LECTURE NOTES ON STATISTICAL MECHANICS

Scott Pratt

Department of Physics and Astronomy
Michigan State University
PHY 831 – 2007-2016

FORWARD

These notes are for the one-semester graduate level statistical mechanics class taught at Michigan State University. Although they are more terse than a typical text book, they do cover all the material used in PHY 831. The notes presume a familiarity with basic undergraduate concepts in statistical mechanics, and with some basic concepts from first-year graduate quantum, such as harmonic oscillators and raising and lowering operators. Some of the material in Chapter 3 involves time-dependent perturbation theory, which is described in the notes here, but the terse manner would undoubtedly intimidate students who have never used it before.

Anybody is welcome to use the notes to their heart's content, though the text should be treated with the usual academic respect when it comes to copying material. If anyone is interested in the LaTeX source files, they should contact me (prattsc@msu.edu). Solutions to the end-of-chapter problems are also provided on the course web site http://www.pa.msu.edu/ pratts/phy831. Please beware that this is a web manuscript, and is thus alive and subject to change at any time.

The quality, such as it is, would have been FAR lower had it not been for the generous and diligent assistance with which students provided feedback.

Contents

1   Foundation of Statistical Physics

"I know nothing ... nothing" - John Banner

1.1   Ignorance, Entropy and the Ergodic Theorem

Statistical physics is a beautiful subject. Pretty much everything derives from the simple statement that entropy is maximized. Here, we describe the meaning of entropy, and show how the tenet of maximum entropy is related to time-reversal via the ergodic theorem.

Consider a large number of systems $N_s \to \infty$, each of which can be in some specific quantum state. Let n_i be the number of systems that are in the state i. We will define the *ignorance I* as a measure of the number of ways to arrange the systems given $n_0, n_1 \cdots$.

$$I = \frac{N_s!}{n_0! n_1! \cdots}, \tag{1.1}$$

with the constraint that $n_0 + n_1 + \cdots = N_s$. Our immediate goal is to find n_i that maximizes ignorance while satisfying the constraint. If the observer knows nothing about the population probabilities, the values n_i should be chosen to maximize I. However, before doing so, we will define S as:

$$S \equiv \frac{1}{N_s} \ln(I), \tag{1.2}$$

which will be maximized when I is maximized. By defining it as the log of the ignorance, S will have the convenient property that the entropy of two independent systems will be the sum of the entropy of the two systems, because the combined ignorance of two systems would be the product of the two ignorances, $I = I_1 I_2$, and the log of a product is the sum. Dividing by N_s then makes S independent of the number of systems in the large N_s limit. The quantity S is the *entropy*, the most fundamental quantity of statistical mechanics. Because it was divided by the number of systems, one can consider S to be the entropy of an individual system. Using Stirling's expansion,

$$\lim_{N \to \infty} \ln N! = N \ln N - N + (1/2) \ln N + (1/2) \ln(2\pi) + 1/(12N) + \cdots, \tag{1.3}$$

we keep the first two terms to see that

$$\begin{aligned} S &= \frac{1}{N_s} \left(N_s \ln N_s - \sum_i n_i \ln n_i - N_s + \sum_i n_i + \cdots \right) \\ &= -\sum_i p_i \ln p_i \quad \text{as } N_s \to \infty, \end{aligned} \tag{1.4}$$

where $p_i \equiv n_i/N_s$ is the probability a given system is in state i. As $N_s \to \infty$, all terms beyond the $p_i \ln p_i$ term in Eq. (1.4) vanish. Note that if all the probability is confined to one state, the entropy will be zero. Furthermore, because for each probability, $0 < p_i \le 1$, the entropy is always positive.

Our goal is to maximize S. Maximizing a multi-dimensional function (in this case a function of $n_0, n_1 \cdots$) with a constraint is often done with Lagrange multipliers. In that case, one maximizes the quantity, $S - \lambda C(\vec{n})$, with respect to all variables and with respect to λ. Here, the constraint C must be some function of the variables constrained to zero, in our case $C = \sum_i p_i - 1$.

The coefficient λ is called the Lagrange multiplier. Stating the minimization,

$$\frac{\partial}{\partial p_i}\left(-\sum_j p_j \ln p_j - \lambda\left[\sum_j p_j - 1\right]\right) \;=\; 0, \tag{1.5}$$

$$\frac{\partial}{\partial \lambda}\left(-\sum_j p_j \ln p_j - \lambda\left[\sum_j p_j - 1\right]\right) \;=\; 0.$$

The second expression leads directly to the constraint $\sum_j p_j = 1$, while the first expression leads to the following value for p_i,

$$\ln p_i = -\lambda - 1, \;\; \text{or} \;\; p_i = e^{-\lambda-1}. \tag{1.6}$$

The parameter λ is then chosen to normalize the distribution, $e^{-\lambda-1}$ multiplied by the number of states is unity. The important result here is that all states are equally probable. This is the result of stating that you know nothing about which states are populated, i.e., maximizing ignorance is equivalent to stating that all states are equally populated. This can be considered as a fundamental principle – *Disorder (or entropy) is maximized*. **All statistical mechanics derives from this principle**.

ASIDE: REVIEW OF LAGRANGE MULTIPLIERS

Imagine a function $F(x_1 \cdots x_n)$ which one minimizes while satisfying a constraint $C(x_1 \cdots x_n) = 0$. If one considers a small step, starting from a point where the constraint is satisfied, the change in C is

$$\delta C = \nabla C \cdot \delta \vec{x}.$$

The function C will not change for small steps where $\delta \vec{x}$ is perpendicular to ∇C, but will change for steps parallel to ∇C. The gradient of F must vanish for any step normal to ∇C for F to be a minimum on the hyper-surface where C is a constant. Therefore, ∇F must be parallel to ∇C. The constant of proportionality is the Lagrange multiplier λ,

$$\nabla F = \lambda \nabla C.$$

The two gradients are parallel if,

$$\nabla(F - \lambda C) = 0.$$

However, this condition (actually n conditions) on its own is not completely sufficient. If one fixes λ to an arbitrary value, then solves for $\vec{x}$ by solving the parallel-gradients constraint, one will find a solution to the minimization constraint with $C(\vec{x}) = \text{some constant}$, but not zero. Fixing $C = 0$ can be accomplished by additionally requiring the condition,

$$\frac{\partial}{\partial \lambda}(F - \lambda C) = 0.$$

Putting these together, it appears that $F - \lambda C$ is a function of $x_1 \cdots x_n$ and λ, and that this must then be minimized. Thus, the n-dimensional minimization problem with a constraint is translated into an $(n + 1)$-dimensional problem, with $n + 1$ variables and $n + 1$ conditions. Here, λ basically appears as the extra variable. It is counter-intuitive to consider "reducing" an

n dimensional minimization problem to a $n + 1$ dimensional problem as progress. However, in this form the first n conditions often become rather trivial to solve in terms of λ. One is then left with one unknown λ, though that one unknown may be difficult to determine.

1.2   The Ergodic Theorem

The principle of maximizing entropy is related to the Ergodic theorem, which provides the way to understand why all states are equally populated from the perspective of dynamics. The Ergodic theorem is built on the symmetry of time-reversal, i.e., the rate at which one changes from state i to state j is the same as the rate at which one changes from state j to state i. Here, we can consider a state as a quantum eigenstate of the entire system. All static systems, no matter how large or how many particles, have eigenstates, even if they are extremely tightly spaced. If a state is particularly difficult to enter, it is equivalently difficult to exit. Thus, a time average of a given system will cycle through all states and, if one waits long enough, the system will spend equal amounts of net time in each state.

This can also be understood by considering an infinite number of systems where each state is populated with equal probability. For every pair of states i and j, if the rate transitioning from j to i is equal to the rate for transitioning from i to j, it is clear the probability distribution will stay uniform. Thus, a uniform probability distribution is stable, if one assumes time reversal. Here "time reversal" simply means that inverse rates are equal.

Satisfaction of time reversal is sometimes rather subtle. As an example, consider two large identical rooms, a left room and a right room, separated by a door manned by a security guard. If the rooms are populated by 1000 randomly oscillating patrons, and if the security guard grants and denies access with equal probability when going right-to-left vs. left-to-right, the population of the two rooms will, on average, be equal. However, if the security guard denies access to the left room while granting exit of the left room, the population will ultimately skew towards the right room. This explicit violation of the principle of maximized entropy derives from the fact that moving left-to-right and right-to-left, i.e. the time reversed motions, are not treated equivalently.

The same security guard could, in principle, police the traversal of gas molecules between two partitions of a box. Such paradoxes were discussed by Maxwell, and the security guard is referred to as Maxwell's demon. As described by Maxwell,

> ... if we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. For we have seen that molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the

temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

This apparent violation of the second law of thermodynamics was explained by Leó Szilárd in 1929, who argued that the demon would have to expend energy to measure the speed of the molecules, and thus increase entropy somewhere, perhaps in his brain, thus ensuring that the entropy of the entire system (gas + demon) increased. Check out
http://en.wikipedia.org/wiki/Maxwell's_demon.

Previously, we showed that if all one maximizes entropy all states will be equally probable. If one accepts the converse, that if all states are equally probable one maximizes entropy, then the Ergodic theorem demonstrates how time reversal leads to the conclusion that entropy should be maximized.

1.3 Statistical Ensembles

The previous section discussed the manifestations of maximizing ignorance, or equivalently entropy, without regard to any constraints aside from the normalization constraint that probabilities sum to unity. In this section, we discuss the effects of fixing energy and/or particle number or charge. These additional constraints can be easily incorporated by applying additional Lagrange multipliers. For instance, conserving the average energy can be enforced by adding an extra Lagrange multiplier β. Maximizing the entropy per system with respect to the probability p_i for being in state i,

$$\frac{\partial}{\partial p_i} \left(-\sum_j p_j \ln p_j - \lambda[\sum_j p_j - 1] - \beta[\sum_j p_j \epsilon_j - \bar{E}] \right) = 0, \tag{1.7}$$

gives

$$p_i = \exp(-1 - \lambda - \beta \epsilon_i). \tag{1.8}$$

Thus, the states are populated proportional to the factor $e^{-\beta \epsilon_i}$, which is the Boltzmann distribution, with β being identified as the inverse temperature. Again, the parameter λ is chosen to normalize the probability. However, a given β only enforces the constraint that the average energy is some constant, not the particular energy one might wish. Thus, one must adjust β to find the desired energy, a sometimes time-consuming process.

For any quantity which is conserved on average, one need only add a corresponding Lagrange multiplier. For instance, a multiplier α could be used to restrict the average particle number or charge. The probability for being in state i would then be:

$$p_i = \exp(-1 - \lambda - \beta \epsilon_i - \alpha Q_i). \tag{1.9}$$

Typically, the *chemical potential* μ is used to reference the multiplier,

$$\alpha = -\mu/T. \tag{1.10}$$

The charge Q_i could refer to the baryon number, electric charge, or any other conserved quantity. It could be either positive or negative. If there are many conserved charges, Q can be replaced by $\vec{Q}$ and μ can be replaced by $\vec{\mu}$.

Rather than enforcing the last Lagrange multiplier constraint, that derivatives w.r.t. the multiplier are zero, we are often happy with knowing the solution for a given temperature and chemical potential. Inverting the relation to find values of T and μ that yield specific values of the energy and particle number is often difficult, and usually it is the temperature and chemical potentials that are effectively fixed in many physical situations. In the coming subsections we show how this definition of temperature and chemical potential as Lagrange multipliers are equivalent to the familiar thermodynamic definitions.

In most textbooks, the charge is replaced by a number N. This is fine if the number of particles is conserved such as a gas of Argon atoms. However, the system could include complicated chemical reactions, or multiple conserved charges. For instance, both electric charge and baryon number are conserved in the hadronic medium in the interior of a star. The charges Q_i and the associated Lagrange multipliers, α_i, might then be thought of as components of a vector. Positrons and electrons clearly contribute to the charge with opposite signs.

In the early universe there were nearly equal numbers of particles and anti-particles. This leads to the chemical potentials vanishing, but the average number of positrons, or of electrons, would be large. If the system is strongly interacting, counting the number of electrons or positrons can be problematic, because states can be composed of coherent combinations of electrons and positrons. For example, a photon has e^+e^- contributions to its composition once it becomes off-shell. Thus counting electrons becomes a poorly defined enterprise, but the net charge, i.e. the number of positive quanta minus the number of negative quanta, is conserved and is indeed well defined. One should only associate Lagrange multipliers with conserved quantities, with the definition of "conservation" depending on the relevant time scale. For example a low-temperature gas of lead atoms can be treated as if the number of lead atoms is conserved. However if one waits for time much greater than the age of the universe, those lead atoms will undergo radioactive decay. But for observations at reasonable time scales the number of lead atoms can be considered a conserved quantity.

One could imagine a system with different molecules undergoing chemical reactions. Each molecule would have a different number of atoms of a given type. The number of each type of atom would represent an independently observed charge, but the number of molecules would not be conserved. Without apology, these notes will switch between using N or Q depending on the context.

In statistical mechanics one first considers which quantities one wishes to fix, and which quantities one wishes to allow to vary (but with the constraint that the average is some value). This choice defines the *ensemble*. The three most common ensembles are the micro-canonical, canonical and grand-canonical. The ensembles differ by which quantities vary, as seen in Table 1. If a quantity is allowed to vary, then a Lagrange multiplier determines the average quantity for that ensemble.

Ensemble	Energy	Charges
micro-canonical	fixed	fixed
canonical	varies	fixed
grand canonical	varies	varies

Table 1: Ensembles vary by what quantity is fixed and what varies.

1.4   Partition Functions

Whereas maximizing entropy represents the central concept of statistical physics, partition functions provide the principal calculational workhorse.

Because probabilities are proportional to $e^{-\beta\epsilon_i - \alpha Q_i}$, the normalization requirement can be absorbed by a factor $1/Z$:

$$
\begin{aligned}
p_i &= \frac{1}{Z} e^{-\beta\epsilon_i - \alpha Q_i}, \\
Z &= \sum_i e^{-\beta\epsilon_i - \alpha Q_i},
\end{aligned}
\tag{1.11}
$$

where Z is referred to as the partition function. Partition functions are convenient for calculating the average energy or charge,

$$
\begin{aligned}
\langle E \rangle &= \frac{\sum_i \epsilon_i e^{-\beta\epsilon_i - \alpha Q_i}}{Z} \tag{1.12} \\
&= -\frac{\partial}{\partial\beta} \ln Z, \\
&= T^2 \frac{\partial}{\partial T} \ln Z \quad (\text{fixed } \alpha = -\mu/T), \\
\langle Q \rangle &= \frac{\sum_i Q_i e^{-\beta\epsilon_i - \alpha Q_i}}{Z} \tag{1.13} \\
&= -\frac{\partial}{\partial\alpha} \ln Z, \\
&= T \frac{\partial}{\partial\mu} \ln Z \quad (\text{fixed } T).
\end{aligned}
$$

The partition function can also be related to the entropy, but in a way that varies with the choice of ensemble. In the grand canonical ensemble,

$$
\begin{aligned}
S &= -\sum_i p_i \ln p_i = \sum_i p_i(\ln Z + \beta\epsilon_i + \alpha Q_i) \tag{1.14} \\
&= \ln Z + \beta\langle E \rangle + \alpha\langle Q \rangle,
\end{aligned}
$$

which is derived using the normalization condition, $\sum p_i = 1$, along with Eq. (1.11).

Example 1.1:

Consider a 3-level system with energies $-\epsilon, 0$ and ϵ. As a function of T find:
(a) the partition function, (b) the average energy, (c) the entropy
The partition function is:

$$
Z = e^{\epsilon/T} + 1 + e^{-\epsilon/T},
\tag{1.15}
$$

and the average energy is:

$$
\langle E \rangle = \frac{\sum_i \epsilon_i e^{-\epsilon_i/T}}{Z} = \epsilon \frac{-e^{\epsilon/T} + e^{-\epsilon/T}}{Z} .
\tag{1.16}
$$

At $T = 0$, $\langle E \rangle = -\epsilon$, and at $T = \infty$, $\langle E \rangle = 0$.
The entropy is given by:

$$S = \ln Z + \langle E \rangle / T = \ln(e^{\epsilon/T} + 1 + e^{-\epsilon/T}) + \frac{\epsilon}{T} \frac{-e^{\epsilon/T} + e^{-\epsilon/T}}{e^{\epsilon/T} + 1 + e^{-\epsilon/T}}, \tag{1.17}$$

and is zero for $T = 0$ and becomes $\ln 3$ for $T = \infty$. The infinite temperature limit can be understood by remembering that, in the limit that states are equally populated, the entropy is the logarithm of the number of states.

Example 1.2:

Consider two single-particle levels whose energies are $-\epsilon$ and ϵ. Into these levels, we place two electrons (no more than one electron of the same spin per level). As a function of T find:
(a) the partition function, (b) the average energy, (c) the entropy.

First, we enumerate the system states, where "system state" refers to a configuration of the entire system. We can have one state with both electrons in the lower level, one state where both electrons are in the higher level, and four states with one electron in the lower level and one in the higher level. These four states differ based on whether the electrons are in the ↑↑, ↑↓, ↓↑, or ↓↓ configurations. The partition function is:

$$Z = e^{2\epsilon/T} + 4 + e^{-2\epsilon/T}, \tag{1.18}$$

and the average energy is:

$$\langle E \rangle = \frac{\sum_i \epsilon_i e^{-\epsilon_i/T}}{Z} = 2\epsilon \frac{-e^{2\epsilon/T} + e^{-2\epsilon/T}}{Z}. \tag{1.19}$$

At $T = 0$, $\langle E \rangle = -2\epsilon$, and at $T = \infty$, $\langle E \rangle = 0$.
The entropy is given by:

$$S = \ln Z + \langle E \rangle / T = \ln(e^{2\epsilon/T} + 4 + e^{-2\epsilon/T}) + \frac{2\epsilon}{T} \frac{-e^{2\epsilon/T} + e^{-2\epsilon/T}}{e^{2\epsilon/T} + 4 + e^{-2\epsilon/T}}, \tag{1.20}$$

and is zero for $T = 0$ and becomes $\ln 6$ for $T = \infty$.

