 = \frac{L}{24} \cdot 3^L \quad . \tag{6.33}$$

To establish this bound, consider any site on such a loop γ. Initially we have 4 possible directions to proceed to the next site, but thereafter there are only 3 possibilities for each subsequent step, since the loop cannot run into itself. This gives $4 \cdot 3^{L-1}$ possibilities. But we are clearly overcounting, since any

point on the loop could have been chosen as the initial point, and moreover we could have started by proceeding either clockwise or counterclockwise. So we are justified in dividing this by $2L$. We are still overcounting, because we have not accounted for the constraint that γ is a closed loop, nor that $\gamma \cap \Gamma = \emptyset$. We won't bother trying to improve our estimate to account for these constraints. However, we are clearly undercounting due to the fact that a given loop can be translated in space so long as the origin remains within it. To account for this, we multiply by the area of a square of side length $L/4$, which is the maximum area that can be enclosed by a loop of perimeter L. We therefore arrive at eqn. 6.33. Finally, we note that the smallest possible value of L is $L = 4$, corresponding to a square enclosing the central site alone. Therefore

$$\sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < \frac{1}{12} \sum_{k=2}^{\infty} k \cdot \left(3\, e^{-2\beta J}\right)^{2k} = \frac{x^4\,(2-x^2)}{12\,(1-x^2)^2} \equiv r \quad , \tag{6.34}$$

where $x = 3\, e^{-2\beta J}$. Note that we have accounted for the fact that the perimeter L of each loop γ must be an even integer. The sum is smaller than unity provided $x < x_0 = 0.869756\ldots$, hence the system is ordered provided

$$\frac{k_{\rm B} T}{J} < \frac{2}{\log(3/x_0)} = 1.61531 \quad , \tag{6.35}$$

which establishes a rigorous lower bound for the critical temperature. The exact result, which we will establish in §6.2.6 below, is $k_{\rm B} T_{\rm c}/J = 2/\sinh^{-1}(1) \simeq 2.26918$. Peierls' argument has been generalized to higher dimensional lattices as well[10].

With a little more work we can derive a bound for the magnetization. We have shown that

$$P_-(0) = \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} \sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < r \cdot \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} = r\, P_+(0) \quad . \tag{6.36}$$

Thus,

$$1 = P_+(0) + P_-(0) < (1+r)\, P_+(0) \tag{6.37}$$

and therefore

$$m = P_+(0) - P_-(0) > (1-r)\, P_+(0) > \frac{1-r}{1+r} \quad , \tag{6.38}$$

where $r(T)$ is given in eqn. 6.34.

Importance of interaction range

We showed that the one-dimensional Ising model has no finite temperature phase transition, and is disordered at any finite temperature T, but in two dimensions on the square lattice there is a finite critical temperature $T_{\rm c}$ below which there is long-ranged order. Consider now the construction depicted in fig. 6.3, where the sites of a two-dimensional square lattice are mapped onto those of a linear chain[11]. Clearly we can elicit a one-to-one mapping between the sites of a two-dimensional square lattice and

[10] See. *e.g.* J. L. Lebowitz and A. E. Mazel, *J. Stat. Phys.* **90**, 1051 (1998).

[11] A corresponding mapping can be found between a cubic lattice and the linear chain as well.

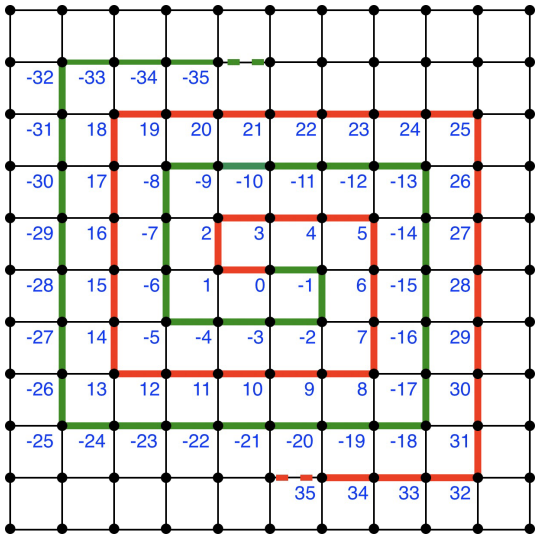


Figure 6.3: A two-dimensional square lattice mapped onto a one-dimensional chain.

those of a one-dimensional chain. That is, the two-dimensional square lattice Ising model may be written as a one-dimensional Ising model, *i.e.*

$$\hat{H} = -J \sum_{\langle ij \rangle}^{\substack{\text{square} \\ \text{lattice}}} \sigma_i \, \sigma_j = -\sum_{n,n'}^{\substack{\text{linear} \\ \text{chain}}} J_{nn'} \, \sigma_n \, \sigma_{n'} \quad . \tag{6.39}$$

How can this be consistent with the results we have just proven?

The fly in the ointment here is that the interaction along the chain $J_{n,n'}$ is long-ranged. This is apparent from inspecting the site labels in fig. 6.3. Note that site $n = 15$ is linked to sites $n' = 14$ and $n' = 16$, but also to sites $n' = -6$ and $n' = -28$. With each turn of the concentric spirals in the figure, the range of the interaction increases. To complicate matters further, the interactions are no longer translationally invariant, *i.e.* $J_{nn'} \neq J(n - n')$. But it is the long-ranged nature of the interactions on our contrived one-dimensional chain which spoils our previous energy-entropy argument, because now the domain walls themselves interact via a long-ranged potential. Consider the linear chain with $J_{n,n'} = J \, |n - n'|^{-\alpha}$, where $\alpha > 0$. Let us compute the energy of a domain wall configuration where $\sigma_n = +1$ if $n > 0$ and $\sigma_n = -1$ if $n \leq 0$. The domain wall energy is then

$$\Delta = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{2J}{|m+n|^\alpha} \quad . \tag{6.40}$$

Here we have written one of the sums in terms of $m = -n'$. For asymptotically large m and n, we can write $\boldsymbol{R} = (m, n)$ and we obtain an integral over the upper right quadrant of the plane:

$$\int_1^\infty dR \, R \int_0^{\pi/2} d\phi \, \frac{2J}{R^\alpha \, (\cos\phi + \sin\phi)^\alpha} = 2^{-\alpha/2} \int_{-\pi/4}^{\pi/4} \frac{d\phi}{\cos^\alpha \phi} \int_1^\infty \frac{dR}{R^{\alpha - 1}} \quad . \tag{6.41}$$

The ϕ integral is convergent, but the R integral diverges for $\alpha \leq 2$. For a finite system, the upper bound on the R integral becomes the system size L. For $\alpha > 2$ the domain wall energy is finite in the thermodynamic limit $L \to \infty$. In this case, entropy again wins. *I.e.* the entropy associated with a single domain wall is $k_{\rm B} \log L$, and therefore $F = E - TS$ is always lowered by having a finite density of domain walls. For $\alpha < 2$, the energy of a single domain wall scales as $L^{2-\alpha}$. It was first proven by F. J. Dyson in 1969 that this model has a finite temperature phase transition provided $1 < \alpha < 2$. There is no transition for $\alpha < 1$ or $\alpha > 2$. The case $\alpha = 2$ is special, and is discussed in the beautiful renormalization group analysis by J. M. Kosterlitz in *Phys. Rev. Lett.* **37**, 1577 (1976).

6.2.5 High temperature expansion

Consider once again the ferromagnetic Ising model in zero field ($H = 0$), but on an arbitrary lattice. The partition function is

$$Z = \mathsf{Tr}\, e^{\beta J \sum_{\langle ij \rangle} \sigma_i \sigma_j} = \big(\cosh \beta J\big)^{N_{\rm nn}} \, \mathsf{Tr} \left\{ \prod_{\langle ij \rangle} \big(1 + x\, \sigma_i \sigma_j\big) \right\} \quad , \tag{6.42}$$

where $x = \tanh \beta J$ and $N_{\rm nn}$ is the number of links (*i.e.* bonds). For regular lattices, $N_{\rm nn} = \frac{1}{2} z N$, where N is the number of lattice sites and z is the lattice coordination number, *i.e.* the number of nearest neighbors for each site. We have used

$$e^{\beta J \sigma \sigma'} = \cosh \beta J \cdot \Big\{ 1 + \sigma \sigma' \tanh \beta J \Big\} = \begin{cases} e^{+\beta J} & \text{if } \sigma \sigma' = +1 \\ e^{-\beta J} & \text{if } \sigma \sigma' = -1 \end{cases} \quad . \tag{6.43}$$

We expand eqn. 6.42 in powers of x, resulting in a sum of $2^{N_{\rm nn}}$ terms, each of which can be represented graphically in terms of so-called *lattice animals*. A lattice animal is a distinct (including reflections and rotations) arrangement of adjacent plaquettes on a lattice. In order that the trace not vanish, only such configurations and their compositions are permitted. This is because each σ_i for every given site i must occur an even number of times in order for a given term in the sum not to vanish. For all such terms, the trace is 2^N. Let $\{\Gamma\}$ represent a collection of lattice animals, and g_Γ the multiplicity of Γ. Then

$$Z = 2^N \big(\cosh \beta J\big)^{N_{\rm nn}} \sum_\Gamma g_\Gamma \, \big(\tanh \beta J\big)^{L_\Gamma} \quad , \tag{6.44}$$

where L_Γ is the total number of links in the diagram Γ, and g_Γ is the multiplicity of Γ. Since x vanishes as $T \to \infty$, this procedure is known as the *high temperature expansion* (HTE).

For the square lattice, the enumeration of all lattice animals with up to order eight is given in fig. 6.4. For the diagram represented as a single elementary plaquette, there are N possible locations for the lower left vertex. For the 2×1 plaquette animal, one has $g = 2N$, because there are two inequivalent orientations as well as N translations. For two disjoint elementary squares, one has $g = \frac{1}{2} N(N-5)$, which arises from subtracting $5N$ 'illegal' configurations involving double lines (remember each link in the partition sum appears only once!), shown in the figure, and finally dividing by two because the individual squares are identical. Note that $N(N-5)$ is always even for any integer value of N. Thus, to lowest interesting order on the square lattice,

$$Z = 2^N \big(\cosh \beta J\big)^{2N} \Big\{ 1 + N x^4 + 2N x^6 + \big(7 - \tfrac{5}{2}\big) N x^8 + \tfrac{1}{2} N^2 x^8 + \mathcal{O}(x^{10}) \Big\} \quad . \tag{6.45}$$

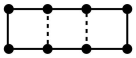
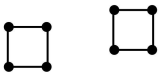
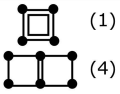
diagram Γ	L_Γ	g_Γ	remarks
	0	1	empty lattice
	4	N	N translations
	6	$2N$	2 rotations, N translations
	8	$2N$	2 rotations, N translations
	8	N	N translations
	8	$4N$	4 rotations, N translations
	8	$\frac{1}{2}(N^2-5N)$	exclude five invalid diagrams; halve for equivalent squares  (1) (4)

Figure 6.4: HTE diagrams on the square lattice and their multiplicities.

The free energy is therefore

$$\begin{aligned} F &= -k_{\rm B}T\log 2 + Nk_{\rm B}T\log(1-x^2) - Nk_{\rm B}T\Big[x^4 + 2\,x^6 + \tfrac{9}{2}\,x^8 + \mathcal{O}(x^{10})\Big] \\ &= Nk_{\rm B}T\log 2 - Nk_{\rm B}T\Big\{x^2 + \tfrac{3}{2}\,x^4 + \tfrac{7}{3}\,x^6 + \tfrac{19}{4}\,x^8 + \mathcal{O}(x^{10})\Big\}, \end{aligned} \tag{6.46}$$

again with $x = \tanh\beta J$. Here we have substituted $\cosh^2\beta J = 1/(1-x^2)$ to write the final result as a power series in x. Notice that the $\mathcal{O}(N^2)$ factor in Z has cancelled upon taking the logarithm, so the free energy is properly extensive.

For the one-dimensional chain or ring, the high temperature expansion yields

$$Z_{\rm OBC}(T,N) = 2^N\cosh^{N-1}\beta J \qquad , \qquad Z_{\rm PBC}(T,N) = 2^N\big(\cosh^N\beta J + \sinh^N\beta J\big) \quad , \tag{6.47}$$

in agreement with the transfer matrix calculations. Notice how for OBC there is only one lattice animal which contributes, *i.e.* the empty lattice. For PBC there are two contributing animals, the empty lattice and the entire ring. In higher dimensions, where there is a finite temperature phase transition, one typically computes the specific heat $c(T)$ and tries to extract its singular behavior in the vicinity of $T_{\rm c}$, where $c(T) \sim A\,(T-T_{\rm c})^{-\alpha}$. Since $x(T) = \tanh(J/k_{\rm B}T)$ is analytic in T, we have $c(x) \sim A'\,(x-x_{\rm c})^{-\alpha}$, where $x_{\rm c} = x(T_{\rm c})$. One assumes $x_{\rm c}$ is the singularity closest to the origin and corresponds to the radius of convergence of the high temperature expansion. If we write

$$c(x) = \sum_{n=0}^{\infty} a_n\,x^n \sim A''\left(1-\frac{x}{x_{\rm c}}\right)^{-\alpha} \quad , \tag{6.48}$$

then according to the binomial theorem we should expect

$$\frac{a_n}{a_{n-1}} = \frac{1}{x_{\rm c}}\left[1 - \frac{1-\alpha}{n}\right] \quad . \tag{6.49}$$

Thus, by plotting a_n/a_{n-1} *versus* $1/n$, one extracts $1/x_{\rm c}$ as the intercept, and $(\alpha - 1)/x_{\rm c}$ as the slope.

High temperature expansion for correlation functions

Can we also derive a high temperature expansion for the spin-spin correlation function $C_{kl} = \langle \sigma_k \, \sigma_l \rangle$? Yes we can. We have

$$C_{kl} = \frac{{\rm Tr}\left[\sigma_k \, \sigma_l \, e^{\beta J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j}\right]}{{\rm Tr}\left[e^{\beta J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j}\right]} \equiv \frac{Y_{kl}}{Z} \quad . \tag{6.50}$$

Recall our analysis of the partition function Z. We concluded that in order for the trace not to vanish, the spin variable σ_i on each site i must occur an even number of times in the expansion of the product. Similar considerations hold for Y_{kl}, except now due to the presence of σ_k and σ_l, those variables now must occur an *odd* number of times when expanding the product. It is clear that the only nonvanishing diagrams will be those in which there is a finite string connecting sites k and l, in addition to the usual closed HTE loops. See fig. 6.5 for an instructive sketch. One then expands both Y_{kl} as well as Z in powers of $x = \tanh \beta J$, taking the ratio to obtain the correlator C_{kl}. At high temperatures ($x \to 0$), both numerator and denominator are dominated by the configurations Γ with the shortest possible total perimeter. For Z, this means the trivial path $\Gamma = \{\emptyset\}$, while for Y_{kl} this means finding the shortest length path from k to l. (If there is no straight line path from k to l, there will in general be several such minimizing paths.) Note, however, that the presence of the string between sites k and l complicates the analysis of g_Γ for the closed loops, since none of the links of Γ can intersect the string. It is worth stressing that this does not mean that the string and the closed loops cannot intersect at isolated *sites*, but only that they share no common *links*; see once again fig. 6.5.

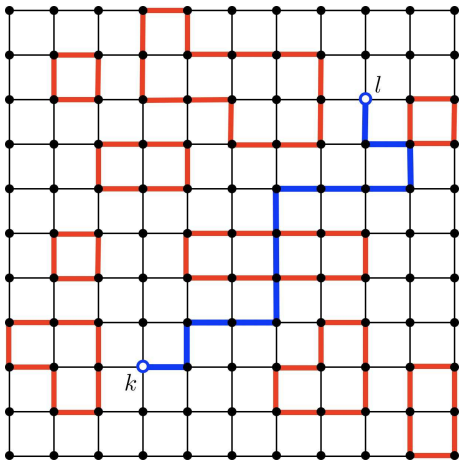


Figure 6.5: HTE diagrams for the numerator Y_{kl} of the correlation function C_{kl}. The blue path connecting sites k and l is the string. The remaining red paths are all closed loops.

6.2.6 Kramers-Wannier duality

Consider an Ising model on a square lattice of size $L_x \times L_y$, with periodic boundary conditions, and let us set $H = 0$, so there is no explicit symmetry breaking. We label each plaquette on the square lattice by a vector $\boldsymbol{R}$ which may be taken to be the position of its center. We write the Ising model Hamiltonian as

$$\hat{H}(J_x, J_y) = -J_x \sum_{\boldsymbol{R}} \sigma_{\boldsymbol{R}} \sigma_{\boldsymbol{R}+\hat{\boldsymbol{x}}} - J_y \sum_{\boldsymbol{R}} \sigma_{\boldsymbol{R}} \sigma_{\boldsymbol{R}+\hat{\boldsymbol{y}}} \quad , \tag{6.51}$$

taking $J_{x,y} > 0$.[12] The ground state $|\Uparrow\rangle$ satisfies $\hat{H}\,|\Uparrow\rangle = E_0\,|\Uparrow\rangle$, where $E_0 = -N(J_x + J_y)$. Note that $|\Downarrow\rangle$ is also a ground state, obtained from $|\Uparrow\rangle$ by a $\mathbb{Z}_2$ transformation which flips all the plaquette spins and leaves the Hamiltonian invariant. Starting from $|\Uparrow\rangle$, let the state $|\,\Gamma\,\rangle$ represent the configuration where all spins within each of the individual loops comprising Γ have been flipped, as discussed in §6.2.4 and depicted in fig. 6.2. Let $n^\Gamma_{x,y}$ denote the total number of x- and y-directed links in Γ, respectively. The partition function $Z(T, J_x, J_y)$ may be written as

$$Z(T, J_x, J_y) = 2 \sum_\Gamma e^{(N-2n^\Gamma_x)\beta J_y}\, e^{(N-2n^\Gamma_y)\beta J_x} \quad , \tag{6.52}$$

since the spins on the plaquettes located at $\boldsymbol{R}$ and $\boldsymbol{R} + \hat{\boldsymbol{x}}$ on opposite sides of each $\hat{\boldsymbol{y}}$-directed link in Γ are anti-aligned, according to the construction of $|\,\Gamma\,\rangle$, thereby raising the energy by $2J_x$ relative to E_0. Similarly, the anti-aligned plaquette spins at $\boldsymbol{R}$ and $\boldsymbol{R} + \hat{\boldsymbol{y}}$ on opposite sides of an $\hat{\boldsymbol{x}}$-directed link raise the energy by $2J_y$. The overall factor of 2 is because the set of states Γ is only 2^{N-1} in number. For example, the state $|\Downarrow\rangle$ cannot be constructed according to the rules. For each state $|\,\Gamma\,\rangle$, we can obtain a unique symmetry-related state $|\,\Gamma'\,\rangle$ by applying a $\mathbb{Z}_2$ global spin flip, which completes the full set of 2^N spin configurations.

Now define the *dual* couplings $K_{x,y}$ as follows[13]:

$$e^{-2\beta J_x} = \tanh(\beta K_y) \quad , \quad e^{-2\beta J_y} = \tanh(\beta K_x) \quad \Rightarrow \quad \beta K_{x,y} = \tfrac{1}{2} \log\, \mathrm{ctnh}\,(\beta J_{y,x}) \quad . \tag{6.53}$$

In the isotropic case $J_x = J_y = J$ we have $\exp(-2\beta J) = \tanh(\beta K)$ and thus for $J \to 0$ we have $K \to \infty$ while for $J \to \infty$ we have $K \to 0$. Thus, in the phase diagram of the dual couplings, *the temperature is inverted*. We now have

$$Z(T, J_x, J_y) = \big[\tanh(\beta K_y)\tanh(\beta K_x)\big]^{-N/2} \cdot 2 \sum_\Gamma \big[\tanh(\beta K_y)\big]^{n^\Gamma_x} \big[\tanh(\beta K_x)\big]^{n^\Gamma_y} \quad . \tag{6.54}$$

We next invoke the identity

$$2 \sum_\Gamma \big[\tanh(\beta K_y)\big]^{n^\Gamma_x} \big[\tanh(\beta K_x)\big]^{n^\Gamma_y} = 2^{-N}\, \mathrm{Tr} \prod_{\boldsymbol{r}} \big[1 + \tanh(\beta K_x)\, \mu_{\boldsymbol{r}} \mu_{\boldsymbol{r}+\hat{\boldsymbol{x}}}\big] \big[1 + \tanh(\beta K_y)\, \mu_{\boldsymbol{r}} \mu_{\boldsymbol{r}+\hat{\boldsymbol{y}}}\big] \quad , \tag{6.55}$$

[12] Thus we assume the model is ferromagnetic. If $J_{x,y} < 0$ the model is antiferromagnetic, which we may rewrite as a ferromagnet defining $\sigma'_{\boldsymbol{R}} = -\sigma_{\boldsymbol{R}}$ on one sublattice (*i.e.* the white squares on a chessboard) and retaining $\sigma'_{\boldsymbol{R}} = \sigma_{\boldsymbol{R}}$ on the other, effectively reversing the signs of $J_{x,y}$.

[13] Note that if $\exp(-2y) = \tanh x$ then $\exp(-2x) = \tanh y$.

where $\mu_{\boldsymbol{r}} \in \{+1, -1\}$ is an Ising spin located at the *site* $\boldsymbol{r}$. Note that $\boldsymbol{R}$ denotes a location of a *plaquette* while $\boldsymbol{r}$ denotes the location of a *site*. Each site can be uniquely associated with a plaquette (*e.g.*, the site in the lower left corner), so there are N sites and 2^N configurations $\{\mu_{\boldsymbol{r}}\}$. But from where do we get eqn. 6.55? Observe that the product on the RHS can be expressed as giant sum over 2^{2N} terms. Each such term has a graphical interpretation as a collection of nearest neighbor links $\langle \boldsymbol{r}\boldsymbol{r}' \rangle$, and we label each such collection by Θ. Unless each site $\boldsymbol{r}$ appears an even number of times (0, 2, or 4) in Θ, when we take the trace over the spin configurations the term will vanish. The surviving Θ configurations then correspond to the set of loop configurations $\{\Gamma\}$! We then have from eqn. 6.42 that

$$\mathrm{Tr} \prod_{\boldsymbol{r}} \big[1+\tanh(\beta K_x)\,\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}+\hat{\boldsymbol{x}}}\big]\big[1+\tanh(\beta K_y)\,\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}+\hat{\boldsymbol{y}}}\big] = \big[\cosh(\beta K_x)\big]^{-N}\big[\cosh(\beta K_y)\big]^{-N}\,\mathrm{Tr}\; e^{-\beta \hat{H}^{\mathrm{D}}(K_x,K_y)} \quad , \tag{6.56}$$

where

$$\hat{H}^{\mathrm{D}}(K_x, K_y) = -K_x \sum_{\boldsymbol{r}} \sigma_{\boldsymbol{r}}\,\sigma_{\boldsymbol{r}+\hat{\boldsymbol{x}}} - K_y \sum_{\boldsymbol{r}} \sigma_{\boldsymbol{r}}\,\sigma_{\boldsymbol{r}+\hat{\boldsymbol{y}}} \quad . \tag{6.57}$$

is the *dual Hamiltonian* defined on the *dual lattice* of sites. The plaquette lattice is *dual* to the site lattice (and vice versa). Therefore we have

$$Z(T, J_x, J_y) = \big[\sinh(2\beta K_x)\sinh(2\beta K_y)\big]^{-N/2}\, Z(T, K_x, K_y) \quad . \tag{6.58}$$

This establishes a *duality relation* between the original and inverted temperature regions. Assuming that the ferromagnetic square lattice model has a unique transition separating its high- and low-temperature phases, duality requires that transition to lie on the *self-dual line*. Thus, at the critical point the system must satisfy $J_{x,y} = K_{x,y}$, and we have the relation

$$\sinh(2\beta_{\mathrm{c}} J_x)\sinh(2\beta_{\mathrm{c}} J_y) = 1 \quad . \tag{6.59}$$

This equation defines a critical surface $T_{\mathrm{c}}(J_x, J_y)$. If $J_x = J_y = J$, we have

$$k_{\mathrm{B}} T_{\mathrm{c}} = \frac{2J}{\sinh^{-1}(1)} = \frac{2J}{\log\big(1+\sqrt{2}\big)} \simeq 2.26918\, J \quad . \tag{6.60}$$

Thus we have derived the exact critical temperature for the two-dimensional square lattice Ising model.

6.3 Nonideal Classical Gases

Let's switch gears now and return to the study of continuous classical systems described by a Hamiltonian $\hat{H}\big(\{\boldsymbol{x}_i\}, \{\boldsymbol{p}_i\}\big)$. In the next chapter, we will see how the critical properties of classical fluids can in fact be modeled by an appropriate *lattice gas* Ising model, and we'll derive methods for describing the liquid-gas phase transition in such a model.

6.3.1 The configuration integral

Consider the ordinary canonical partition function for a nonideal system of identical point particles interacting via a central two-body potential $u(r)$. We work in the ordinary canonical ensemble. The

N-particle partition function is

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\int\prod_{i=1}^{N}\frac{d^dp_i\,d^dx_i}{h^d}\,e^{-\hat{H}/k_{\rm B}T} \\ &= \frac{\lambda_T^{-Nd}}{N!}\int\prod_{i=1}^{N}d^dx_i\,\exp\left(-\frac{1}{k_{\rm B}T}\sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big)\right) \quad . \end{aligned} \tag{6.61}$$

Here, we have assumed a many body Hamiltonian of the form

$$\hat{H} = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m} + \sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big) \quad , \tag{6.62}$$

in which massive nonrelativistic particles interact via a two-body central potential. As before, the thermal wavelength is $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$. We can now write

$$Z(T,V,N) = \lambda_T^{-Nd}\,Q_N(T,V) \quad , \tag{6.63}$$

where the *configuration integral* $Q_N(T,V)$ is given by

$$Q_N(T,V) = \frac{1}{N!}\int d^dx_1\cdots\int d^dx_N\,\prod_{i<j}e^{-\beta u(r_{ij})} \quad . \tag{6.64}$$

There are no general methods for evaluating the configurational integral exactly.

6.3.2 One-dimensional Tonks gas

The Tonks gas is a one-dimensional generalization of the hard sphere gas. Consider a one-dimensional gas of indistinguishable particles of mass m interacting via the potential

$$u(x-x') = \begin{cases}\infty & \text{if } |x-x'|<a \\ 0 & \text{if } |x-x'|\geq a\end{cases} \quad . \tag{6.65}$$

Thus, the Tonks gas may be considered to be a gas of *hard rods*. The above potential guarantees that the portion of configuration space in which any rods overlap is *forbidden* in this model. Let the gas be placed in a finite volume L. The hard sphere nature of the particles means that no particle can get within a distance $\frac{1}{2}a$ of the ends at $x=0$ and $x=L$. That is, there is a one-body potential $v(x)$ acting as well, where

$$v(x) = \begin{cases}\infty & \text{if } x<\frac{1}{2}a \\ 0 & \text{if } \frac{1}{2}a\leq x\leq L-\frac{1}{2}a \\ \infty & \text{if } x>L-\frac{1}{2}a\end{cases} \quad . \tag{6.66}$$

The configuration integral of the 1D Tonks gas is given by

$$Q_N(T,L) = \frac{1}{N!}\int_0^L dx_1\cdots\int_0^L dx_N\,\chi(x_1,\ldots,x_N) \quad , \tag{6.67}$$

where $\chi = e^{-U/k_{\rm B}T}$ is zero if any two 'rods' (of length a) overlap, or if any rod overlaps with either boundary at $x = 0$ and $x = L$, and $\chi = 1$ otherwise. Note that χ does not depend on the temperature. Due to permutation symmetry, we may integrate over the subspace where $x_1 < x_2 < \cdots < x_N$ and then multiply the result by $N!$. Clearly x_j must lie to the right of $x_{j-1} + a$ as well as to the left of $Y_j \equiv L - Na + (j - \frac{1}{2})a$. Note that since $Y_j - a = Y_{j-1}$. Thus, the configurational integral is

$$Q_N(T,L) = \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-1}+a}^{Y_N} dx_N = \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-2}+a}^{Y_{N-1}} dx_{N-1} \left(Y_{N-1} - x_{N-1}\right) \tag{6.68}$$

$$= \frac{1}{2} \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-3}+a}^{Y_{N-2}} dx_{N-2} \left(Y_{N-2} - x_{N-2}\right)^2 = \cdots = \frac{1}{k!} \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-k-1}+a}^{Y_{N-k}} dx_{N-k} \left(Y_{N-k} - x_{N-k}\right)^k$$

$$= \frac{1}{N!} \left(Y_1 - \tfrac{1}{2}a\right)^N = \frac{1}{N!} \left(L - Na\right)^N \quad .$$

The partition function is $Z(T,L,N) = \lambda_T^{-N} \, Q_N(T,L)$, and so the free energy is

$$F = -k_{\rm B}T \log Z = -Nk_{\rm B}T \left\{ -\log \lambda_T + 1 + \log\left(\frac{L}{N} - a\right) \right\} \quad , \tag{6.69}$$

where we have used Stirling's rule to write $\log N! \approx N \log N - N$. The pressure is

$$p = -\frac{\partial F}{\partial L} = \frac{nk_{\rm B}T}{1 - na} \quad , \tag{6.70}$$

where $n = N/L$ is the one-dimensional density. Note that the pressure diverges as n approaches $1/a$. The usual one-dimensional ideal gas law, $pL = Nk_{\rm B}T$, is replaced by $pL_{\rm eff} = Nk_{\rm B}T$, where $L_{\rm eff} = L - Na$ is the 'free' volume obtained by subtracting the total "excluded volume" Na from the original volume L. Note the similarity here to the van der Waals equation of state, $(p + av^{-2})(v - b) = RT$, where $v = N_{\rm A}V/N$ is the molar volume. Defining $\tilde{a} \equiv a/N_{\rm A}^2$ and $\tilde{b} \equiv b/N_{\rm A}$, we have

$$p + \tilde{a}n^2 = \frac{nk_{\rm B}T}{1 - \tilde{b}n} \quad , \tag{6.71}$$

where $n = N_{\rm A}/v$ is the number density. The term involving the constant $\tilde{a}$ is due to the long-ranged attraction of atoms due to their mutual polarizability. The term involving $\tilde{b}$ is an excluded volume effect. The Tonks gas models only the latter.

6.3.3 Mayer cluster expansion

Let us return to the general problem of computing the configuration integral. Consider the function $e^{-\beta u_{ij}}$, where $u_{ij} \equiv u(|\boldsymbol{x}_i - \boldsymbol{x}_j|)$. We assume that at very short distances there is a strong repulsion between particles, *i.e.* $u_{ij} \to \infty$ as $r_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j| \to 0$, and that $u_{ij} \to 0$ as $r_{ij} \to \infty$. Thus, $e^{-\beta u_{ij}}$ vanishes as $r_{ij} \to 0$ and approaches unity as $r_{ij} \to \infty$. For our purposes, it will prove useful to define the function

$$f(r) = e^{-\beta u(r)} - 1 \quad , \tag{6.72}$$

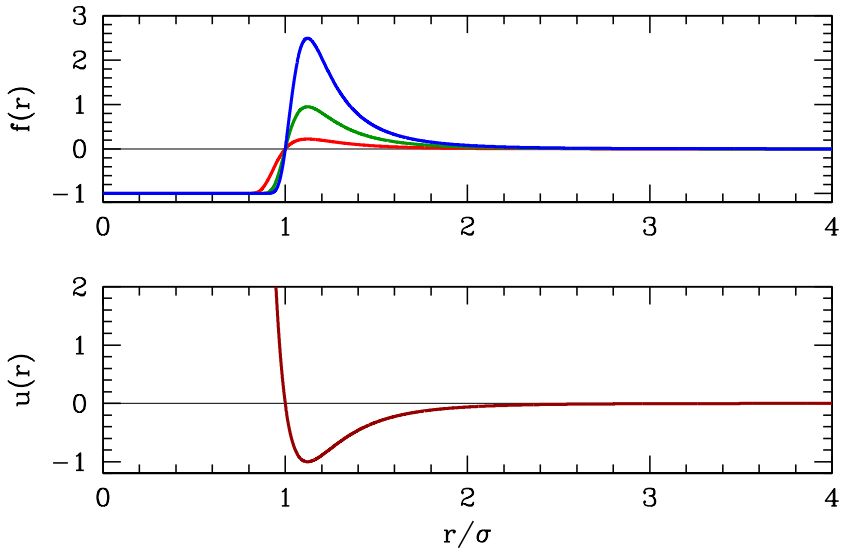


Figure 6.6: Bottom panel: Lennard-Jones potential $u(r) = 4\epsilon\left(x^{-12} - x^{-6}\right)$, with $x = r/\sigma$ and $\epsilon = 1$. Note the weak attractive tail and the strong repulsive core. Top panel: Mayer function $f(r,T) = e^{-u(r)/k_{\rm B}T} - 1$ for $k_{\rm B}T = 0.8\,\epsilon$ (blue), $k_{\rm B}T = 1.5\,\epsilon$ (green), and $k_{\rm B}T = 5\,\epsilon$ (red).

called the *Mayer function* after Josef Mayer. We may now write

$$Q_N(T,V) = \frac{1}{N!}\int d^dx_1 \cdots \int d^dx_N \prod_{i<j}\left(1 + f_{ij}\right) \quad . \tag{6.73}$$

A typical potential we might consider is the semi-phenomenological Lennard-Jones potential,

$$u(r) = 4\,\epsilon\left\{\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right\} \quad . \tag{6.74}$$

This accounts for a long-distance attraction due to mutually induced electric dipole fluctuations, and a strong short-ranged repulsion, phenomenologically modelled with a r^{-12} potential, which mimics a hard core due to overlap of the atomic electron distributions. Setting $u'(r) = 0$ we obtain $r^* = 2^{1/6}\,\sigma \approx 1.12246\,\sigma$ at the minimum, where $u(r^*) = -\epsilon$. In contrast to the Boltzmann weight $e^{-\beta u(r)}$, the Mayer function $f(r)$ vanishes as $r \to \infty$, behaving as $f(r) \sim -\beta u(r)$. The Mayer function also depends on temperature. Sketches of $u(r)$ and $f(r)$ for the Lennard-Jones model are shown in fig. 6.6.

The Lennard-Jones potential[14] is realistic for certain simple fluids, but it leads to a configuration integral which is in general impossible to evaluate. Indeed, even a potential as simple as that of the hard sphere gas is intractable in more than one space dimension. We can however make progress by deriving a series expansion for the equation of state in powers of the particle density. This is known as the *virial*

[14]Disambiguation footnote: Take care not to confuse Philipp Lenard (Hungarian-German, cathode ray tubes, Nazi), Alfred-Marie Liénard (French, Liénard-Wiechert potentials, not a Nazi), John Lennard-Jones (British, molecular structure, also not a Nazi), and Lynyrd Skynyrd (American, "Free Bird"). I thank my colleague Oleg Shpyrko for setting me straight on this.

$$(f_{1,4}\, f_{4,7}\, f_{4,9}\, f_{7,9})\,(f_{2,5}\, f_{2,6})\,(f_{3,10})\,(f_{8,11})$$

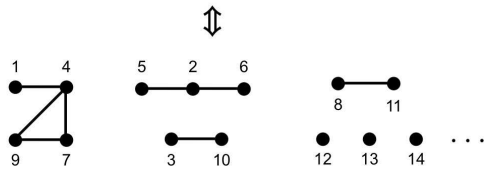


Figure 6.7: Diagrammatic interpretation of a term involving a product of eight Mayer functions.

expansion. As was the case when we investigated noninteracting quantum statistics, it is convenient to work in the grand canonical ensemble and to derive series expansions for the density $n(T,z)$ and the pressure $p(T,z)$ in terms of the fugacity z, then solve for $z(T,n)$ to obtain $p(T,n)$. These expansions in terms of fugacity have a nifty diagrammatic interpretation, due to Mayer.

We begin by expanding the product in eqn. 6.73 as

$$\prod_{i<j}\big(1+f_{ij}\big) = 1 + \sum_{i<j} f_{ij} + \sum_{\substack{i<j\,,\,k<l\\(ij)\neq(kl)}} f_{ij}\, f_{kl} + \ldots \quad . \tag{6.75}$$

As there are $\frac{1}{2}N(N-1)$ possible pairings, there are $2^{N(N-1)/2}$ terms in the expansion of the above product. Each such term may be represented by a graph, as shown in fig. 6.7. For each such term, we draw a connection between dots representing different particles i and j if the factor f_{ij} appears in the term under consideration. The contribution for any given graph may be written as a product over contributions from each of its disconnected component clusters. For example, in the case of the term in fig. 6.7, the contribution to the configurational integral would be

$$\begin{aligned}\Delta Q = \frac{V^{N-11}}{N!}\int d^dx_1\, d^dx_4\, d^dx_7\, d^dx_9\; f_{1,4}\, f_{4,7}\, f_{4,9}\, f_{7,9} \\ \times \int d^dx_2\, d^dx_5\, d^dx_6\; f_{2,5}\, f_{2,6} \times \int d^dx_3\, d^dx_{10}\; f_{3,10} \times \int d^dx_8\, d^dx_{11}\; f_{8,11} \quad .\end{aligned} \tag{6.76}$$

We will refer to a given product of Mayer functions which arises from this expansion as a *term*.

The particular labels we assign to each vertex of a given graph don't affect the overall value of the graph. Now a given unlabeled graph consists of a certain number of connected subgraphs. For a system with N particles, we may then write

$$N = \sum_\gamma m_\gamma\, n_\gamma \quad , \tag{6.77}$$

where γ ranges over all possible connected subgraphs, and

$$\begin{aligned} m_\gamma &= \text{number of connected subgraphs of type } \gamma \text{ in the unlabeled graph} \\ n_\gamma &= \text{number of vertices in the connected subgraph } \gamma \quad . \end{aligned}$$

Note that the single vertex $\bullet$ counts as a connected subgraph, with $n_\bullet = 1$. We now ask: how many ways are there of assigning the N labels to the N vertices of a given unlabeled graph? One might first think the answer is simply $N!$, however this is too big, because different assignments of the labels to the vertices

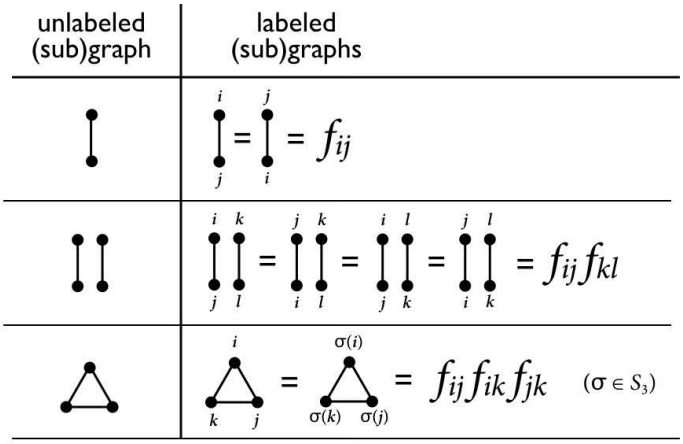


Figure 6.8: Different assignments of labels to vertices may not result in a distinct term in the expansion of the configuration integral.

may not result in a distinct graph. To see this, consider the examples in fig. 6.8. In the first example, an unlabeled graph with four vertices consists of two identical connected subgraphs. Given any assignment of labels to the vertices, then, we can simply exchange the two subgraphs and get the same term. So we should divide $N!$ by the product $\prod_\gamma m_\gamma!$. But even this is not enough, because within each connected subgraph γ there may be permutations which leave the integrand unchanged, as shown in the second and third examples in fig. 6.8. We define the *symmetry factor* s_γ as the number of permutations of the labels which leaves a given connected subgraphs γ invariant. Examples of symmetry factors are shown in fig. 6.9. Consider, for example, the third subgraph in the top row. Clearly one can rotate the figure about its horizontal symmetry axis to obtain a new labeling which represents the same term. This twofold axis is the only symmetry the diagram possesses, hence $s_\gamma = 2$. For the first diagram in the second row, one can rotate *either* of the triangles about the horizontal symmetry axis. One can also rotate the figure in the plane by 180° so as to exchange the two triangles. Thus, there are $2 \times 2 \times 2 = 8$ symmetry operations which result in the same term, and $s_\gamma = 8$. Finally, the last subgraph in the second row consists of five vertices each of which is connected to the other four. Therefore any permutation of the labels results in the same term, and $s_\gamma = 5! = 120$. In addition to dividing by the product $\prod_\gamma m_\gamma!$, we must then also divide by $\prod_\gamma s_\gamma^{m_\gamma}$.

We can now write the partition function as

$$\begin{aligned} Z &= \frac{\lambda_T^{-Nd}}{N!} \sum_{\{m_\gamma\}} \frac{N!}{\prod m_\gamma!\, s_\gamma^{m_\gamma}} \cdot \prod_\gamma \left(\int d^d x_1 \cdots d^d x_{n_\gamma} \prod_{i<j}^{\gamma} f_{ij} \right)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \\ &= \sum_{\{m_\gamma\}} \prod_\gamma \frac{1}{m_\gamma!} \left(\frac{V b_\gamma(T)}{\lambda_T^d} \right)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \end{aligned} \tag{6.78}$$

where the product $\prod_{i<j}^{\gamma} f_{ij}$ is over all links in the subgraph γ. The final Kronecker delta enforces the

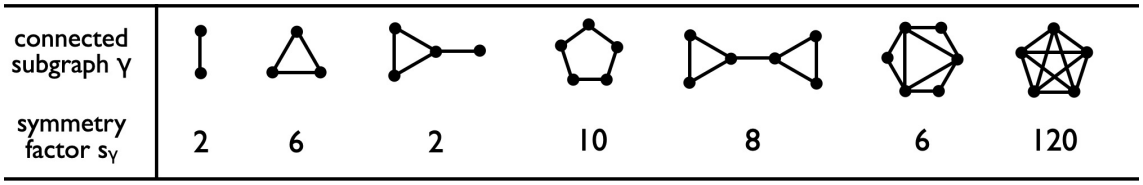


Figure 6.9: The symmetry factor s_γ for a connected subgraph γ is the number of permutations of its indices which leaves the term $\prod_{(ij)\in\gamma} f_{ij}$ invariant.

constraint $N = \sum_\gamma m_\gamma\, n_\gamma$. We have defined the *dimensionless cluster integrals* b_γ as

$$b_\gamma(T) \equiv \frac{1}{s_\gamma}\int\frac{d^dx_1}{\lambda_T^d}\cdots\int\frac{d^dx_{n_\gamma-1}}{\lambda_T^d}\prod_{i<j}^{\gamma} f_{ij} \quad , \tag{6.79}$$

where we assume the limit $V\to\infty$. Since $f_{ij} = f\big(|\boldsymbol{x}_i - \boldsymbol{x}_j|\big)$, the product $\prod_{i<j}^{\gamma} f_{ij}$ is invariant under simultaneous translation of all the coordinate vectors by any constant vector, and hence the integral over the n_γ position variables contains exactly one factor of the volume, which yields factor of V within the round brackets in the second line of eqn. 6.78. Thus, each cluster integral is *intensive*[15], scaling as V^0.

If we compute the grand partition function, then the fixed N constraint is relaxed, and we can do the sums:

$$\begin{aligned}\Xi(T,V,\mu) &= \sum_{\{m_\gamma\}}\big(e^{\beta\mu}\big)^{\sum m_\gamma n_\gamma}\prod_\gamma\frac{1}{m_\gamma!}\left(\frac{Vb_\gamma(T)}{\lambda_T^d}\right)^{m_\gamma}\\ &= \prod_\gamma\sum_{m_\gamma=0}^{\infty}\frac{1}{m_\gamma!}\left(\frac{Vz^{n_\gamma}\,b_\gamma(T)}{\lambda_T^d}\right)^{m_\gamma} = \exp\left(V\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma\right) \quad ,\end{aligned} \tag{6.80}$$

where $z = \exp(\beta\mu)$ is the fugacity. Thus, since $\Omega = -k_{\rm B}T\log\Xi$,

$$\Omega(T,V,\mu) = -\frac{Vk_{\rm B}T}{\lambda_T^d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T) \quad , \tag{6.81}$$

and we can write

$$\begin{aligned}p &= k_{\rm B}T\,\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T)\\ n &= \lambda_T^{-d}\sum_\gamma n_\gamma\,z^{n_\gamma}\,b_\gamma(T) \quad ,\end{aligned} \tag{6.82}$$

where $b_\bullet \equiv 1$. As in the case of ideal quantum gas statistical mechanics, we can systematically invert the relation $n = n(z,T)$ to obtain $z = z(n,T)$, and then insert this into the equation for $p(z,T)$ to obtain the equation of state $p = p(n,T)$. This yields the *virial expansion* of the equation of state,

$$p = nk_{\rm B}T\Big\{1 + B_2(T)\,n + B_3(T)\,n^2 + \ldots\Big\} \quad . \tag{6.83}$$

[15]We assume that the long-ranged behavior of $f(r)\approx -\beta u(r)$ is integrable.

It is useful to define the dimensionless quantities $\nu \equiv n\lambda_T^d$ and $\pi = p\lambda_T^d/k_{\rm B}T$, as well as the dimensionless cluster integral sums

$$\mathsf{b}_k \equiv \sum_\gamma b_\gamma\,\delta_{k,n_\gamma} \quad , \tag{6.84}$$

which is the sum of all cluster integrals b_γ with $n_\gamma = k$ vertices. Then

$$\nu(z) = \sum_{k=1}^\infty k\,\mathsf{b}_k\,z^k \qquad , \qquad \pi(z) = \sum_{k=1}^\infty \mathsf{b}_k\,z^k \quad . \tag{6.85}$$

The virial expansion of the dimensionless equation of state is then

$$\pi(\nu) = \sum_{k=1}^\infty \mathsf{B}_k\,\nu^k \tag{6.86}$$

We may again apply the Lagrange method introduced in §5.3.2 for the quantum virial coefficients, writing

$$\mathsf{B}_k = \oint \frac{d\nu}{2\pi i}\,\frac{\pi(\nu)}{\nu^{k+1}} = \oint \frac{dz}{2\pi i}\,\frac{\nu'(z)\,\pi(z)}{\big[\nu(z)\big]^{k+1}} = -\frac{1}{k}\oint\frac{dz}{2\pi i}\,\pi(z)\,\frac{d}{dz}\big[\nu(z)\big]^{-k} \quad , \tag{6.87}$$

where the contour encloses the origin in the complex plane. Integrating by parts, and using the relation $\pi'(z) = \nu(z)/z$, we obtain[16]

$$\begin{aligned}\mathsf{B}_k &= \frac{1}{k}\oint\frac{dz}{2\pi i}\,\pi'(z)\,\big[\nu(z)\big]^{-k} = \frac{1}{k}\oint\frac{dz}{2\pi i}\,\frac{1}{z}\,\big[\nu(z)\big]^{1-k} \\ &= \frac{1}{k}\oint\frac{dz}{2\pi i}\,\frac{1}{z^k}\,\Big(1 + 2\mathsf{b}_2 z + 3\mathsf{b}_3 z^2 + 4\mathsf{b}_4 z^4 + \ldots\Big)^{1-k} \quad ,\end{aligned} \tag{6.88}$$

where the contour is a small circle enclosing the origin. Working out the first two virial coefficients, we find

$$\mathsf{B}_2 = -\mathsf{b}_2 \qquad , \qquad \mathsf{B}_3 = 4\mathsf{b}_2^2 - 2\mathsf{b}_3 \quad . \tag{6.89}$$

The dimensionful virial coefficients in eqn. 6.83 are then given by $B_k = \mathsf{B}_k\,\lambda_T^{(k-1)d}$.

Lowest order expansion

We have

$$b_-(T) = \tfrac{1}{2}\int\frac{d^dx_1}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big) = \tfrac{1}{2}\int\frac{d^dr}{\lambda_T^d}\,f(r) \tag{6.90}$$

and

$$\begin{aligned}b_\wedge(T) &= \tfrac{1}{2}\int\frac{d^dx_1}{\lambda_T^d}\int\frac{d^dx_2}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big)\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_3|\big) \\ &= \tfrac{1}{2}\int\frac{d^dr}{\lambda_T^d}\int\frac{d^dr'}{\lambda_T^d}\,f(r)\,f(r') = 2\big(b_-\big)^2\end{aligned} \tag{6.91}$$

[16]Since there is no term proportional to $\log w$ in the Laurent expansion of $\pi(w)\big[n(w)\big]^{-k}$, there is no residue arising from integrating its derivative around the unit circle.

and

$$\begin{aligned} b_\triangle(T) &= \tfrac{1}{6}\int\frac{d^dx_1}{\lambda_T^d}\int\frac{d^dx_2}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1-\boldsymbol{x}_2|\big)\,f\big(|\boldsymbol{x}_1-\boldsymbol{x}_3|\big)\,f\big(|\boldsymbol{x}_2-\boldsymbol{x}_3|\big) \\ &= \tfrac{1}{6}\int\frac{d^dr}{\lambda_T^d}\int\frac{d^dr'}{\lambda_T^d}\,f(r)\,f(r')\,f\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) \quad . \end{aligned} \tag{6.92}$$

Thus we have $\mathsf{b}_2 = b_-$ and $\mathsf{b}_3 = b_\wedge + b_\triangle = 2b_-^2 + b_\triangle$. From eqn. 6.89 we now have

$$\begin{aligned} \mathsf{B}_2(T) &= -\mathsf{b}_2(T) = -b_-(T) \\ \mathsf{B}_3(T) &= \big[4\mathsf{b}_2^2(T) - 2\mathsf{b}_3(T)\big] = -2b_\triangle(T) \quad . \end{aligned} \tag{6.93}$$

Note that $b_\wedge$ does not contribute to B_3, even though the graph $\wedge$ has three vertices, and only $b_\triangle$ appears. This is because the virial coefficients B_j involve only cluster integrals b_γ for *one-particle irreducible* clusters, *i.e.* those clusters which remain connected and don't fall into multiple pieces if any of its vertices is removed, as depicted in fig. 6.10.

Cookbook recipe

Just follow these simple steps:

- The pressure and number density are written as sums over unlabeled connected clusters γ, *viz.*

$$\begin{aligned} p &= k_{\rm B}T\,\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T) \\ n &= \lambda_T^{-d}\sum_\gamma n_\gamma\,z^{n_\gamma}\,b_\gamma(T) \quad , \end{aligned} \tag{6.94}$$

  where $z = \exp(\beta\mu)$ is the fugacity.

- To compute the dimensionless cluster integral $b_\gamma(T)$, first draw the connected cluster γ with *unlabeled* vertices.

- Next, assign labels $1, 2, \ldots, n_\gamma$ to the vertices, where n_γ is the total number of vertices in the cluster γ. It doesn't matter how you assign the labels.

- Write down the product $\prod_{i<j}^\gamma f_{ij}$. The factor f_{ij} appears in the product if there is a link in your (now labeled) cluster between sites i and j.

- The symmetry factor s_γ is the number of elements of the symmetric group S_{n_γ} which leave the product $\prod_{i<j}^\gamma f_{ij}$ invariant. The identity permutation leaves the product invariant, so $s_\gamma \geq 1$.

- The dimensionless cluster integral $b_\gamma(T)$ is given by

$$b_\gamma(T) \equiv \frac{1}{s_\gamma}\int\frac{d^dx_1}{\lambda_T^d}\cdots\int\frac{d^dx_{n_\gamma-1}}{\lambda_T^d}\prod_{i<j}^\gamma f_{ij} \quad , \tag{6.95}$$

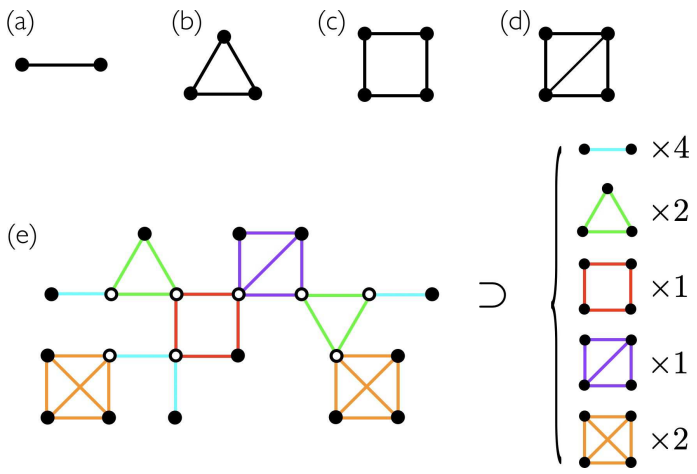


Figure 6.10: Connected *versus* irreducible clusters. Clusters (a) through (d) are *irreducible* in that they remain connected if any component site and its connecting links are removed. Cluster (e) is connected, but is *reducible*. Its integral b_γ is proportional to a product over its irreducible components, each shown in a unique color, and occurring with various multiplicities. The open circles denote *articulation points*. Removal of an articulation point and all the links connected to it results in a disconnected diagram. Removal of any of the closed circles and its associated links does not result in a disconnected diagram.

Due to translation invariance, $b_\gamma(T) \propto V^0$. One can therefore set $\boldsymbol{x}_{n_\gamma} \equiv 0$, eliminate the volume factor from the denominator, and perform the integral over the remaining $n_\gamma - 1$ coordinates.

- This procedure generates expansions for $p(T,z)$ and $n(T,z)$ in powers of the fugacity $z = \exp(\beta\mu)$. To obtain something useful like $p(T,n)$, we must invert the equation $n = n(T,z)$ to find $z = z(T,n)$, and then substitute into the equation $p = p(T,z)$ to obtain $p = p\big(T, z(T,n)\big) = p(T,n)$. The result is the virial expansion,

$$p = nk_{\rm B}T\left\{1 + B_2(T)\,n + B_3(T)\,n^2 + \ldots\right\} \quad , \tag{6.96}$$

where

$$B_k(T) = -(k-1)\,\lambda_T^{(k-1)d} \sum_{\gamma\in\Gamma_k} b_\gamma(T) \quad , \tag{6.97}$$

with Γ_k the set of all *irreducible* k-site clusters. An irreducible cluster is one which contains no articulation points, *i.e.* it does not fall apart into multiple pieces if any given vertex and all its connecting links are removed.

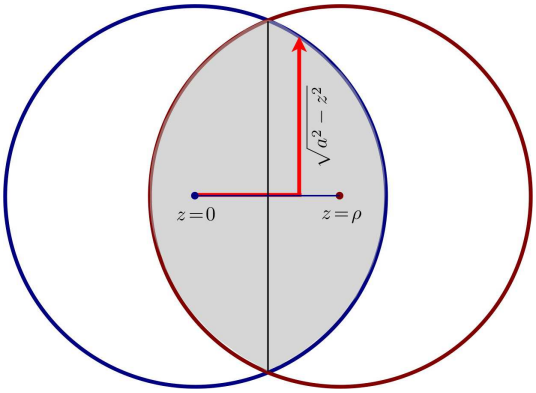


Figure 6.11: The overlap of hard sphere Mayer functions. The shaded volume is $\mathcal{V}$.