With problems such as these, you need to very carefully differentiate between single-particle levels and system levels, as only the latter appear in the sum for the partition function.

1.5   Thermal Traces

The examples thus far are based on the assumption that one knows the eigenstates of the Hamiltonian and the Charge operators. Each state in the sum had an energy E_i and a fixed charge q_i. The average of some operator, which in quantum mechanics would be expressed as a matrix, is

$$\langle A \rangle = \frac{1}{Z} \sum_i A_{ii} e^{-\beta E_i - \alpha q_i}, \tag{1.21}$$

$$Z = \sum_i e^{-\beta E_i - \alpha q_i}.$$

However, there are times when the eigenstates of the Hamiltonian H and the charge operator Q are not known. In that case we write $H - \mu Q$ as an operator and

$$\langle A \rangle \; = \; \frac{1}{Z} \mathrm{Tr} \; A e^{-\beta(H-\mu Q)}, \tag{1.22}$$

$$Z \; = \; \mathrm{Tr} \; e^{-\beta(H-\mu Q)}.$$

The trace of an operator $\mathcal{O}$,

$$\mathrm{Tr} \; \mathcal{O} \; = \; \sum_a \mathcal{O}_{aa}, \tag{1.23}$$

is unchanged by unitary transformations,

$$\begin{aligned}
\mathrm{Tr} \; \mathcal{O} \; &= \; \sum_{abcde} U_{ab}^{-1} U_{bc} \mathcal{O}_{cd} U_{de}^{-1} U_{ea} \tag{1.24} \\
&= \; \sum_{bcde} U_{bc} \mathcal{O}_{cd} U_{de}^{-1} \sum_a U_{ea} U_{ab}^{-1} \\
&= \; \mathrm{Tr} \; U \mathcal{O} U^{-1} = \mathrm{Tr} \; \mathcal{O}',
\end{aligned}$$

where $\mathcal{O}'$ is simply $\mathcal{O}$ in the new basis, and in our case $\mathcal{O} = A e^{-\beta(H-\mu Q)}$. Because $\mathcal{O}$ is Hermitian, there is some primed basis where it is diagonalized and the partition function is clearly equal to the trace. Then, because the trace can be performed in any basis, as long as $H - \mu Q$ is expressed in the new basis, the partition function can be written as a trace, independent of basis. This does not imply that taking into account non-diagonal elements of $H - \mu Q$ is easy! In the next chapter, we consider non-interacting particles and diagonalization is relatively trivial because eigenstates are products of momentum states. For interactions, one introduces non-diagonal terms into $H - \mu Q$ and accounting for these terms is usually extremely difficult and requires some strategy such as perturbation theory or lattice gauge theory. The array of such approaches are generally referred to as "many-body theory".

1.6   Thermodynamic Potentials and Free Energies

Because logs of the partition functions are useful quantities, they are often directly referred to as thermodynamic potentials or as free energies. For the case where both the charge and energy are allowed to vary, the potential is referred to as the *grand canonical* potential,

$$\Omega_{GC} \equiv -T \ln Z_{GC} = \langle E \rangle - \mu \langle Q \rangle - TS, \tag{1.25}$$

which is merely a restating of Eq. (1.14). Here, the subscript GC denotes that it is the grand canonical potential, sometimes called the "grand potential". This differs from the *canonical* case, where the energy varies but the charge is fixed (only states of the same charge are considered). For that case there is no $\mu Q/T$ term in the argument of the exponentials used for calculating the partition funciton. For the canonical case, the analog of the potential is the Helmholtz free energy,

$$F \equiv -T \ln Z_C = \langle E \rangle - TS. \tag{1.26}$$

For the grand potential and the Helmholtz free energy, the signs and temperature factors are unfortunate as the definitions are historical in nature. Additionally, one might consider the *microcanonical* case where energy is strictly conserved. In this case, one only considers the entropy.

The choice of canonical vs. grand canonical vs. micro-canonical ensembles depends on the specific problem. Micro-canonical treatments tend to be more difficult and are only necessary for small systems where the fluctuations of the energy in a heat bath would be of the same order as the average energy. Microcanonical treatments are sometimes used for modestly excited atomic nuclei. Canonical treatments are useful when the particle number is small, so that fluctuations of the particle number are similar to the number. Highly excited light nuclei are often treated with canonical approaches, allowing the energy to fluctuate, but enforcing the absolute conservation of particle number. Usually, calculations are simplest in the grand canonical ensemble. In principle, the grand canonical ensemble should only be used for a system in contact with both a heat bath and a particle bath. However, in practice it becomes well justified for a system of a hundred particles or more, where the excitation energy exceeds several times the temperature. Heavier nuclei can often be justifiably treated from the perspective of the grand canonical ensemble.

The various ensembles can be related to one another. For instance, if one calculates the entropy in the microcanonical ensemble, $S(E, Q)$, one can derive the Helmholtz free energy $F(T, Q)$, by first calculating the density of states $\rho(E)$. If the microcanonical entropy was calculated for states within a range δE, the density of states would be expressed as

$$\rho(Q, E) = \frac{e^{S(Q,E)}}{\delta E}, \tag{1.27}$$

from which one could calculate the Helmholtz free energy,

$$F(Q, T) = -T \ln Z_C = -T \ln \int dE \, \rho(Q, E) e^{-E/T}. \tag{1.28}$$

Finally, the grand canonical potential can be generated from F,

$$\Omega_{GC}(\mu/T, T) = T \ln Z_{GC} = T \ln \left\{ \sum_Q e^{-F(Q,T)/T} e^{\mu Q/T} \right\}. \tag{1.29}$$

It is tempting to relate the various ensembles through the entropy. For instance, one could calculate $\langle E \rangle$ and $\langle Q \rangle$ in a grand canonical ensemble from $\ln Z_{GC}$ as a function of μ and T, then calculate the entropy from Eq. (1.25). One could then use the value of S to generate the F in Eq. (1.26). However, the entropies in Eq.s (1.25) and (1.26) are not identical for small systems, but do become identical for large systems. This difference derives from the fact that more states are available if the charge or energy is allowed to fluctuate.

In addition to charge and energy, thermodynamic quantities can also depend on the volume V of the system, which is assumed to be fixed for all three of the ensembles discussed above. The dependence of the thermodynamic potentials on the volume defines the pressure. In the grand canonical ensemble, the pressure is

$$P(\mu, T, V) \equiv -\frac{\partial}{\partial V} \Omega_{GC}(\mu/T, T, V). \tag{1.30}$$

For large systems, those where the dimensions are much larger than the range of any interaction or correlations, the potential Ω_{GC} is proportional to the volume and $P = -\Omega_{GC}/V$. In this case,

$$PV = T \ln Z_{GC} = TS - E + \mu Q, \tag{1.31}$$

and the pressure plays the role of a thermodynamic potential for the grand canonical ensemble. However, one should remember that this is only true for the bulk limit with no long-range interactions or correlations.

1.7   Thermodynamic Relations

It is straightforward to derive thermodynamic relations involving expressions for the entropy in the grand canonical ensemble in Eq. (1.25) and the definition of the pressure, Eq. (1.31),

$$
\begin{aligned}
S &= \ln Z_{GC}(\mu/T, T, V) + E/T - (\mu/T)Q, \\
dS &= \left(\frac{\partial \ln Z_{GC}}{\partial T} - \frac{\langle E \rangle}{T^2} \right) dT + \left(\frac{\partial \ln Z_{GC}}{\partial(\mu/T)} - \langle Q \rangle \right) d(\mu/T) + \frac{P}{T}dV + \frac{dE}{T} - \frac{\mu dQ}{T}, \\
TdS &= PdV + dE - \mu dQ.
\end{aligned}
$$

Here, Eq.s (1.12) and (1.13) were used to eliminate the first two terms in the second line. This is probably the most famous thermodynamic relation. For one thing, it demonstrates that if two systems a and b are in contact and can trade energy or charge, or if one can expand at the other's expense, i.e.,

$$dE_a = -dE_b, \quad dQ_a = -dQ_b, \quad dV_a = -dV_b, \tag{1.32}$$

that entropy can be increased by the transfer of energy, particle number or volume, if the temperatures, chemical potentials or pressures are not identical.

$$dS_a + dS_b = \left(\frac{P_a}{T_a} - \frac{P_b}{T_b} \right) dV_a + \left(\frac{1}{T_a} - \frac{1}{T_b} \right) dE_a - \left(\frac{\mu_a}{T_a} - \frac{\mu_b}{T_b} \right) dQ_a. \tag{1.33}$$

If not blocked, such transfers will take place until the temperatures, pressures and chemical potentials are equal in the two systems. At this point equilibrium is attained.

From the expression for the energy in terms of partition functions,

$$\langle E \rangle = \frac{\sum_i \epsilon_i e^{-\epsilon_i/T + \mu_i Q_i/T}}{\sum_i e^{-\epsilon_i/T + \mu_i Q_i/T}}, \tag{1.34}$$

one can see that energy rises monotonically with temperature because higher temperatures give higher relative weights to states with higher energy. Thus, as energy moves from a higher temperature system to a lower temperature system, the temperatures will approach one another. This is no less than the second law of thermodynamics, i.e., stating that heat only moves spontaneously from hot to cold is equivalent to stating that entropy must increase. Similar arguments can be made for the charge and pressure, as charge will spontaneously move to regions with lower chemical potentials, and higher pressure regions will expand into lower density regions.

Equation (1.32) also shows why the definition of pressure as $\partial_V(\ln T Z_{GC})$ indeed agrees with the usual definition of pressure, $dE = -PdV$ at fixed entropy.

1.8 What to Minimize... What to Maximize

For many problems in physics, there are parameters which adjust themselves according to thermodynamic considerations such as maximizing the entropy. For instance, in a neutron star the nuclear mean field might adjust itself to minimize the free energy, or a liquid might adjust its chemical concentrations. Here, we will refer to such a parameter as x, and discuss how thermodynamics can be used to determine x.

For a system at fixed energy, one would plot the entropy as a function of x, $S(x, E, V, Q)$ and find the value for x at which the entropy would be maximized. If the system is connected to a heat bath, where energy was readily exchanged, one would also have to consider the change of the entropy in the heat bath. For a system a,

$$dS_{\text{tot}} = dS_a + dS_{\text{bath}} = dS_a - dE_a/T, \tag{1.35}$$

where the term dE_a/T describes the entropy change of the heat bath. The total entropy would be maximized when dS_{tot} would be zero for changes dx. Stating that $dS_{\text{tot}} = 0$ is equivalent to stating that $dF_a = 0$ at fixed T,

$$dF_a = d(E_a - TS_a) = -T(dS_a - dE_a/T), \tag{1.36}$$

where $F_a = E_a - TS_a$ is the Helmholtz free energy of system a,

$$dS_{\text{tot}}|_{T,Q,V} = -\frac{1}{T}dF_a|_{T,Q,V} = 0. \tag{1.37}$$

Thus, if a system is in contact with a heat bath at temperature T, the value of x adjusts itself to minimize (because the entropy comes in with the opposite sign) $F(x, T, V, Q)$. If charges are also allowed to enter/exit the system with a bath at chemical potential μ, the change in entropy is

$$dS_{\text{tot}} = dS_a - dE_a/T + \mu_a dQ_a/T = 0. \tag{1.38}$$

Here, stating that $dS_{\text{tot}} = 0$ is equivalent to stating,

$$dS_{\text{tot}}|_{T,V,\mu} = d(S_a - \beta E_a + \beta\mu Q_a)|_{T,V,\mu} = d\left.\frac{PV}{T}\right|_{T,V,\mu} = 0, \tag{1.39}$$

Thus, at fixed chemical potential and temperature a system adjusts itself to maximize the pressure.

The calculations above assumed that the volume was fixed. In some instances, the pressure and particle number might be fixed rather than the chemical potential and volume. In that case,

$$dS_{\text{tot}} = dS_a - dE_a/T - \beta P dV_a = 0. \tag{1.40}$$
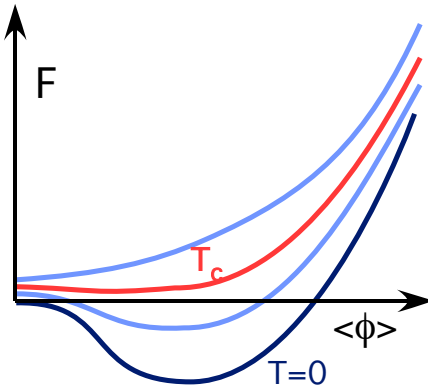
This can then be stated as:

$$d(S_a - \beta E_a - \beta P V_a)|_{T,Q,P} = -\beta d(\mu Q) = 0. \tag{1.41}$$

Here, the term μQ is referred to as the Gibb's free energy,

$$G \equiv \mu Q = E + PV - TS. \tag{1.42}$$

Summarizing the four cases discussed above:

Figure 1.1: The free energy has a minimum at non-zero $\langle \phi \rangle$ at $T = 0$. In such circumstances a condensed field forms, as is the case for the Higg's condensate for the electroweak transition or the scalar quark condensate for the QCD chiral transition. At high temperature, the minimum moves back to $\langle \phi \rangle = 0$. If $\langle \phi \rangle$ is discontinuous as a function of T, a phase transition ensues. If $\langle \phi \rangle$ is multidimensional, i.e., ϕ has two components with rotational symmetry, the transition violates symmetry and is referred to a spontaneous symmetry breaking. Such potentials are often referred to as "mexican hat" potentials.

- If E, V and Q are fixed, x is chosen to maximize $S(x, E, V, Q)$.

- If Q and V are fixed, and the system is connected to a heat bath, x is chosen to minimize $F(x, T, V, Q)$.

- If the system is allowed to exchange charge and energy with a bath at chemical potential μ and temperature T, x is chosen to maximize $P(x, T, \mu, V)$ (or Ω_{GC} in the case that there are long range correlations or interactions).

- If the charge is fixed, but the volume is allowed to adjust itself at fixed pressure, the system will attempt to minimize the Gibb's free energy, $G(x, T, P, Q)$.

Perhaps the most famous calculation of this type concerns the electroweak phase transition which purportedly took place in the very early universe, 10^{-12} s after the big bang. In that transition the Higg's condensate $\langle \phi \rangle$ became non-zero in a process called spontaneous symmetry breaking. The free energy can be calculated as a function of $\langle \phi \rangle$ (Search literature of 1970s by Jackiw, Coleman and Linde). The resulting free energy depends on T and $\langle \phi \rangle$, as illustrated qualitatively in Fig. 1.1.

1.9   More about Free Energies

Perusing the literature, one is likely to find mention of several free energies. For macroscopic systems without long-range interactions one can write the grand potential as $-PV$. The other common free energies can also be expressed in terms of the pressure as listed below:

In the grand canonical ensemble, one finds the quantities E and Q by taking derivatives of the grand potential w.r.t. the Lagrange multipliers. Similarly, one can find the Lagrange multipliers by taking derivatives of some of the other free energies with respect to Q or E. To see this we begin with the fundamental thermodynamic relation,

$$TdS = dE + PdV - \mu dQ. \tag{1.43}$$

Because the Helmholtz free energy is usually calculated in the canonical ensemble,

$$F = -T \ln Z_C(Q, T, V) = E - TS, \tag{1.44}$$

Grand Potential	$\Omega = -PV = E - \mu Q - TS$
Helmholtz Free Energy	$F = E - TS = \mu Q - PV$
Gibb's Free Energy	$G = \mu Q = E + PV - TS$
Enthalpy*	$H = E + PV = TS + \mu Q$

Table 2: Various free energies
*Enthalpy is conserved in what is called a Joule-Thomson process, see http://en.wikipedia.org/wiki/Joule-Thomson_effect#Proof_that_enthalpy_remains_constant_in_a_Joule-Thomson_process.

it is typically given as a function of Q and T. Nonetheless, one can still find μ by considering,

$$dF = dE - TdS - SdT = -PdV - SdT + \mu dQ. \tag{1.45}$$

This shows that the chemical potential is

$$\mu = \left. \frac{\partial F}{\partial Q} \right|_{V,T}. \tag{1.46}$$

Similarly, in the microcanonical ensemble one typically calculates the entropy as a function of E and Q. In this case the relation

$$dS = \beta dE + \beta P dV - \beta \mu dQ, \tag{1.47}$$

allows one to read off the temperature as

$$\beta = \left. \frac{\partial S}{\partial E} \right|_{V,Q}. \tag{1.48}$$

In many text books, thermodynamics begins from this perspective. We have instead begun with the entropy defined as $-\sum p_i \ln p_i$, as this approach does not require the continuum limit, and allows the thermodynamic relations to be generated from general statistical concepts. The question "When is statistical mechanics thermodynamics?" is not easily answered. A more meaningful question concerns the validity of the fundamental thermodynamic relation, $TdS = dE + PdV - \mu dQ$. This becomes valid as long as one can justify differentials, which is always the case for grand canonical ensemble, where E and Q are averages. However, for the other ensembles one must be careful because the discreteness of particle number or energy levels can invalidate the use of differentials. This is never a problem for macroscopic systems because the level densities are so high that differential considerations are justified.