6.3.4 Examples

Hard sphere gas in three dimensions

The hard sphere potential is given by

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.98}$$

Here a is the *diameter* of the spheres. The corresponding Mayer function is then temperature independent, and given by

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.99}$$

We can change variables to obtain

$$\mathsf{b}_2(T) = \tfrac{1}{2}\int \frac{d^3r}{\lambda_T^3}\, f(r) = -\tfrac{2}{3}\pi a^3 \lambda_T^{-3} \quad . \tag{6.100}$$

The calculation of b_3 is more challenging. We have

$$\mathsf{b}_3 = \tfrac{1}{6}\int \frac{d^3\rho}{\lambda_T^3}\int \frac{d^3r}{\lambda_T^3}\, f(\rho)\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \quad . \tag{6.101}$$

We must first compute the volume of overlap for spheres of *radius* a (recall a is the *diameter* of the constituent hard sphere particles) centered at 0 and at $\boldsymbol{\rho}$:

$$\begin{aligned} \mathcal{V} &= \int d^3r\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \\ &= 2\int_{\rho/2}^{a} dz\, \pi(a^2-z^2) = \tfrac{4\pi}{3}a^3 - \pi a^2\rho + \tfrac{\pi}{12}\rho^3 \quad . \end{aligned} \tag{6.102}$$

We then integrate over region $|\boldsymbol{\rho}| < a$, to obtain

$$\mathsf{b}_3 = -\tfrac{1}{6} \cdot 4\pi\lambda_T^{-6} \int_0^a d\rho\, \rho^2 \cdot \left\{ \tfrac{4\pi}{3} a^3 - \pi a^2 \rho + \tfrac{\pi}{12}\, \rho^3 \right\} = -\tfrac{5\pi^2}{36}\, a^6 \lambda_T^{-6} \quad . \tag{6.103}$$

Thus, we have

$$B_2 = -\lambda_T^3\, \mathsf{b}_2(T) = \tfrac{2\pi}{3} a^3 = 4\, v_0 \qquad , \qquad B_3 = -2\lambda_T^6\, \mathsf{b}_3(T) = \tfrac{5\pi^2}{18} a^6 = 10\, v_0^2 \quad , \tag{6.104}$$

where $v_0 = \pi a^3/6$ is the hard sphere volume. The equation of state is then

$$p = n k_{\rm B} T \left\{ 1 + B_2\, n + B_3\, n^2 + \mathcal{O}(n^3) \right\} \quad . \tag{6.105}$$

B_4 for the three-dimensional hard sphere gas was calculated by van Laar[17] and is given by

$$B_4 = \left(\tfrac{2707}{70} + \tfrac{219\sqrt{2}}{35\pi} - \tfrac{4131}{70\pi} \cos^{-1}(1/3) \right) v_0^3 \quad . \tag{6.106}$$

The coefficients beyond B_4 require numerical evaluation. The *packing fraction* of hard spheres is given by $\varphi = n v_0$. There is a first order, entropy-driven phase transition as a function of pressure at $\beta p_{\rm c} a^3 = 11.6$, *i.e.* $p_{\rm c}(T) = 11.6\, k_{\rm B} T a^{-3}$. The coexistence region (see fig. 6.12) lies between the packing fraction for freezing, $\varphi_{\rm f} = 0.492$, and the packing fraction for melting, $\varphi_{\rm m} = 0.543$. The critical volume fraction for dense random packing is $\varphi_{\rm drp} \approx 0.64$, and the maximum possible packing occurs for crystalline close packing (face-centered cubic or hexagonal close-packed structures) at $\varphi_{\rm ccp} = \frac{\pi}{3\sqrt{2}} \approx 0.740$.

For the hard disk gas in $d = 2$, one has $v_0 = \frac{1}{4}\pi a^2$ where a again is the diameter, and

$$B_2 = 2\, v_0 \qquad , \qquad B_3 = \left(\tfrac{16}{3} - \tfrac{4\sqrt{3}}{\pi} \right) v_0^2 \qquad , \qquad B_4 = \left(16 - \tfrac{36\sqrt{3}}{\pi} + \tfrac{80}{\pi^2} \right) v_0^3 \quad . \tag{6.107}$$

Weakly attractive tail

Suppose

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -u_0(r) & \text{if } r > a \end{cases} \quad . \tag{6.108}$$

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta u_0(r)} - 1 & \text{if } r > a \end{cases} \quad . \tag{6.109}$$

Thus,

$$\mathsf{b}_2(T) = \tfrac{1}{2} \int \frac{d^3 r}{\lambda_T^3}\, f(r) = -\tfrac{2\pi}{3} a^3 \lambda_T^{-3} + 2\pi\lambda_T^{-3} \int_a^\infty dr\, r^2 \left[e^{\beta u_0(r)} - 1 \right] \quad . \tag{6.110}$$

[17] J. J. Van Laar, Koninklijke Nederlandse Akademie van Wetenschappen Amsterdam *Prov. Sec. Sci.* **1**, 273 (1899).

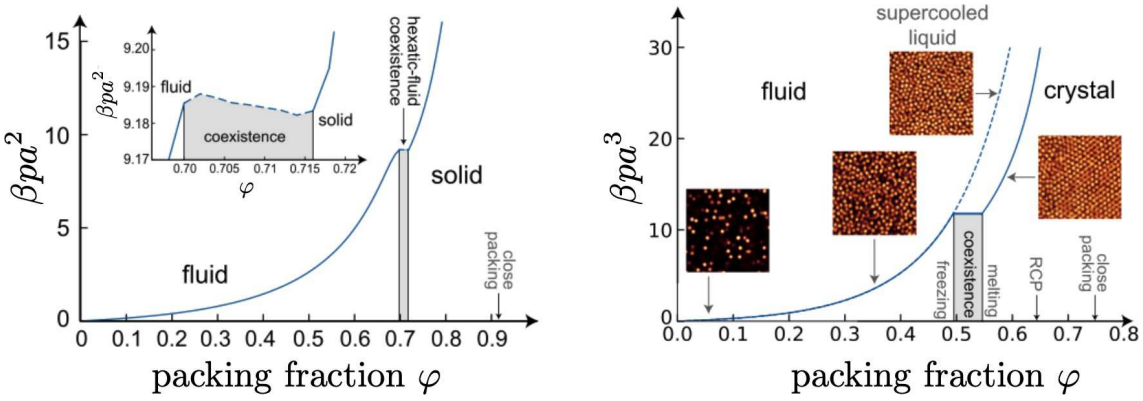


Figure 6.12: Phase diagram of the hard disk (left) and hard sphere gas (right), from G. P. Royall *et al.*, *Rev. Mod. Phys.* **96**, 045003 (2024).

Thus, the second virial coefficient is

$$B_2(T) = -\lambda_T^3 \, \mathsf{b}_2(T) \approx \tfrac{2\pi}{3} a^3 - \frac{2\pi}{k_{\rm B}T} \int\limits_a^\infty dr \, r^2 \, u_0(r) \quad , \tag{6.111}$$

where we have assumed $|u_0(r)| \ll k_{\rm B}T$. We see that the second virial coefficient *changes sign* at some temperature T_0, from a negative low temperature value to a positive high temperature value.

Spherical potential well

Consider an attractive spherical well potential with an infinitely repulsive core,

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -\epsilon & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.112}$$

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta\epsilon} - 1 & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.113}$$

Writing $s \equiv R/a$, we have

$$\begin{aligned} B_2(T) = -\lambda_T^3 \, \mathsf{b}_2(T) &= -\tfrac{1}{2} \int d^3r \, f(r) \\ &= -\frac{1}{2} \left\{ (-1) \cdot \tfrac{4\pi}{3} a^3 + \left(e^{\beta\epsilon} - 1 \right) \cdot \tfrac{4\pi}{3} a^3 (s^3 - 1) \right\} \\ &= \tfrac{2\pi}{3} a^3 \left\{ 1 - (s^3 - 1)\left(e^{\beta\epsilon} - 1 \right) \right\} \quad . \end{aligned} \tag{6.114}$$

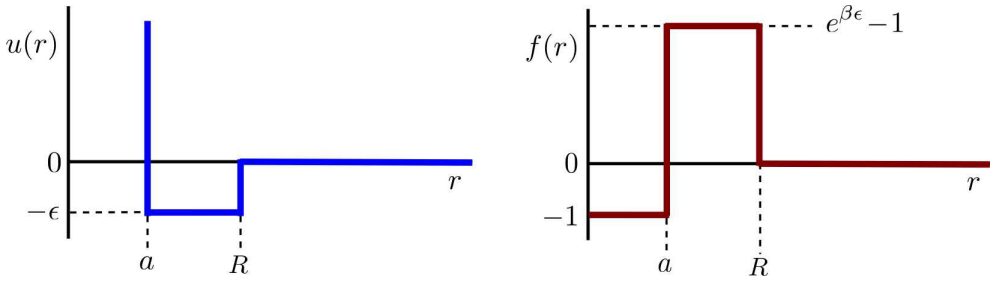


Figure 6.13: An attractive spherical well with a repulsive core $u(r)$ and its associated Mayer function $f(r)$.

To find the temperature T_0 where $B_2(T)$ changes sign, we set $B_2(T_0) = 0$ and obtain

$$k_{\rm B}T_0 = \epsilon \Big/ \log\left(\frac{s^3}{s^3-1}\right) \quad . \tag{6.115}$$

Recall in our study of the thermodynamics of the Joule-Thomson effect in §2.10.7 that the throttling process is *isenthalpic*. The temperature change, when a gas is pushed (or escapes) through a porous plug from a high pressure region to a low pressure one is

$$\Delta T = \int\limits_{p_1}^{p_2} dp \, \left(\frac{\partial T}{\partial p}\right)_H \quad , \tag{6.116}$$

where

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{C_p}\left[T\left(\frac{\partial V}{\partial T}\right)_p - V\right] \quad . \tag{6.117}$$

Appealing to the virial expansion, and working to lowest order in corrections to the ideal gas law, we have

$$p = \frac{N}{V}\,k_{\rm B}T + \frac{N^2}{V^2}\,k_{\rm B}T\,B_2(T) + \ldots \tag{6.118}$$

and we compute $\left(\frac{\partial V}{\partial T}\right)_p$ by setting

$$0 = dp = -\frac{Nk_{\rm B}T}{V^2}\,dV + \frac{Nk_{\rm B}}{V}\,dT - \frac{2N^2}{V^3}\,k_{\rm B}T\,B_2(T)\,dV + \frac{N^2}{V^2}\,d\big(k_{\rm B}T\,B_2(T)\big) + \ldots \quad . \tag{6.119}$$

Dividing by dT, we find

$$T\left(\frac{\partial V}{\partial T}\right)_p - V = N\left[T\,\frac{\partial B_2}{\partial T} - B_2\right] \quad . \tag{6.120}$$

The temperature where $\left(\frac{\partial T}{\partial p}\right)_H$ changes sign is called the *inversion temperature* T^*. To find the inversion point, we set $T^*\,B_2'(T^*) = B_2(T^*)$, *i.e.*

$$\left.\frac{d\log B_2}{d\log T}\right|_{T^*} = 1 \quad . \tag{6.121}$$

If we approximate $B_2(T) \approx A - \frac{B}{T}$, then the inversion temperature follows simply:

$$\frac{B}{T^*} = A - \frac{B}{T^*} \qquad \Longrightarrow \qquad T^* = \frac{2B}{A} \quad . \tag{6.122}$$

Hard spheres with a hard wall

Consider a hard sphere gas in three dimensions in the presence of a hard wall at $z = 0$. The gas is confined to the region $z > 0$. The total potential energy is now

$$W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \sum_i v(\boldsymbol{x}_i) + \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad , \tag{6.123}$$

where

$$v(\boldsymbol{r}) = v(z) = \begin{cases} \infty & \text{if} \quad z \le \frac{1}{2}a \\ 0 & \text{if} \quad z > \frac{1}{2}a \end{cases} \quad , \tag{6.124}$$

and $u(\boldsymbol{r})$ is given in eqn. 6.98. The grand potential is written as a series in the total particle number N, and is given by

$$\Xi = e^{-\beta\Omega} = 1 + \xi \int d^3r\, e^{-\beta v(z)} + \tfrac{1}{2}\xi^2 \int d^3r \int d^3r'\, e^{-\beta v(z)}\, e^{-\beta v(z')}\, e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} + \ldots \quad , \tag{6.125}$$

where $\xi = z\,\lambda_T^{-3}$, with $z = e^{\mu/k_{\rm B}T}$ the fugacity. Taking the logarithm, and invoking the Taylor series $\log(1+\delta) = \delta - \frac{1}{2}\delta^2 + \frac{1}{3}\delta^3 - \ldots$, we obtain

$$-\beta\Omega = \xi \int\limits_{z>\frac{a}{2}} d^3r + \tfrac{1}{2}\xi^2 \int\limits_{z>\frac{a}{2}} d^3r \int\limits_{z'>\frac{a}{2}} d^3r' \left[e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} - 1 \right] + \ldots \tag{6.126}$$

The volume is $V = \int\limits_{z>0} d^3r$. Dividing by V, we have, in the thermodynamic limit,

$$\begin{aligned} -\frac{\beta\Omega}{V} = \beta p &= \xi + \tfrac{1}{2}\xi^2 \frac{1}{V} \int\limits_{z>\frac{a}{2}} d^3r \int\limits_{z'>\frac{a}{2}} d^3r' \left[e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} - 1 \right] + \ldots \\ &= \xi - \tfrac{2}{3}\pi a^3\,\xi^2 + \mathcal{O}(\xi^3) \quad . \end{aligned} \tag{6.127}$$

The number density is

$$n = \xi\,\frac{\partial}{\partial\xi}\,(\beta p) = \xi - \tfrac{4}{3}\pi a^3\,\xi^2 + \mathcal{O}(\xi^3) \quad , \tag{6.128}$$

and inverting to obtain $\xi(n)$ and then substituting into the pressure equation, we obtain the lowest order virial expansion for the equation of state,

$$p = k_{\rm B}T \left\{ n + \tfrac{2}{3}\pi a^3\, n^2 + \ldots \right\} \quad . \tag{6.129}$$

As expected, the presence of the wall does not affect a bulk property such as the equation of state.

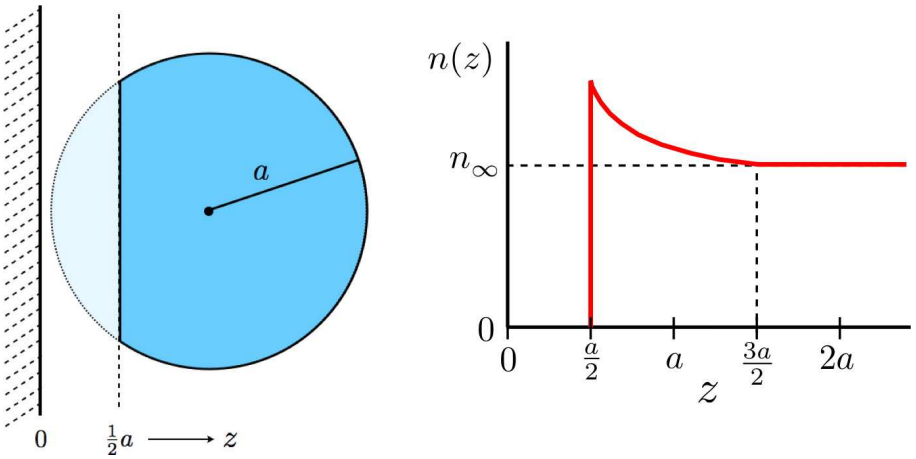


Figure 6.14: In the presence of a hard wall, the Mayer sphere is cut off on the side closest to the wall. The resulting density $n(z)$ vanishes for $z < \frac{1}{2}a$ since the center of each sphere must be at least one radius ($\frac{1}{2}a$) away from the wall. Between $z = \frac{1}{2}a$ and $z = \frac{3}{2}a$ there is a density *enhancement*. If the calculation were carried out to higher order, $n(z)$ would exhibit damped spatial oscillations with wavelength $\lambda \sim a$.

Next, let us compute the number density $n(z)$, given by

$$n(z) = \Big\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \Big\rangle \quad . \tag{6.130}$$

Due to translational invariance in the (x, y) plane, we know that the density must be a function of z alone. The presence of the wall at $z = 0$ breaks translational symmetry in the z direction. The number density is

$$\begin{aligned} n(z) &= \mathrm{Tr}\left[e^{\beta(\mu\hat{N} - \hat{H})} \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \right] \Big/ \mathrm{Tr}\, e^{\beta(\mu\hat{N} - \hat{H})} \\ &= \Xi^{-1} \left\{ \xi\, e^{-\beta v(z)} + \xi^2\, e^{-\beta v(z)} \int d^3r'\, e^{-\beta v(z')}\, e^{-\beta u(\boldsymbol{r} - \boldsymbol{r}')} + \ldots \right\} \\ &= \xi\, e^{-\beta v(z)} + \xi^2\, e^{-\beta v(z)} \int d^3r'\, e^{-\beta v(z')} \left[e^{-\beta u(\boldsymbol{r} - \boldsymbol{r}')} - 1 \right] + \ldots \quad . \end{aligned} \tag{6.131}$$

Note that the term in square brackets in the last line is the Mayer function $f(\boldsymbol{r} - \boldsymbol{r}') = e^{-\beta u(\boldsymbol{r} - \boldsymbol{r}')} - 1$. Consider the function

$$e^{-\beta v(z)}\, e^{-\beta v(z')}\, f(\boldsymbol{r} - \boldsymbol{r}') = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \text{ or } z' < \frac{1}{2}a \\ 0 & \text{if } |\boldsymbol{r} - \boldsymbol{r}'| > a \\ -1 & \text{if } z > \frac{1}{2}a \text{ and } z' > \frac{1}{2}a \text{ and } |\boldsymbol{r} - \boldsymbol{r}'| < a \end{cases} \quad . \tag{6.132}$$

Now consider the integral of the above function with respect to $\boldsymbol{r}'$. Clearly the result depends on the value of z. If $z > \frac{3}{2}a$, then there is no excluded region in $\boldsymbol{r}'$ and the integral is (-1) times the full Mayer sphere volume, *i.e.* $-\frac{4}{3}\pi a^3$. If $z < \frac{1}{2}a$ the integral vanishes due to the $e^{-\beta v(z)}$ factor. For z infinitesimally

larger than $\frac{1}{2}a$, the integral is (-1) times half the Mayer sphere volume, *i.e.* $-\frac{2}{3}\pi a^3$. For $z \in \left[\frac{a}{2}\,,\,\frac{3a}{2}\right]$ the integral interpolates between $-\frac{2}{3}\pi a^3$ and $-\frac{4}{3}\pi a^3$. Explicitly, one finds by elementary integration,

$$\int d^3r'\, e^{-\beta v(z)}\, e^{-\beta v(z')}\, f(\boldsymbol{r}-\boldsymbol{r}') = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \\ \left[-1 - \frac{3}{2}\left(\frac{z}{a}-\frac{1}{2}\right) + \frac{1}{2}\left(\frac{z}{a}-\frac{1}{2}\right)^3\right]\cdot\frac{2}{3}\pi a^3 & \text{if } \frac{1}{2}a < z < \frac{3}{2}a \\ -\frac{4}{3}\pi a^3 & \text{if } z > \frac{3}{2}a \quad . \end{cases} \tag{6.133}$$

After substituting $\xi = n + \frac{4}{3}\pi a^3 n^2 + \mathcal{O}(n^3)$ to relate ξ to the bulk density $n = n_\infty$, we obtain the desired result:

$$n(z) = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \\ n + \left[1 - \frac{3}{2}\left(\frac{z}{a}-\frac{1}{2}\right) + \frac{1}{2}\left(\frac{z}{a}-\frac{1}{2}\right)^3\right]\cdot\frac{2}{3}\pi a^3\, n^2 & \text{if } \frac{1}{2}a < z < \frac{3}{2}a \\ n & \text{if } z > \frac{3}{2}a \quad . \end{cases} \tag{6.134}$$

A sketch is provided in the right hand panel of fig. 6.14. Note that the density $n(z)$ vanishes identically for $z < \frac{1}{2}a$ due to the exclusion of the hard spheres by the wall. For z between $\frac{1}{2}a$ and $\frac{3}{2}a$, there is a density *enhancement*, the origin of which has a simple physical interpretation. Since the wall excludes particles from the region $z < \frac{1}{2}a$, there is an empty slab of thickness $\frac{1}{2}a$ coating the interior of the wall. There are then no particles in this region to exclude neighbors to their right, hence the density builds up just on the other side of this slab. The effect vanishes to the order of the calculation past $z = \frac{3}{2}a$, where $n(z) = n$ returns to its bulk value. Had we calculated to higher order, we'd have found damped oscillations with spatial period $\lambda \sim a$.

6.4 Liquid State Physics

6.4.1 The many-particle distribution function

The virial expansion is typically applied to low-density systems. When the density is high, *i.e.* when $na^3 \sim 1$, where a is a typical molecular or atomic length scale, the virial expansion is impractical. There are too many terms to compute, and to make progress one must use sophisticated resummation techniques to investigate the high density regime.

To elucidate the physics of liquids, it is useful to consider the properties of various *correlation functions*. These objects are derived from the general N-body Boltzmann distribution for identical particles,

$$\varrho_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N) = \frac{1}{Z(T,V,N)}\,\frac{1}{N!}\,e^{-\beta \hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} \quad , \tag{6.135}$$

here given in the ordinary canonical ensemble (OCE), where

$$Z(T,V,N) = \frac{1}{N!}\,\mathsf{Tr}\; e^{-\beta \hat{H}_N} = \frac{1}{N!}\int \prod_{j=1}^{N} \frac{d^d x_j\, d^d p_j}{h^d}\, e^{-\beta \hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} \tag{6.136}$$

We assume a Hamiltonian of the form

$$\hat{H}_N = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad . \tag{6.137}$$

The quantity

$$\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\,\frac{d^dx_1\,d^dp_1}{h^d}\cdots\frac{d^dx_N\,d^dp_N}{h^d} \tag{6.138}$$

is the probability of finding particle #1 lying within d^3x_1 of $\boldsymbol{x}_1$ and having momentum within d^dp_1 of $\boldsymbol{p}_1$, *etc.* Note $\mathsf{Tr}\;\varrho_N = 1$. Note that the definition of the trace (Tr) includes a factor $1/N!$ in order to account for particle indistinguishability, and that ϱ_N is normalized according to

$$\int\prod_{j=1}^N\frac{d^dx_j\,d^dp_j}{h^d}\;\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) = 1 \quad . \tag{6.139}$$

Thus, the canonical average of any function $\Phi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n)$ is given by

$$\langle\,\Phi\,\rangle_{T,V,N} = \int\prod_{j=1}^N\frac{d^dx_j\,d^dp_j}{h^d}\;\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\,\Phi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) \quad . \tag{6.140}$$

In the grand canonical ensemble (GCE), the probability that there are N particles present is

$$\mathsf{P}_N(T,V,\mu) = \Xi^{-1}\,\mathsf{Tr}\,\left(e^{-\beta\hat{K}}\,\delta_{\hat{N},N}\right) = \frac{Z(T,V,N)\,e^{N\mu/k_{\rm B}T}}{\Xi(T,V,\mu)} \quad , \tag{6.141}$$

with $\hat{K} = \hat{H} - \mu\hat{N}$ and

$$\Xi(T,V,\mu) = \mathsf{Tr}\;e^{-\beta\hat{K}} = \sum_{N=0}^\infty Z(T,V,N)\,e^{\beta\mu N} = \sum_{N=0}^\infty\frac{e^{\beta\mu N}}{N!}\int\prod_{j=1}^N\frac{d^dx_j\,d^dp_j}{h^d}\,e^{-\beta\hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} \quad . \tag{6.142}$$

The grand canonical average of Φ is given by

$$\langle\,\Phi\,\rangle_{T,V,\mu} = \sum_{N=0}^\infty\mathsf{P}_N(T,V,\mu)\,\langle\,\Phi\,\rangle_{T,V,N} \quad . \tag{6.143}$$

If we compute averages of quantities which only depend on the positions $\{\boldsymbol{x}_j\}$ and not on the momenta $\{\boldsymbol{p}_j\}$, then we may integrate out the momenta to obtain, in the OCE,

$$\begin{aligned}P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) &= \int\prod_{j=1}^N\frac{d^dp_j}{h^d}\;\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\\ &\equiv Q_N^{-1}\cdot\frac{1}{N!}\,e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad ,\end{aligned} \tag{6.144}$$

where W is the total potential energy,

$$W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \sum_i v(\boldsymbol{x}_i) + \sum_{i<j}u(\boldsymbol{x}_i-\boldsymbol{x}_j) + \sum_{i<j<k}w(\boldsymbol{x}_i-\boldsymbol{x}_j\,,\,\boldsymbol{x}_j-\boldsymbol{x}_k) + \ldots \quad , \tag{6.145}$$

and Q_N is the configuration integral,

$$Q_N(T,V) = \frac{1}{N!}\int d^d x_1 \cdots \int d^d x_N \, e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad . \tag{6.146}$$

We will, for the most part, consider only two-body central potentials as contributing to W, which is to say we will only retain the middle term on the RHS. Note that $P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$ is invariant under any permutation of the particle labels, and is normalized according to $\int \prod_{j=1}^N d^d x_j \, P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = 1$.

6.4.2 Averages over the distribution

To compute an average, one integrates over the distribution:

$$\big\langle F(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\big\rangle = \int d^d x_1 \cdots \int d^d x_N \, P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)\, F(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad . \tag{6.147}$$

The overall N-particle probability density is normalized, with $\int \prod_{j=1}^N d^d x_j \, P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = 1$.

The average local density is

$$n_1(\boldsymbol{r}) = \big\langle \sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i)\big\rangle = N\int d^d x_2 \cdots \int d^d x_N \, P(\boldsymbol{r},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N) \quad . \tag{6.148}$$

Note that the local density obeys the sum rule $\int d^d r \, n_1(\boldsymbol{r}) = N$. In a translationally invariant system, $n_1 = n = \frac{N}{V}$ is a constant independent of position. The boundaries of a system will in general break translational invariance, so in order to maintain the notion of a translationally invariant system of finite total volume, one must impose periodic boundary conditions.