1.10 Forces at Finite Temperature

Let's assume a particle feels a force dependent on its position r, and that the force depends on how the thermal medium reacts to the particle's position. Normally, one would consider the force from the gradient of the energy, $f_i = -\partial_i E(\mathbf{r})$. Indeed, this is the case in a thermal system,

$$f_i = - \left. \frac{\partial E(\mathbf{r})}{\partial r_i} \right|_{S,N,V}, \tag{1.49}$$

where the partial derivative is taken at constant entropy, not constant temperature. This is based on the idea that the particle moves slowly, and that the system remains in the given, but shifting, energy level of the system as the coordinate is slowly varied. However, thermal calculations are often performed more readily at fixed temperature, rather than fixed entropy. For that reason, one would like to know how to calculate the force using derivatives where the temperature is fixed, and then correct for the fact that the temperature must be altered to maintain constant entropy. For simplicity, consider a one-dimensional system. The fundamental thermodynamic relation becomes

$$dE = TdS - PdV + \mu dQ - fdx, \tag{1.50}$$

and f manifestly matches the definition in Eq. (1.49). One can alter this equation by adding $-d(TS)$ to both sides, which givs

$$d(E - TS) \;=\; -SdT - PdV + \mu dQ - fdx, \tag{1.51}$$

where this is the same force as before, only in this case inspection reveals

$$f \;=\; -\left.\frac{\partial(E - TS)}{\partial x}\right|_{T,V,Q}. \tag{1.52}$$

Thus, the Helmholtz free energy, $F = E - TS$, plays the role of the potential. This result should not be interpreted as the particle feels a force that pushes it towards minimizing the potential energy and maximizing the entropy. Particles don't feel entropy. The $-TS$ term accounts for taking gradients at fixed temperature, whereas the definition of the force is the derivative of the energy with respect to x at fixed entropy.

1.11   Maxwell Relations

Maxwell relations are a set of equivalences between derivatives of thermodynamic variables. They can prove useful in several cases, for example in the Clausius-Clapeyron equation, which will be discussed later. Derivations of Maxwell relations also tends to appear in written examinations.

All Maxwell relations can be derived from the fundamental thermodynamic relation,

$$dS = \beta dE + (\beta P)dV - (\beta\mu)dQ, \tag{1.53}$$

from which one can readily identify β, P and μ with partial derivatives,

$$\beta \;=\; \left.\frac{\partial S}{\partial E}\right|_{V,Q}, \tag{1.54}$$

$$\beta P \;=\; \left.\frac{\partial S}{\partial V}\right|_{E,Q},$$

$$\beta\mu \;=\; -\left.\frac{\partial S}{\partial Q}\right|_{E,V}.$$

Here, the subscripts refer to the quantities that are fixed. Because for any continuous function $f(x, y)$, $\partial^2 f/(\partial x \partial y) = \partial^2 f/(\partial y \partial x)$, one can state the following identities which are known as Maxwell relations,

$$\left.\frac{\partial \beta}{\partial V}\right|_{E,Q} = \left.\frac{\partial(\beta P)}{\partial E}\right|_{V,Q}, \tag{1.55}$$

$$\left.\frac{\partial \beta}{\partial Q}\right|_{E,V} = -\left.\frac{\partial(\beta \mu)}{\partial E}\right|_{V,Q},$$

$$\left.\frac{\partial(\beta P)}{\partial Q}\right|_{E,V} = -\left.\frac{\partial(\beta \mu)}{\partial V}\right|_{E,Q}.$$

An entirely different set of relations can be derived by beginning with quantities other than dS. For instance, if one begins with dF,

$$F = E - TS, \tag{1.56}$$

$$dF = dE - SdT - TdS,$$

$$= -SdT - PdV - \mu dQ,$$

three more Maxwell relations can be derived by considering second derivatives of F with respect to T, V and Q. Other sets can be derived by considering the Gibbs free energy, $G = E + PV - TS$, or the pressure, $PV = TS - E + \mu Q$.

Example 1.3:

Validate the following Maxwell relation:

$$\left.\frac{\partial V}{\partial Q}\right|_{P,T} = \left.\frac{\partial \mu}{\partial P}\right|_{Q,T} \tag{1.57}$$

Start with the standard relation,

$$dS = \beta dE + (\beta P)dV - (\beta \mu)dQ. \tag{1.58}$$

This will yield Maxwell's relations where E, V and Q are held constant. Because we want a relation where T, N and P/T are held constant, we should consider:

$$d(S - (\beta P)V - \beta E) = -Ed\beta - Vd(\beta P) - (\beta \mu)dQ. \tag{1.59}$$

This allows one to realize:

$$V = -\left.\frac{\partial(S - \beta PV - \beta E)}{\partial(\beta P)}\right|_{\beta,Q}, \quad \beta \mu = -\left.\frac{\partial(S - \beta PV - \beta E)}{\partial Q}\right|_{\beta,\beta P}. \tag{1.60}$$

Using the fact that $\partial^2 f/(\partial x \partial y) = \partial^2 f/(\partial y \partial x)$,

$$\left.\frac{\partial V}{\partial Q}\right|_{\beta P,T} = \left.\frac{\partial \beta \mu}{\partial \beta P}\right|_{\beta,Q}. \tag{1.61}$$

On the l.h.s., fixing βP is equivalent to fixing P because T is also fixed. On the r.h.s., the factors of β in $\partial(\beta\mu)/\partial(\beta P)$ cancel because T is fixed. This then leads to the desired result, Eq. (1.57).

Example 1.4:

Show that

$$\left.\frac{\partial E}{\partial Q}\right|_{S,P} = \mu - P\left.\frac{\partial\mu}{\partial P}\right|_{S,Q}.$$

One way to start is to look at the relation and realize that V does not appear. Thus, let's rewrite the fundamental thermodynamic relation with $dV = \cdots$. This ensures that we won't have any Vs appearing on the r.h.s. of the equations going forward.

$$dV = \frac{T}{P}dS - \frac{1}{P}dE + \frac{\mu}{P}dQ.$$

Now, because the relation will involve S, P and Q being constant add $d(E/P)$ from each side to get

$$
\begin{aligned}
d(V + E/P) &= \frac{T}{P}dS + Ed(1/P) + \frac{\mu}{P}dQ \\
\left.\frac{\partial E}{\partial Q}\right|_{S,P} &= \left.\frac{\partial(\mu/P)}{\partial(1/P)}\right|_{S,Q} \\
&= \mu - P\left.\frac{\partial\mu}{\partial P}\right|_{S,Q}.
\end{aligned}
$$

Some students find a nmemonic useful for deriving some of the Maxwell relations. See http://en.wikipedia.org/wiki/Thermodynamic_square. However, this is more cute than useful, as you will see with some of the more difficult problems at the end of this chapter.

1.12 Fluctuations

Statistical averages of higher order terms in the energy or the density can also be extracted from the partition function. For instance, in the grand canonical ensemble (using $\beta = 1/T$ and $\alpha = -\mu/T$ as the variables),

$$
\begin{aligned}
\partial_\beta^2 \ln Z(\beta, \alpha) &= \partial_\beta\frac{\partial_\beta Z}{Z} \\
&= \frac{1}{Z}\partial_\beta^2 Z - \left(\frac{1}{Z}\partial_\beta Z\right)^2 \\
&= \langle E^2\rangle - \langle E\rangle^2 \\
&= \langle(E - \langle E\rangle)^2\rangle.
\end{aligned}
\tag{1.62}
$$

Thus, the second derivative of $\ln Z$ gives the fluctuation of the energy from its average value.

Similarly, one can see that

$$\partial_\beta \partial_\alpha \ln Z = \langle (E - \langle E \rangle)(Q - \langle Q \rangle) \rangle \tag{1.63}$$
$$\partial_\alpha^2 \ln Z = \langle (Q - \langle Q \rangle)^2 \rangle . \tag{1.64}$$

The specific heat at constant α and μ is quickly related to the fluctuation,

$$\left. \frac{dE}{dT} \right|_{V,\alpha} = -\frac{1}{T^2} \left. \frac{dE}{d\beta} \right|_{V,\alpha} = \frac{1}{T^2} \langle (E - \langle E \rangle)^2 \rangle .$$

Fluctuations play a critical role in numerous areas of physics, most notably in critical phenomena.

1.13   Problems

1. Using the methods of Lagrange multipliers, find x and y that minimize the following function,
$$f(x, y) = 3x^2 - 4xy + y^2,$$
subject to the constraint,
$$3x + y = 0.$$

2. Consider 2 identical bosons (A given level can have an arbitrary number of particles) in a 2-level system, where the energies are 0 and ϵ. In terms of ϵ and the temperature T, calculate:

    (a) The partition function Z_C

    (b) The average energy $\langle E \rangle$. Also, give $\langle E \rangle$ in the $T = 0, \infty$ limits.

    (c) The entropy S. Also give S in the $T = 0, \infty$ limits.

    (d) Now, connect the system to a particle bath with chemical potential μ. Calculate $Z_{GC}(\mu, T)$. Find the average number of particles, $\langle N \rangle$ as a function of μ and T. Also, give the $T = 0, \infty$ limits.
    Hint: For a grand-canonical partition function of non-interacting particles, one can state that $Z_{GC} = Z_1 Z_2 \cdots Z_n$, where Z_i is the partition function for one single-particle level, $Z_i = 1 + e^{-\beta(\epsilon_i - \mu)} + e^{-2\beta(\epsilon_i - \mu)} + e^{-3\beta(\epsilon_i - \mu)} \cdots$, where each term refers to a specific number of bosons in that level.

3. Repeat the problem above assuming the particles are identical Fermions (No level can have more than one particle, e.g., both are spin-up electrons).

4. Beginning with the expression,
$$TdS = dE + PdV - \mu dQ,$$
show that the pressure can be derived from the Helmholtz free energy, $F = E - TS$, with
$$P = -\left. \frac{\partial F}{\partial V} \right|_{Q,T}.$$

5. Assuming that the pressure P is independent of V when written as a function of μ and T, i.e., $\ln Z_{GC} = PV/T$ (true if the system is much larger than the range of interaction),

    (a) Find expressions for E/V and Q/V in terms of P, T, and partial derivatives of P or P/T w.r.t. $\alpha \equiv -\mu/T$ and $\beta \equiv 1/T$. Here, assume the chemical potential is associated with the conserved number Q.

    (b) Find an expression for $C_V = dE/dT|_{Q,V}$ in terms of P/T, E, Q, V and the derivatives of P, P/T, E and Q w.r.t. β and α.

    (c) Show that the entropy density $s = \partial_T P|_\mu$.

6. Beginning with:
$$dE = TdS - PdV + \mu dN,$$

derive the Maxwell relation,
$$\left.\frac{\partial V}{\partial \mu}\right|_{S,P} = -\left.\frac{\partial N}{\partial P}\right|_{S,\mu}.$$

7. Beginning with the definition,
$$C_P = T\left.\frac{\partial S}{\partial T}\right|_{N,P},$$

Show that
$$\left.\frac{\partial C_p}{\partial P}\right|_{T,N} = -T\left.\frac{\partial^2 V}{\partial T^2}\right|_{P,N}$$

Hint: Find a quantity Y for which both sides of the equation become $\partial^3 Y/\partial^2 T \partial P$.

8. Beginning with
$$TdS = dE + PdV - \mu dQ, \quad \text{and } G \equiv E + PV - TS,$$

    (a) Show that
$$S = -\left.\frac{\partial G}{\partial T}\right|_{N,P}, \quad V = \left.\frac{\partial G}{\partial P}\right|_{N,T}.$$

    (b) Beginning with
$$\delta S(P, N, T) = \frac{\partial S}{\partial P}\delta P + \frac{\partial S}{\partial N}\delta N + \frac{\partial S}{\partial T}\delta T,$$

    Show that the specific heats,
$$C_P \equiv T\left.\frac{\partial S}{\partial T}\right|_{N,P}, \quad C_V \equiv T\left.\frac{\partial S}{\partial T}\right|_{N,V},$$

    satisfy the relation:
$$C_P = C_V - T\left(\left.\frac{\partial V}{\partial T}\right|_{P,N}\right)^2 \left(\left.\frac{\partial V}{\partial P}\right|_{T,N}\right)^{-1}$$

    Note that the compressibility, $\equiv -\partial V/\partial P$, is positive (unless the system is unstable), therefore $C_P > C_V$.

9. From Sec. 1.12, it was shown how to derive fluctuations in the grand canonical ensemble. Thus, it is straightforward to find expressions for the following fluctuations, $\phi_{EE} \equiv \langle \delta E \delta E \rangle / V$, $\phi_{QQ} \equiv \langle \delta Q \delta Q \rangle / V$ and $\phi_{QE} \equiv \langle \delta E \delta Q \rangle / V$. In terms of the 3 fluctuations above, calculated in the grand canonical ensemble, and in terms of the volume and the temperature T, express the specific heat at constant volume and charge,

$$C_V = \left. \frac{dE}{dT} \right|_{Q,V}.$$

Note that the fluctuation observables, ϕ_{ij}, are intrinsic quantities, assuming that the correlations in energy density occur within a finite range and that the overall volume is much larger than that range.

2   Statistical Mechanics of Non-Interacting Particles

"Its a gas! gas! gas!" - M. Jagger & K. Richards

When students think of gases, they reminisce fondly back to high school physics or chemistry, and remember the ideal gas law. This simple physics applies for gases which are sufficiently dilute that interactions between particles can be neglected and that identical-particle statistics can be ignored. For this chapter, we will indeed confine ourselves to the assumption that interactions can be neglected, but will consider in detail the effects of quantum statistics. Quantum statistics play a dominant role in situations across all fields of physics. For instance, Fermi gas concepts play the essential role in understanding nuclear structure, stellar structure and dynamics, and the properties of metals and superconductors. For Bose systems, quantum degeneracy can drive the creation of super-fluids, such as liquid Helium or in many atom/molecule trap systems. Even in the presence of strong interactions between constituents, the role of quantum degeneracy can still play a dominant role in determining the properties and behaviors of numerous systems.

2.1   Non-Interacting Gases

A non-interacting gas can be considered as a set of independent momentum modes, labeled by the momentum p. For particles in a box defined by $0 < x < L_x$, $0 < y < L_y$, $0 < z < L_z$, the single-particle wave functions have the form,

$$\psi_{p_x,p_y,p_z}(x,y,z) \sim \sin(p_x x/\hbar)\sin(p_y y/\hbar)\sin(p_z z/\hbar) \tag{2.1}$$

with the quantized momenta $\vec{p}$ satisfying the boundary conditions,

$$\begin{aligned}
p_x L_x/\hbar &= n_x \pi, \quad n_x = 1, 2, \cdots, \\
p_y L_y/\hbar &= n_y \pi, \quad n_y = 1, 2, \cdots, \\
p_z L_z/\hbar &= n_z \pi, \quad n_z = 1, 2, \cdots
\end{aligned} \tag{2.2}$$

The number of states in d^3p is $dN = V d^3p/(\pi\hbar)^3$. Here p_x, p_y and p_z are positive numbers as the momenta are labels of the static wavefunctions, each of which can be considered as a coherent combination of a forward and a backward moving plane wave. The states above are not eigenstates of the momentum operator, as $\sin px/\hbar$ is a mixture of left-going and right-going states. If one wishes to consider plane waves, the signs of the momenta can be either positive or negative. Because one doubles the range of momentum for each dimension when one includes both positive and negative momenta, one needs to introduce a factor of $1/2$ to the density of states for each dimension when including the negative momentum states. The density of states in three dimensions then becomes:

$$dN = (2s+1)V\frac{d^3p}{(2\pi\hbar)^3}. \tag{2.3}$$

For one or two dimensions, d^3p is replaced by $d^D p$ and $(2\pi\hbar)^3$ is replaced by $(2\pi\hbar)^D$, where D is the number of dimensions. For two dimensions, the volume is replaced by the area, and in one

dimension, the length fulfills that role. The number of single-particle states also increases with spin, thus the $(2s+1)$ factor. It is critical to distinguish that dN refers to the number of modes, or single-particle states, not the number of particles, and not the number of system states.

In the grand canonical ensemble, each momentum mode can be considered as being an independent system, described by its own partition function, z_p. Thus, the partition function of the entire system is the product of partition functions of each mode,

$$Z_{GC} = \prod_p z_p, \quad z_p = \sum_{n_p} e^{-n_p \beta(\epsilon_p - \mu q)}. \tag{2.4}$$

Here n_p is the number of particles of charge q in the mode, and n_p can be 0 or 1 for Fermions, or can be $0, 1, 2, \cdots$ for bosons. The classical limit corresponds to the case where the probability of having more than 1 particle would be so small that the Fermi and Bose cases become indistinguishable. Later on it will be shown that this limit requires that μ is much smaller (in units of T) than the energy of the lowest mode.

For thermodynamic quantities, only $\ln Z_{GC}$ plays a role, which allows the product in Eq. (2.4) to be replaced by a sum,

$$\begin{aligned}
\ln Z_{GC} &= \sum_p \ln z_p \\
&= (2s+1) \int \frac{V d^3 p}{(2\pi\hbar)^3} \ln z_p \\
&= (2s+1) \int \frac{V d^3 p}{(2\pi\hbar)^3} \ln(1 + e^{-\beta(\epsilon - \mu q)}) \quad \text{(Fermions)}, \\
&= (2s+1) \int \frac{V d^3 p}{(2\pi\hbar)^3} \ln\left(\frac{1}{1 - e^{-\beta(\epsilon - \mu q)}}\right) \quad \text{(Bosons)}, \\
&= (2s+1) \int \frac{V d^3 p}{(2\pi\hbar)^3} (\mp) \ln(1 \mp e^{-\beta(\epsilon - \mu q)}) \quad \text{(Bosons/Fermions)}.
\end{aligned} \tag{2.5}$$

Here, the intrinsic spin of the particles is s, and each of the $(2s+1)$ spin projections m_s contributes independently to the partition function, i.e. the independent modes are labeled by the momenta and spin projection. For the Fermi expression, where $z_p = 1 + e^{-\beta(\epsilon_p - \mu)}$, the first term represents the contribution for having zero particles in the state and the second term accounts for having one particle in the level. For bosons, which need not obey the Pauli exclusion principle, one would include additional terms for having $2, 3 \cdots$ particles in the mode, and the Bosonic relation has made use of the fact that $1 + x + x^2 + x^3 \cdots = 1/(1-x)$.

Example 2.1:

For a classical two-dimensional, non-interacting, non-relativistic gas of fermions of mass m, and charge $q = 1$ and spin s, find the charge density and energy density of a gas in terms of μ and T.