The two-particle density matrix $n_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is defined by

$$\begin{aligned} n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) &= \big\langle \sum_{i\neq j} \delta(\boldsymbol{r}_1-\boldsymbol{x}_i)\,\delta(\boldsymbol{r}_2-\boldsymbol{x}_j)\big\rangle \\ &= N(N-1)\int d^d x_3 \cdots \int d^d x_N \, P(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{x}_3,\ldots,\boldsymbol{x}_N) \quad . \end{aligned} \tag{6.149}$$

As in the case of the one-particle density matrix, *i.e.* the local density $n_1(\boldsymbol{r})$, the two-particle density matrix also satisfies a sum rule:

$$\int d^d r_1 \int d^d r_2 \, n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) = N(N-1) \quad . \tag{6.150}$$

Generalizing further, one defines the *k-particle density matrix* as

$$\begin{aligned} n_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) &= \big\langle {\sum_{i_1\cdots i_k}}' \delta(\boldsymbol{r}_1-\boldsymbol{x}_{i_1})\cdots\delta(\boldsymbol{r}_k-\boldsymbol{x}_{i_k})\,\big\rangle \\ &= \frac{N!}{(N-k)!}\int d^d x_{k+1}\cdots\int d^d x_N \, P(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k,\boldsymbol{x}_{k+1},\ldots,\boldsymbol{x}_N) \quad , \end{aligned} \tag{6.151}$$

where the prime on the sum indicates that all the indices $i_1, \ldots, i_k$ are distinct. The corresponding sum rule is then

$$\int d^d r_1 \cdots \int d^d r_k \, n_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k) = \frac{N!}{(N-k)!} \quad . \tag{6.152}$$

The average potential energy can be expressed in terms of the distribution functions. Assuming only two-body interactions, we have

$$\begin{aligned} \langle W \rangle &= \Big\langle \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \Big\rangle \\ &= \tfrac{1}{2} \int d^d r_1 \int d^d r_2 \, u(\boldsymbol{r}_1 - \boldsymbol{r}_2) \, \Big\langle \sum_{i \neq j} \delta(\boldsymbol{r}_1 - \boldsymbol{x}_i) \, \delta(\boldsymbol{r}_2 - \boldsymbol{x}_j) \Big\rangle \\ &= \tfrac{1}{2} \int d^d r_1 \int d^d r_2 \, u(\boldsymbol{r}_1 - \boldsymbol{r}_2) \, n_2(\boldsymbol{r}_1, \boldsymbol{r}_2) \quad . \end{aligned} \tag{6.153}$$

As the separations $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ get large, we expect the correlations to vanish, in which case

$$\begin{aligned} n_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k) &= \Big\langle {\sum_{i_1 \cdots i_k}}' \delta(\boldsymbol{r}_1 - \boldsymbol{x}_{i_1}) \cdots \delta(\boldsymbol{r}_k - \boldsymbol{x}_{i_k}) \Big\rangle \xrightarrow[r_{ij} \to \infty]{} {\sum_{i_1 \cdots i_k}}' \big\langle \delta(\boldsymbol{r}_1 - \boldsymbol{x}_{i_1}) \big\rangle \cdots \big\langle \delta(\boldsymbol{r}_k - \boldsymbol{x}_{i_k}) \big\rangle \\ &= \frac{N!}{(N-k)!} \cdot \frac{1}{N^k} \, n_1(\boldsymbol{r}_1) \cdots n_1(\boldsymbol{r}_k) \\ &= \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{k-1}{N}\right) n_1(\boldsymbol{r}_1) \cdots n_1(\boldsymbol{r}_k) \quad . \end{aligned} \tag{6.154}$$

The *k-particle distribution function* is defined as the ratio

$$g_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k) \equiv \frac{n_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k)}{n_1(\boldsymbol{r}_1) \cdots n_1(\boldsymbol{r}_k)} \quad . \tag{6.155}$$

For large separations, then,

$$g_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k) \xrightarrow[r_{ij} \to \infty]{} \prod_{j=1}^{k-1} \left(1 - \frac{j}{N}\right) \quad . \tag{6.156}$$

For isotropic systems, the two-particle distribution function $g_2(\boldsymbol{r}_1, \boldsymbol{r}_2)$ depends only on the magnitude $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$. As a function of this scalar separation, the function is known as the *radial distribution function*:

$$g(r) \equiv g_2(\boldsymbol{r}) = \frac{1}{n^2} \, \Big\langle \sum_{i \neq j} \delta(\boldsymbol{r} - \boldsymbol{x}_i) \, \delta(\boldsymbol{x}_j) \Big\rangle = \frac{1}{Vn^2} \, \Big\langle \sum_{i \neq j} \delta(\boldsymbol{r} - \boldsymbol{x}_i + \boldsymbol{x}_j) \Big\rangle \quad . \tag{6.157}$$

The radial distribution function is of great importance in the physics of liquids because

- thermodynamic properties of the system can be related to $g(r)$
- $g(r)$ is directly measurable by scattering experiments

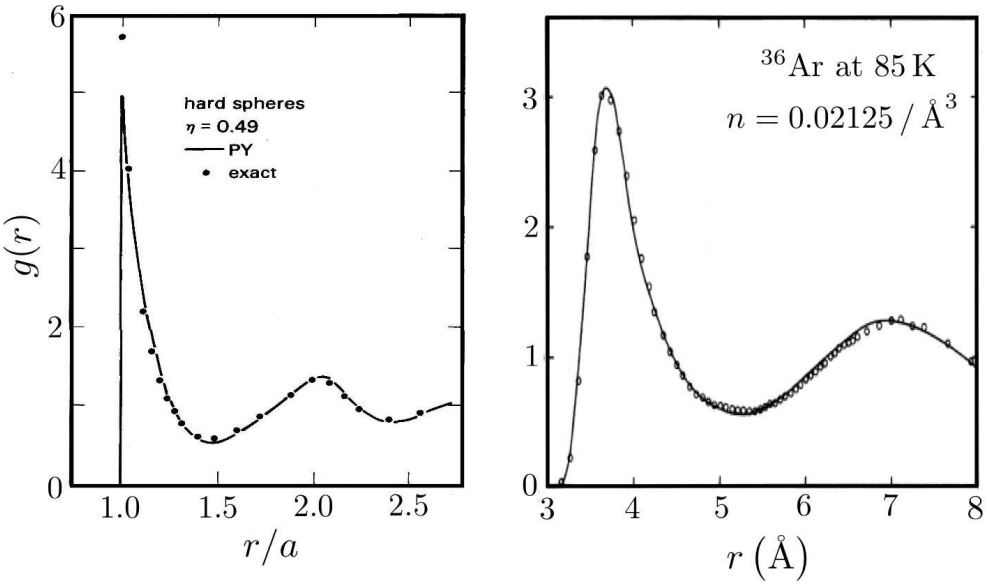


Figure 6.15: Pair distribution functions for hard spheres of diameter a at filling fraction $\varphi = \frac{\pi}{6}a^3 n = 0.49$ (left) and for liquid Argon at $T = 85\,\mathrm{K}$ (right). Molecular dynamics data for hard spheres (points) is compared with the result of the Percus-Yevick approximation (see below in §6.4.8). Reproduced (without permission) from J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, fig 5.5. Experimental data on liquid argon are from the neutron scattering work of J. L. Yarnell *et al.*, *Phys. Rev. A* **7**, 2130 (1973). The data (points) are compared with molecular dynamics calculations by Verlet (1967) for a Lennard-Jones fluid.

For example, in an isotropic system the average potential energy is given by

$$\begin{aligned}\langle W\rangle &= \tfrac{1}{2}\int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1-\boldsymbol{r}_2)\, n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) \\ &= \tfrac{1}{2}n^2 \int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1-\boldsymbol{r}_2)\, g\big(|\boldsymbol{r}_1-\boldsymbol{r}_2|\big) = \frac{N^2}{2V}\int d^d r\, u(r)\, g(r) \quad .\end{aligned} \tag{6.158}$$

For a three-dimensional system, the average internal (*i.e.* potential) energy per particle is

$$\frac{\langle W\rangle}{N} = 2\pi n \int_0^\infty dr\, r^2\, g(r)\, u(r) \quad . \tag{6.159}$$

Intuitively, $f(r)\,dr \equiv 4\pi r^2\, n\, g(r)\,dr$ is the average number of particles lying at a radial distance between r and $r+dr$ from a given reference particle. The total potential energy of interaction with the reference particle is then $f(r)\,u(r)\,dr$. Now integrate over all r and divide by two to avoid double-counting. This recovers eqn. 6.159.

In the OCE, $g(r)$ obeys the sum rule

$$\int d^d r\, g(r) = \frac{V}{N^2}\cdot N(N-1) = V - \frac{V}{N} \quad , \tag{6.160}$$

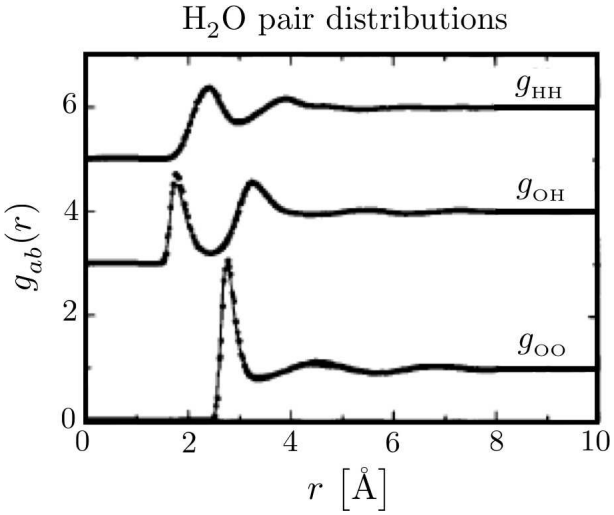


Figure 6.16: Monte Carlo pair distribution functions for liquid water. From A. K. Soper, *Chem Phys.* **202**, 295 (1996).

hence

$$n\int d^dr\,\big[g(r)-1\big] = -1 \qquad \text{(OCE)} \quad . \tag{6.161}$$

The function $h(r) \equiv g(r) - 1$ is called the *pair correlation function*.

In the grand canonical formulation, we have

$$\begin{aligned} n\int d^3r\,h(r) &= \frac{\langle N\rangle}{V}\cdot\left[\frac{\big\langle N(N-1)\big\rangle}{\langle N\rangle^2}\,V - V\right] \\ &= \frac{\big\langle N^2\big\rangle - \big\langle N\big\rangle^2}{\big\langle N\big\rangle} - 1 = nk_{\rm B}T\kappa_T - 1 \qquad \text{(GCE)} \quad , \end{aligned} \tag{6.162}$$

where κ_T is the isothermal compressibility. Note that in an ideal gas we have $h(r) = 0$ and $\kappa_T = \kappa_T^0 \equiv 1/nk_{\rm B}T$. Self-condensed systems, such as liquids and solids far from criticality, are nearly incompressible, hence $0 < nk_{\rm B}T\,\kappa_T \ll 1$, and therefore $n\int d^3r\,h(r) \approx -1$. For incompressible systems, where $\kappa_T = 0$, this becomes an equality.

As we shall see below in §6.4.4, the function $h(r)$, or rather its Fourier transform $\hat{h}(\boldsymbol{k})$, is directly measured in a scattering experiment. The question then arises as to which result applies: the OCE result from eqn. 6.161 or the GCE result from eqn. 6.162. The answer is that under almost all experimental conditions it is the GCE result which applies. The reason for this is that the scattering experiment typically illuminates only a subset of the entire system. This subsystem is in particle equilibrium with the remainder of the system, hence it is appropriate to use the grand canonical ensemble. The OCE results would only apply if the scattering experiment were to measure the entire system.

6.4.3 Virial equation of state

The *virial* of a mechanical system is defined to be

$$G = \sum_i \boldsymbol{x}_i \cdot \boldsymbol{F}_i \quad , \tag{6.163}$$

where $\boldsymbol{F}_i$ is the total force acting on particle i. If we average G over time, we obtain

$$\begin{aligned} \langle G \rangle &= \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \, \sum_i \boldsymbol{x}_i \cdot \boldsymbol{F}_i \\ &= - \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \, \sum_i m \, \dot{\boldsymbol{x}}_i^2 = -3Nk_{\rm B}T \quad . \end{aligned} \tag{6.164}$$

Here, we have made use of

$$\boldsymbol{x}_i \cdot \boldsymbol{F}_i = m \, \boldsymbol{x}_i \cdot \ddot{\boldsymbol{x}}_i = - m \, \dot{\boldsymbol{x}}_i^2 + \frac{d}{dt}\Big(m \, \boldsymbol{x}_i \cdot \dot{\boldsymbol{x}}_i \Big) \quad , \tag{6.165}$$

as well as ergodicity and equipartition of kinetic energy. We have also assumed three space dimensions. In a bounded system, there are two contributions to the force $\boldsymbol{F}_i$. One contribution is from the surfaces which enclose the system. This is given by[18]

$$\langle G \rangle_{\rm surfaces} = \Big\langle \sum_i \boldsymbol{x}_i \cdot \boldsymbol{F}_i^{(\rm surf)} \Big\rangle = -3pV \quad . \tag{6.166}$$

The remaining contribution is due to the interparticle forces. Thus,

$$\frac{p}{k_{\rm B}T} = \frac{N}{V} - \frac{1}{3Vk_{\rm B}T} \Big\langle \sum_i \boldsymbol{x}_i \cdot \boldsymbol{\nabla}_i W \Big\rangle \quad . \tag{6.167}$$

Invoking the definition of $g(r)$, we have

$$p = nk_{\rm B}T \left\{ 1 - \frac{2\pi n}{3k_{\rm B}T} \int_0^\infty dr \, r^3 \, g(r) \, u'(r) \right\} \quad . \tag{6.168}$$

As an alternate derivation, consider the First Law of Thermodynamics,

$$d\Omega = -S \, dT - p \, dV - N \, d\mu \quad , \tag{6.169}$$

[18] To derive this expression, note that $\boldsymbol{F}^{(\rm surf)}$ is directed inward and vanishes away from the surface. Each Cartesian direction $\alpha = (x, y, z)$ then contributes $-F_\alpha^{(\rm surf)} L_\alpha$, where L_α is the corresponding linear dimension. But $F_\alpha^{(\rm surf)} = pA_\alpha$, where A_α is the area of the corresponding face and p is the pressure. Summing over the three possibilities for α, one obtains eqn. 6.166.

from which we derive

$$p = -\left(\frac{\partial\Omega}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad . \tag{6.170}$$

Now let $V \to \ell^3 V$, where ℓ is a scale parameter. Then

$$p = -\frac{\partial\Omega}{\partial V} = -\frac{1}{3V}\,\frac{\partial}{\partial\ell}\bigg|_{\ell=1} \Omega(T, \ell^3 V, \mu) \quad . \tag{6.171}$$

Now

$$\begin{aligned}\Xi(T,\ell^3V,\mu) &= \sum_{N=0}^{\infty} \frac{1}{N!}\, e^{\beta\mu N}\, \lambda_T^{-3N} \int\limits_{\ell^3 V} d^3x_1 \cdots \int\limits_{\ell^3 V} d^3x_N\, e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \\ &= \sum_{N=0}^{\infty} \frac{1}{N!}\, \left(e^{\beta\mu}\,\lambda_T^{-3}\right)^N \ell^{3N} \int\limits_{V} d^3x_1 \cdots \int\limits_{V} d^3x_N\, e^{-\beta W(\ell\boldsymbol{x}_1,\ldots,\ell\boldsymbol{x}_N)}\end{aligned} \tag{6.172}$$

Thus,

$$\begin{aligned} p &= -\frac{1}{3V}\,\frac{\partial\Omega(\ell^3V)}{\partial\ell}\bigg|_{\ell=1} = \frac{k_{\rm B}T}{3V}\,\frac{1}{\Xi}\,\frac{\partial\Xi(\ell^3V)}{\partial\ell}\bigg|_{\ell=1} \\ &= \frac{k_{\rm B}T}{3V}\,\frac{1}{\Xi}\sum_{N=0}^{\infty}\frac{1}{N!}\big(z\lambda_T^{-3}\big)^N \left\{\int\limits_V d^3x_1\cdots\int\limits_V d^3x_N\, e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)}\left[3N - \beta\sum_i \boldsymbol{x}_i\cdot\frac{\partial W}{\partial\boldsymbol{x}_i}\right]\right\} \\ &= nk_{\rm B}T - \frac{1}{3V}\left\langle\frac{\partial W}{\partial\ell}\right\rangle_{\ell=1} \quad . \end{aligned} \tag{6.173}$$

Finally, from $W = \sum_{i<j} u(\ell\boldsymbol{x}_{ij})$ we have

$$\left\langle\frac{\partial W}{\partial\ell}\right\rangle_{\ell=1} = \left\langle\sum_{i<j}\boldsymbol{x}_{ij}\cdot\boldsymbol{\nabla}u(\boldsymbol{x}_{ij})\right\rangle = \frac{2\pi N^2}{V}\int\limits_0^\infty dr\, r^3 g(r)\, u'(r) \quad , \tag{6.174}$$

and hence

$$p = nk_{\rm B}T - \tfrac{2}{3}\pi n^2 \int\limits_0^\infty dr\, r^3\, g(r)\, u'(r) \quad . \tag{6.175}$$

Note that the density n enters the equation of state explicitly on the RHS of the above equation, but also implicitly through the pair distribution function $g(r)$, which has implicit dependence on both n and T.

6.4.4 Correlations and scattering

Consider the scattering of a light or particle beam (*i.e.* photons or neutrons) from a liquid. We label the states of the beam particles by their wavevector $\boldsymbol{k}$ and we assume a general dispersion $\varepsilon_{\boldsymbol{k}}$. For photons,

$\varepsilon_{\boldsymbol{k}} = \hbar c|\boldsymbol{k}|$, while for neutrons $\varepsilon_{\boldsymbol{k}} = \hbar^2\boldsymbol{k}^2/2m_{\rm n}$. We assume a single scattering process with the liquid, during which the total momentum and energy of the liquid plus beam are conserved. We write

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q} \qquad , \qquad \varepsilon_{\boldsymbol{k}'} = \varepsilon_{\boldsymbol{k}} + \hbar\omega \quad , \tag{6.176}$$

where $\boldsymbol{k}'$ is the final state of the scattered beam particle. Thus, the fluid transfers momentum $\Delta\boldsymbol{p} = \hbar\boldsymbol{q}$ and energy $\hbar\omega$ to the beam.

Now consider the scattering process between an initial state $|\, i, \boldsymbol{k}\,\rangle$ and a final state $|\, j, \boldsymbol{k}'\,\rangle$, where these states describe both the beam and the liquid. According to Fermi's Golden Rule, the scattering rate is

$$\Gamma_{i\boldsymbol{k}\to j\boldsymbol{k}'} = \frac{2\pi}{\hbar}\left|\langle\, j, \boldsymbol{k}' \,|\, \mathcal{V} \,|\, i, \boldsymbol{k}\,\rangle\right|^2 \delta(E_j - E_i + \hbar\omega) \quad , \tag{6.177}$$

where $\mathcal{V}$ is the scattering potential and E_i is the initial internal energy of the liquid. Note that overall energy conservation requires $E_j + \varepsilon_{\boldsymbol{k}'} = E_i + \varepsilon_{\boldsymbol{k}}$ and therefore $E_j = E_i - \hbar\omega$. If $\boldsymbol{r}$ is the position of the beam particle and $\{\boldsymbol{x}_l\}$ are the positions of the liquid particles, then

$$\mathcal{V}(\boldsymbol{r}) = \sum_{l=1}^{N} v(\boldsymbol{r} - \boldsymbol{x}_l) \quad . \tag{6.178}$$

For elastic scattering, the *differential scattering cross section* $d\sigma$ is defined to be

$$d\sigma(\hat{\Omega}) = \frac{\text{rate at which particles scattered into solid angle } d\Omega}{\text{incident flux}} \quad . \tag{6.179}$$

For inelastic scattering,

$$d^2\sigma(\hat{\Omega}, \omega) = \frac{\text{rate at which particles scattered into solid angle } d\Omega \text{ and energy change within } \hbar\, d\omega}{\text{incident flux}} \quad . \tag{6.180}$$

The inelastic differential scattering cross section (per unit frequency per unit solid angle) is

$$\frac{\partial^2\sigma}{\partial\Omega\,\partial\omega} = \frac{\hbar}{4\pi}\,\frac{g(\varepsilon_{\boldsymbol{k}'})}{|\boldsymbol{v}_{\boldsymbol{k}}|}\sum_{i,j} P_i\,\Gamma_{i\boldsymbol{k}\to j\boldsymbol{k}'} \quad , \tag{6.181}$$

where $g(\varepsilon) = \int \frac{d^dk}{(2\pi)^d}\,\delta(\varepsilon - \varepsilon_{\boldsymbol{k}})$ is the density of states for the beam particles, and $P_i = Z^{-1}\,e^{-\beta E_i}$ is the Boltzmann weight.

Consider now the matrix element

$$\begin{aligned}\left\langle\, j, \boldsymbol{k}' \,\middle|\, \mathcal{V} \,\middle|\, i, \boldsymbol{k}\,\right\rangle &= \left\langle\, j \,\middle|\, \frac{1}{V}\sum_{l=1}^{N}\int d^dr\; e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\, v(\boldsymbol{r} - \boldsymbol{x}_l) \,\middle|\, i\,\right\rangle \\ &= \frac{1}{V}\,\hat{v}(\boldsymbol{q})\left\langle\, j \,\middle|\, \sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l} \,\middle|\, i\,\right\rangle \quad ,\end{aligned} \tag{6.182}$$

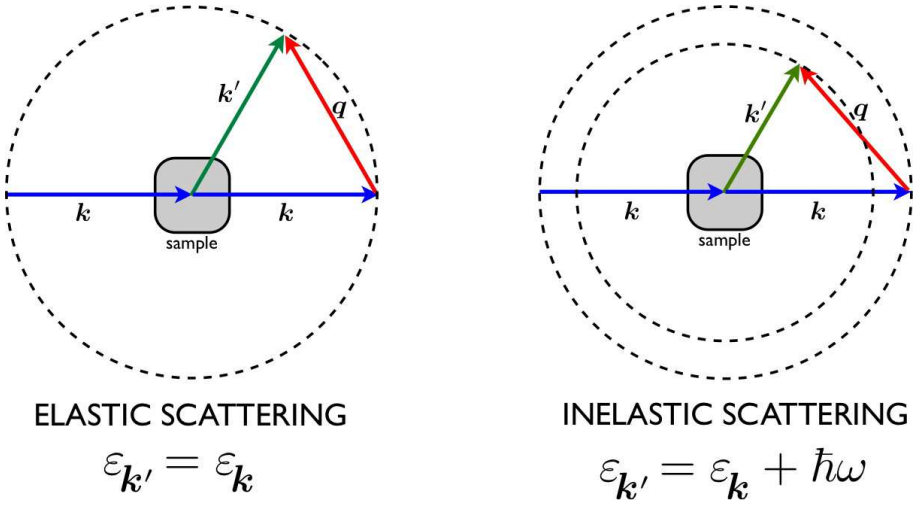


Figure 6.17: In a scattering experiment, a beam of particles interacts with a sample and the beam particles scatter off the sample particles. A momentum $\hbar\boldsymbol{q}$ and energy $\hbar\omega$ are transferred to the beam particle during such a collision. If $\omega = 0$, the scattering is said to be *elastic*. For $\omega \neq 0$, the scattering is *inelastic*.

where we have assumed that the incident and scattered beams are plane waves. We then have

$$\begin{aligned}\frac{\partial^2\sigma}{\partial\Omega\,\partial\omega} &= \frac{\hbar}{2}\,\frac{g(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}})}{|\boldsymbol{\nabla}_{\boldsymbol{k}}\varepsilon_{\boldsymbol{k}}|}\,\frac{|\hat{v}(\boldsymbol{q})|^2}{V^2}\sum_i P_i\sum_j \left|\left\langle\, j\,\right|\sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l}\left|\,i\,\right\rangle\right|^2\delta(E_j - E_i + \hbar\omega) \\ &= \frac{g(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}})}{4\pi\,|\boldsymbol{\nabla}_{\boldsymbol{k}}\varepsilon_{\boldsymbol{k}}|}\,\frac{N}{V^2}\,|\hat{v}(\boldsymbol{q})|^2\,S(\boldsymbol{q},\omega) \quad ,\end{aligned} \tag{6.183}$$

where $S(\boldsymbol{q},\omega)$ is the *dynamic structure factor*,

$$S(\boldsymbol{q},\omega) = \frac{2\pi\hbar}{N}\sum_i P_i\sum_j \left|\left\langle\, j\,\right|\sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l}\left|\,i\,\right\rangle\right|^2\delta(E_j - E_i + \hbar\omega) \tag{6.184}$$

Note that for an arbitrary operator A,

$$\begin{aligned}\sum_j \left|\left\langle\, j\,\right|A\left|\,i\,\right\rangle\right|^2\delta(E_j - E_i + \hbar\omega) &= \frac{1}{2\pi\hbar}\sum_j\int_{-\infty}^{\infty} dt\; e^{i(E_j - E_i + \hbar\omega)\,t/\hbar}\left\langle\, i\,\right|A^\dagger\left|\,j\,\right\rangle\left\langle\, j\,\right|A\left|\,i\,\right\rangle \\ &= \frac{1}{2\pi\hbar}\sum_j\int_{-\infty}^{\infty} dt\; e^{i\omega t}\left\langle\, i\,\right|A^\dagger\left|\,j\,\right\rangle\left\langle\, j\,\right|e^{i\hat{H}t/\hbar}\,A\,e^{-i\hat{H}t/\hbar}\left|\,i\,\right\rangle \\ &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dt\; e^{i\omega t}\left\langle\, i\,\right|A^\dagger(0)\,A(t)\left|\,i\,\right\rangle \quad .\end{aligned} \tag{6.185}$$

Thus,

$$\begin{aligned} S(\boldsymbol{q},\omega) &= \frac{1}{N}\int\limits_{-\infty}^{\infty} dt\, e^{i\omega t}\, \sum_i P_i \left\langle\, i \,\right|\, \sum_{l,l'} e^{i\boldsymbol{q}\cdot\boldsymbol{x}_l(0)}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_{l'}(t)} \,\left|\, i \,\right\rangle \\ &= \frac{1}{N}\int\limits_{-\infty}^{\infty} dt\, e^{i\omega t}\, \Big\langle \sum_{l,l'} e^{i\boldsymbol{q}\cdot\boldsymbol{x}_l(0)}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_{l'}(t)} \Big\rangle \quad , \end{aligned} \tag{6.186}$$

where the angular brackets in the last line denote a thermal expectation value of a quantum mechanical operator. If we integrate over all frequencies, we obtain the equal time correlator,

$$\begin{aligned} S(\boldsymbol{q}) &= \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, S(\boldsymbol{q},\omega) = \frac{1}{N}\sum_{l,l'} \left\langle e^{i\boldsymbol{q}\cdot(\boldsymbol{x}_l - \boldsymbol{x}_{l'})} \right\rangle \\ &= N\,\delta_{\boldsymbol{q},0} + 1 + n\int d^dr\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\, \big[g(r) - 1\big] \quad . \end{aligned} \tag{6.187}$$

known as the *static structure factor*[19]. Note that $S(\boldsymbol{q}=0) = N$, since all the phases $e^{i\boldsymbol{q}\cdot(\boldsymbol{x}_i - \boldsymbol{x}_j)}$ are then unity. As $q \to \infty$, the phases oscillate rapidly with changes in the distances $|\boldsymbol{x}_i - \boldsymbol{x}_j|$, and average out to zero. However, the 'diagonal' terms in the sum, *i.e.* those with $i = j$, always contribute a total of 1 to $S(\boldsymbol{q})$. Therefore in the $q \to \infty$ limit we have $S(q \to \infty) = 1$.

In general, the detectors used in a scattering experiment are sensitive to the energy of the scattered beam particles, although there is always a finite experimental resolution, both in $\boldsymbol{q}$ and ω. This means that what is measured is actually something like

$$S_{\rm meas}(\boldsymbol{q},\omega) = \int d^dq' \int d\omega'\, F(\boldsymbol{q}-\boldsymbol{q}')\, G(\omega-\omega')\, S(\boldsymbol{q}',\omega') \quad , \tag{6.188}$$

where F and G are essentially Gaussian functions of their argument, with width given by the experimental resolution. If one integrates over all frequencies ω, *i.e.* if one simply counts scattered particles as a function of $\boldsymbol{q}$ but without any discrimination of their energies, then one measures the static structure factor $S(\boldsymbol{q})$. Elastic scattering is determined by $S(\boldsymbol{q},\omega=0)$, *i.e.* at no energy transfer.

6.4.5 Correlation and response

Suppose an external potential $v(\boldsymbol{x})$ is also present. Then

$$P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \frac{1}{Q_N[v]}\cdot\frac{1}{N!}\, e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)}\, e^{-\beta\sum_i v(\boldsymbol{x}_i)} \quad , \tag{6.189}$$

where

$$Q_N[v] = \frac{1}{N!}\int d^dx_1 \cdots \int d^dx_N\, e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)}\, e^{-\beta\sum_i v(\boldsymbol{x}_i)} \quad . \tag{6.190}$$

[19] We may write $\delta_{\boldsymbol{q},0} = \frac{1}{V}\,(2\pi)^d\,\delta(\boldsymbol{q})$.

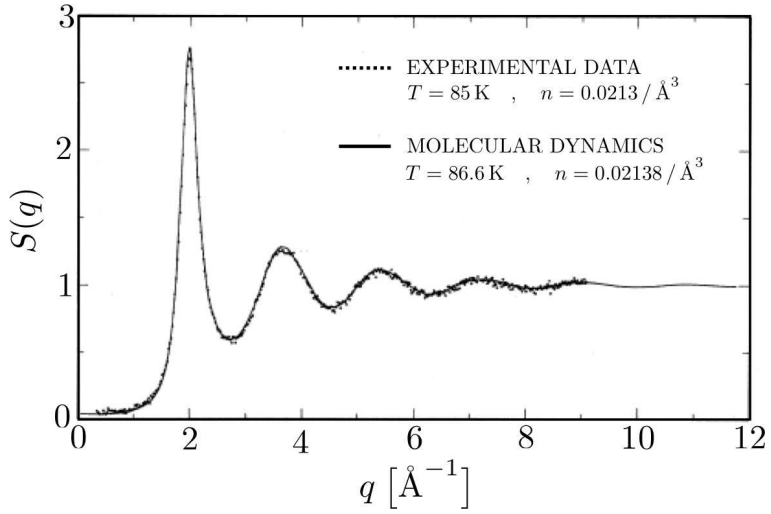


Figure 6.18: Comparison of the static structure factor of Argon as determined by neutron scattering work of J. L. Yarnell *et al., Phys. Rev. A* **7**, 2130 (1973) with molecular dynamics calculations by Verlet (1967) for a Lennard-Jones fluid.