The charge can be found by taking derivatives of $\ln Z_{GC}$ with respect to $\alpha = -\beta\mu$,

$$Q = \frac{\partial}{\partial(\beta\mu)}\ln Z_{GC}, \tag{2.6}$$

$$= (2s+1)A\int_0^\infty \frac{2\pi p\,dp}{(2\pi\hbar)^2}\cdot\frac{e^{-\beta(p^2/2m-\mu)}}{1+e^{-\beta(p^2/2m-\mu)}}$$

$$= (2s+1)A\frac{mT}{2\pi\hbar^2}\int_0^\infty dx\,\frac{e^{-(x-\beta\mu)}}{1+e^{-(x-\beta\mu)}}.$$

The above expression applies to Fermions. To apply the classical limit, one takes the limit that $\beta\mu \to -\infty$,

$$\frac{Q}{A} = (2s+1)\frac{mT}{2\pi\hbar^2}\int_0^\infty dx\,e^{-(x-\beta\mu)} \tag{2.7}$$

$$= (2s+1)\frac{mT}{2\pi\hbar^2}e^{\beta\mu}.$$

In calculating the energy per unit area, one follows a similar procedure beginning with,

$$E = -\frac{\partial}{\partial\beta}\ln Z_{GC}, \tag{2.8}$$

$$\frac{E}{A} = (2s+1)\int_0^\infty \frac{2\pi p\,dp}{(2\pi\hbar)^2}\frac{p^2}{2m}\frac{e^{-\beta(p^2/2m-\mu)}}{1+e^{-\beta(p^2/2m-\mu q)}}$$

$$= (2s+1)\frac{mT^2}{2\pi\hbar^2}\int_0^\infty dx\,x\frac{e^{-(x-\beta\mu)}}{1+e^{-(x-\beta\mu)}}$$

$$\approx (2s+1)\frac{mT^2}{2\pi\hbar^2}e^{\beta\mu}, \tag{2.9}$$

where in the last step we have taken the classical limit. If we had begun with the Bosonic expression, we would have came up with the same answer after assuming $\beta\mu \to -\infty$.

As can be seen in working the last example, the energy and charge can be determined by the following integrals,

$$\frac{Q}{V} = (2s+1)\int\frac{d^Dp}{(2\pi\hbar)^D}\,f(\epsilon), \tag{2.10}$$

$$\frac{E}{V} = (2s+1)\int\frac{d^Dp}{(2\pi\hbar)^D}\,\epsilon(p)f(\epsilon),$$

where D is again the number of dimensions and $f(p)$ is the occupation probability, a.k.a. the phase-space occupancy or the phase-space filling factor,

$$f(\epsilon) = \begin{cases} e^{-\beta(\epsilon-\mu)}/(1+e^{-\beta(\epsilon-\mu)}), & \text{Fermions, Fermi} - \text{Dirac distribution} \\ e^{-\beta(\epsilon-\mu)}/(1-e^{-\beta(\epsilon-\mu)}), & \text{Bosons, Bose} - \text{Einstein distribution} \\ e^{-\beta(\epsilon-\mu)}, & \text{Classical, Boltzmann distribution.} \end{cases} \tag{2.11}$$

Here $f(\epsilon)$ can be identified as the average number of particles in mode p. These expressions are valid for both relativistic or non-relativisitc gases, the only difference being that for relativistic gases $\epsilon(p) = \sqrt{m^2 + p^2}$.

The classical limit is valid whenever the occupancies are always much less than unity, i.e., $e^{\beta(\mu - \epsilon_0)} \ll 1$, or equivalently $\beta(\mu - \epsilon_0) \to -\infty$. Here, ϵ_0 refers to the energy of the lowest momentum mode, which is zero for a massless gas, or for a non-relativistic gas where the rest-mass energy is ignored, $\epsilon = p^2/2m$. If one preserves the rest-mass energy in $\epsilon(p)$, then $\epsilon_0 = mc^2$. The decision to ignore the rest-mass in $\epsilon(p)$ is accompanied by translating μ by the same amount. For relativistic systems, one usually keeps the rest mass energy because one can not ignore the contribution from anti-particles. For instance, if one has a gas of electrons and positrons, the chemical potential modifies the phase space occupancy through a factor $e^{\beta\mu}$, while the anti-particle is affected by $e^{-\beta\mu}$, and absorbing the rest-mass into μ would destroy the symmetry between the expressions for the densities of particles and anti-particles.

2.2   Equipartition and Virial Theorems

Particles undergoing interaction with an external field, as opposed to with each other, also behave independently. The phase space expressions from the previous section are again important here, as long as the external potential does not confine the particles to such small volumes that the uncertainty principle plays a role. The structure of the individual quantum levels will become important only if the energy spacing is not much smaller than the temperature.

The equipartion theorem applies for the specific case where the quantum statistics can be neglected, i.e. $f(\epsilon) = e^{-\beta(\epsilon - \mu)}$, and one of the variables, p or q, contributes to the energy quadratically. The equipartion theory states that the average energy associated with that variable is then $(1/2)T$. This is the case for the momentum of non-relativistic particles being acted on by potentials depending only on the position. In that case,

$$E = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + V(r). \tag{2.12}$$

When calculating the average of $p_x^2/2m$ (in the classical limit),

$$\left\langle \frac{p_x^2}{2m} \right\rangle = \frac{(2\pi\hbar)^{-3} \int dp_x dp_y dp_z d^3r \; (p_x^2/2m) e^{-(p_x^2 + p_y^2 + p_z^2)/2mT - V(\mathbf{r})/T}}{(2\pi\hbar)^{-3} \int dp_x dp_y dp_z d^3r \; e^{-(p_x^2 + p_y^2 + p_z^2)/2mT - V(\mathbf{r})/T}}, \tag{2.13}$$

the integrals over p_y, p_z and r factorize and then cancel leaving only,

$$\left\langle \frac{p_x^2}{2m} \right\rangle = \frac{\int dp_x \; (p_x^2/2m) e^{-p_x^2/2mT}}{\int dp_x \; e^{-p_x^2/2mT}} = \frac{T}{2}. \tag{2.14}$$

The average energy associated with p_x is then independent of m.

The same would hold for any other degree of freedom that appears in the Hamiltonian quadratically. For instance, if a particle moves in a three-dimensional harmonic oscillator,

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + \frac{1}{2}m\omega_x^2 x^2 + \frac{1}{2}m\omega_y^2 y^2 + \frac{1}{2}m\omega_z^2 z^2, \tag{2.15}$$

the average energy is

$$\langle H \rangle = 3T, \tag{2.16}$$

with each of the six degrees of freedom contributing $T/2$.

The *generalized equipartion theorem* works for any variable that is confined to a finite region, and states

$$\left\langle q \frac{\partial H}{\partial q} \right\rangle = T. \tag{2.17}$$

For the case where the potential is quadratic, this merely states the ungeneralized form of the equipartion theorem. To prove the generalized form, one need only integrate by parts,

$$
\begin{aligned}
\left\langle q \frac{\partial H}{\partial q} \right\rangle &= \frac{\int dq \; q(\partial H/\partial q) \, e^{-\beta H(q)}}{\int dq \; e^{-\beta H(q)}} \\
&= \frac{-T \int dq \; q(\partial/\partial q) \, e^{-\beta H(q)}}{\int dq \; e^{-\beta H(q)}} \\
&= \frac{T \left[q e^{-\beta H(q)} \big|_{-\infty}^{\infty} + \int dq \; e^{-\beta H(q)} \right]}{\int dq \; e^{-\beta H(q)}} \\
&= T \quad \text{if } H(q \to \pm\infty) = \infty.
\end{aligned}
\tag{2.18}
$$

Note that disposing of the term with the limits $\pm\infty$ requires that $H(q \to \pm\infty) \to \infty$, or equivalently, that the particle is confined to a finite region of phase space. The rule also works if p takes the place of q.

The virial theorem has an analog in classical mechanics, and is often referred to as the "Clausius Virial Theorem". It states,

$$\left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle = \left\langle q_i \frac{\partial H}{\partial q_i} \right\rangle. \tag{2.19}$$

To prove the theorem we use Hamilton's equations of motion,

$$\left\langle \frac{d}{dt}(q_i p_i) \right\rangle = \langle p_i \dot{q}_i \rangle + \langle q_i \dot{p}_i \rangle = \left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle - \left\langle q_i \frac{\partial H}{\partial q_i} \right\rangle. \tag{2.20}$$

The theorem will be proved if the time average of $d/dt(p_i q_i) = 0$. Defining the time average as the average over an infinite interval τ,

$$\left\langle \frac{d}{dt}(p_i q_i) \right\rangle \equiv \frac{1}{\tau} \int_0^\tau dt \, \frac{d}{dt}(p_i q_i) = \frac{1}{\tau} \, (p_i q_i)\big|_0^\tau, \tag{2.21}$$

which will equal zero because $\tau \to \infty$ and $p_i q_i$ is finite. Again, this finite-ness requires that the Hamiltonian confines both p_i and q_i to a finite region in phase space. For cyclic coordinates, Hamiltonian equations do not apply at the discontinuity, e.g. $\phi \to \phi - 2\pi$, so this relation is invalid for cyclic coordinates. It is valid for reflective potentials, e.g. a square well, even in the limit the reflective potential becomes infinite. In the first line of the expression above, the equivalence between the statistical average and the time-weighted average required the the Ergodic theorem.

Example 2.2:

Using the virial theorem and the non-generalized equipartion theorem, show that for a non-relativistic particle moving in a one-dimensional potential,

$$V(x) = \kappa x^6,$$

that the average potential energy is

$$\langle V(x) \rangle = T/6.$$

To proceed, we first apply the virial theorem to relate an average of the potential to an average of the kinetic energy which is quadratic in p and can thus be calculated with the equipartion theorem,

$$\left\langle x \frac{\partial V}{\partial x} \right\rangle = \left\langle p \frac{\partial H}{\partial p} \right\rangle = \left\langle \frac{p^2}{m} \right\rangle = 2 \left\langle \frac{p^2}{2m} \right\rangle = T. \tag{2.22}$$

Next, one can identify

$$\left\langle x \frac{\partial V}{\partial x} \right\rangle = 6 \langle V(x) \rangle,$$

to see that

$$\langle V(x) \rangle = \frac{T}{6}.$$

One could have obtained this result in one step with the generalized equipartion theorem.

2.3   Degenerate Bose Gases

The term *degenerate* refers to the fact there are some modes for which the occupation probability $f(p)$ is not much smaller than unity. In such cases, Fermions and Bosons behave very differently. For bosons, the occupation probability,

$$f(\epsilon) = \frac{e^{-\beta(\epsilon - \mu)}}{1 - e^{-\beta(\epsilon - \mu)}}, \tag{2.23}$$

will diverge if μ reaches ϵ. For the non-relativistic case ($\epsilon = p^2/2m$ ignoring the rest-mass energy) this means that a divergence will ensue if $\mu \geq 0$. For the relativistic case ($\epsilon = \sqrt{p^2 + m^2}$ includes the rest-mass in the energy) a divergence will ensue when $\mu \geq m$. This divergence leads to the phenomena of *Bose condensation*.

To find the density required for Bose condensation, one needs to calculate the density for $\mu \to 0^-$. For lower densities, μ will be negative and there are no singularities in the phase space

density. The calculation is a bit tedious but straight-forward,

$$
\begin{aligned}
\rho(\mu = 0^-, T) &= (2s+1) \int \frac{d^D p}{(2\pi\hbar)^D} \frac{e^{-p^2/2mT}}{1 - e^{-p^2/2mT}} \\
&= (2s+1) \int \frac{d^D p}{(2\pi\hbar)^D} \sum_{n=1}^{\infty} e^{-np^2/2mT}, \\
&= (2s+1) \frac{(mT)^{D/2}}{(2\pi\hbar)^D} \left(\int d^D x\, e^{-x^2/2} \right) \sum_{n=1}^{\infty} n^{-D/2}
\end{aligned}
\tag{2.24}
$$

For $D = 1, 2$, the sum over n diverges. Thus for one or two dimensions, Bose condensation never occurs (at least for massive particles where $\epsilon_p \sim p^2$) as one can attain an arbitrarily high density without having μ reach zero. This can also be seen by expanding $f(\epsilon)$ for ϵ_p and μ near zero,

$$
f(p \to 0) \approx \frac{1}{\beta p^2/m - \beta\mu}.
\tag{2.25}
$$

Again, for $D = 1, 2$, the phase space weight $p^{D-1} dp$ will not kill the divergence of $f(p \to 0)$ in the limit $\mu \to 0$. Thus, one can obtain arbitrarily high densities without μ actually reaching 0. Because μ never reaches zero, even as the density approaches infinity, Bose condensation never occurs.

For $D = 3$, the condensation density is finite,

$$
\begin{aligned}
\rho(\mu = 0) &= (2s+1) \frac{(mT)^{3/2}}{(2\pi\hbar)^3} \left(\int d^3 x\, e^{-x^2/2} \right) \sum_{n=1}^{\infty} n^{-3/2} \\
&= (2s+1) \frac{(mT)^{3/2}}{(2\pi)^{3/2}\hbar^3} \zeta(3/2),
\end{aligned}
\tag{2.26}
$$

where $\zeta(n)$ is the Riemann-Zeta function,

$$
\zeta(n) \equiv \sum_{j=1}^{\infty} \frac{1}{j^n}
\tag{2.27}
$$

The Riemann-Zeta function appears often in statistical mechanics, and especially often in graduate written exams. Some values are:

$$
\zeta(3/2) = 2.612375348685... \quad \zeta(2) = \pi^2/6, \quad \zeta(3) = 1.202056903150..., \zeta(4) = \pi^4/90
\tag{2.28}
$$

Once the density exceeds $\rho_c = \rho(\mu = 0^-)$, μ no longer continues to rise. Instead, the occupation probability of the ground state,

$$
f = \frac{e^{-\beta(\epsilon_0 - \mu)}}{1 - e^{-\beta(\epsilon_0 - \mu)}},
\tag{2.29}
$$

becomes undefined, and any additional density above ρ_c is carried by particles in the ground state. The gas has two components. For the normal component, the momentum distribution is described by the normal Bose-Einstein distribution with $\mu = 0$ and has a density of ρ_c, while

the condensation component has density $\rho - \rho_c$. Similarly, one can fix the density and find the critical temperature T_c, below which the system condenses.

Having a finite fraction of the particles in the ground state does not explain superfluidity by itself. However, particles at very small relative momentum (in this case zero) will naturally form ordered systems if given some kind of interaction (For dilute gases the interaction must be repulsive, as otherwise the particles would form droplets). An ordered conglomeration of bosons will also move with remarkably little friction. This is due to the fact that a "gap" energy is required to remove any of the particles from the ordered system, thus allowing the macroscopic condensate to move coherently as a single object carrying a macroscopic amount of current even though it moves very slowly. Because particles bounce elastically off the condensate, the drag force is proportional to the square of the velocity (just like the drag force on a car moving through air) and is negligible for a slow moving condensate. Before the innovations in atom traps during the last 15 years, the only known Bose condensate was liquid Helium-4, which condenses at atmospheric pressure at a temperature of 2.17 °K, called the lambda point. Even though liquid Helium-4 is made of tightly packed, and therefore strongly interacting, atoms, the prediction for the critical temperature using the calculations above is only off by a degree or so.

Example 2.3:

a) Photons are bosons and have no charge, thus no chemical potential. Find the energy density and pressure of a two-dimensional gas of photons ($\epsilon = p$) as a function of the temperature T. Note: whereas in three dimensions photons have two polarizations, they would have only one polarization in two dimensions. Also, in two dimensions the pressure is a force per unity length, not a force per unit area as it is in three dimensions.

First, write down an expression for the pressure (See HW problem),

$$P = \frac{1}{(2\pi\hbar)^2} \int 2\pi p dp \, \frac{p^2}{2\epsilon} \frac{e^{-\epsilon/T}}{1 - e^{-\epsilon/T}}, \tag{2.30}$$

which after replacing $\epsilon = p$,

$$
\begin{aligned}
P &= \frac{1}{4\pi\hbar^2} \int p^2 dp \frac{e^{-p/T}}{1 - e^{-p/T}}. \tag{2.31} \\
&= \frac{T^3}{4\pi\hbar^2} \int x^2 dx \left\{ e^{-x} + e^{-2x} + e^{-3x} + \cdots \right\}, \\
&= \frac{T^3}{2\pi\hbar^2} \left\{ 1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots \right\} = \frac{T^3}{2\pi\hbar^2}\zeta(3),
\end{aligned}
$$

where the Riemann-Zeta function, $\zeta(3)$, is 1.202056903150... Finally, the energy density, by inspection of the expression for the pressure, is $E/V = 2P$.

b) If these were charged particles, could they form a condensate?

Yes, if you consider the term $1/(1 - e^{-p/T})$, you will see it behaves as $1/p$ as $p \to 0$. The factor pdp cancels the divergence and the integral for the density,

$$\rho_c = \frac{1}{(2\pi\hbar)^2} \int 2\pi p dp \, \frac{p^2}{2\epsilon} \frac{e^{-\epsilon/T}}{1 - e^{-\epsilon/T}},$$

is finite.

2.4 Degenerate Fermi Gases

The goal of this section is to derive the properties of a non-interacting Fermi gas at low temperature. By "low", we mean that the temperature is much less than the Fermi energy, so that the thermal excitations only affect those levels near the Fermi "surface". The Fermi energy, ϵ_f, refers to the single-particle energy below which the single-particle levels are filled, $f \to$ ', at zero temperature, whereas all the levels with energies greater than ϵ_f are empty. The Fermi "surface" refers to momentum space, and the momentum p_f corresponding to $\epsilon_f = p_f^2/2m$. The density of particles at zero temperature is

$$\rho(T=0) = \frac{2}{(2\pi\hbar)^3}\frac{4}{3}\pi p_f^3, \tag{2.32}$$

where the last factor is the volume of momentum space inside the Fermi surface. When the temperature is raised from zero, particles in levels close to the Fermi surface can be excited above ϵ_f, but only those levels in the neighborhood $\epsilon_f - T \lesssim \epsilon \lesssim \epsilon_f + T$. The number of such states affected in $\sim D(\epsilon_f)T$, where $D(\epsilon)$ is the density of single-particle states per energy. The typical exciation is $\sim 2T$, so the excitation energy should behave as $E^* \sim DT^2$. Deriving this proportionality factor is our immediate goal, along with the finding expressions for the change in the density at fixed chemical potential.