The Helmholtz free energy is then

$$F = -\frac{1}{\beta}\,\log\left(\lambda_T^{-dN}\,Q_N[v]\right) \quad . \tag{6.191}$$

Now consider the functional derivative

$$\frac{\delta F}{\delta v(\boldsymbol{r})} = -\frac{1}{\beta}\cdot\frac{1}{Q_N}\cdot\frac{\delta Q_N}{\delta v(\boldsymbol{r})} \quad . \tag{6.192}$$

Using

$$\sum_i v(\boldsymbol{x}_i) = \int d^d r\, v(\boldsymbol{r})\,\sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i) \quad , \tag{6.193}$$

hence

$$\frac{\delta F}{\delta v(\boldsymbol{r})} = \int d^d x_1 \cdots \int d^d x_N\, P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)\,\sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i) = n_1(\boldsymbol{r}) \quad , \tag{6.194}$$

which is the local density at $\boldsymbol{r}$.

Next, consider the *response function*,

$$\begin{aligned}\chi(\boldsymbol{r},\boldsymbol{r}') &\equiv \frac{\delta n_1(\boldsymbol{r})}{\delta v(\boldsymbol{r}')} = \frac{\delta^2 F[v]}{\delta v(\boldsymbol{r})\,\delta v(\boldsymbol{r}')} \\ &= \frac{1}{\beta}\cdot\frac{1}{Q_N^2}\,\frac{\delta Q_N}{\delta v(\boldsymbol{r})}\,\frac{\delta Q_N}{\delta v(\boldsymbol{r}')} - \frac{1}{\beta}\cdot\frac{1}{Q_N}\,\frac{\delta^2 Q_N}{\delta v(\boldsymbol{r})\,\delta v(\boldsymbol{r}')} \\ &= \beta\, n_1(\boldsymbol{r})\, n_1(\boldsymbol{r}') - \beta\, n_1(\boldsymbol{r})\,\delta(\boldsymbol{r}-\boldsymbol{r}') - \beta\, n_2(\boldsymbol{r},\boldsymbol{r}') \quad .\end{aligned} \tag{6.195}$$

In an isotropic system, $\chi(\boldsymbol{r},\boldsymbol{r}') = \chi(\boldsymbol{r}-\boldsymbol{r}')$ is a function of the coordinate separation, and

$$\begin{aligned} -k_{\rm B}T\,\chi(\boldsymbol{r}-\boldsymbol{r}') &= -n^2 + n\,\delta(\boldsymbol{r}-\boldsymbol{r}') + n^2 g\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) \\ &= n^2\,h\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) + n\,\delta(\boldsymbol{r}-\boldsymbol{r}') \quad . \end{aligned} \tag{6.196}$$

Taking the Fourier transform,

$$-k_{\rm B}T\,\hat\chi(\boldsymbol{q}) = n + n^2\,\hat h(\boldsymbol{q}) = n\,S(\boldsymbol{q}) \quad . \tag{6.197}$$

We may also write

$$-\frac{k_{\rm B}T}{n}\,\hat\chi(0) = 1 + n\,\hat h(0) = \frac{\kappa_T}{\kappa_T^0} \quad , \tag{6.198}$$

i.e. $n^2\kappa_T = -\hat\chi(0)$, since $\kappa_T^0 = 1/nk_{\rm B}T$.

What does this all mean? Suppose we have an isotropic system which is subjected to a weak, spatially inhomogeneous potential $v(\boldsymbol{r})$. We expect that the density $n(\boldsymbol{r})$ in the presence of the inhomogeneous potential itself to be inhomogeneous. The first corrections to the $v=0$ value $n=n_0$ are linear in v, and given by

$$\begin{aligned} \delta n(\boldsymbol{r}) &= \int d^d r'\,\chi(\boldsymbol{r},\boldsymbol{r}')\,v(\boldsymbol{r}') \\ &= -\beta n_0\,v(\boldsymbol{r}) - \beta n_0^2 \int d^d r'\,h(\boldsymbol{r}-\boldsymbol{r})\,v(\boldsymbol{r}') \quad . \end{aligned} \tag{6.199}$$

Note that if $v(\boldsymbol{r}) > 0$ it becomes energetically more costly for a particle to be at $\boldsymbol{r}$. Accordingly, the density response is negative, and proportional to the ratio $v(\boldsymbol{r})/k_{\rm B}T$ – this is the first term in the above equation. If there were no correlations between the particles, then $h=0$ and this would be the entire story. However, the particles in general *are* correlated. Consider, for example, the case of hard spheres of diameter a, and let there be a repulsive potential at $\boldsymbol{r}=0$. This means that it is less likely for a particle to be centered anywhere within a distance a of the origin. But then it will be *more* likely to find a particle in the next 'shell' of radial thickness a.

6.4.6 BBGKY hierarchy

The distribution functions satisfy a hierarchy of integro-differential equations known as the *BBGKY hierarchy*[20]. In homogeneous systems, we have

$$g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = \frac{N!}{(N-k)!}\,\frac{1}{n^k}\int d^d x_{k+1}\cdots\int d^d x_N\,P(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{x}_{k+1}\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad , \tag{6.200}$$

where

$$P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \frac{1}{Q_N}\cdot\frac{1}{N!}\,e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)} \quad . \tag{6.201}$$

[20] So named after Bogoliubov, Born, Green, Kirkwood, and Yvon.

Taking the gradient with respect to $\boldsymbol{r}_1$, we have

$$\frac{\partial}{\partial \boldsymbol{r}_1} g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = \frac{1}{Q_N}\cdot\frac{n^{-k}}{(N-k)!}\int d^dx_{k+1}\cdots\int d^dx_N\, e^{-\beta\sum_{k<i<j}u(\boldsymbol{x}_{ij})} \times \frac{\partial}{\partial \boldsymbol{r}_1}\left[\, e^{-\beta\sum_{i<j\le k}u(\boldsymbol{r}_{ij})}\cdot e^{-\beta\sum_{i\le k<j}u(\boldsymbol{r}_i-\boldsymbol{x}_j)}\right] \quad , \tag{6.202}$$

where $\sum_{k<i<j}$ means to sum on indices i and j such that $i<j$ and $k<i$, *i.e.*

$$\begin{aligned}\sum_{k<i<j} u(\boldsymbol{x}_{ij}) &\equiv \sum_{i=k+1}^{N-1}\sum_{j=i+1}^{N} u(\boldsymbol{x}_i-\boldsymbol{x}_j)\\ \sum_{i<j\le k} u(\boldsymbol{r}_{ij}) &\equiv \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} u(\boldsymbol{r}_i-\boldsymbol{r}_j)\\ \sum_{i\le k<j} u(\boldsymbol{r}_i-\boldsymbol{x}_j) &= \sum_{i=1}^{k}\sum_{j=k+1}^{N} u(\boldsymbol{r}_i-\boldsymbol{x}_j) \quad .\end{aligned} \tag{6.203}$$

Now

$$\frac{\partial}{\partial \boldsymbol{r}_1}\left[e^{-\beta\sum_{i<j\le k}u(r_{ij})}\cdot e^{-\beta\sum_{i\le k<j}u(\boldsymbol{r}_i-\boldsymbol{x}_j)}\right] = \tag{6.204}$$
$$-\,\beta\left\{\sum_{j=1}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1} + \sum_{j=k+1}^{N}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_j)}{\partial \boldsymbol{r}_1}\right\}\cdot\left[\, e^{-\beta\sum_{i<j\le k}u(r_{ij})}\cdot e^{-\beta\sum_{i\le k<j}u(\boldsymbol{r}_i-\boldsymbol{x}_j)}\right] \quad ,$$

hence

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{r}_1} g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) &= -\beta\sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\, g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k)\\ &\qquad -\beta(N-k)\int d^dx_{k+1}\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_{k+1})}{\partial \boldsymbol{r}_1}\, P(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{x}_{k+1}\,,\,\ldots\,,\,\boldsymbol{x}_N)\\ &= -\beta\sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\, g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k)\\ &\qquad -\beta n\int d^dx_{k+1}\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_{k+1})}{\partial \boldsymbol{r}_1}\, g_{k+1}(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,x_{k+1}) \quad .\end{aligned} \tag{6.205}$$

Thus, we obtain the BBGKY hierarchy:

$$\begin{aligned}-k_{\rm B}T\,\frac{\partial}{\partial \boldsymbol{r}_1} g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) &= \sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\, g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k)\\ &\qquad + n\int d^dr'\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}')}{\partial \boldsymbol{r}_1}\, g_{k+1}(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{r}') \quad .\end{aligned} \tag{6.206}$$

The BBGKY hierarchy is an infinite tower of coupled integro-differential equations, relating g_k to g_{k+1} for all k. If we approximate g_k at some level k in terms of equal or lower order distributions, then we obtain a closed set of equations which in principle can be solved, at least numerically. For example, the *Kirkwood approximation* closes the hierarchy at order $k = 2$ by imposing the condition

$$g_3(\boldsymbol{r}_1\,,\,\boldsymbol{r}_2\,,\,\boldsymbol{r}_3) \equiv g(\boldsymbol{r}_1 - \boldsymbol{r}_2)\,g(\boldsymbol{r}_1 - \boldsymbol{r}_3)\,g(\boldsymbol{r}_2 - \boldsymbol{r}_3) \quad . \tag{6.207}$$

This results in the single integro-differential equation

$$-k_{\rm B}T\,\boldsymbol{\nabla} g(\boldsymbol{r}) = g(\boldsymbol{r})\,\boldsymbol{\nabla} u + n\int d^dr'\,g(\boldsymbol{r})\,g(\boldsymbol{r}')\,g(\boldsymbol{r}-\boldsymbol{r}')\,\boldsymbol{\nabla} u(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{6.208}$$

This is known as the Born-Green-Yvon (BGY) equation. In practice, the BGY equation, which is solved numerically, gives adequate results only at low densities.

6.4.7 Ornstein-Zernike theory

The *direct correlation function* $c(\boldsymbol{r})$ is defined by the equation

$$h(\boldsymbol{r}) = c(\boldsymbol{r}) + n\int d^3r'\,h(\boldsymbol{r}-\boldsymbol{r}')\,c(\boldsymbol{r}') \quad , \tag{6.209}$$

where $h(r) = g(r) - 1$ and we assume an isotropic system. This is called the *Ornstein-Zernike equation*. The first term, $c(\boldsymbol{r})$, accounts for local correlations, which are then propagated in the second term to account for long-ranged correlations.

The OZ equation is an integral equation, but it becomes a simple algebraic one upon Fourier transforming:

$$\hat{h}(\boldsymbol{q}) = \hat{c}(\boldsymbol{q}) + n\,\hat{h}(\boldsymbol{q})\,\hat{c}(\boldsymbol{q}) \quad , \tag{6.210}$$

the solution of which is

$$\hat{h}(\boldsymbol{q}) = \frac{\hat{c}(\boldsymbol{q})}{1 - n\,\hat{c}(\boldsymbol{q})} \quad . \tag{6.211}$$

The static structure factor is then

$$S(\boldsymbol{q}) = 1 + n\,\hat{h}(\boldsymbol{q}) = \frac{1}{1 - n\,\hat{c}(\boldsymbol{q})} \quad . \tag{6.212}$$

In the grand canonical ensemble, we can write

$$\kappa_T = \frac{1 + n\,\hat{h}(0)}{nk_{\rm B}T} = \frac{1}{nk_{\rm B}T}\cdot\frac{1}{1 - n\,\hat{c}(0)} \qquad \Longrightarrow \qquad n\,\hat{c}(0) = 1 - \frac{\kappa_T^0}{\kappa_T} \quad , \tag{6.213}$$

where $\kappa_T^0 = 1/nk_{\rm B}T$ is the ideal gas isothermal compressibility.

Ornstein-Zernike approximation at long wavelengths

Let's expand the direct correlation function $\hat{c}(\boldsymbol{q})$ in powers of the wavevector $\boldsymbol{q}$, *viz.*

$$\hat{c}(\boldsymbol{q}) = \hat{c}(0) + c_2\, q^2 + c_4\, q^4 + \dots \quad . \tag{6.214}$$

Here we have assumed spatial isotropy. Then

$$\begin{aligned} 1 - n\,\hat{c}(\boldsymbol{q}) = \frac{1}{S(\boldsymbol{q})} &= 1 - n\,\hat{c}(0) - n\,c_2\,q^2 + \dots \\ &\equiv \xi^{-2} R^2 + q^2 R^2 + \mathcal{O}(q^4) \quad , \end{aligned} \tag{6.215}$$

where

$$R^2 = -n\,c_2 = \frac{2\pi n}{3} \int\limits_0^\infty dr\; r^4\, c(r) \tag{6.216}$$

and

$$\xi^{-2} = \frac{1 - n\,\hat{c}(0)}{R^2} = \frac{3}{2\pi n} \cdot \frac{1 - 4\pi n \int_0^\infty dr\, r^2\, c(r)}{\int_0^\infty dr\, r^4\, c(r)} \quad . \tag{6.217}$$

The quantity $R(T)$ tells us something about the effective range of the interactions, while $\xi(T)$ is the *correlation length*. As we approach a critical point, the correlation length diverges as a power law:

$$\xi(T) \sim A\,|T - T_{\rm c}|^{-\nu} \quad . \tag{6.218}$$

From eqns. 6.199 and 6.212, the susceptibility is given by

$$\hat{\chi}(\boldsymbol{q}) = -n\beta\, S(\boldsymbol{q}) = -\frac{n\beta R^{-2}}{\xi^{-2} + q^2 + \mathcal{O}(q^4)} \tag{6.219}$$

Note that $\hat{\chi}(\boldsymbol{q}) < 0$ because the susceptibility is the functional derivative of the density with respect to the added potential, *i.e.* $\chi(\boldsymbol{r}, \boldsymbol{r}') = \delta n(\boldsymbol{r})/\delta v(\boldsymbol{r}')$. A local positive potential is repulsive, causing the density to decrease in its vicinity. In the *Ornstein-Zernike approximation*, one drops the $\mathcal{O}(q^4)$ terms in the denominator and retains only the long-wavelength behavior in the direct correlation function. Thus,

$$\hat{\chi}^{\rm OZ}(\boldsymbol{q}) = -\frac{n\beta R^{-2}}{\xi^{-2} + q^2} \quad . \tag{6.220}$$

We now apply the inverse Fourier transform back to real space to obtain $\chi^{\rm OZ}(\boldsymbol{r})$. In $d = 1$ dimension the result can be obtained exactly:

$$\chi^{\rm OZ}_{d=1}(x) = -\frac{n}{k_{\rm B}TR^2} \int\limits_{-\infty}^{\infty} \frac{dq}{2\pi}\, \frac{e^{iqx}}{\xi^{-2} + q^2} = -\frac{n\xi}{2k_{\rm B}TR^2}\, e^{-|x|/\xi} \quad . \tag{6.221}$$

In $d = 2$ dimensions,

$$\begin{aligned} \chi^{\rm OZ}_{d=2}(\boldsymbol{x}) &= -\frac{n}{k_{\rm B}TR^2} \int \frac{d^2q}{(2\pi)^2}\, \frac{e^{i\boldsymbol{q}\cdot\boldsymbol{x}}}{\xi^{-2} + \boldsymbol{q}^2} \\ &= -\frac{n}{2\pi k_{\rm B}TR^2} \int\limits_0^\infty dq\; \frac{q\, J_0(q|\boldsymbol{x}|)}{\xi^{-2} + q^2} = -\frac{n}{2\pi k_{\rm B}TR^2}\, K_0(|\boldsymbol{x}|/\xi) \quad , \end{aligned} \tag{6.222}$$

where $K_0(z)$ is the modified Bessel function of the second kind. As $z \to 0$ we have[21]

$$K_0(z) = -\log z + \log 2 - \gamma + \mathcal{O}(z^2 \log z) \quad , \tag{6.223}$$

where $\gamma = 0.57721566\ldots$ is Euler's constant.

In higher dimensions $d > 2$ we can obtain the result asymptotically in two limits:

- Take $r \to \infty$ with ξ fixed. Then

$$\chi_d^{\rm OZ}(\boldsymbol{r}) \simeq -C_d\, n \cdot \frac{\xi^{(3-d)/2}}{k_{\rm B}T\, R^2} \cdot \frac{e^{-r/\xi}}{r^{(d-1)/2}} \cdot \left\{1 + \mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad , \tag{6.224}$$

  where the C_d are dimensionless constants.

- Take $\xi \to \infty$ with r fixed; this is the limit $T \to T_{\rm c}$ at fixed r. In dimensions $d > 2$ we obtain

$$\chi_d^{\rm OZ}(\boldsymbol{r}) \simeq -\frac{C_d'\, n}{k_{\rm B}T R^2} \cdot \frac{e^{-r/\xi}}{r^{d-2}} \cdot \left\{1 + \mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad . \tag{6.225}$$

At criticality, $\xi \to \infty$, and clearly our results in $d = 1$ and $d = 2$ dimensions are nonsensical, as they are divergent. To correct this behavior, M. E. Fisher in 1963 suggested that the OZ correlation functions in the $r \ll \xi$ limit be replaced by

$$\chi(\boldsymbol{r}) \simeq -\frac{C_d''\, n}{k_{\rm B}T R^2} \cdot \frac{e^{-r/\xi}}{r^{d-2+\eta}} \quad , \tag{6.226}$$

a result known as *anomalous scaling*. Here, η is the *anomalous scaling exponent*.

Recall that the isothermal compressibility is given by $\kappa_T = -\hat{\chi}(0)$. Near criticality, the integral in $\hat{\chi}(0)$ is dominated by the $r \ll \xi$ part, since $\xi \to \infty$. Thus, using Fisher's anomalous scaling,

$$\begin{aligned} \kappa_T = -\hat{\chi}(0) &= -\int d^d r\, \chi(\boldsymbol{r}) \\ &\sim A \int d^d r\, \frac{e^{-r/\xi}}{r^{d-2+\eta}} \sim B\, \xi^{2-\eta} \sim C\, \big|T - T_{\rm c}\big|^{-(2-\eta)\nu} \quad , \end{aligned} \tag{6.227}$$

where A, B, and C are temperature-dependent constants which are nonsingular at $T = T_{\rm c}$. Thus, since $\kappa_T \propto |T - T_{\rm c}|^{-\gamma}$, we conclude

$$\gamma = (2 - \eta)\, \nu \quad , \tag{6.228}$$

a result known as *Fisher's scaling relation*.

In the above derivation of OZ theory, we have merely substituted one unknown function, $h(\boldsymbol{r})$, for another, namely $c(\boldsymbol{r})$. To close the system of equations, we need to relate $c(\boldsymbol{r})$ to $h(\boldsymbol{r})$ again in some way. There are various approximation schemes which do just this.

[21] For a systematic small z expansion, see formula 10.31.2 in the NIST Handbook of Mathematical Functions.

6.4.8 Percus-Yevick and hypernetted chains approximations

In the Percus-Yevick approximation, we take

$$c(\boldsymbol{r}) = \left[1 - e^{\beta u(\boldsymbol{r})}\right] \cdot g(\boldsymbol{r}) \quad . \tag{6.229}$$

Note that $c(\boldsymbol{r})$ vanishes whenever the potential $u(\boldsymbol{r})$ itself vanishes. This results in the following integro-differential equation for the pair distribution function $g(\boldsymbol{r})$:

$$g(\boldsymbol{r}) = e^{-\beta u(\boldsymbol{r})} + n\, e^{-\beta u(\boldsymbol{r})} \int d^3r' \, \left[g(\boldsymbol{r} - \boldsymbol{r}') - 1\right] \cdot \left[1 - e^{\beta u(\boldsymbol{r}')}\right] g(\boldsymbol{r}') \quad . \tag{6.230}$$

This is the *Percus-Yevick equation*. Remarkably, the Percus-Yevick (PY) equation can be solved analytically for the case of hard spheres, where $u(r) = \infty$ for $r \le a$ and $u(r) = 0$ for $r > a$, where a is the hard sphere diameter. Define the function $y(\boldsymbol{r}) = e^{\beta u(\boldsymbol{r})} g(\boldsymbol{r})$, in which case

$$c(\boldsymbol{r}) = y(\boldsymbol{r})\, f(\boldsymbol{r}) = \begin{cases} -y(\boldsymbol{r}) & , \quad r \le a \\ 0 & , \quad r > a \end{cases} \quad . \tag{6.231}$$

Here, $f(\boldsymbol{r}) = e^{-\beta u(\boldsymbol{r})} - 1$ is the Mayer function. We remark that the definition of $y(\boldsymbol{r})$ may cause some concern for the hard sphere system, because of the $e^{\beta u(\boldsymbol{r})}$ term, which diverges severely for $r \le a$. However, $g(r)$ vanishes in this limit, and their product $y(\boldsymbol{r})$ is in fact finite! The PY equation may then be written for the function $y(\boldsymbol{r})$ as

$$y(\boldsymbol{r}) = 1 + n \int\limits_{r'<a} d^3r' \, y(\boldsymbol{r}') - n \int\limits_{\substack{r'<a \\ |\boldsymbol{r}-\boldsymbol{r}'|>a}} d^3r' \, y(\boldsymbol{r}')\, y(\boldsymbol{r} - \boldsymbol{r}') \quad . \tag{6.232}$$

This has been solved using Laplace transform methods by M. S. Wertheim, *J. Math. Phys.* **5**, 643 (1964). The final result for $c(\boldsymbol{r})$ is

$$c(\boldsymbol{r}) = -\left\{\lambda_1 + 6\varphi\,\lambda_2 \left(\frac{r}{a}\right) + \tfrac{1}{2}\varphi\,\lambda_1 \left(\frac{r}{a}\right)^3\right\} \Theta(a - r) \quad , \tag{6.233}$$

where $\varphi = \frac{1}{6}\pi a^3 n$ is the packing fraction[22] and

$$\lambda_1 = \frac{(1 + 2\varphi)^2}{(1 - \varphi)^4} \qquad , \qquad \lambda_2 = -\frac{(1 + \frac{1}{2}\varphi)^2}{(1 - \varphi)^4} \quad . \tag{6.234}$$

This leads to the equation of state

$$p = n k_{\rm B} T \, \frac{1 + \varphi + \varphi^2}{(1 - \varphi)^3} \quad . \tag{6.235}$$

This gets B_2 and B_3 exactly right. The accuracy of the PY approximation for higher order virial coefficients is shown in table 6.1.

[22] Thus $\varphi = \frac{4}{3}\pi s^3 n$ is the mean number of particles within a hard sphere volume, with the hard sphere radius given by $s = \frac{1}{2}a$.

quantity	exact	PY	HNC
B_4/B_2^3	0.28695	0.2969	0.2092
B_5/B_2^4	0.1103	0.1211	0.0493
B_6/B_2^5	0.0386	0.0281	0.0449
B_7/B_2^6	0.0138	0.0156	–

Table 6.1: Comparison of exact (Monte Carlo) results to those of the Percus-Yevick (PY) and hypernetted chains approximation (HNC) for hard spheres in three dimensions. *Sources*: Hansen and McDonald (1990) and Reichl (1998)

To obtain the equation of state from eqn. 6.233, we invoke the compressibility equation,

$$nk_{\rm B}T\,\kappa_T = \left(\frac{\partial n}{\partial p}\right)_T = \frac{1}{1-n\,\hat{c}(0)} \quad . \tag{6.236}$$

We therefore need

$$\begin{aligned}\hat{c}(0) = \int d^3r\, c(\boldsymbol{r}) &= -4\pi a^3 \int_0^1 dx\, x^2\, \left[\lambda_1 + 6\,\varphi\,\lambda_2\,x + \tfrac{1}{2}\,\varphi\,\lambda_1\,x^3\right] \\ &= -4\pi a^3 \left[\tfrac{1}{3}\,\lambda_1 + \tfrac{3}{2}\,\varphi\,\lambda_2 + \tfrac{1}{12}\,\varphi\,\lambda_1\right] \quad .\end{aligned} \tag{6.237}$$

With $\varphi = \frac{1}{6}\pi a^3 n$ and using the definitions of $\lambda_{1,2}$ in eqn. 6.234, one finds

$$1 - n\,\hat{c}(0) = \frac{1+4\varphi+4\varphi^2}{(1-\varphi)^4} \quad . \tag{6.238}$$

We then have, from the compressibility equation,

$$\frac{6k_{\rm B}T}{\pi a^3}\,\frac{\partial p}{\partial \varphi} = \frac{1+4\varphi+4\varphi^2}{(1-\varphi)^4} \quad . \tag{6.239}$$

Integrating, we obtain $p(\varphi)$ up to a constant. The constant is set so that $p = 0$ when $n = 0$. The result is eqn. 6.235.

Another commonly used scheme is the *hypernetted chains* (HNC) approximation, for which

$$c(\boldsymbol{r}) = -\beta u(\boldsymbol{r}) + h(\boldsymbol{r}) - \log\Big(1 + h(\boldsymbol{r})\Big) \quad . \tag{6.240}$$

The rationale behind the HNC and other such approximation schemes is rooted in diagrammatic approaches, which are extensions of the Mayer cluster expansion to the computation of correlation functions. For details and references to their application in the literature, see Hansen and McDonald (1990) and Reichl (1998).

6.5 Coulomb Systems : Plasmas and the Electron Gas

6.5.1 Electrostatic potential

Coulomb systems are particularly interesting in statistical mechanics because of their long-ranged forces, which result in the phenomenon of *screening*. Long-ranged forces wreak havoc with the Mayer cluster expansion, since the Mayer function is no longer integrable. Thus, the virial expansion fails, and new techniques need to be applied to reveal the physics of plasmas.

The potential energy of a Coulomb system is

$$U = \tfrac{1}{2}\int d^d r \int d^d r' \, \rho(\boldsymbol{r})\, u(\boldsymbol{r}-\boldsymbol{r}')\, \rho(\boldsymbol{r}') \quad , \tag{6.241}$$

where $\rho(\boldsymbol{r})$ is the charge density and $u(\boldsymbol{r})$, which has the dimensions of (energy)/(charge)^2, satisfies

$$\nabla^2 u(\boldsymbol{r}-\boldsymbol{r}') = -4\pi\,\delta(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{6.242}$$

Thus,

$$u(\boldsymbol{r}) = \begin{cases} -2\pi\,|x-x'| & , \quad d=1 \\ -2\,\log|\boldsymbol{r}-\boldsymbol{r}'| & , \quad d=2 \\ |\boldsymbol{r}-\boldsymbol{r}'|^{-1} & , \quad d=3 \quad . \end{cases} \tag{6.243}$$

For discrete particles, the charge density $\rho(\boldsymbol{r})$ is given by

$$\rho(\boldsymbol{r}) = \sum_i q_i\,\delta(\boldsymbol{r}-\boldsymbol{x}_i) \quad , \tag{6.244}$$

where q_i is the charge of the $i^{\rm th}$ particle. We will assume two types of charges: $q = \pm e$, with $e > 0$. Such a system is called a *plasma*. The electric potential is

$$\phi(\boldsymbol{r}) = \int d^d r'\, u(\boldsymbol{r}-\boldsymbol{r}')\,\rho(\boldsymbol{r}') = \sum_i q_i\, u(\boldsymbol{r}-\boldsymbol{x}_i) \quad . \tag{6.245}$$

This satisfies the Poisson equation, $\nabla^2\phi(\boldsymbol{r}) = -4\pi\rho(\boldsymbol{r})$. The total potential energy can be written as

$$U = \tfrac{1}{2}\int d^d r\,\phi(\boldsymbol{r})\,\rho(\boldsymbol{r}) = \tfrac{1}{2}\sum_i q_i\,\phi(\boldsymbol{x}_i) \quad , \tag{6.246}$$

where it is understood that we omit self-interaction terms.

6.5.2 Debye-Hückel theory

A simple 'derivation'

For a gas of noninteracting classical particles with dispersion $\varepsilon(\boldsymbol{k})$, the number density in the GCE is given by

$$n(T,z) = z\int \frac{d^d k}{(2\pi)^d}\, e^{-\varepsilon(\boldsymbol{k})/k_{\rm B}T} \equiv z\lambda_T^{-d} \quad . \tag{6.247}$$

For the ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$ one has $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$. Suppose the particles each have charge q, and are in the presence of a constant external potential ϕ. Then the internal energy of each particle is $q\phi$ and $n(T,z) = z\lambda_T^{-d}\exp(-q\phi/k_{\rm B}T)$, which is obvious after replacing $\varepsilon(\boldsymbol{k}) \to \varepsilon(\boldsymbol{k}) + q\phi$. Now let us suppose $\phi(\boldsymbol{r})$ varies smoothly in space. We then have that the *local* number density is

$$n(\boldsymbol{r}) = z\lambda_T^{-d}\exp(-q\,\phi(\boldsymbol{r})/k_{\rm B}T) \quad . \tag{6.248}$$

If there are two species of charges, with $q = \pm e$, then the local charge density is

$$\rho(\boldsymbol{r}) = en_+(\boldsymbol{r}) - en_-(\boldsymbol{r}) + \rho_{\rm ext}(\boldsymbol{r}) \quad , \tag{6.249}$$

where $n_\pm(\boldsymbol{r}) = z_\pm\lambda_\pm^{-d}\exp(\mp e\phi(\boldsymbol{r})/k_{\rm B}T)$ and where $\rho_{\rm ext}$ is a fixed external charge density[23]. Now we close the loop and acknowledge that the total charge density $\rho(\boldsymbol{r})$ is the *source* of the electric potential $\phi(\boldsymbol{r})$, which entails Poisson's equation. We then arrive at the coupled system of equations,

$$\begin{aligned}\rho(\boldsymbol{r}) &= ez_+\lambda_+^{-d}\,e^{-e\phi(\boldsymbol{r})/k_{\rm B}T} - ez_-\lambda_-^{-d}\,e^{+e\phi(\boldsymbol{r})/k_{\rm B}T} + \rho_{\rm ext}(\boldsymbol{r}) \\ \nabla^2\phi(\boldsymbol{r}) &= -4\pi\rho(\boldsymbol{r}) \quad .\end{aligned} \tag{6.250}$$

These are the nonlinear Poisson-Boltzmann equations. In the weak potential limit, their linearization yields what is known as Debye-Hückel theory (see below).

Mean field approach

Formally, the grand partition function for our plasma is given by

$$\Xi(T,V,\mu_+,\mu_-) = \sum_{N_+=0}^{\infty}\sum_{N_-=0}^{\infty}\frac{(z_+\lambda_+^{-d})^{N_+}}{N_+!}\cdot\frac{(z_-\lambda_-^{-d})^{N_-}}{N_-!}\cdot\int d^dx_1\cdots\int d^dx_N\,e^{-\beta U(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad , \tag{6.251}$$

where $z_\pm = \exp(\mu_\pm/k_{\rm B}T)$ are the fugacities, $\lambda_\pm = (2\pi\hbar^2/m_\pm k_{\rm B}T)^{1/2}$ are the thermal wavelengths, $N = N_+ + N_-$ is the total number of particles in the (N_+,N_-) sector of the double sum, and

$$U(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \sum_{i<j}^{N} q_i\,q_j\,u(\boldsymbol{x}_i - \boldsymbol{x}_j) \tag{6.252}$$

is the total potential energy. The many body integral in position space is in general intractable. What's more, the Mayer function

$$f_{ij} = e^{-\beta\,q_iq_ju(r_{ij})} - 1 \tag{6.253}$$

behaves as $f_{ij} \approx -\beta\,q_iq_ju(r_{ij})$ as $r_{ij}\to\infty$, which is not integrable. Thus it would prove a tricky business indeed to attempt to derive a cluster expansion as we did in §6.3.3 for the equation of state[24].

Here we adopt a *mean field* approach, writing

$$\rho(\boldsymbol{r}) = \bar\rho(\boldsymbol{r}) + \delta\rho(\boldsymbol{r}) \qquad , \qquad \phi(\boldsymbol{r}) = \bar\phi(\boldsymbol{r}) + \delta\phi(\boldsymbol{r}) \quad . \tag{6.254}$$

[23]Note that there are two separate chemical potentials $\mu_\pm$ and in principle two independent thermal wavelengths $\lambda_\pm$. But there is one temperature T to rule them all.

[24]One can make progress in this direction, but it requires resumming infinite classes of diagrams.

We then have

$$\begin{aligned} U &= \tfrac{1}{2}\int d^dr\,\big[\bar\rho(\boldsymbol{r})+\delta\rho(\boldsymbol{r})\big]\cdot\big[\bar\phi(\boldsymbol{r})+\delta\phi(\boldsymbol{r})\big] \\ &= \tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\delta\rho(\boldsymbol{r}) + \tfrac{1}{2}\int d^dr\,\bar\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r})}^{\text{these two terms are the same}} + \tfrac{1}{2}\int d^dr\,\delta\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r}) \\ &= -\tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r}) + \int d^dr\,\bar\phi(\boldsymbol{r})\,\rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\int d^dr\,\delta\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r})}^{\text{ignore fluctuation term}} \\ &= -U_0 + \int d^dr\,\bar\phi(\boldsymbol{r})\,\rho(\boldsymbol{r}) + (\text{flucts})^2 \quad , \end{aligned} \tag{6.255}$$

where $U_0 \equiv \frac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r})$, and where we have used $\delta\rho = \rho - \bar\rho$. Thus we have

$$\Xi = e^{U_0/k_{\rm B}T}\exp\left(z_+\lambda_+^{-d}\int d^dr_+\,e^{-e\,\bar\phi(\boldsymbol{r}_+)/k_{\rm B}T}\right)\exp\left(z_-\lambda_-^{-d}\int d^dr_-\,e^{e\,\bar\phi(\boldsymbol{r}_-)/k_{\rm B}T}\right) \tag{6.256}$$

whence

$$\Omega(T,V,\mu_+,\mu_-) = -U_0 - k_{\rm B}Tz_+\lambda_+^{-d}\int d^dr\,\exp\left(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) - k_{\rm B}Tz_-\lambda_-^{-d}\int d^dr\,\exp\left(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad , \tag{6.257}$$

where

$$\lambda_\pm = \left(\frac{2\pi\hbar^2}{m_\pm k_{\rm B}T}\right) \qquad , \qquad z_\pm = \exp\left(\frac{\mu_\pm}{k_{\rm B}T}\right) \quad . \tag{6.258}$$

Note that since $\bar\phi(\boldsymbol{r}) = \int d^dr'\,u(\boldsymbol{r}-\boldsymbol{r}')\,\bar\rho(\boldsymbol{r}')$ is a linear functional of $\bar\rho(\boldsymbol{r})$, we have

$$\frac{\delta U_0}{\delta\bar\phi(\boldsymbol{r})} = \bar\rho(\boldsymbol{r}) \quad . \tag{6.259}$$

We next demand that the free energy Ω is extremized with respect to the mean field $\bar\phi(\boldsymbol{r})$, *viz.*

$$0 = \frac{\delta\Omega}{\delta\bar\phi(\boldsymbol{r})} = -\bar\rho(\boldsymbol{r}) + e\,\lambda_+^{-d}\,z_+\,\exp\left(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) - e\,\lambda_-^{-d}\,z_-\,\exp\left(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad . \tag{6.260}$$

This yields a mean field charge density $\bar\rho(\boldsymbol{r})$ which agrees with that in eqn. 6.250. At $r\to\infty$, we assume charge neutrality and $\phi(\infty) = 0$. Thus

$$\lambda_+^{-d}\,z_+ = n_+(\infty) = \lambda_-^{-d}\,z_- = n_-(\infty) \equiv n_\infty \quad , \tag{6.261}$$

where n_∞ is the ionic density of either species at infinity. Therefore,

$$\rho(\boldsymbol{r}) = -2e\,n_\infty\sinh\left(\frac{e\,\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad , \tag{6.262}$$

where we have dropped the bars on $\bar\phi$ and $\bar\rho$ for convenience. We now invoke Poisson's equation,

$$\nabla^2\phi = 8\pi e n_\infty\sinh(\beta e\phi) - 4\pi\rho_{\rm ext} \quad , \tag{6.263}$$

where $\rho_{\rm ext}$ is an externally imposed charge density.

If $|e\phi| \ll k_{\rm B}T$, we can expand the $\sinh$ function and obtain

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi\rho_{\rm ext} \quad , \tag{6.264}$$

where

$$\kappa_{\rm D} = \left(\frac{8\pi n_\infty e^2}{k_{\rm B}T}\right)^{1/2} \qquad , \qquad \lambda_{\rm D} = \left(\frac{k_{\rm B}T}{8\pi n_\infty e^2}\right)^{1/2} \quad . \tag{6.265}$$

The quantity $\lambda_{\rm D}$ is known as the *Debye screening length*. Consider, for example, a point charge Q located at the origin. We then solve Poisson's equation in the weak field limit,

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi Q\,\delta(\boldsymbol{r}) \quad . \tag{6.266}$$

Fourier transforming, we obtain

$$-\boldsymbol{q}^2\,\hat\phi(\boldsymbol{q}) = \kappa_{\rm D}^2\,\hat\phi(\boldsymbol{q}) - 4\pi Q \qquad \Rightarrow \qquad \hat\phi(\boldsymbol{q}) = \frac{4\pi Q}{\boldsymbol{q}^2 + \kappa_{\rm D}^2} \quad . \tag{6.267}$$

Transforming back to real space, we obtain, in three dimensions, the Yukawa potential,

$$\phi(\boldsymbol{r}) = \int\frac{d^3q}{(2\pi)^3}\,\frac{4\pi Q\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}}}{\boldsymbol{q}^2+\kappa_{\rm D}^2} = \frac{Q}{r}\,e^{-\kappa_{\rm D}r} \quad . \tag{6.268}$$

For the charge density,

$$\hat\rho(\boldsymbol{q}) = \frac{\boldsymbol{q}^2}{4\pi}\,\hat\phi(\boldsymbol{q}) = \frac{Q\,\boldsymbol{q}^2}{\boldsymbol{q}^2+\kappa_{\rm D}^2} \quad \Rightarrow \quad \rho(\boldsymbol{r}) = Q\left(\delta(\boldsymbol{r}) - \frac{\kappa_{\rm D}^2}{4\pi r}\,e^{-\kappa_{\rm D}r}\right) \quad . \tag{6.269}$$

Note that $\hat\rho(0) = 0$, which says there is total charge neutrality: $\int d^3r\,\rho(\boldsymbol{r}) = \hat\rho(0) = 0$. Thus, the plasma perfectly screens the external charge. This solution must however break down sufficiently close to $r = 0$, where the assumption $e\phi(r) \ll k_{\rm B}T$ is not valid. For larger r, the Yukawa form is increasingly accurate.

For another example, consider an electrolyte held between two conducting plates, one at an electric potential (*i.e.* voltage) $\phi(x=0) = 0$ and the other at potential $\phi(x=L) = V$, where $\hat{\boldsymbol{x}}$ is normal to the plane of the plates. Again assuming a weak field $e\phi \ll k_{\rm B}T$, we solve $\nabla^2\phi = \kappa_{\rm D}^2\,\phi$ and obtain

$$\phi(x) = A\,e^{\kappa_{\rm D}x} + B\,e^{-\kappa_{\rm D}\,x} \quad . \tag{6.270}$$

We fix the constants A and B by invoking the boundary conditions, which results in

$$\phi(x) = \frac{V\sinh(\kappa_{\rm D}x)}{\sinh(\kappa_{\rm D}L)} \quad . \tag{6.271}$$

Debye-Hückel theory is valid provided $n_\infty\,\lambda_{\rm D}^3 \gg 1$, so that the statistical assumption of many charges in a screening volume is justified.

6.5.3 The electron gas : Thomas-Fermi screening

Assuming $k_{\rm B}T \ll \varepsilon_{\rm F}$, thermal fluctuations are unimportant and we may assume $T = 0$. In the same spirit as the Debye-Hückel approach, we assume a slowly varying mean electrostatic potential $\phi(\boldsymbol{r})$. Locally, we can write

$$\varepsilon_{\rm F} = \frac{\hbar^2 k_{\rm F}^2}{2m} - e\phi(\boldsymbol{r}) \quad . \tag{6.272}$$

Thus, the Fermi wavevector $k_{\rm F}$ is spatially varying, according to the relation

$$k_{\rm F}(\boldsymbol{r}) = \left[\frac{2m}{\hbar^2}\Big(\varepsilon_{\rm F} + e\phi(\boldsymbol{r})\Big)\right]^{1/2} \quad . \tag{6.273}$$

The local electron number density is

$$n(\boldsymbol{r}) = \frac{k_{\rm F}^3(\boldsymbol{r})}{3\pi^2} = n_\infty \left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} \quad . \tag{6.274}$$

In the presence of a uniform compensating positive background charge $\rho_+ = en_\infty$, Poisson's equation takes the form

$$\nabla^2\phi = 4\pi e\, n_\infty \left[\left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} - 1\right] - 4\pi\rho_{\rm ext}(\boldsymbol{r}) \quad . \tag{6.275}$$

If $|e\phi| \ll \varepsilon_{\rm F}$, we may expand in powers of the ratio, obtaining

$$\nabla^2\phi = \frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\,\phi - 4\pi\rho_{\rm ext} \equiv \kappa_{\rm TF}^2\,\phi - 4\pi\rho_{\rm ext} \quad , \tag{6.276}$$

where $\kappa_{\rm TF}$ is the *Thomas-Fermi wavevector*,

$$\kappa_{\rm TF} = \left(\frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\right)^{1/2} \quad . \tag{6.277}$$

Thomas-Fermi theory is valid provided $n_\infty\,\lambda_{\rm TF}^3 \gg 1$, where $\lambda_{\rm TF} = \kappa_{\rm TF}^{-1}$, so that the statistical assumption of many electrons in a screening volume is justified.

One important application of Thomas-Fermi screening is to the theory of metals. In a metal, the outer, valence electrons of each atom are stripped away from the positively charged ionic core and enter into itinerant, plane-wave-like states. These states disperse with some $\varepsilon(\boldsymbol{k})$ function (that is periodic in the Brillouin zone, *i.e.* under $\boldsymbol{k} \to \boldsymbol{k} + \boldsymbol{G}$, where $\boldsymbol{G}$ is a reciprocal lattice vector), and at $T = 0$ this *energy band* is filled up to the Fermi level $\varepsilon_{\rm F}$, as Fermi statistics dictates. (In some cases, there may be several bands at the Fermi level, as we saw in the case of yttrium.) The set of ionic cores then acts as a neutralizing positive background. In a perfect crystal, the ionic cores are distributed periodically, and the positive background is approximately uniform. A charged impurity in a metal, such as a zinc atom in a copper matrix, has a different nuclear charge and a different valency than the host. The charge of the ionic core, when valence electrons are stripped away, differs from that of the host ions, and therefore the impurity acts as a *local charge impurity*. For example, copper has an electronic configuration of $[{\rm Ar}]\,3d^{10}\,4s^1$. The $4{\rm s}$ electron forms an energy band which contains the Fermi surface. Zinc has a configuration of

$[\mathrm{Ar}]\,3d^{10}\,4s^2$, and in a Cu matrix the Zn gives up its two $4s$ electrons into the $4s$ conduction band, leaving behind a charge $+2$ ionic core. The Cu cores have charge $+1$ since each copper atom contributed only one $4s$ electron to the conduction band. The conduction band electrons neutralize the uniform positive background of the Cu ion cores. What is left is an extra $Q = +e$ nuclear charge at the Zn site, and one extra $4s$ conduction band electron. The $Q = +e$ impurity is, however, *screened* by the electrons, and at distances greater than an atomic radius the potential that a given electron sees due to the Zn core is of the Yukawa form,

$$\phi(\boldsymbol{r}) = \frac{Q}{r}\, e^{-\kappa_{\rm TF} r} \quad . \tag{6.278}$$

We should take care, however, that the dispersion $\varepsilon(\boldsymbol{k})$ for the conduction band in a metal is not necessarily of the free electron form $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$. To linear order in the potential, however, the change in the local electronic density is

$$\delta n(\boldsymbol{r}) = e\phi(\boldsymbol{r})\, g(\varepsilon_{\rm F}) \quad , \tag{6.279}$$

where $g(\varepsilon_{\rm F})$ is the density of states at the Fermi energy. Thus, in a metal, we should write

$$\nabla^2\phi = (-4\pi)(-e\,\delta n) = 4\pi e^2 g(\varepsilon_{\rm F})\,\phi = \kappa_{\rm TF}^2\,\phi \quad , \tag{6.280}$$

where

$$\kappa_{\rm TF} = \sqrt{4\pi e^2\, g(\varepsilon_{\rm F})} \quad . \tag{6.281}$$

The value of $g(\varepsilon_{\rm F})$ will depend on the form of the dispersion. For parabolic bands with an effective mass m^*, the formula in eqn. 6.276 still applies.

The Thomas-Fermi atom

Consider an ion formed of a nucleus of charge $+Ze$ and an electron cloud of charge $-Ne$. The net ionic charge is then $(Z-N)e$. Since we will be interested in atomic scales, we can no longer assume a weak field limit and we must retain the full nonlinear screening theory, for which

$$\nabla^2\phi(\boldsymbol{r}) = 4\pi e\cdot\frac{(2m)^{3/2}}{3\pi^2\hbar^3}\Big(\varepsilon_{\rm F} + e\phi(\boldsymbol{r})\Big)^{3/2} - 4\pi Ze\,\delta(\boldsymbol{r}) \quad . \tag{6.282}$$

We assume an isotropic solution. It is then convenient to define

$$\varepsilon_{\rm F} + e\phi(\boldsymbol{r}) = \frac{Ze^2}{r}\cdot\chi(r/r_0) \quad , \tag{6.283}$$

where r_0 is yet to be determined. As $r\to 0$ we expect $\chi\to 1$ since the nuclear charge is then unscreened. We then have

$$\nabla^2\left\{\frac{Ze^2}{r}\cdot\chi(r/r_0)\right\} = \frac{1}{r_0^2}\,\frac{Ze^2}{r}\,\chi''(r/r_0) \quad , \tag{6.284}$$

thus we arrive at the *Thomas-Fermi equation*,

$$\chi''(t) = \frac{1}{\sqrt{t}}\,\chi^{3/2}(t) \quad , \tag{6.285}$$

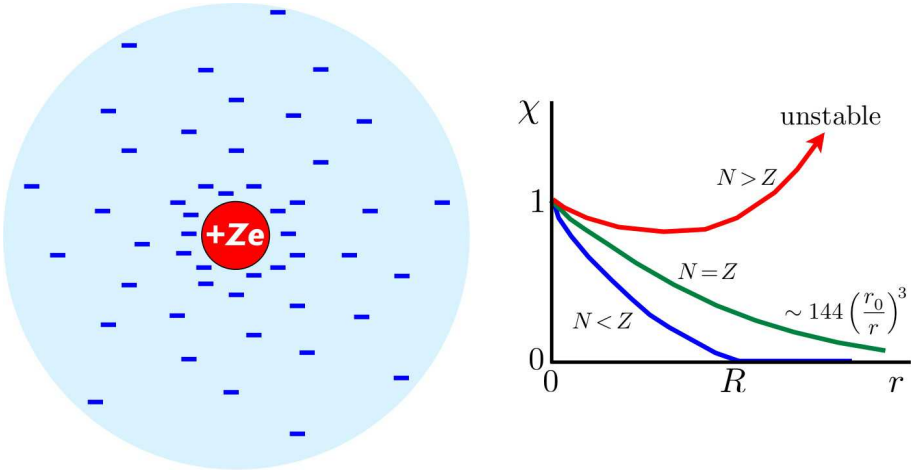


Figure 6.19: The Thomas-Fermi atom consists of a nuclear charge $+Ze$ surrounded by N electrons distributed in a cloud. The electric potential $\phi(\boldsymbol{r})$ felt by any electron at position $\boldsymbol{r}$ is screened by the electrons within this radius, resulting in a self-consistent potential $\phi(\boldsymbol{r}) = \phi_0 + (Ze/r)\,\chi(r/r_0)$.

with $r = t\,r_0$, provided we take

$$r_0 = \frac{\hbar^2}{2me^2}\left(\frac{3\pi}{4\sqrt{Z}}\right)^{2/3} = 0.885\,Z^{-1/3}\,a_{\rm B} \quad , \tag{6.286}$$

where $a_{\rm B} = \hbar^2/me^2 = 0.529\,\text{Å}$ is the Bohr radius. The TF equation is subject to the following boundary conditions:

- At short distances, the nucleus is unscreened, *i.e.* $\chi(0) = 1$.
- For positive ions, with $N < Z$, there is perfect screening at the ionic boundary $R = t^*\,r_0$, where $\chi(t^*) = 0$. This requires

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi = \left[-\frac{Ze}{R^2}\,\chi(R/r_0) + \frac{Ze}{R\,r_0}\,\chi'(R/r_0)\right]\hat{\boldsymbol{r}} = \frac{(Z-N)\,e}{R^2}\,\hat{\boldsymbol{r}} \quad . \tag{6.287}$$

  This requires

$$-t^*\,\chi'(t^*) = 1 - \frac{N}{Z} \quad . \tag{6.288}$$

For an atom, with $N = Z$, the asymptotic solution to the TF equation is a power law, and by inspection is found to be $\chi(t) \sim C\,t^{-3}$, where C is a constant. The constant follows from the TF equation, which yields $12\,C = C^{3/2}$, hence $C = 144$. Thus, a neutral TF atom has a density with a power law tail, with $\rho \sim r^{-6}$. TF ions with $N > Z$ are unstable.

6.6 Polymers

6.6.1 Basic concepts

Linear chain polymers are repeating structures with the chemical formula $(A)_x$, where A is the *formula unit* and x is the *degree of polymerization*. In many cases (*e.g.* polystyrene), $x \gtrsim 10^5$ is not uncommon. For a very readable introduction to the subject, see P. G. de Gennes, *Scaling Concepts in Polymer Physics*.

Quite often a given polymer solution will contain a distribution of x values; this is known as *polydispersity*. Various preparation techniques, such as chromatography, can mitigate the degree of polydispersity. Another morphological feature of polymers is *branching*, in which the polymers do not form linear chains.

Polymers exhibit a *static flexibility* which can be understood as follows. Consider a long chain hydrocarbon with a $-\mathrm{C}-\mathrm{C}-\mathrm{C}-$ backbone. The angle between successive $\mathrm{C}-\mathrm{C}$ bonds is fixed at $\theta \approx 68^\circ$, but the azimuthal angle φ can take one of three possible low-energy values, as shown in the right panel of fig. 6.21. Thus, the relative probabilities of *gauche* and *trans* orientations are

$$\frac{\mathrm{Prob\,(gauche)}}{\mathrm{Prob\,(trans)}} = 2\,e^{-\Delta\varepsilon/k_{\mathrm{B}}T} \tag{6.289}$$

where $\Delta\varepsilon$ is the energy difference between *trans* and *gauche* configurations. This means that the polymer chain is in fact a *random coil* with a *persistence length*

$$\ell_{\mathrm{p}} = \ell_0\,e^{\Delta\varepsilon/k_{\mathrm{B}}T} \tag{6.290}$$

where ℓ_0 is a microscopic length scale, roughly given by the length of a formula unit, which is approximately a few Ångstroms (see fig. 6.22). Let L be the total length of the polymer when it is stretched into a straight line. If $\ell_{\mathrm{p}} > L$, the polymer is *rigid*. If $\ell_{\mathrm{p}} \ll L$, the polymer is rigid on the length scale ℓ_{p} but flexible on longer scales. We have

$$\frac{\ell_{\mathrm{p}}}{L} = \frac{1}{N}\,e^{\Delta\varepsilon/k_{\mathrm{B}}T} \quad , \tag{6.291}$$

where we now use N (rather than x) for the degree of polymerization.

In the time domain, the polymer exhibits a *dynamical flexibility* on scales longer than a *persistence time*. The persistence time τ_{p} is the time required for a *trans-gauche* transition. The rate for such transitions is set by the energy barrier B separating *trans* from *gauche* configurations:

$$\tau_{\mathrm{p}} = \tau_0\,e^{B/k_{\mathrm{B}}T} \tag{6.292}$$

where $\tau_0 \sim 10^{-11}$ s. On frequency scales $\omega \ll \tau_{\mathrm{p}}^{-1}$ the polymer is dynamically flexible. If $\Delta\varepsilon \sim k_{\mathrm{B}}T \ll B$ the polymer is flexible from a static point of view, but dynamically rigid. That is, there are many *gauche* orientations of successive carbon bonds which reflect a quenched disorder. The polymer then forms a frozen random coil, like a twisted coat hanger.

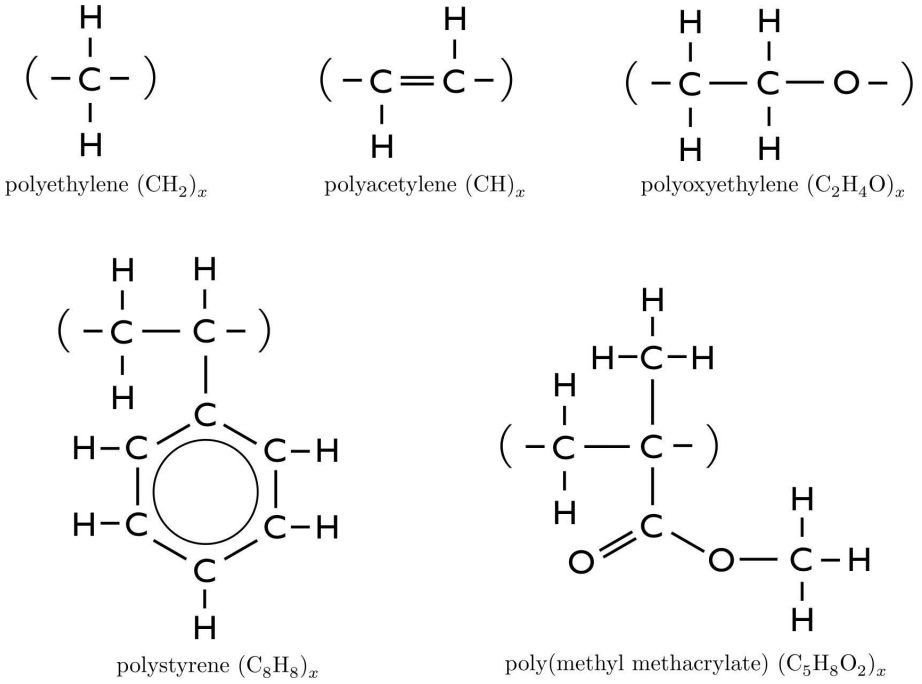


Figure 6.20: Some examples of linear chain polymers.

6.6.2 Polymers as random walks

A polymer can be modeled by a *self-avoiding random walk* (SAW). That is, on scales longer than $\ell_{\rm p}$, it twists about randomly in space subject to the constraint that it doesn't overlap itself. Before we consider the mathematics of SAWs, let's first recall some aspects of ordinary random walks which are not self-avoiding.

We'll simplify matters further by considering random walks on a hypercubic lattice of dimension d. Such a lattice has coordination number $2d$, *i.e.* there are $2d$ nearest neighbor separation vectors, given by $\boldsymbol{\delta} = \pm a\,\hat{\mathbf{e}}_1\,,\,\pm a\,\hat{\mathbf{e}}_2\,,\,\ldots\,,\,\pm a\,\hat{\mathbf{e}}_d$, where a is the lattice spacing. Consider now a random walk of N steps starting at the origin. After N steps the position is $\sum_{j=1}^N \boldsymbol{\delta}_j$, where each $\boldsymbol{\delta}_j$ takes on one of $2d$ possible values. Now N is no longer the degree of polymerization, but something approximating $L/\ell_{\rm p}$, which is the number of persistence lengths in the chain. We assume each step is independent, hence

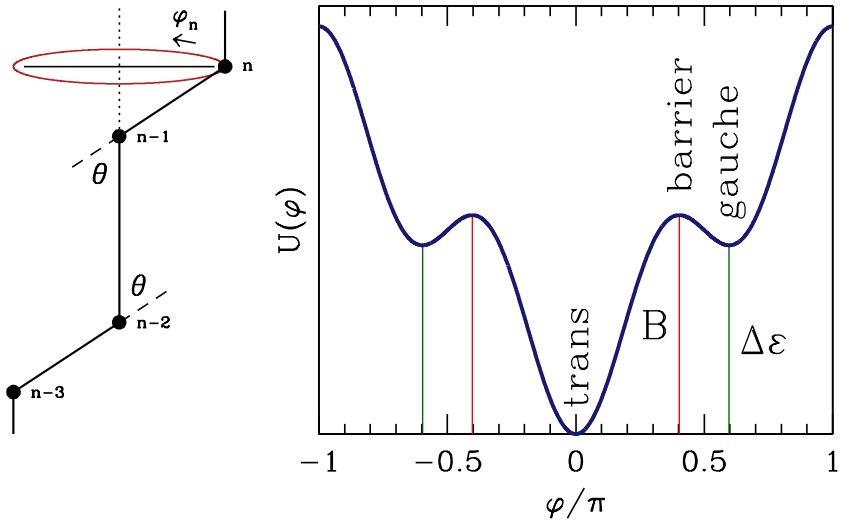


Figure 6.21: Left: *trans* and *gauche* orientations in carbon chains. Right: energy as a function of azimuthal angle φ. There are three low energy states: *trans* ($\varphi = 0$) and *gauche* ($\varphi = \pm\varphi_0$).