At zero temperature, the Fermi-Dirac distribution becomes,

$$f(T \to 0) = \left.\frac{e^{-\beta(\epsilon-\mu)}}{1 + e^{-\beta(\epsilon-\mu)}}\right|_{\beta\to\infty} = \Theta(\mu - \epsilon), \tag{2.33}$$

a step-function, which is unity for energies below the chemical potential and zero for $\epsilon > \mu$. The density is then,

$$\rho(\mu, T=0) = \int_0^\mu d\epsilon \, D(\epsilon)/V. \tag{2.34}$$

Here, one can see that the chemical potential and Fermi energy are synonymous when $T = 0$ and there are no interactions.

For three dimensions, the density of states is:

$$D(\epsilon)/V = \frac{1}{V}\frac{dN_{\text{states}}}{dp}\frac{dp}{d\epsilon} = \frac{(2s+1)}{(2\pi\hbar)^3}(4\pi p^2)\frac{dp}{d\epsilon}. \tag{2.35}$$

For small temperatures (much smaller than μ) one can expand f in the neighborhood of μ,

$$
\begin{aligned}
f(\epsilon) &= f_0 + \delta f, \tag{2.36}\\
\delta f &= \begin{cases} e^{-\epsilon'/T}/(1 + e^{-\epsilon'/T}) - 1, & \epsilon' \equiv \epsilon - \mu < 0 \\ e^{-\epsilon'/T}/(1 + e^{-\epsilon'/T}), & \epsilon' > 0 \end{cases}\\
&= \begin{cases} -e^{\epsilon'/T}/(1 + e^{\epsilon'/T}), & \epsilon' < 0 \\ e^{-\epsilon'/T}/(1 + e^{-\epsilon'/T}), & \epsilon' > 0 \end{cases}.
\end{aligned}
$$
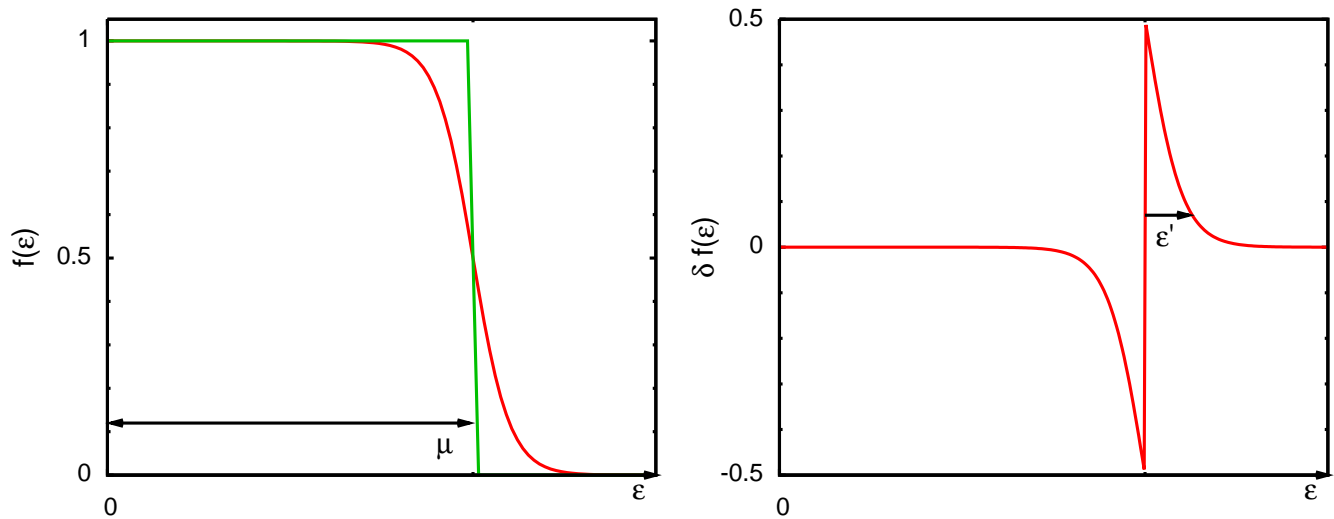
Figure 2.1: The phase space density at small temperature is shown in red, while the phase space density at zero temperature is shown in green.

Here, δf is the difference of the phase space occupancy compared to its value at $T = 0$.

The function δf is illustrated in Fig. 2.1. The function $\delta f(\epsilon')$ is $\mp 1/2$ at $\epsilon' = 0$ and returns to zero for $|\epsilon'| >> T$. Also, δf is an odd function in ϵ', which will be important below.

One can expand $D(\mu + \epsilon')$ in powers of ϵ', and if the temperature is small, one need only keep the first term in the Taylor expansion if one is considering the limit of small temperature.

$$\rho(\mu, T) \approx \rho(\mu, T = 0) + \frac{1}{V} \int_{-\infty}^{\infty} d\epsilon' \left[D(\mu) + \epsilon' \frac{dD}{d\epsilon} \right] \delta f(\epsilon'), \tag{2.37}$$

$$\frac{E}{V} \approx \frac{E}{V}(\mu, T = 0) + \frac{1}{V} \int_{-\infty}^{\infty} d\epsilon' \, (\mu + \epsilon') \left[D(\mu) + \epsilon' \frac{dD}{d\epsilon} \right] \delta f(\epsilon').$$

Because one integrates from $-\infty \rightarrow \infty$, one need only keep terms in the integrand which are even in ϵ'. Given that δf is odd in ϵ', the following terms remain,

$$\delta \rho(\mu, T) = \frac{2}{V} \frac{dD}{d\epsilon} \mathcal{I}(T), \tag{2.38}$$

$$\delta \left(\frac{E}{V} \right) = \mu \delta \rho + \frac{2}{V} D(\mu) \mathcal{I}(T), \tag{2.39}$$

$$\mathcal{I}(T) \equiv \int_0^{\infty} d\epsilon' \, \epsilon' \frac{e^{-\epsilon'/T}}{1 + e^{-\epsilon'/T}}$$

$$= T^2 \int_0^{\infty} dx \, x \frac{e^{-x}}{1 + e^{-x}}.$$

With some trickery, $\mathcal{I}(T)$ can be written in terms of $\zeta(2)$,

$$
\mathcal{I}(T) = T^2 \int_0^\infty dx\, x \left(e^{-x} - e^{-2x} + e^{3x} - e^{-4x} + \cdots \right), \tag{2.40}
$$

$$
= T^2 \int_0^\infty dx\, x \left\{ \left(e^{-x} + e^{-2x} + e^{3x} + \cdots \right) - 2 \left(e^{-2x} + e^{-4x} + e^{6x} + e^{-8x} + \cdots \right)\right\},
$$

$$
= T^2 \zeta(2) - 2(T/2)^2 \zeta(2) = T^2 \zeta(2)/2,
$$

where $\zeta(2) = \pi^2/6$. This gives the final expressions:

$$
\delta\rho(\mu, T) = \left.\frac{\pi^2 T^2}{6V} \frac{dD}{d\epsilon}\right|_{\epsilon=\mu}, \tag{2.41}
$$

$$
\delta\left(\frac{E}{V}\right) = \mu\delta\rho + \frac{\pi^2 T^2}{6V} D(\mu). \tag{2.42}
$$

These expressions are true for any number of dimensions, as the dimensionality affects the answer by changing the values of D and $dD/d\epsilon$. Both expressions assume that μ is fixed. If ρ if fixed instead, the $\mu\delta\rho$ term in the expression for δE is neglected. It is important to note that the excitation energy rises as the square of the temperature for any number of dimensions. One power of the temperature can be thought of as characterizing the range of energies $\delta\epsilon$, over which the Fermi distribution is affected, while the second power comes from the change of energy associated with moving a particle from $\mu - \delta\epsilon$ to $\mu + \delta\epsilon$.

Example 2.4:

Find the specific heat at fixed density for a relativistic three-dimensional gas of Fermions of mass m, spin s and density ρ at low temperature T.

Even though the temperature is small, the system can still be relativistic if μ is not much smaller than the rest mass. Such is the case for the electron gas inside a neutron star. In that case, $\epsilon = \sqrt{m^2 + p^2}$ which leads to $d\epsilon/dp = p/\epsilon$, and the density of states and the derivative from Eq. (2.35) becomes

$$
D(\epsilon) = V\frac{(2s+1)}{(2\pi\hbar)^3} 4\pi p\epsilon \,. \tag{2.43}
$$

The excitation energy is then:

$$
\delta E_{\text{fixed } N} = (2s+1)\frac{p_f \epsilon_f V}{12\hbar^3} T^2, \tag{2.44}
$$

where p_f and ϵ_f are the Fermi momentum and Fermi energy, i.e., the momentum where $\epsilon(p_f) = \epsilon_f = \mu$ at zero temperature. For the specific heat this yields:

$$
C_V = \frac{dE/dT}{\rho V} = (2S+1)\frac{2p_f \epsilon_f}{12\hbar^3 \rho} T. \tag{2.45}
$$

To express the answer in terms of the density, one can use the fact that at zero temperature,

$$
\rho = \frac{(2s+1)}{(2\pi\hbar)^3} \frac{4\pi p_f^3}{3}, \tag{2.46}
$$

$$
p_f = \left\{\frac{(6\pi^2\hbar^3\rho)}{(2s+1)}\right\}^{1/3}.
$$

to substitute for p_f in the expression for C_V, and using the fact that $\epsilon_f = \sqrt{p_f^2 + m^2}$.

One final note: Rather than using the chemical potential or the density, author's often refer to the Fermi energy ϵ_f, the Fermi momentum p_f, or the Fermi wave-number $k_f = p_f/\hbar$. These quantities usually refer to the density, rather than to the chemical potential, though not necessarily always. That is, because the chemical potential changes as the temperature rises from zero to maintain a fixed density, it is somewhat arbitrary as to whether ϵ_f refers to the chemical potential, or whether it refers to the chemical potential one would have used if T were zero. Usually, the latter criterium is used for defining ϵ_f, p_f and k_f. That is, ϵ_f is always measured relative to the energy for which $\mathbf{p} = 0$, whereas μ is measured relative to the total energy which would include the potential. For instance, if there was a gas of density ρ, the Fermi momentum would be defined by:

$$\rho = (2s + 1)\frac{1}{(2\pi\hbar)^3}\frac{4\pi}{3}p_f^3, \tag{2.47}$$

and the Fermi energy (for the non-relativistic case) would be given by:

$$\epsilon_f = \frac{p_f^2}{2m}. \tag{2.48}$$

If there were an attractive potential of strength V, the chemical potential at zero temperature would then be $\epsilon_f - V$. If the temperature were raised while keeping ρ fixed, p_f and ϵ_f would keep constant, but μ would change to keep ρ fixed, consistent with Eq. (2.41).

Example 2.5:

Consider a two-dimensional non-relativistic gas of spin-1/2 Fermions of mass m at fixed chemical potential μ.
(A) For a small temperature T, find the change in density to order T^2.
(B) In terms of ρ and m, what is the energy per particle at $T = 0$
(C) In terms of ρ, m and T, what is the change in energy per particle to order T^2.

First, calculate the density of states for two dimensions,

$$D(\epsilon) = 2\frac{A}{(2\pi\hbar)^2}2\pi p\frac{dp}{d\epsilon} = \frac{A}{\pi\hbar^2}m, \tag{2.49}$$

which is independent of p_f. Thus, $dD/d\epsilon = 0$, and as can be seen from Eq. (2.41), the density does not change to order T^2.
The number of particles in area A is:

$$N = 2A\frac{1}{(2\pi\hbar)^2}\pi p_f^2 = \frac{A}{2\pi\hbar^2}p_f^2, \tag{2.50}$$

and the total energy of those particles at zero temperature is:

$$E = 2A\frac{1}{(2\pi\hbar)^2}\int_0^{p_f} 2\pi p_f dp_f \frac{p_f^2}{2m} = \frac{A}{8m\pi\hbar^2}p_f^4, \tag{2.51}$$

and the energy per particle becomes

$$\frac{E}{N}\bigg|_{T=0} = \frac{p_f^2}{4m} = \frac{\epsilon_f}{2}, \; p_f = \sqrt{2\pi\hbar^2\rho}. \tag{2.52}$$

From Eq. (2.42), the change in the energy is:

$$\delta E = D(\mu)\frac{\pi^2}{6}T^2 = A\frac{\pi m}{6\hbar^2}T^2, \tag{2.53}$$

and the energy per particle changes by an amount,

$$\frac{\delta E}{N} = \frac{\pi m}{6\hbar^2\rho}T^2 \tag{2.54}$$

2.5   Rotating Gases and Polarization

Every conserved quantity can be associated with a Lagrange multiplier. Angular momentum is no different. A rotating gas has angular momentum, which must choose how to partition itself between orbital angular momentum and spin angular momentum. Here, we derive the spin polarization for a gas rotating around the z axis with angular frequency $\vec{\omega} = \omega\hat{z}$. In such a system the collective velocity is

$$\vec{v}^{(\text{coll})} = \vec{\omega} \times \vec{r}. \tag{2.55}$$

Let the Lagrange multiplier related to angular momentum be γ. The single particle distribution (neglecting degeneracy effects) is

$$f(\vec{p}, \vec{r}) = \exp\left\{-\beta\frac{p^2}{2m} + \beta\mu - \gamma\left[S_z + \vec{r} \times \vec{p}\right]\right\}. \tag{2.56}$$

Using cylindrical coordinates, where $r_\perp \equiv \sqrt{x^2 + y^2}$, $v_\perp \equiv (x\dot{x} + y\dot{y})/r_\perp$, and $v_\phi = (x\dot{y} - y\dot{x})/r_\perp$,

$$\begin{aligned}
f(\vec{p}, \vec{r}) &= \exp\left\{-\beta m\frac{v_r^2 + v_\phi^2 + v_z^2}{2} + \beta\mu - \gamma\left[S_z + mr_\perp v_\phi\right]\right\} \\
&= \exp\left\{-\beta m\frac{v_r^2 + v_z^2}{2} - \beta m\frac{(v_\phi + (\gamma/\beta)r_\perp)^2}{2} + \beta m\frac{(\gamma/\beta)^2 r_\perp^2}{2} - \gamma S_z + \beta\mu\right\}.
\end{aligned} \tag{2.57}$$

For the distribution of velocities to be centered around the desired collective velocity, $v_\phi = \omega r_\perp$, one must choose γ so that

$$\frac{\gamma}{\beta} = -\omega. \tag{2.58}$$

Inserting this into Eq. (2.57),

$$f(\vec{p}, \vec{r}) = \exp\left\{-\frac{\beta m}{2}\left[v_r^2 + v_z^2 + (v_\phi - \omega r_\perp)^2\right] + \beta m\frac{\omega^2 r_\perp^2}{2} + \omega\beta S_z + \beta\mu\right\}. \tag{2.59}$$

The distribution depends on the spin projection through the factor

$$f(\vec{p}, \vec{r}) \sim e^{\vec{\omega} \cdot \vec{S}/T}, \tag{2.60}$$

and the spin aligns itself with $\vec{\omega}$. Finally, in addition to the distribution being centered around $v^{(\text{coll})}(\vec{r})$, there is an additional factor $e^{m\omega^2 r_\perp^2/2T}$. This is the centrifugal term, representing the fact that the rotation pushes the matter outward.

2.6   Grand Canonical vs. Canonical

In the macroscopic limit, the grand canonical ensemble is the easiest ensemble to analyze. However, when working with only a few particles, there are differences between the restriction that only those states with exactly N particles are considered, as opposed to the restriction that the average number is N. This difference can matter when the fluctuation of the number, which usually goes as $\sqrt{N}$, is not much less than N itself. To understand how this comes about, consider the grand canonical partition function,

$$Z_{GC}(\mu, T) = \sum_N e^{\mu N/T} Z_C(N, T). \tag{2.61}$$

Thermodynamic quantities depend on the log of the partition function. If the contributions to Z_{GC} come from a range of $\bar{N} \pm \delta N$, we can approximate Z_{GC} by,

$$\begin{aligned} Z_{GC}(\mu, T) &\approx \delta N e^{\mu \bar{N}/T} Z_C(\bar{N}, T) \\ \ln Z_{GC} &\approx \mu \bar{N}/T + \ln Z_C(\bar{N}, T) + \ln \delta N, \end{aligned} \tag{2.62}$$

and if one calculates the entropy from the two ensembles, $\ln Z_{GC} + \beta \bar{E} - \mu \bar{N}/T$ in the grand canonical ensemble and $\ln Z + \beta \bar{E}$ in the canonical ensemble, one will see that the entropy is greater in the grand canonical case by an amount,

$$S_{GC} \approx S_C + \ln \delta N. \tag{2.63}$$

The entropy is greater in the grand canonical ensemble because more states are being considered. However, in the limit that the system becomes macroscopic, S_C will rise proportional to N, and if one calculates the entropy per particle, S/N, the canonical and grand canonical results will become identical. For 100 particles, $\delta N \approx \sqrt{N} = 10$, and the entropy per particle differs only by $\ln 10/100 \approx 0.023$. As a comparison, the entropy per particle of a gas at room temperature is on the order of 20 and the entropy of a quantum gas of quarks and gluons at high temperature is between three and four units per particle.

2.7   Gibb's Paradox

Next, we consider the manifestations of particles being identical in the classical limit, i.e., the limit when there is little chance two particles are in the same single-particle level. We consider the partition function $Z_C(N, T)$ for N identical spinless particles,

$$Z_C(N, T) = \frac{z(T)^N}{N!}, \tag{2.64}$$

where $z(T)$ is the partition function of a single particle,

$$z(T) = \sum_p e^{-\beta \epsilon_p}. \tag{2.65}$$

In the macroscopic limit, the sum over momentum states can be replaced by an integral, $\sum_p \rightarrow [V/(2\pi\hbar)^3] \int d^3p$. The factor $1/N!$ accounts for the fact that if you multiply the N sums, you will get cross terms where the N particles are in N different states. For example, for $N = 4$ one cross term would be $e^{-\beta\epsilon_4 - \beta\epsilon_{11} - \beta\epsilon_2 - \beta\epsilon_{93}}$. However, there are $4!$ terms with the same combination of states, but with the numbers $4, 11, 2, 93$ rearranged. This counting assumes that one can neglect the cross terms where the same state appears more than once, which is justified if there are many more states (relevant states that have significant contributions) than there are particles, or equivalently, if the occupation probability of any given state is much less than unity.

If the particles were not identical, there would be no $1/N!$ factor in Eq. (2.64). This factor originates from the fact that if particle of type a is in level i and particle of type b is in level j, it is distinguishable from the state where a is in j and b is in i. This is related to Gibb's paradox, which concerns the entropy of two gases of non-identical particles separated by a partition. If the partition is lifted, the entropy of each particle, which has a contribution proportional to $\ln V$, will increase by an amount $\ln(V) - \ln(V/2) = \ln 2$. The total entropy will then increase by $N \ln 2$. However, if one states that the particles are identical and thus includes the $1/N!$ factor in Eq. (2.64), the change in the total entropy due to the $1/N!$ factor is, using Stirling's approximation,

$$\Delta S \text{ (from } 1/N! \text{ term)} = -\ln N! + 2\ln(N/2)! \approx -N \ln 2. \tag{2.66}$$

Here, the first term comes from the $1/N!$ term in the denominator of Eq. (2.64), and the second term comes from the consideration of the same terms for each of the two sub-volumes. This exactly cancels the $N \ln 2$ factors inherent to doubling the volume, and the total entropy is not affected by the removal of the partition. This underscores the importance of taking into account the indistinguishability of particles even when there is no significant quantum degeneracy.

To again demonstrate the effect of the particles being identical, we calculate the entropy for a three-dimensional classical gas of N particles in the canonical ensemble.

$$S = \ln Z_C + \beta E = N \ln z(T) - N \log(N) + N + \beta E, \tag{2.67}$$

$$z(T) = \frac{V}{(2\pi\hbar)^3} \int d^3p\, e^{-p^2/2mT} = V \left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}, \tag{2.68}$$

$$S/N = \ln\left[\frac{V}{N}\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\right] + \frac{5}{2}, \tag{2.69}$$

where the equipartition theorem was applied, $E/N = 3T/2$. The beauty of the expression is that the volume of the system does not matter, only V/N, the volume per identical particle. This means that when calculating the entropy per particle in a dilute gas, one can calculate the entropy of one particle in a volume equal to the subvolume required so that the average number of particles of the specific species in the volume is unity. Note that if the components of a gas have spin s, the volume per identical particle would be $(2s + 1)V/N$. Also, if there was a macroscopic volume with all identical particles except one, e.g. one oxygen molecule with a weird isotope, that one particle would contribute an amount of entropy proportional to $\ln(V)$, the logarithm of the macroscopic volume.

2.8   Iterative Techniques for the Canonical and Microcanonical Ensembles

The $1/N!$ factor must be changed if there is a non-negligible probability for two particles to be in the same state p. For instance, the factor z^N in Eq. (2.64) contains only one contribution for all N particles being in the ground state, so there should be no need to divide that contribution by $N!$ (if the particles are bosons, if they are Fermions one must disregard any contribution with more than 1 particle per state). To include the degeneracy for a gas of bosons or fermions, one can use an iterative procedure. First, assume you have correctly calculated the partition function for $N - 1$ particles. The partition function for N particles can be written as:

$$Z(N) = \frac{1}{N}\left\{ z_1 Z(N-1) \pm z_2 Z(N-2) + z_3 Z(N-3) \pm z_4 Z(N-4) \cdots \right\},$$
$$z_n \equiv \sum_p e^{-n\beta\epsilon_p}. \tag{2.70}$$

Here, the $\pm$ signs refer to Bosons/Fermions. If only the first term is considered, one recreates Eq. (2.64) after beginning with $Z(N = 0) = 1$ (only one way to arrange zero particles). Proof of this iterative relation can be accomplished by considering the trace of $e^{-\beta H}$ using symmetrized/anti-symmetrized wave functions (See Pratt, PRL**84**, p.4255, 2000).

If energy and particle number are conserved, one then works with the microcanonical ensemble. Because energy is a continuous variable, constraining the energy is usually somewhat awkward. However, for the harmonic oscillator energy levels are evenly spaced, which allows iterative treatments to be applied. If the energy levels are $0, \epsilon, 2\epsilon \cdots$, the number of ways, $N(A, E)$ to arrange A identical particles so that the total energy is $E = n\epsilon$ can be found by an iterative relation,

$$N(A, E) = \frac{1}{A}\sum_{a=1}^{A}\sum_i (\pm 1)^{a-1} N(A - a, E - a\epsilon_i), \tag{2.71}$$

where i denotes a specific quantum state. For the Fermionic case this relation can be used to solve a problem considered by Euler, "How many ways can a integers be arranged to sum to E without using the same number twice?"

Iterative techniques similar to the ones mentioned here can also incorporate conservation of multiple charges, angular momentum, and in the case of QCD, can even include the constraint that the overall state is a coherent color singlet (Pratt and Ruppert, PRC68, 024904, 2003).

2.9   Enforcing Canonical and Microcanonical Constraints through Integrating over Complex Chemical Potentials

For systems that are interactive, or for systems where quantum degeneracy plays an important role, grand canonical partition functions are usually easier to calculate. However, knowing the grand canonical partition function for all imaginary chemical potentials can lead one to the canonical partition function. First, consider the grand canonical partition function with an imaginary chemical potential, $\mu/T = i\theta$,

$$Z_{GC}(\mu = iT\theta, T) = \sum_i e^{-\beta E_i + i\theta Q_i}. \tag{2.72}$$

Using the fact that,

$$\delta_{Q,Q_i} = \frac{1}{2\pi} \int_0^{2\pi} d\theta\, e^{i(Q-Q_i)\theta},$$

(2.73)

one can see that,

$$\frac{1}{2\pi} \int_0^{2\pi} d\theta\, Z_{GC}(\mu = iT\theta, T)e^{-iQ\theta} = \sum_i e^{-\beta E_i}\delta_{Q,Q_i}$$

$$= Z_C(Q,T).$$

(2.74)

A similar expression can be derived for the microcanonical ensemble,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dz\, Z_C(T = i/z)e^{-izE} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz \sum_i e^{iz(\epsilon_i - E)}$$

$$= \sum_i \delta(E - \epsilon_i) = \rho(E).$$

(2.75)

As stated previously, the density of states plays the role of the partition function in the micro-canonical ensemble.

2.10   Problems

1. Consider classical non-relativistic particles acting through a spherically symmetric potential,

$$V(r) = V_0 \exp(r/\lambda).$$

Using the equipartion and virial theorems, show that

$$\left\langle \frac{r}{\lambda}V(r)\right\rangle = 3T.$$

2. Consider a relativistic ($\epsilon = \sqrt{m^2 + p^2}$) particle moving in one dimension,

   (a) Using the generalized equipartion theorem, show that

   $$\left\langle \frac{p^2}{\epsilon}\right\rangle = T.$$

   (b) Show the same result by explicitly performing the integrals in

   $$\left\langle \frac{p^2}{\epsilon}\right\rangle = \frac{\int dp\,(p^2/\epsilon)e^{-\epsilon/T}}{\int dp\, e^{-\epsilon/T}}.$$

   HINT: Integrate the numerator by parts.

3. Beginning with the expression for the pressure for a non-interacting gas of bosons,

$$\frac{PV}{T} = \ln Z_{GC} = \sum_p \ln \left(1 + e^{-\beta(\epsilon_p - \mu)} + e^{-2\beta(\epsilon_p - \mu)} + \cdots\right), \sum_p \to (2s+1)\frac{V}{(2\pi\hbar)^3}\int d^3p,$$

   show that
$$P = \frac{(2s+1)}{(2\pi\hbar)^3}\int d^3p\,\frac{p^2}{3\epsilon}f(\mathrm{p}), \text{ where } f = \frac{e^{-\beta(\epsilon - \mu)}}{1 - e^{-\beta(\epsilon - \mu)}}.$$

   Here, the energy is relativistic, $\epsilon = \sqrt{p^2 + m^2}$.

4. Derive the corresponding expression for Fermions in the previous problem.

5. Derive the corresponding expression for Bosons/Fermions in two dimensions in the previous problem. Note that in two dimension, P describes the work done per expanding by a unit area, $dW = PdA$.

6. For the two-dimensional problem above, show that P gives the rate at which momentum is transferred per unit length of the boundary of a 2-d confining box. Use simple kinematic considerations.

7. Consider a massless three-dimensional gas of bosons with spin degeneracy N_s. Assuming zero chemical potential, find the coefficients A and B for the expressions for the pressure and energy density,

$$P = AN_sT^4, \quad \left(\frac{E}{V}\right) = BN_sT^4$$

8. Show that if the previous problem is repeated for Fermions that:

$$A_{Fermions} = \frac{7}{8}A_{Bosons}, \quad B_{Fermions} = \frac{7}{8}B_{Bosons}.$$

9. Consider a three-dimensional solid at low temperature where both the longitudinal and transverse sound speeds are given by $c_s = 3000$ m/s. Calculate the ratio of specific heats,

$$\frac{C_V(\text{due to phonons})}{C_V(\text{due to photons})},$$

   where the photon calculation assumes the photons move in a vacuum of the same volume. Note that for sound waves the energy is $\epsilon = \hbar\omega = c_s\hbar k = c_s p$. For phonons, there are three polarizations (two transverse and one longitudinal), and because the temperature is low, one can ignore the Debye cutoff which excludes high momentum nodes, as their wavelengths are below the spacing of the crystal.
   Data: $c = 3.0 \times 10^8$ m/s, $\hbar = 1.05457266 \times 10^{-34}$ Js.

10. For a one-dimensional non-relativistic gas of spin-1/2 Fermions of mass m, find the change of the chemical potential $\delta\mu(T, \rho)$ necessary to maintain a constant density per unity length, ρ, while the temperature is raised from zero to T. Give answer to order T^2.

11. For a two-dimensional gas of spin-1/2 non-relativistic Fermions of mass m at low temperature, find both the quantities below:

$$\frac{d(E/A)}{dT}\bigg|_{\mu,V}, \qquad \frac{d(E/A)}{dT}\bigg|_{N,V}$$

Give both answers to the lowest non-zero order in T, providing the constants of proportionality in terms of the chemical potential at zero temperature and m.

12. The neutron star, PSR J1748-2446ad, discovered in 2004, spins at 716 times a second. Spin half particles have a spin angular momentum of $S_z = \pm\hbar/2$. If the neutron star has a temperature of 10^5 K, what is the polarization due to the spinning? $P = (n_\uparrow - n_\downarrow)/(n_\uparrow + n_\downarrow)$. Note that this neglects the polarization due to the magnetic field.

13. Consider harmonic oscillator levels, $0, \epsilon, 2\epsilon \cdots$, populated by A Fermions of the same spin.

    (a) Using the iterative relation in Eq. (2.71) solve for the number of ways, $N(A, E)$, to arrange A particles to a total energy E for all $A \leq 3$ and all $E \leq 6\epsilon$.

    (b) For $N(A = 3, E = 6\epsilon)$, list all the individual ways to arrange the particles.

14. Consider a two-dimensional gas of spinless non-degenerate non-relativistic particles of mass m.

    (a) Show that the grand canonical partition function $Z_{GC}(\mu, T)$ is

    $$Z_{GC}(\mu, T) = \exp\left\{e^{\mu/T}\frac{AmT}{2\pi\hbar^2}\right\}$$

    (b) Using Eq. (2.74) in Sec. 2.9 and the expression above, show that the canonical partition function $Z_C(N, T)$ is

    $$Z_C(N, T) = \frac{\left(\frac{AmT}{2\pi\hbar^2}\right)^N}{N!}$$

3   Interacting Gases and the Liquid-Gas Phase Transition

"Life is just a phase you're going through ... you'll get over it" – anonymous

3.1   Virial Expansion

The virial expansion is defined as an expansion of the pressure in powers of the density,

$$P = \rho T \left[A_1 + \sum_{n=2}^{\infty} A_n \left(\frac{\rho}{\rho_0} \right)^{n-1} \right], \quad \rho_0 \equiv \frac{(2s+1)}{(2\pi\hbar)^3} \int d^3p \, e^{-\epsilon_p/T}. \tag{3.1}$$

The leading term is always $A_1 = 1$, as low density matter always behaves as an ideal gas (here we assume non-relativistic gases). The contribution from interactions to the second-order term, A_2, is negative for attractive interactions and positive for repulsive interactions. Quantum statistics also affects A_2. For bosons, the contribution is negative, while for fermions it is positive. The negative contribution for bosons can be understood by considering the Bose-Einstein form for the phase-space occupancy, which provides a relatively stronger enhancement to low-momentum particles. This lowers the average momentum of particles colliding with the boundaries, which lowers the pressure relative to $P = \rho T$. Similarly, the contribution is positive for fermions, as the average momentum of a particle in a Fermi gas is raised by quantum statistics.

Like perturbation theory, virial expansions are only justifiable in the limit that they are unimportant, i.e., if the corrections are large you need more terms than you can calculate. For that reason, the second-order coefficient is often calculated carefully from first principles, but subsequent coefficients are often inserted phenomenologically. The virial expansion will be considered more rigorously in Section 3.5, where coefficients will be expressed in terms of phase shifts (only second order) and perturbation theory.

3.2   The Van der Waals Eq. of State

All equations of state, $P(\rho, T)$, should behave as $P = \rho T$ at low density. The second-order virial coefficient is usually negative, as long-range interactions are usually attractive and dominate at low density. For molecules, the long range interaction usually originates from induced dipoles attracting one another and falls off as $1/r^6$ at large r. At high densities, particles tend to run out of space and repel one another as molecules begin to physically overlap, and the contribution to the pressure from interactions tends to rise once again.

The Van der Waals parameterization of the equation of state incorporates both an attractive component for low density and repulsion at high density,

$$P = \frac{\rho T}{1 - \rho/\rho_s} - a\rho^2. \tag{3.2}$$

Here, the second order virial coefficient is

$$A_2 = \frac{\rho_0}{\rho_s} - \frac{a\rho_0}{T}, \tag{3.3}$$

and as long as a is large enough, the correction to the pressure to order ρ^2 is negative.

At high density, the first term in Eq. (3.2) will dominate, and in fact $P \to \infty$ as $\rho \to \rho_s$. The quantity ρ_s is the saturation density, as it is the highest density one can obtain before the pressure jumps to infinity. As the temperature goes to zero, a system will approach the saturation density if there is no pressure. The inverse of ρ_s is often referred to as the excluded volume. The Van der Waals equation of state thus implies a hard-core interaction because densities can never surpass ρ_s.

Example 3.1:

Using the Maxwell relation,

$$\frac{\partial(P/T)}{\partial\beta}\bigg|_{N,V} = -\frac{\partial E}{\partial V}\bigg|_{N,T},$$

derive the energy per particle E/N for a Van der Waals equation of state, as a function of ρ and T.

First,

$$\frac{\partial}{\partial\beta}\left(\frac{P}{T}\right) = \frac{\partial}{\partial\beta}\left(\frac{\rho}{1 - \rho/\rho_s} - \beta a\rho^2\right) = -a\rho^2.$$

Beggining with $E/N = 3/2T$ at $\rho = 0$,

$$\frac{E}{N} = \frac{3}{2}T + \frac{1}{N}\int_\infty^V dV\, a\rho^2 = \frac{3}{2}T - \int_0^\rho \frac{1}{\rho^2}d\rho\, a\rho^2 = \frac{3}{2}T - a\rho.$$

Remarkably, the energy per particle is independent of ρ_s.

If one plots P vs. V, the behavior is no longer monotonic for low T as seen in Fig. 3.1. For temperatures below the critical temperature,

$$T_c = \frac{8a\rho_s}{27}, \tag{3.4}$$

multiple densities can provide the same pressure. We refer to the inflection point as the critical point, at which $V_c = 3/\rho_s$, and $P_c = a\rho_s^2/27$. We will derive this expression shortly.