$\langle \delta_j^\alpha \, \delta_{j'}^\beta \rangle = (a^2/d)\, \delta_{jj'} \delta^{\alpha\beta}$ and $\big\langle \boldsymbol{R}_N^2 \big\rangle = Na^2$. The full distribution $P_N(\boldsymbol{R})$ is given by

$$\begin{aligned} P_N(\boldsymbol{R}) &= (2d)^{-N} \sum_{\boldsymbol{\delta}_1} \cdots \sum_{\boldsymbol{\delta}_N} \delta_{\boldsymbol{R},\sum_j \boldsymbol{\delta}_j} \\ &= a^d \int\limits_{-\pi/a}^{\pi/a} \frac{dk_1}{2\pi} \cdots \int\limits_{-\pi/a}^{\pi/a} \frac{dk_d}{2\pi}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \left[\frac{1}{d} \sum_{\mu=1}^{d} \cos(k_\mu a) \right]^N \\ &= a^d \int\limits_{\hat{\Omega}} \frac{d^d k}{(2\pi)^d}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \exp\left[N \log\left(1 - \frac{1}{2d}\, \boldsymbol{k}^2 a^2 + \ldots \right) \right] \\ &\approx \left(\frac{a}{2d} \right)^d \int d^d k\, e^{-N\boldsymbol{k}^2 a^2/2d}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} = \left(\frac{d}{2\pi N} \right)^{d/2} e^{-d\boldsymbol{R}^2/2Na^2} \quad . \end{aligned} \tag{6.293}$$

This is a simple Gaussian, with width $\big\langle \boldsymbol{R}^2 \big\rangle = d \cdot (Na^2/d) = Na^2$, as we have already computed. The quantity $\boldsymbol{R}$ defined here is the *end-to-end vector* of the chain. The RMS end-to-end distance is then $\langle \boldsymbol{R}^2 \rangle^{1/2} = \sqrt{N} a \equiv R_0$.

A related figure of merit is the *radius of gyration*, $R_{\rm g}$, defined by

$$R_{\rm g}^2 = \frac{1}{N} \Big\langle \sum_{n=1}^{N} \big(\boldsymbol{R}_n - \boldsymbol{R}_{\rm CM} \big)^2 \Big\rangle \quad , \tag{6.294}$$

where $R_{\rm CM} = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{R}_j$ is the center of mass position. A brief calculation yields

$$R_g^2 = \big(N + 3 - 4N^{-1} \big)\, a^2 \sim \frac{Na^2}{6} \quad , \tag{6.295}$$

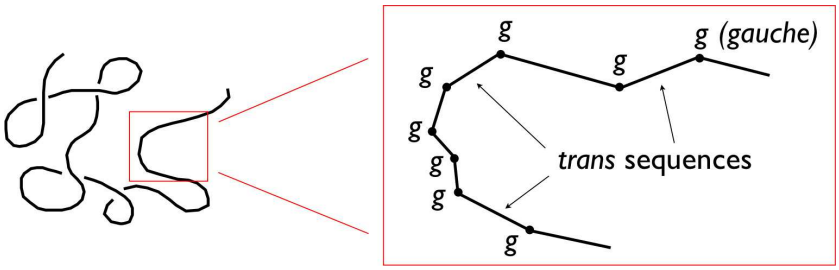


Figure 6.22: The polymer chain as a random coil.

in all dimensions.

The total number of random walk configurations with end-to-end vector $\boldsymbol{R}$ is then $(2d)^N P_N(\boldsymbol{R})$, so the entropy of a chain at fixed elongation is

$$S(\boldsymbol{R}, N) = k_{\rm B} \log \Big[(2d)^N P_N(\boldsymbol{R}) \Big] = S(0, N) - \frac{dk_{\rm B}\boldsymbol{R}^2}{2Na^2} \quad . \tag{6.296}$$

If we assume that the energy of the chain is conformation independent, then $E = E_0(N)$ and

$$F(\boldsymbol{R}, N) = F(0, N) + \frac{dk_{\rm B}T\boldsymbol{R}^2}{2Na^2} \quad . \tag{6.297}$$

In the presence of an external force $F_{\rm ext}$, the Gibbs free energy is the Legendre transform

$$G(\boldsymbol{F}_{\rm ext}, N) = F(\boldsymbol{R}, N) - \boldsymbol{F}_{\rm ext} \cdot \boldsymbol{R} \quad , \tag{6.298}$$

and $\partial G/\partial \boldsymbol{R} = 0$ then gives the relation

$$\big\langle \boldsymbol{R}(\boldsymbol{F}_{\rm ext}, N) \big\rangle = \frac{Na^2}{dk_{\rm B}T}\, \boldsymbol{F}_{\rm ext} \quad . \tag{6.299}$$

This may be considered an equation of state for the polymer.

Following de Gennes, consider a chain with charges $\pm e$ at each end, placed in an external electric field of magnitude $E = 30,000\,{\rm V/cm}$. Let $N = 10^4$, $a = 2\,\text{Å}$, and $d = 3$. What is the elongation? From the above formula, we have

$$\frac{R}{R_0} = \frac{eER_0}{3k_{\rm B}T} = 0.8 \quad , \tag{6.300}$$

with $R_0 = \sqrt{N}a$ as before.

Structure factor

We can also compute the structure factor,

$$S(\boldsymbol{k}) = \frac{1}{N} \left\langle \sum_{m=1}^{N} \sum_{n=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \right\rangle = 1 + \frac{2}{N} \sum_{m=1}^{N} \sum_{n=1}^{m-1} \left\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \right\rangle \quad . \tag{6.301}$$

For averages with respect to a Gaussian distribution,

$$\left\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\right\rangle = \exp\left\{-\frac{1}{2}\left\langle\big(\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)\big)^2\right\rangle\right\} \quad . \tag{6.302}$$

Now for $m > n$ we have $\boldsymbol{R}_m - \boldsymbol{R}_n = \sum_{j=n+1}^{m} \boldsymbol{\delta}_j$, and therefore

$$\left\langle\big(\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)\big)^2\right\rangle = \sum_{j=n+1}^{m}\left\langle(\boldsymbol{k}\cdot\boldsymbol{\delta}_j)^2\right\rangle = \frac{1}{d}\,(m-n)\,\boldsymbol{k}^2a^2 \quad , \tag{6.303}$$

since $\langle\delta_j^\alpha\,\delta_{j'}^\beta\rangle = (a^2/d)\,\delta_{jj'}\delta^{\alpha\beta}$. We then have

$$S(\boldsymbol{k}) = 1 + \frac{2}{N}\sum_{m=1}^{N}\sum_{n=1}^{m-1} e^{-(m-n)\,\boldsymbol{k}^2a^2/2d} = \frac{N\,(e^{2\mu_{\boldsymbol{k}}}-1) - 2\,e^{\mu_{\boldsymbol{k}}}\,(1-e^{-N\mu_{\boldsymbol{k}}})}{N\big(e^{\mu_{\boldsymbol{k}}}-1\big)^2} \quad , \tag{6.304}$$

where $\mu_{\boldsymbol{k}} = \boldsymbol{k}^2a^2/2d$. In the limit where $N\to\infty$ and $a\to 0$ with $Na^2 = R_0^2$ constant, the structure factor has a *scaling form*, $S(\boldsymbol{k}) = Nf(N\mu_{\boldsymbol{k}}) = (R_0/a)^2\,f(k^2R_0^2/2d)$, where

$$f(x) = \frac{2}{x^2}\left(e^{-x}-1+x\right) = 1 - \frac{x}{3} + \frac{x^2}{12} + \ldots \quad . \tag{6.305}$$

6.6.3 The Rouse model

Consider next a polymer chain subjected to stochastic forcing. We model the chain as a collection of mass points connected by springs, with a potential energy $U = \frac{1}{2}k\sum_n\big(\boldsymbol{x}_{n+1}-\boldsymbol{x}_n\big)^2$. This reproduces the distribution of eqn. 6.293 if we take the spring constant to be $k = 3k_{\rm B}T/a^2$ and set the equilibrium length of each spring to zero. The equations of motion are then

$$M\ddot{\boldsymbol{x}}_n + \gamma\,\dot{\boldsymbol{x}}_n = -k\big(2\boldsymbol{x}_n - \boldsymbol{x}_{n-1} - \boldsymbol{x}_{n+1}\big) + \boldsymbol{f}_n(t) \quad , \tag{6.306}$$

where $n\in\{1,\ldots,N\}$ and $\{f_n^\mu(t)\}$ a set of Gaussian white noise forcings, each with zero mean, and

$$\left\langle f_n^\mu(t)\,f_{n'}^\nu(t')\right\rangle = 2\gamma k_{\rm B}T\,\delta_{nn'}\,\delta^{\mu\nu}\,\delta(t-t') \quad . \tag{6.307}$$

We define $\boldsymbol{x}_0\equiv\boldsymbol{x}_1$ and $\boldsymbol{x}_{N+1}\equiv\boldsymbol{x}_N$ so that the end mass points $n=1$ and $n=N$ experience a restoring force from only one neighbor. We assume the chain is overdamped and set $M\to 0$. We then have

$$\gamma\,\dot{\boldsymbol{x}}_n = -k\sum_{n'=1}^{N} A_{nn'}\,\boldsymbol{x}_{n'} + \boldsymbol{f}_n(t) \quad , \tag{6.308}$$

where

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ 0 & 0 & -1 & \ddots & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix} \quad . \tag{6.309}$$

The matrix A is real and symmetric. Its eigenfunctions are labeled $\psi_j(n)$, with $j \in \{0, \ldots, N-1\}$. The $j = 0$ eigenfunction is a constant, $\psi_0(n) = 1/\sqrt{N}$, and the others are given by

$$\psi_j(n) = \sqrt{\frac{2}{N}}\,\cos\left(\frac{(2n-1)j\pi}{2N}\right) \qquad , \qquad j \in \{1, \ldots, N-1\} \tag{6.310}$$

The completeness and orthonormality relations are

$$\sum_{j=0}^{N-1} \psi_j(n)\,\psi_j(n') = \delta_{nn'} \qquad , \qquad \sum_{n=1}^{N} \psi_j(n)\,\psi_{j'}(n) = \delta_{jj'} \quad , \tag{6.311}$$

with eigenvalues $\lambda_j = 4\sin^2\!\big(\pi j/2N\big)$. Note that $\lambda_0 = 0$.

We now work in the basis of normal modes $\{\eta_j^\mu\}$, where

$$\eta_j^\mu(t) = \sum_{n=1}^{N} \psi_j(n)\,x_n^\mu(t) \qquad , \qquad x_n^\mu(t) = \sum_{j=0}^{N-1} \psi_j(n)\,\eta_j^\mu(t) \quad . \tag{6.312}$$

We then have

$$\frac{d\boldsymbol{\eta}_j}{dt} = -\frac{1}{\tau_j}\,\boldsymbol{\eta}_j + \boldsymbol{g}_j(t) \quad , \tag{6.313}$$

where the $j^{\rm th}$ relaxation time is

$$\tau_j = \frac{\gamma}{4k\sin^2\!\big(\pi j/2N\big)} \tag{6.314}$$

and

$$g_j^\mu(t) = \gamma^{-1}\sum_{n=1}^{N} \psi_j(n)\,f_n^\mu(t) \quad . \tag{6.315}$$

Note that

$$\big\langle g_j^\mu(t)\,g_{j'}^\nu(t')\big\rangle = 2\gamma^{-1}k_{\rm B}T\,\delta_{jj'}\,\delta^{\mu\nu}\,\delta(t-t') \quad . \tag{6.316}$$

Integrating eqn. 6.313 we have, for $j = 0$,

$$\boldsymbol{\eta}_0(t) = \boldsymbol{\eta}_0(0) + \int_0^t dt'\,\boldsymbol{g}_0(t') \quad . \tag{6.317}$$

For the $j > 0$ modes,

$$\boldsymbol{\eta}_j(t) = \boldsymbol{\eta}_j(0)\,e^{-t/\tau_j} + \int_0^t dt'\,\boldsymbol{g}_j(t')\,e^{(t'-t)/\tau_j} \quad . \tag{6.318}$$

Thus,

$$\begin{aligned}\big\langle \eta_0^\mu(t)\,\eta_0^\nu(t')\big\rangle_{\rm c} &= 2\gamma^{-1}k_{\rm B}T\,\delta^{\mu\nu}\,\min(t,t') \\ \big\langle \eta_j^\mu(t)\,\eta_j^\nu(t')\big\rangle_{\rm c} &= \gamma^{-1}k_{\rm B}T\,\delta^{\mu\nu}\,\tau_j\left(e^{-|t-t'|/\tau_j} - e^{-(t+t')/\tau_j}\right) \quad ,\end{aligned} \tag{6.319}$$

where the 'connected average' is defined to be $\langle A(t)\,B(t')\rangle_{\rm c} \equiv \langle A(t)\,B(t')\rangle - \langle A(t)\rangle\langle B(t')\rangle$. Transforming back to the original real space basis, we then have

$$\left\langle x^\mu_n(t)\,x^\nu_{n'}(t')\right\rangle_{\rm c} = \frac{2k_{\rm B}T}{N\gamma}\,\delta^{\mu\nu}\mathsf{min}(t,t') + \frac{k_{\rm B}T}{\gamma}\,\delta^{\mu\nu}\sum_{j=1}^{N-1}\tau_j\,\psi_j(n)\,\psi_j(n')\left(e^{-|t-t'|/\tau_j} - e^{-(t+t')/\tau_j}\right) \quad . \tag{6.320}$$

In particular, the 'connected variance' of $\boldsymbol{x}_n(t)$ is

$$\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] \equiv \left\langle\big[\boldsymbol{x}_n(t)\big]^2\right\rangle_{\rm c} = \frac{6k_{\rm B}T}{N\gamma}\,t + \frac{3k_{\rm B}T}{\gamma}\sum_{j=1}^{N-1}\tau_j\,\big[\psi_j(n)\big]^2\left(1-e^{-2t/\tau_j}\right) \quad . \tag{6.321}$$

From this we see that at long times, *i.e.* when $t \gg \tau_1$, the motion of $\boldsymbol{x}_n(t)$ is diffusive, with diffusion constant $D = k_{\rm B}T/N\gamma \propto B^{-1}$, which is inversely proportional to the chain length. Recall the Stokes result $\gamma = 6\pi\eta R$ for a sphere of radius R moving in a fluid of dynamical viscosity η. From $D = k_{\rm B}T/\gamma$, shouldn't we expect the diffusion constant to be $D = k_{\rm B}T/6\pi\eta R \propto N^{-1/2}$, since the radius of gyration of the polymer is $R_{\rm g} \propto N^{1/2}$? This argument smuggles in the assumption that the only dissipation is taking place at the *outer surface* of the polymer, modeled as a ball of radius $R_{\rm g}$. In fact, for a Gaussian random walk in three space dimensions, the density for $r < R_{\rm g}$ is $\rho \propto N^{-1/2}$ since there are N monomers inside a region of volume $N^{3/2}$. Accounting for Flory swelling due to steric interactions (see below), the density is $\rho \sim N^{-4/5}$, which is even smaller. So as $N \to \infty$, the density within the $r = R_{\rm g}$ effective sphere gets small, which means water molecules can easily penetrate, in which case the entire polymer chain should be considered to be in a dissipative environment, which is what the Rouse model says – each monomer executes overdamped motion.

A careful analysis of eqn. 6.321 reveals that there is a subdiffusive regime[25] where $\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] \propto t^{1/2}$. To see this, first take the $N \gg 1$ limit, in which case we may write $\tau_j = N^2\tau_0/j^2$, where $\tau_0 \equiv \gamma/\pi^2 k$ and $j \in \{1,\ldots,N-1\}$. Let $s \equiv (n-\frac{1}{2})/N \in [0,1]$ be the scaled coordinate along the chain. The second term in eqn. 6.321 is then

$$S(s,t) \equiv \frac{6k_{\rm B}T}{\gamma}\cdot\frac{\tau_1}{N}\sum_{j=1}^{N-1}\frac{\cos^2(\pi j s)}{j^2}\left(1-e^{-2j^2t/\tau_1}\right) \quad . \tag{6.322}$$

Let $\sigma \equiv (t/\tau_1)^{1/2}$. When $t \ll \tau_1$, *i.e.* $\sigma \ll 1$, we have

$$S(s,t) \simeq \frac{6k_{\rm B}T}{\gamma}\cdot\frac{\tau_1}{N}\,\sigma\int_0^{N\sigma} du\,\frac{\cos^2(\pi u s/\sigma)}{u^2}\left(1-e^{-2u^2}\right) \quad . \tag{6.323}$$

Since $s/\sigma \gg 1$, we may replace the cosine squared term by its average $\frac{1}{2}$. If we further assume $N\sigma \gg 1$, which means we are in the regime $1 \ll t/\tau_0 \ll N^2$, after performing the integral we obtain the result

$$S(s,t) = \frac{3k_{\rm B}T}{\gamma}\sqrt{2\pi\tau_0 t} \quad , \tag{6.324}$$

provided $s = \mathcal{O}(1)$, *i.e.* the site n is not on either end of the chain. The result in eqn. 6.324 dominates the first term on the RHS of eqn. 6.321 since $\tau_0 \ll t \ll \tau_1$. This is the subdiffusive regime.

[25] I am grateful to Jonathan Lam and Olga Dudko for explaining this to me.

When $t \gg \tau_1 = N^2\tau_0$, the exponential on the RHS of eqn. 6.322 is negligible, and if we again approximate $\cos^2(\pi js) \simeq \frac{1}{2}$, and we extend the upper limit on the sum to infinity, we obtain

$$S(t) = \frac{3k_{\rm B}T}{\gamma} \cdot \frac{\tau_1}{N} \cdot \frac{\pi^2}{6} \propto t^0 \quad , \tag{6.325}$$

which is dominated by the leading term on the RHS of eqn. 6.321. This is the diffusive regime, with $D = k_{\rm B}T/N\gamma$.

Finally, when $t \ll \tau_0$, the factor $1 - \exp(-2t/\tau_j)$ may be expanded to first order in t. One then obtains $\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] = (6k_{\rm B}T/\gamma)\,t$, which is independent of the force constant k. In this regime, the monomers don't have time to respond to the force from their neighbors, hence they each diffuse independently. On such short time scales, however, one should check to make sure that inertial effects can be ignored, *i.e.* that $t \gg M/\gamma$.

One serious defect of the Rouse model is its prediction of the relaxation time of the $j = 1$ mode, $\tau_1 \propto N^2$. The experimentally observed result is $\tau_1 \propto N^{3/2}$. We should stress here that the Rouse model applies to ideal chains. In the theory of polymer solutions, a *theta solvent* is one in which polymer coils act as ideal chains. An extension of the Rouse model, due to my former UCSD colleague Bruno Zimm, accounts for hydrodynamically-mediated interactions between any pair of 'beads' along the chain. Specifically, the Zimm model is given by

$$\frac{dx_n^\mu}{dt} = \sum_{n'} H^{\mu\nu}(\boldsymbol{x}_n - \boldsymbol{x}_{n'})\Big[k\big(x^\nu_{n'+1} + x^\nu_{n'-1} - 2x^\nu_{n'}\big) + f^\nu_{n'}(t)\Big] \quad , \tag{6.326}$$

where

$$H^{\mu\nu}(\boldsymbol{R}) = \frac{1}{6\pi\eta R}\big(\delta^{\mu\nu} + \hat{R}^\mu \hat{R}^\nu\big) \tag{6.327}$$

is known as the Oseen hydrodynamic tensor (1927) and arises when computing the velocity in a fluid at position $\boldsymbol{R}$ when a point force $\boldsymbol{F} = \boldsymbol{f}\,\delta(\boldsymbol{r})$ is applied at the origin. Typically one replaces $H(\boldsymbol{R})$ by its average over the equilibrium distribution of polymer configurations. Zimm's model more correctly reproduces the behavior of polymers in θ-solvents.

6.6.4 Flory theory of self-avoiding walks

What is missing from the random walk free energy is the effect of *steric interactions*. An argument due to Flory takes these interactions into account in a mean field treatment. Suppose we have a chain of radius R. Then the average monomer density within the chain is $c = N/R^d$. Assuming short-ranged interactions, we should then add a term to the free energy which effectively counts the number of near self-intersections of the chain. This number should be roughly Nc. Thus, we write

$$F(\boldsymbol{R}, N) = F_0 + u(T)\,\frac{N^2}{R^d} + \tfrac{1}{2}dk_{\rm B}T\,\frac{R^2}{Na^2} \quad . \tag{6.328}$$

The effective interaction $u(T)$ is positive in the case of a so-called 'good solvent'.

The free energy is minimized when

$$0 = \frac{\partial F}{\partial R} = -\frac{duN^2}{R^{d+1}} + \frac{dRk_{\rm B}T}{Na^2} \quad , \tag{6.329}$$

which yields the result

$$R_{\rm F}(N) = \left(\frac{ua^2}{k_{\rm B}T}\right)^{1/(d+2)} N^{3/(d+2)} \propto N^{\nu} \quad . \tag{6.330}$$

Thus, we obtain $\nu = 3/(d+2)$. In $d = 1$ this says $\nu = 1$, which is exactly correct because a SAW in $d = 1$ has no option but to keep going in the same direction. In $d = 2$, Flory theory predicts $\nu = \frac{3}{4}$, which is also exact. In $d = 3$, we have $\nu_{d=3} = \frac{3}{5}$, which is extremely close to the numerical value $\nu = 0.5880$. Flory theory is again exact above the SAW upper critical dimension, which is $d = 4$, where $\nu = \frac{1}{2}$, corresponding to a Gaussian random walk[26]. Best. Mean. Field. Theory. Ever.

How well are polymers described as SAWs? Fig. 6.23 shows the radius of gyration $R_{\rm g}$ *versus* molecular weight M for polystyrene chains in a toluene and benzene solvent. The slope is $\nu = d\log R_{\rm g}/d\log M = 0.5936$. Experimental results can vary with concentration and temperature, but generally confirm the validity of the SAW model.

For a SAW under an external force, we compute the Gibbs partition function,

$$Y(\boldsymbol{F}_{\rm ext}, N) = \int d^dR\, P_N(\boldsymbol{R})\, e^{\boldsymbol{F}_{\rm ext}\cdot\boldsymbol{R}/k_{\rm B}T} = \int d^dx\, f(x)\, e^{s\hat{\boldsymbol{n}}\cdot\boldsymbol{x}} \quad , \tag{6.331}$$

where $x = R/R_{\rm F}$ and $s = R_{\rm F}F_{\rm ext}/k_{\rm B}T$ and $\hat{\boldsymbol{n}} = \hat{\boldsymbol{F}}_{\rm ext}$. One then has $R(F_{\rm ext}) = R_{\rm F}\,\Phi(R_{\rm F}/\xi)$, where $\xi = k_{\rm B}T/F_{\rm ext}$ and, in the small $F_{\rm ext}$ limit, $R(F_{\rm ext}) = F_{\rm ext}R_{\rm F}^2/k_{\rm B}T$. For small values of its argument one has $\Phi(u) \propto u$. For large u it can be shown that $R(F_{\rm ext}) \propto (F_{\rm ext}R_{\rm F}/k_{\rm B}T)^{2/3}$.

On a lattice of coordination number z, the number of N-step random walks starting from the origin is $\Omega_N = z^N$. If we constrain our random walks to be self-avoiding, the number is reduced to

$$\Omega_N^{\rm SAW} = \mathcal{C}\, N^{\gamma-1}\, y^N \quad , \tag{6.332}$$

where $\mathcal{C}$ and γ are dimension-dependent constants, and we expect $y \lesssim z-1$, since at the very least a SAW cannot immediately double back on itself. In fact, on the cubic lattice one has $z = 6$ but $y = 4.68$, slightly less than $z-1$. One finds $\gamma_{d=2} \simeq \frac{4}{3}$ and $\gamma_{d=3} \simeq \frac{7}{6}$. The RMS end-to-end distance of the SAW is

$$R_{\rm F} = a\, N^{\nu} \quad , \tag{6.333}$$

where a and ν are d-dependent constants, with $\nu_{d=1} = 1$, $\nu_{d=2} \simeq \frac{3}{4}$, and $\nu_{d=3} \simeq \frac{3}{5}$. The distribution $P_N(\boldsymbol{R})$ has a scaling form,

$$P_N(R) = \frac{1}{R_{\rm F}^d}\, f\!\left(\frac{R}{R_{\rm F}}\right) \qquad (a \ll R \ll Na) \quad . \tag{6.334}$$

One finds

$$f(x) \sim \begin{cases} x^g & x \ll 1 \\ \exp(-x^\delta) & x \gg 1 \end{cases} \quad , \tag{6.335}$$

with $g = (\gamma-1)/\nu$ and $\delta = 1/(1-\nu)$.

[26] There are logarithmic corrections to the SAW result exactly at $d = 4$, but for all $d > 4$ one has $\nu = \frac{1}{2}$.

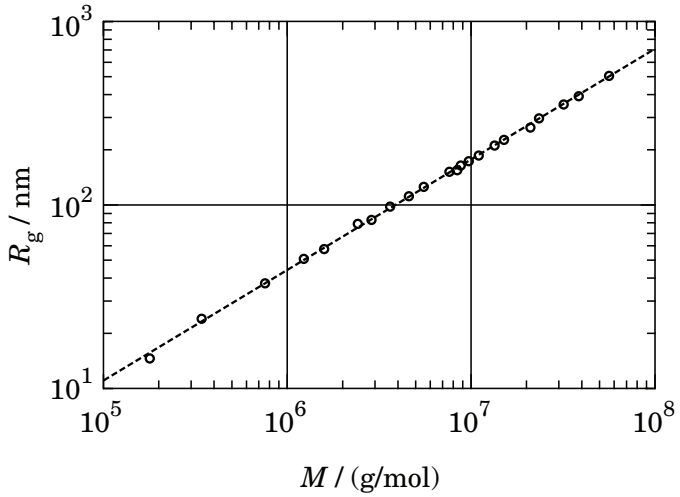


Figure 6.23: Radius of gyration $R_{\rm g}$ of polystyrene in a toluene and benzene solvent, plotted as a function of molecular weight of the polystyrene. The best fit corresponds to a power law $R_{\rm g} \propto M^{\nu}$ with $\nu = 0.5936$. From J. Des Cloizeaux and G. Jannink, *Polymers in Solution: Their Modeling and Structure* (Oxford, 1990).

6.6.5 Polymers and solvents

Consider a solution of monodisperse polymers of length N in a solvent. Let ϕ be the dimensionless monomer concentration, so ϕ/N is the dimensionless polymer concentration and $\phi_{\rm s} = 1 - \phi$ is the dimensionless solvent concentration. (Dimensionless concentrations are obtained by dividing the corresponding dimensionful concentration by the overall density.) The entropy of mixing for such a system is given by eqn. 2.352. We have

$$S_{\rm mix} = -\frac{V k_{\rm B}}{v_0} \cdot \left\{ \frac{1}{N} \phi \log \phi + (1-\phi) \log(1-\phi) \right\} \quad , \tag{6.336}$$

where $v_0 \propto a^3$ is the volume per monomer. Accounting for an interaction between the monomer and the solvent, we have that the free energy of mixing is

$$\frac{v_0 \, F_{\rm mix}}{V k_{\rm B} T} = \frac{1}{N} \, \phi \log \phi + (1-\phi) \log(1-\phi) + \chi \, \phi(1-\phi) \quad . \tag{6.337}$$

where χ is the dimensionless polymer-solvent interaction, called the *Flory parameter*. This provides a mean field theory of the polymer-solvent system.