A region of densities for $T < T_c$ is unstable to phase separation. To find the region we consider the conditions for phase co-existence.

$$T_{\text{gas}} = T_{\text{liq}}, \quad P_{\text{gas}} = P_{\text{liq}}, \quad \mu_{\text{gas}} = \mu_{\text{liq}}. \tag{3.5}$$

The first two conditions can be satisfied by taking any two points on the same isotherm in Fig. 3.1 that are at the same pressure. The third condition, that the chemical potentials are equal, can also be illustrated with the isotherm. Beginning with

$$TdS = dE + PdV - \mu dN, \quad TS = PV + E - \mu N, \tag{3.6}$$
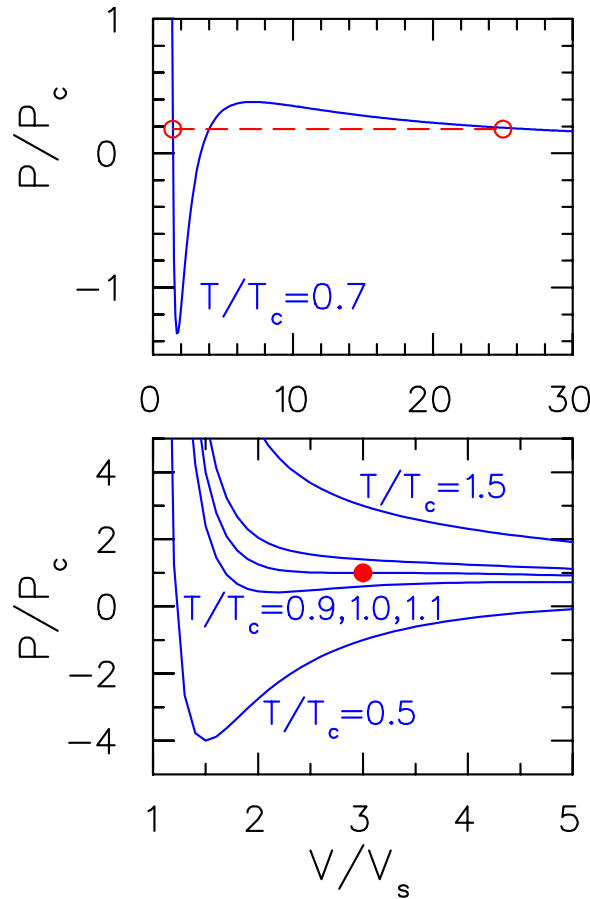
Figure 3.1: The lower panel illustrates P vs V for several isotherms for the Van der Waals equation of state. For $T < T_c$, the behavior is no longer monotonic and phase separation can occur. For $T = T_c$, there is an inflection point, where $d^2P/dV^2 = dP/dV = 0$. This is the critical point. In the upper panel, the $T = 0.7 \cdot T_c$ isotherm is shown along with the two points denoting the liquid and gas densities. Graphically, these points are found by requiring that the area between the isotherm and a horizontal line connecting between two points on the isotherm integrates to zero.

one can solve for $Nd\mu$,

$$
\begin{aligned}
d(E - TS + \mu N) &= -PdV - SdT - Nd\mu \\
&= -d(PV) \\
Nd\mu &= -SdT + VdP.
\end{aligned}
\tag{3.7}
$$

Following along an isotherm, $dT = 0$, and between two points illustrated in the upper panel of Fig. 3.1 the change in the chemical potential is

$$
\mu_{\text{gas}} - \mu_{\text{liq}} = \int vdP = \int d(Pv) - \int Pdv = P_{\text{gas}}v_{\text{gas}} - P_{\text{liq}}v_{\text{liq}} - \int_{v_{\text{liq}}}^{v_{\text{gas}}} Pdv = 0. \tag{3.8}
$$

Here, $v = V/N$ is the volume per particle. The condition that $\mu_{\text{liq}} = \mu_{\text{gas}}$ can then be stated graphically by stating that the integrated area between the P vs. V isotherm in the upper panel of Fig. 3.1 and the line between the two points is zero, i.e. there are equal amounts of area above and below the line.

As an example of coexistence, Fig. 3.2 shows how the coexistence region is remarkably similar for a variety of transitions when viewed after scaling the variables. Here, we make only a quick remark that the actual form of the coexistence curve, and the behavior of the matter, in the region near the critical point defies the simple Van der Waals picture. In that region, there are no longer two distinct phases and the density fluctuates wildly. Critical phenomena represents a field in itself, and will be discussed later in the course.
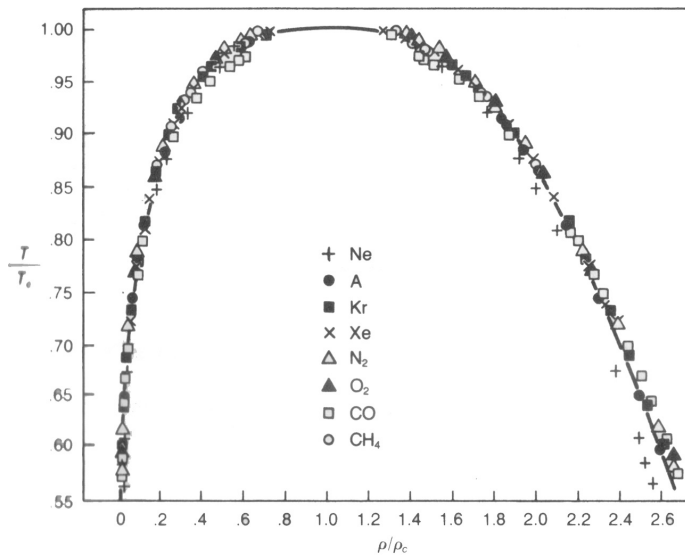
Figure 3.2: Phase coexistence for a variety of substances shows a remarkable universality after being scaled. The Van der Waals along with Maxwell constructions fails when working very close to T_c, but explains the behavior well for $T < T_c$. This figure was compiled by Guggenheim in 1945 (J.Chem).

Example 3.2:

Solve for T_c, P_c and V_c in the Van der Waals equation of state using the scaled variables, $p = P/a\rho_s^2$, $t = T/a\rho_s$, $v = \rho_s V/N = \rho_s/\rho$.

In terms of these variables,

$$p = \frac{t}{v - 1} - \frac{1}{v^2}.$$

For $T < T_c$ there are two points where $dP/dV = 0$. At one of these points, $d^2P/dV^2 < 0$. At the critical point,

$$\frac{dp}{dv} = \frac{d^2p}{dv^2} = 0.$$

If the second derivative were not zero, P vs. V would not be monotonic. These relations become:

$$\frac{t_c}{(v_c - 1)^2} = 2\frac{1}{v_c^3}, \quad 2\frac{t_c}{(v_c - 1)^3} = 6\frac{1}{v_c^4}.$$

This gives, $v_c = 3$ and $t_c = 8/27$. Plugging this into the expression for p, $p_c = 1/27$. Back in the usual variables,

$$\rho_c = \rho_s/3, \quad T_c = \frac{8}{27}a\rho_s, \quad P_c = \frac{1}{27}a\rho_s^2.$$

3.3   Clausius-Clapeyron Equation

The Clausius-Clapeyron equation allows us to plot the coexistence line in a P-V diagram if given the latent heat and the change of the volume per particle as one goes across the phase line on a P vs. T representation of the phase boundary:

$$\frac{dP}{dT} = \frac{L/T}{(V/N)_{\text{gas}} - (V/N)_{\text{liq}}} \tag{3.9}$$

The phase line can be found by inspecting the coexistence curve, e.g. Fig. 3.1, for the value of the coexistence pressure for a given isotherm.
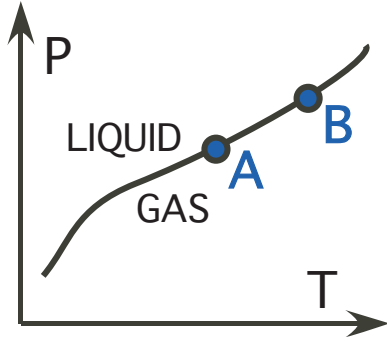
Figure 3.3: For phase coexistence, P, T and μ are continuous across the coexistence line. The requirement that $\mu^{(B)} - \mu^{(A)}$ is the same on both sides of the coexistence line is the starting point for proving the Clausius-Clapeyron equation.

To prove the relation consider the coexistence line in the P-V diagram sketched in Fig. 3.3. Because P, T and μ are continuous across the line,

$$\mu_{\text{liq}}^{(B)} - \mu_{\text{liq}}^{(A)} = \mu_{\text{gas}}^{(B)} - \mu_{\text{gas}}^{(A)}. \tag{3.10}$$

Because $N d\mu = -S dT + V dP$,

$$-\left(\frac{S}{N}\right)_{\text{liq}} dT + \left(\frac{V}{N}\right)_{\text{liq}} dP = -\left(\frac{S}{N}\right)_{\text{gas}} dT + \left(\frac{V}{N}\right)_{\text{gas}} dP, \tag{3.11}$$

$$\frac{dP}{dT} = \frac{(S/N)_{\text{gas}} - (S/N)_{\text{liq}}}{(V/N)_{\text{gas}} - (V/N)_{\text{liq}}}.$$

Given that the latent heat is $L = T\Delta S$, this proves the relation. Because L is $T\Delta S$ between the two adjacent points across the phase boundary on the P vs. T coexistence line, it can be identified with the heat added to change the phase if done either at constant T or at constant P.

3.4   Virial Expansion, Revisited

The grand canonical partition function, or equivalently the pressure, can then be expanded in powers of $e^{\beta\mu}$,

$$P/T = B_1 e^{\beta\mu} + B_2 e^{2\beta\mu} + \cdots B_n e^{n\beta\mu} + \cdots \tag{3.12}$$

Given the coefficients for this expansion, B_n, one can then find the coefficient A_n for the virial expansion, which is based on powers of the density.

The density, $\rho = \partial P/\partial\mu$, can also be expanded in powers of $e^{\beta\mu}$,

$$\rho = B_1 e^{\beta\mu} + 2B_2 e^{2\beta\mu} + \cdots n B_n e^{n\beta\mu} + \cdots \tag{3.13}$$

Repeating the virial expansion,

$$P = \rho T \left[A_1 + \sum_{n=2}^{\infty} A_n \left(\frac{\rho}{\rho_0}\right)^{n-1} \right], \quad \rho_0 \equiv B_1 = \frac{(2s+1)}{(2\pi\hbar)^3} \int d^3p \, e^{-\epsilon_p/T}. \tag{3.14}$$

By inspection, one can determine the first two virial coefficients, A_1 and A_2, in terms of the coefficients B_i,

$$A_1 = 1, \quad A_2 = -B_2/\rho_0. \tag{3.15}$$

The fact that A_2 and B_2 have opposite signs is easy to understand. An attractive interaction will increase both the density and the pressure by increasing the Boltzmann factor which behaves as $e^{-V/T}$. Because the density increase more due to the extra powers of n in Eq. (3.13), the ratio of P/ρ will fall for attractive interactions. Subsequent coefficients are more difficult to calculate. The next two are,

$$A_3 = -2B_3/\rho_0 + 4B_2^2/\rho_0^2, \quad A_4 = -3B_4/\rho_0 + 18B_2B_3/\rho_0^2 - 20B_2^3/\rho_0^3. \tag{3.16}$$

In Mayer's cluster theory, one can make a further reduction and show that the complex terms in $A_3, A_4 \cdots$ that include products of several powers of B can all be neglected if one only uses "multiply connected" graphs (See Pathria). This is also true for perturbation theory, see Chapter 7.

Like perturbation theory, virial expansions are only justifiable in the limit that they are unimportant, i.e., either only the first few terms matter. For that reason, second-order coefficient are often calculated carefully from first principles, but systems where the subsequent coefficients are necessary, such as liquids, are often treated phenomenologically, e.g. the Van der Waals equation of state.

3.5   Virial Coefficients from Phase Shifts

Quantum scattering theory can provide a simple means for calculating the second virial coefficient in terms of phase shifts. In scattering theory, the relative wave function is a distorted plane wave, where the plane wave $e^{iq \cdot r}$ is written as an expansion in partial waves,

$$e^{iq \cdot r} = \frac{1}{qr} \sum_{\ell} i^{\ell}(2\ell + 1)\phi_{\ell}(r)P_{\ell}(\cos\theta), \tag{3.17}$$

each of which is an eigenstate of angular momentum. Each partial wave can then be written as an incoming piece plus an outgoing piece. The effect of a spherically symmetric potential is to shift the phase of the outgoing part, leading to the following form at large r,

$$\phi_{\ell}(r \to \infty) \sim \frac{e^{i\delta_{\ell}}}{qr} \sin[qr/\hbar - \ell\pi + \delta_{\ell}(q)]. \tag{3.18}$$

If one considers the relative wave function to be confined to a spherical volume of radius R by an infinite potential, the boundary condition is:

$$\frac{qR}{\hbar} - \ell\pi + \delta_{\ell} = n\pi, \quad n = 1, 2, \cdots \tag{3.19}$$

The density of states is then,

$$\frac{dn}{dq} = \frac{(2\ell + 1)}{\pi}\left(R/\hbar + \frac{d\delta_{\ell}}{dq}\right), \quad \frac{dn}{d\epsilon} = \frac{(2\ell + 1)}{\pi}\left(Rm/q\hbar + \frac{d\delta_{\ell}}{d\epsilon}\right) \tag{3.20}$$

As compared to no interaction, the density of states is changed by an amount

$$\Delta \frac{dn}{d\epsilon} = \frac{(2\ell+1)}{\pi} \frac{d\delta_\ell}{d\epsilon} \tag{3.21}$$

The most transparent way to calculate the correction to the partition function is to consider the addition to the density of states as a new mass state. If the phase shift jumped by π in a small energy level with relative energy ϵ^* (which is what happens with a narrow resonance), the statistical effect would be identical to having a free particle with spin ℓ and mass $M = m_a + m_b$ and an additional Boltzmann weight $e^{-\epsilon^*/T}$. From that perspective, one can write the correction to $\ln Z$ as

$$\ln Z = \ln Z_0 + \frac{V}{(2\pi\hbar)^3} \int d^3P \, e^{-P^2/2MT + 2\mu/T} \sum_{\epsilon^*} e^{-\epsilon^*/T} \tag{3.22}$$

$$\ln Z = \ln Z_0 + \frac{V}{(2\pi\hbar)^3} \int d^3P \, e^{-P^2/2MT + 2\mu/T} \sum_\ell \int d\epsilon \, \frac{(2\ell+1)}{\pi} \frac{d\delta_\ell}{d\epsilon} e^{-\epsilon/T}$$

$$= \ln Z_0 + e^{2\mu/T} \frac{V(MT)^{3/2}}{(2\pi)^{3/2}\hbar^3} \sum_\ell \int d\epsilon \, \frac{(2\ell+1)}{\pi} \frac{d\delta_\ell}{d\epsilon} e^{-\epsilon/T}.$$

By comparison with the definition of the virial expansion coefficients,

$$A_2 = -2^{3/2} \sum_\ell \int d\epsilon \, \frac{(2\ell+1)}{\pi} \frac{d\delta_\ell}{d\epsilon} e^{-\epsilon/T}. \tag{3.23}$$

For bosons, the wave function is symmetrized, which eliminates any odd values for ℓ in the partial wave expansion. It also doubles the contribution for even values of ℓ. This adds an extra power of 2 to the above expression, and the sum would cover only even ℓ. For fermions, a factor of 2 is also added, but the sum covers only odd ℓ. In the real world, particles have spins, and because only pairs of the same spin projection are identical, both odd and even values of ℓ contribute. An example is nn scattering. The spin half neutrons can have their intrinsic spins combined as either a $S = 1$ triplet or a $S = 0$ singlet for which the spin component of the wave functions are symmetric or antisymmetric respectively. The spatial wave functions are then antisymmetric and symmetric respectively. Phase shifts are measured for both the triplet (only odd ℓ) and for the singlet (only even ℓ). In this case the virial coefficient becomes:

$$A_2 = -2^{5/2} \left\{ \sum_{\ell=0,2,4...} \int d\epsilon \, \frac{(2\ell+1)}{\pi} \frac{d\delta_{\ell,S=1}}{d\epsilon} e^{-\epsilon/T} + 3 \sum_{\ell=1,3,5...} \int d\epsilon \, \frac{(2\ell+1)}{\pi} \frac{d\delta_{\ell,S=0}}{d\epsilon} e^{-\epsilon/T} \right\}. \tag{3.24}$$

The real situation is a bit more complicated if one wants to accurately address the $S = 1$ terms. First, due to spin-orbit coupling, m_ℓ and m_S are not good quantum numbers, and one needs to add the quantum number J to label the phase shifts. Secondly, the tensor interaction mixes states with different ℓ if they have the same J (ℓ must differ by even number to conserve parity). These couplings are typical causes for headaches among nuclear physicists. Phase shifts for nucleon-nucleon scattering are displayed in Fig. 3.4.

Example 3.3:
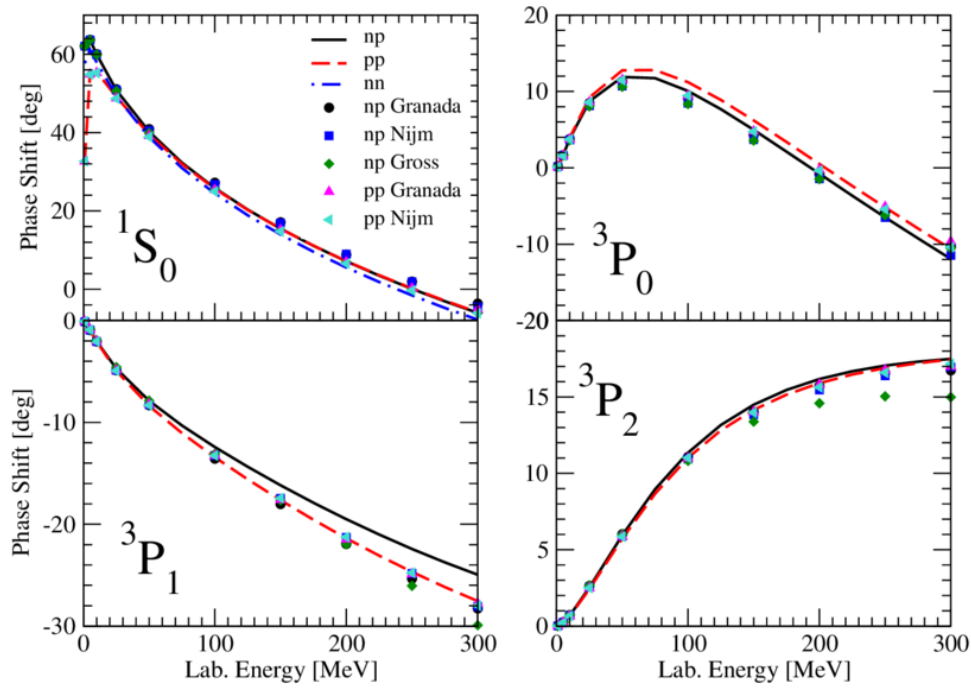
Figure 3.4: Phase shifts for proton-proton and proton-neutron scattering compared to calculations using a variety of nucleon-nucleon potentials, M. Piarulli et al., Phys. Rev. C **91** 024003.