The osmotic pressure Π is defined by

$$\Pi = -\left. \frac{\partial F_{\rm mix}}{\partial V} \right|_{N_{\rm p}} \quad , \tag{6.338}$$

which is the variation of the free energy of mixing with respect to volume *holding the number of polymers constant*. The monomer concentration is $\phi = N N_{\rm p} v_0 / V$, so

$$\left. \frac{\partial}{\partial V} \right|_{N_{\rm p}} = -\frac{\phi^2}{N N_{\rm p} \, v_0} \left. \frac{\partial}{\partial \phi} \right|_{N_{\rm p}} \quad . \tag{6.339}$$

Now we have

$$F_{\rm mix} = N N_{\rm p}\, k_{\rm B} T \left\{ \frac{1}{N}\, \log\phi + (\phi^{-1} - 1)\log(1-\phi) + \chi\,(1-\phi) \right\} \quad , \tag{6.340}$$

and therefore

$$\Pi = \frac{k_{\rm B} T}{v_0} \Big[(N^{-1} - 1)\,\phi - \log(1-\phi) - \chi\,\phi^2 \Big] \quad . \tag{6.341}$$

In the limit of vanishing monomer concentration $\phi \to 0$, we recover

$$\Pi = \frac{\phi\, k_{\rm B} T}{N v_0} \quad , \tag{6.342}$$

which is the ideal gas law for polymers. For $N^{-1} \ll \phi \ll 1$, we expand the logarithm and obtain

$$\begin{aligned} \frac{v_0\,\Pi}{k_{\rm B} T} &= \frac{1}{N}\phi + \tfrac{1}{2}(1-2\chi)\,\phi^2 + \mathcal{O}(\phi^3) \\ &\approx \tfrac{1}{2}(1-2\chi)\,\phi^2 \quad . \end{aligned} \tag{6.343}$$

As $N \to \infty$, $\Pi > 0$ only if $\chi < \frac{1}{2}$, which is the condition for a 'good solvent'. For finite N, the coefficient of the ϕ^2 term, *i.e.* the second virial coefficient, is positive for $\chi < \frac{1}{2}$, corresponding to repulsive effective two-body interactions and good-solvent conditions; it vanishes at the θ point $\chi = \frac{1}{2}$.

In fact, eqn. 6.343 is only qualitatively correct. In the limit where $\chi \ll \frac{1}{2}$, Flory showed that the individual polymer coils behave much as hard spheres of radius $R_{\rm F}$. The osmotic pressure then satisfies something analogous to a virial equation of state:

$$\begin{aligned} \frac{\Pi}{k_{\rm B} T} &= \frac{\phi}{N v_0} + A \left(\frac{\phi}{N v_0} \right)^2 R_{\rm F}^3 + \ldots \\ &= \frac{\phi}{N v_0}\, h(\phi/\phi^*) \quad . \end{aligned} \tag{6.344}$$

This is generalized to a scaling form in the second line, where $h(x)$ is a scaling function, and the quantity ϕ^* is given by $\phi^* = N v_0 / R_{\rm F}^3 \propto N^{-4/5}$, assuming $d = 3$ and $\nu = \frac{3}{5}$ from Flory theory. As $x = \phi/\phi^* \to 0$, we must recover the ideal gas law, so $h(x) = 1 + \mathcal{O}(x)$ in this limit. For $x \to \infty$, we require that the result be independent of the degree of polymerization N. This means $h(x) \propto x^p$ with $\frac{4}{5}p = 1$, *i.e.* $p = \frac{5}{4}$. The result is known as the des Cloiseaux law:

$$\frac{v_0 \Pi}{k_{\rm B} T} = C\,\phi^{9/4} \quad , \tag{6.345}$$

where C is a constant. This is valid for what is known as semi-dilute solutions, where $\phi^* \ll \phi \ll 1$. In the dense limit $\phi \sim 1$, the results do not exhibit this universality, and we must appeal to liquid state theory, which is no fun at all.

6.7 Lee-Yang Theory

6.7.1 Analytic properties of the partition function

How can statistical mechanics describe phase transitions? This question was addressed in some beautiful mathematical analysis by Lee and Yang[27]. Consider the grand partition function Ξ,

$$\Xi(T,V,z) = \sum_{N=0}^{\infty} z^N \, Q_N(T,V) \, \lambda_T^{-dN} \quad , \tag{6.346}$$

where

$$Q_N(T,V) = \frac{1}{N!} \int d^d x_1 \cdots \int d^d x_N \, e^{-U(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)/k_{\rm B}T} \tag{6.347}$$

is the contribution to the N-particle partition function from the potential energy U (assuming no momentum-dependent potentials). For two-body central potentials, we have

$$U(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \sum_{i<j} u\big(|\boldsymbol{x}_i - \boldsymbol{x}_j|\big) \quad . \tag{6.348}$$

Suppose further that these classical particles have hard cores. Then for any *finite* volume, there must be some maximum number N_V such that $Q_N(T,V)$ vanishes for $N > N_V$. This is because if $N > N_V$ at least two spheres must overlap, in which case the potential energy is infinite. One defines the *sphere packing fraction* $\varphi \equiv \frac{1}{6}\pi a^3 n$, where n is the number density of spheres and a is the sphere diameter. Equivalently, $\varphi = v_0/v$ where v_0 is the sphere volume and $v = n^{-1} = V/N$. The maximum packing density for hard spheres is achieved for a hexagonal close packed (HCP) lattice[28], for which $\varphi_{\rm HCP} = \pi/3\sqrt{2} = 0.74048$, in which case $N_V = \sqrt{2}\,V/a^3$ is the maximum particle number.

Thus, if V itself is finite, then $\Xi(T,V,z)$ is a *finite* degree polynomial in z, and may be factorized as

$$\Xi(T,V,z) = \sum_{N=0}^{N_V} z^N \, Q_N(T,V) \, \lambda_T^{-dN} = \prod_{k=1}^{N_V} \left(1 - \frac{z}{z_k}\right) \quad , \tag{6.349}$$

where $z_k(T,V)$ is one of the N_V zeros of the grand partition function. Note that the $\mathcal{O}(z^0)$ term is fixed to be unity. Note also that since the configuration integrals $Q_N(T,V)$ are all positive, $\Xi(z)$ is an increasing function along the positive real z axis. In addition, since the coefficients of z^N in the polynomial $\Xi(z)$ are all real, then $\Xi(z) = 0$ implies $\overline{\Xi(z)} = \Xi(\bar{z}) = 0$, so the zeros of $\Xi(z)$ are either real and negative or else come in complex conjugate pairs.

For finite N_V, the situation is roughly as depicted in the left panel of fig. 6.24, with a set of N_V zeros arranged in complex conjugate pairs (or negative real values). The zeros aren't necessarily distributed along a circle as shown in the figure, though. They could be anywhere, so long as they are symmetrically distributed about the $\mathrm{Re}(z)$ axis, and no zeros occur for z real and nonnegative.

[27] See C. N. Yang and R. D. Lee, *Phys. Rev.* **87**, 404 (1952) and *ibid*, p. 410

[28] See *e.g.* http://en.wikipedia.org/wiki/Close-packing. For *randomly* close-packed hard spheres, one finds, from numerical simulations, $\varphi_{\rm RCP} = 0.644$.

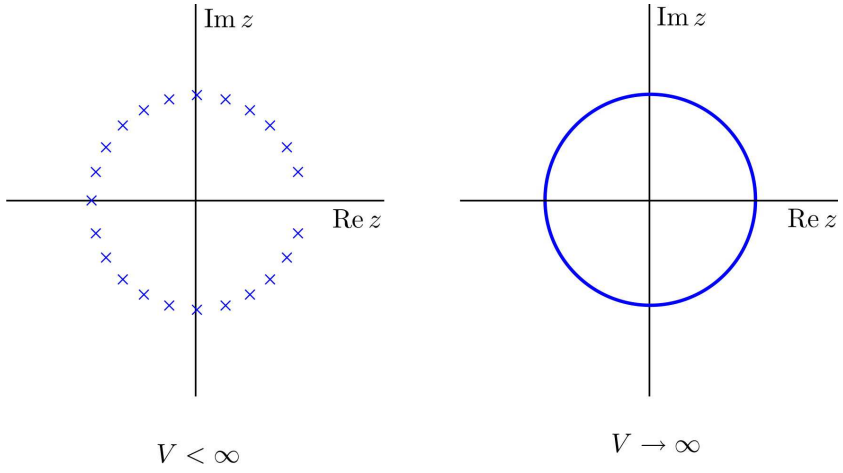


Figure 6.24: In the thermodynamic limit, the grand partition function can develop a singularity at positive real fugacity z. The set of discrete zeros fuses into a branch cut.

Lee and Yang proved the existence of the limits

$$\begin{aligned} \frac{p}{k_{\rm B}T} &= \lim_{V\to\infty} \frac{1}{V}\,\log \Xi(T,V,z) \\ n &= \lim_{V\to\infty} z\,\frac{\partial}{\partial z}\left[\frac{1}{V}\,\log\Xi(T,V,z)\right] \quad , \end{aligned} \tag{6.350}$$

and notably the result

$$n = z\,\frac{\partial}{\partial z}\left(\frac{p}{k_{\rm B}T}\right) \quad , \tag{6.351}$$

which amounts to the commutativity of the thermodynamic limit $V\to\infty$ with the differential operator $z\,\frac{\partial}{\partial z}$. In particular, $p(T,z)$ is a smooth function of z in regions free of roots. If the roots do coalesce and pinch the positive real axis, then the density n can be discontinuous, as in a first order phase transition, or a higher derivative $\partial^j p/\partial z^j$ can be discontinuous or divergent, as in a second order phase transition.

6.7.2 Electrostatic analogy

There is a beautiful analogy to the theory of two-dimensional electrostatics. We write

$$\begin{aligned} \frac{p}{k_{\rm B}T} &= \frac{1}{V}\sum_{k=1}^{N_V}\log\left(1-\frac{z}{z_k}\right) \\ &= -\sum_{k=1}^{N_V}\Big[\phi(z-z_k)-\phi(0-z_k)\Big] \quad , \end{aligned} \tag{6.352}$$

where $\phi(z) = -V^{-1}\log(z)$ is the complex potential due to a line charge of linear density $\lambda = V^{-1}$ located at origin. The number density is then

$$n = z\,\frac{\partial}{\partial z}\left(\frac{p}{k_{\rm B}T}\right) = -z\,\frac{\partial}{\partial z}\sum_{k=1}^{N_V}\phi(z-z_k) \quad , \tag{6.353}$$

to be evaluated for physical values of z, *i.e.* $z \in \mathbb{R}^+$. Since $\phi(z)$ is analytic,

$$\frac{\partial\phi}{\partial\bar{z}} = \frac{1}{2}\,\frac{\partial\phi}{\partial x} + \frac{i}{2}\,\frac{\partial\phi}{\partial y} = 0 \quad . \tag{6.354}$$

If we decompose the complex potential $\phi = \phi_1 + i\phi_2$ into real and imaginary parts, the condition of analyticity is recast as the Cauchy-Riemann equations,

$$\frac{\partial\phi_1}{\partial x} = \frac{\partial\phi_2}{\partial y} \qquad , \qquad \frac{\partial\phi_1}{\partial y} = -\frac{\partial\phi_2}{\partial x} \quad . \tag{6.355}$$

Thus,

$$\begin{aligned} -\frac{\partial\phi}{\partial z} &= -\frac{1}{2}\,\frac{\partial\phi}{\partial x} + \frac{i}{2}\,\frac{\partial\phi}{\partial y} \\ &= -\frac{1}{2}\left(\frac{\partial\phi_1}{\partial x} + \frac{\partial\phi_2}{\partial y}\right) + \frac{i}{2}\left(\frac{\partial\phi_1}{\partial y} - \frac{\partial\phi_2}{\partial x}\right) \\ &= -\frac{\partial\phi_1}{\partial x} + i\,\frac{\partial\phi_1}{\partial y} = E_x - iE_y \quad , \end{aligned} \tag{6.356}$$

where $\boldsymbol{E} = -\boldsymbol{\nabla}\phi_1$ is the electric field. Suppose, then, that as $V \to \infty$ a continuous charge distribution develops, which crosses the positive real z axis at a point $x \in \mathbb{R}^+$. Then

$$\frac{n_+ - n_-}{x} = E_x(x^+) - E_x(x^-) = 2\pi\sigma(x) \quad , \tag{6.357}$$

where σ is the linear charge density (assuming logarithmic two-dimensional potentials), or the two-dimensional charge density (if we extend the distribution along a third axis).

6.7.3 Example

As an example, consider the function

$$\begin{aligned} \Xi(z) &= \frac{(1+z)^M\,(1-z^K)}{1-z} \\ &= (1+z)^M\left(1 + z + z^2 + \ldots + z^{K-1}\right) \quad . \end{aligned} \tag{6.358}$$

The $(M+K-1)$ degree polynomial has an $M^{\rm th}$ order zero at $z=-1$ and $(K-1)$ simple zeros at $z = e^{2\pi ij/K}$, where $j \in \{1,\ldots,K-1\}$. The maximum particle number is the order of $\Xi(z)$, which is

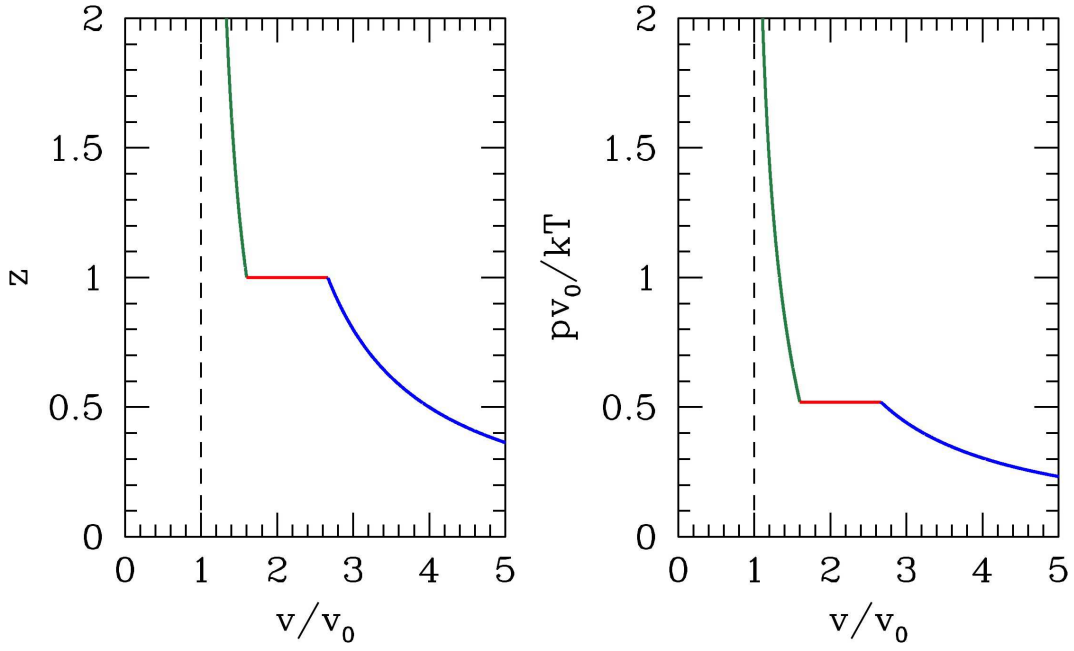


Figure 6.25: Fugacity z and $pv_0/k_{\rm B}T$ *versus* dimensionless specific volume v/v_0 for the example problem discussed in the text. The case $r = 0.75$ is shown.

$N_V = M + K - 1$. We define $v_0 \equiv V/N_V$ and $r \equiv M/N_V \in [0,1]$. The thermodynamic limit is then taken as $N_V \to \infty$ with r finite, in which case $K/N_V = 1 - r$. Thus,

$$\begin{aligned}\frac{pv_0}{k_{\rm B}T} = \lim_{N_V\to\infty}\frac{1}{N_V}\log\Xi(z) &= \lim_{N_V\to\infty}\frac{1}{N_V}\log\Xi(z) \\ &= \lim_{N_V\to\infty}\frac{1}{N_V}\left[M\log(1+z) + \log\left(\frac{1-z^K}{1-z}\right)\right] \quad .\end{aligned} \tag{6.359}$$

The limit depends on whether $|z| > 1$ or $|z| < 1$. For z real and nonnegative we have

$$\frac{pv_0}{k_{\rm B}T} = \begin{cases} r\log(1+z) & \text{if} \quad 0 < z \le 1 \\ r\log(1+z) + (1-r)\log z & \text{if} \quad 1 < z < \infty \end{cases} \quad . \tag{6.360}$$

Thus,

$$\frac{v_0}{v} = nv_0 = z\,\frac{\partial}{\partial z}\left(\frac{pv_0}{k_{\rm B}T}\right) = \begin{cases} \frac{rz}{1+z} & \text{if} \quad 0 < z \le 1 \\ \frac{rz}{1+z} + 1 - r & \text{if} \quad 1 < z < \infty \end{cases} \quad . \tag{6.361}$$

with $v = n^{-1}$. Solving for $z(v)$, we find

$$z = \begin{cases} \frac{v_0 - (1-r)v}{v - v_0} & \text{if} \quad v_0 < v \le \frac{2v_0}{2-r} \\ 1 & \text{if} \quad \frac{2v_0}{2-r} \le v \le \frac{2v_0}{r} \\ \frac{v_0}{rv - v_0} & \text{if} \quad \frac{2v_0}{r} \le v < \infty \end{cases} \quad . \tag{6.362}$$

We then obtain the equation of state,

$$\frac{p v_0}{k_{\rm B} T} = \begin{cases} r \log\left(\frac{rv}{v-v_0}\right) + (1-r)\log\left(\frac{v_0-(1-r)v}{v-v_0}\right) & \text{if} \quad v_0 < v \le \frac{2v_0}{2-r} \\ r \log 2 & \text{if} \quad \frac{2v_0}{2-r} \le v \le \frac{2v_0}{r} \\ r \log\left(\frac{rv}{rv-v_0}\right) & \text{if} \quad \frac{2v_0}{r} \le v < \infty \quad . \end{cases} \tag{6.363}$$

Note that the smallest possible volume is $v_{\rm min} = v_0$. Thus, the specific volume jumps discontinuously from $v_1 = \frac{2v_0}{2-r}$ to $v_2 = \frac{2v_0}{r}$ at a fixed $p/k_{\rm B}T$. This is reminiscent of the liquid-gas transition in §2.11 and in §§7.2,3. Here the interval $v \in [v_1, v_2]$ would correspond to the coexistence region, with $v_{1,2}$ being the specific volumes for the liquid and gas endpoints, respectively.

6.8 Appendix: Potts Model in One Dimension

6.8.1 Hamiltonian

The Potts model is defined by the Hamiltonian

$$H = -J \sum_{\langle ij \rangle} \delta_{\sigma_i, \sigma_j} - h \sum_i \delta_{\sigma_i, 1} \quad . \tag{6.364}$$

Here, the spin variables σ_i take values in the set $\{1, 2, \ldots, q\}$ on each site. The equivalent of an external magnetic field in the Ising case is a field h which prefers a particular value of σ ($\sigma = 1$ in the above Hamiltonian). Once again, it is not possible to compute the partition function on general lattices, however in one dimension we may once again find Z using the transfer matrix method.

6.8.2 Transfer matrix

On a ring of N sites, we have

$$Z = {\rm Tr}\, e^{-\beta H} = \sum_{\{\sigma_n\}} e^{\beta h \delta_{\sigma_1, 1}} \, e^{\beta J \delta_{\sigma_1, \sigma_2}} \, \cdots \, e^{\beta h \delta_{\sigma_N, 1}} \, e^{\beta J \delta_{\sigma_N, \sigma_1}} = {\rm Tr}\left(R^N\right) \quad , \tag{6.365}$$

where the $q \times q$ transfer matrix R is given by

$$R_{\sigma\sigma'} = e^{\beta J \delta_{\sigma\sigma'}} \, e^{\frac{1}{2}\beta h \delta_{\sigma,1}} \, e^{\frac{1}{2}\beta h \delta_{\sigma',1}} = \begin{cases} e^{\beta(J+h)} & \text{if } \sigma = \sigma' = 1 \\ e^{\beta J} & \text{if } \sigma = \sigma' \neq 1 \\ e^{\beta h/2} & \text{if } \sigma = 1 \text{ and } \sigma' \neq 1 \\ e^{\beta h/2} & \text{if } \sigma \neq 1 \text{ and } \sigma' = 1 \\ 1 & \text{if } \sigma \neq 1 \text{ and } \sigma' \neq 1 \text{ and } \sigma \neq \sigma' \quad . \end{cases} \tag{6.366}$$

In matrix form,

$$R = \begin{pmatrix} e^{\beta(J+h)} & e^{\beta h/2} & e^{\beta h/2} & \cdots & & e^{\beta h/2} \\ e^{\beta h/2} & e^{\beta J} & 1 & \cdots & & 1 \\ e^{\beta h/2} & 1 & e^{\beta J} & \cdots & & 1 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ e^{\beta h/2} & 1 & 1 & \cdots & e^{\beta J} & 1 \\ e^{\beta h/2} & 1 & 1 & \cdots & 1 & e^{\beta J} \end{pmatrix} \tag{6.367}$$

The matrix R has q eigenvalues λ_j, with $j = 1, \ldots, q$. The partition function for the Potts chain is then

$$Z = \sum_{j=1}^{q} \lambda_j^N \quad . \tag{6.368}$$

We can actually find the eigenvalues of R analytically. To this end, consider the vectors

$$\phi = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad , \qquad \psi = \big(q - 1 + e^{\beta h}\big)^{-1/2} \begin{pmatrix} e^{\beta h/2} \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad . \tag{6.369}$$

Then R may be written as

$$R = \big(e^{\beta J} - 1\big)\, \mathbb{I} + \big(q - 1 + e^{\beta h}\big)\, |\, \psi \,\rangle\langle\, \psi \,| + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\, |\, \phi \,\rangle\langle\, \phi \,| \quad , \tag{6.370}$$

where $\mathbb{I}$ is the $q \times q$ identity matrix. When $h = 0$, we have a simpler form,

$$R = \big(e^{\beta J} - 1\big)\, \mathbb{I} + q\, |\, \psi \,\rangle\langle\, \psi \,| \quad . \tag{6.371}$$

From this we can read off the eigenvalues:

$$\begin{aligned} \lambda_1 &= e^{\beta J} + q - 1 \\ \lambda_j &= e^{\beta J} - 1 \quad , \quad j \in \{2, \ldots, q\} \quad , \chi \end{aligned} \tag{6.372}$$

since $|\, \psi \,\rangle$ is an eigenvector with eigenvalue $\lambda = e^{\beta J} + q - 1$, and any vector orthogonal to $|\, \psi \,\rangle$ has eigenvalue $\lambda = e^{\beta J} - 1$. The partition function is then

$$Z = \big(e^{\beta J} + q - 1\big)^N + (q-1)\big(e^{\beta J} - 1\big)^N \quad . \tag{6.373}$$

In the thermodynamic limit $N \to \infty$, only the λ_1 eigenvalue contributes, and we have

$$F(T, N, h = 0) = -N k_{\rm B} T \log\big(e^{J/k_{\rm B}T} + q - 1\big) \qquad \text{for } N \to \infty \quad . \tag{6.374}$$

When h is nonzero, the calculation becomes somewhat more tedious, but still relatively easy. The problem is that $|\, \psi \,\rangle$ and $|\, \phi \,\rangle$ are not orthogonal, so we define

$$|\, \chi \,\rangle = \frac{|\, \phi \,\rangle - |\, \psi \,\rangle\langle\, \psi \,|\, \phi \,\rangle}{\sqrt{1 - \langle\, \phi \,|\, \psi \,\rangle^2}} \quad , \tag{6.375}$$

where

$$x \equiv \langle\, \phi \,|\, \psi \,\rangle = \left(\frac{e^{\beta h}}{q - 1 + e^{\beta h}} \right)^{1/2} \quad . \tag{6.376}$$

Now we have $\langle\, \chi \,|\, \psi \,\rangle = 0$, with $\langle\, \chi \,|\, \chi \,\rangle = 1$ and $\langle\, \psi \,|\, \psi \,\rangle = 1$, with

$$|\, \phi \,\rangle = \sqrt{1 - x^2}\; |\, \chi \,\rangle + x \,|\, \psi \,\rangle \quad . \tag{6.377}$$

and the transfer matrix is then

$$\begin{aligned} R &= \big(e^{\beta J} - 1\big)\, \mathbb{I} + \big(q - 1 + e^{\beta h}\big)\, |\, \psi \,\rangle\langle\, \psi \,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \left[(1 - x^2)\, |\, \chi \,\rangle\langle\, \chi \,| + x^2\, |\, \psi \,\rangle\langle\, \psi \,| + x \sqrt{1 - x^2}\, \Big(|\, \chi \,\rangle\langle\, \psi \,| + |\, \psi \,\rangle\langle\, \chi \,| \Big) \right] \\ &= \big(e^{\beta J} - 1\big)\, \mathbb{I} + \left[\big(q - 1 + e^{\beta h}\big) + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \left(\frac{e^{\beta h}}{q - 1 + e^{\beta h}} \right) \right] |\, \psi \,\rangle\langle\, \psi \,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \left(\frac{q - 1}{q - 1 + e^{\beta h}} \right) |\, \chi \,\rangle\langle\, \chi \,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \left(\frac{(q-1)\, e^{\beta h}}{q - 1 + e^{\beta h}} \right)^{1/2} \Big(|\, \chi \,\rangle\langle\, \psi \,| + |\, \psi \,\rangle\langle\, \chi \,| \Big) \quad , \end{aligned} \tag{6.378}$$

which in the two-dimensional subspace spanned by $|\, \chi \,\rangle$ and $|\, \psi \,\rangle$ is of the form

$$R = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \quad . \tag{6.379}$$

Recall that for any 2×2 Hermitian matrix,

$$M = a_0\, \mathbb{I} + \boldsymbol{a} \cdot \boldsymbol{\tau} = \begin{pmatrix} a_0 + a_3 & a_1 - i a_2 \\ a_1 + i a_2 & a_0 - a_3 \end{pmatrix} \quad , \tag{6.380}$$

the characteristic polynomial is

$$P(\lambda) = \det\big(\lambda\, \mathbb{I} - M\big) = (\lambda - a_0)^2 - a_1^2 - a_2^2 - a_3^2 \quad , \tag{6.381}$$

and hence the eigenvalues are

$$\lambda_\pm = a_0 \pm \sqrt{a_1^2 + a_2^2 + a_3^2} \quad . \tag{6.382}$$

For the transfer matrix of eqn. 6.378, we obtain, after a little work,

$$\begin{aligned} \lambda_{1,2} &= e^{\beta J} - 1 + \tfrac{1}{2}\Big[q - 1 + e^{\beta h} + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \Big] \\ &\quad \pm \tfrac{1}{2} \sqrt{ \Big[q - 1 + e^{\beta h} + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) \Big]^2 - 4(q-1)\big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big) } \quad . \end{aligned} \tag{6.383}$$

There are $q-2$ other eigenvalues, however, associated with the $(q-2)$-dimensional subspace orthogonal to $|\, \chi \,\rangle$ and $|\, \psi \,\rangle$. Clearly all these eigenvalues are given by

$$\lambda_j = e^{\beta J} - 1 \qquad , \quad j \in \{3\,, \ldots\,, q\} \quad . \tag{6.384}$$

The partition function is then

$$Z = \lambda_1^N + \lambda_2^N + (q-2)\,\lambda_3^N \quad , \tag{6.385}$$

and in the thermodynamic limit $N \to \infty$ the maximum eigenvalue λ_1 dominates. Note that we recover the correct limit as $h \to 0$.