The scattering length a parameterizes the phase shift at low energy as:

$$\delta_{\ell=0} = -ap/\hbar.$$

For neutrons, $a \approx -2 \times 10^{-14}$ m. Consider a neutron gas at a temperature of $T = 4 \times 10^{-21}$ Joules (room temperature), at what density will the correction to the pressure in a second-order virial expansion be of the same strength as the ρT. DATA: $\hbar = 1.05 \times 10^{-34}$ J·s, $m = 1.67 \times 10^{-27}$ kg. Ignore the spin when making the estimate.

First lets calculate the virial coefficient using Eq. (3.23),

$$
\begin{aligned}
A_2 &= -2^{3/2} \int d\epsilon\, \frac{1}{\pi}\frac{d\delta}{d\epsilon}e^{-\epsilon/T} = -2^{3/2} \int dp\, \frac{1}{\pi}\frac{d\delta}{dp}e^{-\epsilon/T} \\
&= 2^{3/2}\frac{a}{\pi\hbar} \int dp\, e^{-p^2/2mT} = 2^{3/2}\frac{a}{\hbar}\sqrt{\frac{mT}{2\pi}} \approx -6 \times 10^{-4}.
\end{aligned}
$$

Now, the first and second terms will be of the same order when,

$$-A_2\rho = \rho_0 = \frac{(mT)^{3/2}}{(2\pi)^{3/2}\hbar^3} \approx 10^{30} \text{ m}^{-3}.$$

Solving for ρ,

$$\rho \approx 2 \times 10^{33}.$$

The corresponding mass density, 3000 g/cm³, is hundreds of times higher than normal matter densities, but is very low compared to the density of the nucleus, $\approx 2 \times 10^{15}$ g/cm³. A realistic

calculation would have to account for the fact the neutrons are spin-half objects, and also include the effects of quantum degeneracy.

The phase shift expressions discussed here are not applicable once Coulomb is introduced into the problem. Coulomb forces are long range, hence the scattered waves never fully emerge from the influence of the force.

3.6   Problems

1. Consider a low density three-dimensional gas of non-relativistic spin-zero bosons of mass m at temperature $T = 1/\beta$ and chemical potential μ.

   (a) Find ρ_0 as defined in Eq. (3.1) in terms of m and T.

   (b) Expand the density ρ to second order in $e^{\beta\mu}$, i.e., to $e^{2\beta\mu}$. Express your answers for this part and the next two parts in terms of ρ_0.

   (c) Expand ρ^2 to second order in $e^{\beta\mu}$.

   (d) Expand $\delta P \equiv P - \rho T$ to second order in $e^{\beta\mu}$. (Hint: it is easier if you use the expression for P expanded in $f_0 = e^{\beta\mu-\beta\epsilon}$, i.e., $\ln(1 + f_0 + f_0^2 + f_0^3 \cdots) = -\ln(1 - f_0)$, then expand the logarithm in powers of f_0)

   (e) Determine the second virial coefficient defined by Eq. (3.1).

2. Consider the Van der Waals equation of state in scaled variables,

$$p = \frac{t}{v - 1} - \frac{1}{v^2},$$

   where $p = P/a\rho_s^2$, $v = V/V_s$, $t = T/a\rho_s$.

   (a) Derive the Maxwell relation,

$$\left.\frac{\partial(P/T)}{\partial\beta}\right|_{N,V} = -\left.\frac{\partial E}{\partial V}\right|_{N,T}.$$

   (b) Find the scaled energy per particle $e \equiv E/(a\rho_s N)$ as a function of v and t using the Maxwell relation above. Begin with the fact that $e = (3/2)t$ as $v \to \infty$.

   (c) Show that the change of entropy/particle $s = S/N$ between two values of v at a fixed temperature t is:
$$s_b - s_a = \ln[(v_b - 1)/(v_a - 1)].$$

   (d) Using the fact that $ts = e + pv - \mu$, show that

$$\mu_b - \mu_a = -\frac{2}{v_b} + \frac{2}{v_a} + t\left[\frac{v_b}{v_b - 1} - \frac{v_a}{v_a - 1}\right] - t\ln\left(\frac{v_b - 1}{v_a - 1}\right).$$

   (e) Show that as $t \to 0$, p_b will equal p_a if $v_b \to \infty$ and $v_a = 1 + t$. Then, show that in the same limit, μ_a will equal μ_b if $v_b = te^{1/t}$.

(f) Find the latent heat $L = t(s_b - s_a)$ for the small t limit. How does it compare with the minimum of e at $t = 0$?

(g) At $t = 0$, the system will have $p = 0$ in order to minimize the energy. Using the Clausius-Clapeyron equation, find dp/dt along the coexistence line at $t = 0$.

3. Using Eq. (3.23), calculate the second-order virial coefficient for a gas of distinguishable non-relativistic particles of mass m at temperature T that interact through a hard core potential,

$$V(r) = \begin{cases} \infty, & r < a \\ 0, & r > a \end{cases}$$

Consider only the s-wave contribution (valid at low T).

4   Dynamics of Liquids and Gases

Vision without action is a daydream. Action without vision is a nightmare. - Japanese proverb

4.1   "Adiabatic" Expansions

The science of thermodynamics was developed around engines. In particular, the constraints on the efficiency of engines related to entropy are of both intellectual and practical interest.

One begins with

$$dE = TdS - PdV + \mu dN, \tag{4.1}$$

where the term TdS is often referred to as the heat. Heat is not an energy, but is a change in energy, and one should not refer to the amount of heat in a gas or liquid – only the amount of heat added or removed from a gas or liquid. Adding energy without changing the particle number or the volume will raise the entropy. The most common method for adding heat is through thermal contact such as the microscopic transfer of energy between the molecules in a cylinder and the sides of the cylinder. Below, we will that associating the change in entropy depends on the assumption of a "slow" expansion.

An insulated gas is one for which there is no change in the energy aside from the work done on the gas. Thus, for an insulated gas, Eq. (4.1) implies that insulating the gas is equivalent to fixing the entropy and particle number. However, this is clearly not the case if the volume changes suddenly. If a piston is instantly moved so that the volume of a cylinder is doubled, the gas molecules will not have time to change their energy, and dE would be zero. In this case, the entropy of each molecule, which is proportional to $\ln[(mT)^{3/2}V]$, would eventually increase by an amount $\ln 2$ due to the increase in volume once the gas has re-equilibrated in the larger volume. Thus, an insulated system could have a change in entropy if the boundaries are suddenly changed. In fact, Eq. (4.1) assumes that the system changes slowly. Otherwise, the gas will not do the work PdV. Here, "slowly" requires that the boundaries move much more slowly than the molecular velocity of the gas.

One must remain cognizant of the meaning of several terms, some of which are often thrown about rather casually, with disregard for the confusion that might ensue from the carelessness. Some of the terms are:

1. *isentropic*: Entropy is conserved, $\Delta S = 0$.

2. *insulated*: No energy is transferred from thermal contact with the outside.

3. *adiabatic*: Different meanings to different people, and sometimes different meanings to the same person depending on the context. Sometimes means "slowly", sometimes means $\Delta S = 0$, and sometimes means "quickly".

4. *reversible*: Usually means $\Delta S = 0$.

The slowness considerations of the previous paragraph are the source of much confusion related to the definition of the word "adiabatic". In many instances (especially to engineers), adiabatic means that no heat enters a gas, $\Delta S = 0$. Whereas for many problems, adiabatic means that things change slowly. For instance, if reading a quantum mechanics book, one might encounter

the example of a particle in the ground state of a square well that is expanded slowly. The text might simply say that the well expands adiabatically. This might be contrasted with the sudden expansion of the well, and a question might be posed regarding the probability that the particle remains in the ground state after the well changed in both circumstances. For the particle in the well, it will remain in the ground state in a slow (adiabatic) expansion, but will find its probability split between many states in a sudden expansion. In the adiabatic case, the entropy remains zero because the system is in one state, whereas the entropy becomes finite for the sudden expansion. Thus, in this case slowness of the expansion and conserving entropy are synonymous and the term "adiabatic" is not ambiguous. However, this is not always the case. Some introductory physics books describe an adiabatic expansion as being rapid. The intent of such a statement is that in a rapid expansion, heat has no time to leak out of the container, so that entropy is conserved. However, it must not be rapid relative to the thermal velocity of the gas molecules, or one would create entropy because of the lack of slowness.

Example 4.1:

A gas of N_a indistinguishable, non-relativistic, spinless, non-degenerate, weakly interacting 'a' particles of mass m_a are confined to half of an insulated box by a barrier. Each half has a volume V. The other half of the box is a vacuum. The particles are thermalized at a temperature T when the barrier is removed suddenly.

a) What is the average energy and entropy per particle before the barrier is removed?
   – The energy per particle for a non-degenerate non-relativistic gas is $3T/2$. The entropy per particle is:

$$S/N = (\ln Z)/N + \beta E/N$$

$$Z = \frac{1}{N!}\left(\frac{V}{(2\pi\hbar)^3}\int d^3p\, e^{-p^2/2mT}\right)^N = \frac{1}{N!}V^N\left(\frac{mT}{2\pi\hbar^2}\right)^{3N/2},$$

   where Z, E and S are the partition function, average energy and entropy. Using $E/N = (3/2)T$ and $\ln(N!) \approx N\ln N - N$,

$$S/N = \frac{5}{2} + \ln\left[(V/N)\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\right].$$

b) What is the average energy and entropy per particle after the barrier is removed and the particles fill the entire volume?
   – The energy is the same and the entropy per particle increases by an amount $\ln 2$ due to the increase in volume.

c) The barrier is instead slowly pushed to the right like a piston. What is the average energy and entropy per particle be after the particles fill the entire volume?
   – In this case the entropy is fixed. If V/N doubles, $T^{3/2}$ must decrease by a factor of 2, which reduces T and E/N by a factor of $2^{-2/3}$.

d) The experiment is changed so that a separate type of 'b' particles populate the empty half of the box before the barrier is removed. The particles are at the same initial density and

temperature, and the masses and spins of the two types of particles are identical. What is the average energy and entropy per particle after the partition is suddenly removed and the gases have thoroughly mixed?

– The volume that plays a role in the entropy is the volume per **identical** particle. Though the density is the same, the volume per identical particle has doubled, thus increasing the entropy per particle by an amount **ln 2**. The energy per particle is unchanged.

4.2   Molecular Excitations and Thermodynamics

Atoms can be excited by moving electrons into higher-lying states. Molecules also have intrinsic excitations due to vibration, rotation or other internal degrees of freedom. Nuclei also have a rich excitation spectrum, including single-particle-like excitations as well as collective rotational and vibrational states. These states greatly affect the thermodynamic properties of the matter. For a single atom, the average internal energy per particle will be:

$$\langle E_{\text{int}}/N \rangle = -\frac{\partial}{\partial \beta} \ln z_{\text{int}} = \frac{\sum_\ell d_\ell \epsilon_\ell e^{-\epsilon_\ell/T}}{z_{\text{int}}}, \tag{4.2}$$

$$z_{\text{int}} \equiv \sum_\ell d_\ell e^{-\epsilon_\ell/T},$$

where d_ℓ is the degeneracy of the energy level ℓ.

For a spectrum of one-dimensional harmonic oscillator levels separated by $\hbar\omega$,

$$z_{\text{int}} = 1 + e^{-\beta\hbar\omega} + e^{-2\beta\hbar\omega} + e^{-3\beta\hbar\omega} \cdots \tag{4.3}$$

$$= \frac{1}{1 - e^{-\beta\hbar\omega}}.$$

The average internal energy is then,

$$\langle E_{\text{int}}/N \rangle = \hbar\omega \frac{e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}, \tag{4.4}$$

which looks exactly like what we had for bosons in specific energy level because those energies were also evenly spaced, $0, \epsilon, 2\epsilon \cdots$.

Taking the high and low T limits in Eq. (4.4),

$$\langle E_{\text{int}}/N \rangle = \begin{cases} \hbar\omega e^{-\hbar\omega/T}, & T << \hbar\omega \\ T, & T >> \hbar\omega \end{cases} \tag{4.5}$$

Thus, the excitations can be neglected if the temperature is much less than the energy required to excite the lowest level. If the temperature is much higher than $\hbar\omega$, the harmonic excitations contribute just as one would expect from considering the equipartition theorem result for a classical harmonic oscillator, where the equipartition theorem states that for each degree of freedom that appears quadratically in the Hamiltonian, an amount of energy $T/2$ is added to the system. For

a one-dimensional harmonic oscillator, both x and p_x contribute $T/2$. Vibrational excitations of diatomic molecules tend to behave like a one-dimensional harmonic oscillator, contributing one unit of T per particle for temperatures much higher than $\hbar\omega$. For most diatomic molecules, $\hbar\omega$ is thousands of degrees K, which means that vibrational excitations can be largely ignored at room temperature.

Any rotating object whose position can be denoted by one unit vector (or direction θ and ϕ), and whose energy is independent of such energies, will have a rotational spectra,

$$\epsilon_{\text{rot}} = \frac{\hbar^2 \ell(\ell+1)}{2I}, \tag{4.6}$$

where I is the moment of inertia of the molecule. Because the orientation of a diatomic molecule is determined by a single unit vector, they will have such a spectra. Along with the degeneracy of $(2\ell+1)$, the internal partition function is given by the sum,

$$z_{\text{rot}} = \sum_{\ell=0}^{\infty} (2\ell+1) \exp\left\{-\beta\hbar^2\ell(\ell+1)/2I\right\}. \tag{4.7}$$

For $T << \hbar^2/I$, only the $\ell = 0$ contribution matters and $z_{\text{rot}} \to 1$. For $T >> \hbar^2/I$, many terms in the sum contribute, which allows the sum to be approximated by an integral,

$$\begin{aligned}
z_{\text{rot}}(T >> \hbar^2/I) &= 2\int d\ell\, \ell e^{-\ell^2/(2IT/\hbar^2)} \tag{4.8}\\
&= 2IT/\hbar^2.
\end{aligned}$$

The energy $(= -\partial \ln z/\partial\beta)$ then becomes

$$\langle E_{\text{rot}}/N\rangle|_{T>>\hbar^2/I} = T. \tag{4.9}$$

The characteristic rotational energy $\hbar^2/2I$ is below room temperature for most molecules. Thus, for a dilute gas of diatomic molecules, the energy per particle is:

$$\left\langle \frac{E}{N} \right\rangle \approx \begin{cases} 3T/2, & T << \hbar^2/2I \\ 5T/2, & \hbar^2/2I << T << \hbar\omega_{\text{vibrational}} \\ 7T/2, & T >> \hbar\omega_{\text{vibrational}} \end{cases} \tag{4.10}$$

The specific heat per particle is defined in terms of the heat required to raise the temperature,

$$C_V \equiv \frac{1}{N}T\frac{dS}{dT}\bigg|_{N,V}, \quad C_P \equiv \frac{1}{N}T\frac{dS}{dT}\bigg|_{N,P}. \tag{4.11}$$

Using the fundamental thermodynamic relation, $TdS = dE + PdV - \mu dN$,

$$TdS|_{N,V} = dE, \tag{4.12}$$

and because the energy of the gas depends only on the temperature, $E/N = 3T/2, 5T/2, 7T/2$, C_V for a gas will be of the form

$$C_V = 3/2,\ 5/2,\ 7/2. \tag{4.13}$$

Calculating C_P is not quite so easy,

$$TdS|_{P,N} = dE + PdV. \tag{4.14}$$

Again the dE term is no trouble for a gas because the energy per particle is a function of the temperature only, but one needs to re-express PdV in terms of dT. To do this, one differentiates the ideal gas law, $PV = NT$ at fixed P and N to obtain:

$$PdV|_{P,N} = NdT. \tag{4.15}$$

Thus at constant pressure the fundamental thermodynamic relations becomes,

$$TdS|_{P,N} = (dE + PdV) = N(C_V + 1)\,dT, \tag{4.16}$$
$$C_P = C_V + 1. \tag{4.17}$$

For the three limits above, $C_P = 5/2, 7/2, 9/2$. If C_P and C_V are defined as TdS/dT, not per particle, the relation becomes

$$C_P = C_V + N. \tag{4.18}$$

Example 4.2:

Consider an isentropic expansion for a dilute gas of O_2 molecules near room temperature, where the volume changes from V_a to V_b.

- Find the final temperature T_b in terms of the volumes and the original temperature T_a.

  Using conservation of entropy and particle number, the fundamental thermodynamic relation gives:
  $$TdS = 0 = dE + PdV = N(C_V dT + TdV/V),$$
  with $C_V = 5/2$. This becomes

  $$(5/2) \int_{T_a}^{T_b} dT/T = -\int_{V_a}^{V_b} dV/V,$$
  $$T_b = T_a \left(\frac{V_b}{V_a}\right)^{-2/5}.$$

- Find the final pressure P_b in terms of the original pressure P_a and V_a and V_b.

  Using the ideal gas law, $P = NT/V$,

  $$\frac{P_b}{P_a} = \frac{T_b}{T_a}\frac{V_a}{V_b} = \left(\frac{V_b}{V_a}\right)^{-7/5}.$$

  Often the ratio C_P/C_V is referred to as γ, which in this case is $7/5$. The above equation can then be equivalently stated as

  $$PV^\gamma|_{N,S} = \text{constant}, \quad \gamma = C_P/C_V.$$

4.3   Efficiency of Engines: The Third Law of Thermodynamics

The third law of thermodynamics is nothing more than a statement that the entropy of the entire system must increase. For a closed amount of gas that is heated, where the gas does work as it expands, there is a limit on the efficiency of the engine. Because the cycle returns to its original volume and temperature, the net work done by the gas must equal the net heat added during the cycle. Similarly, because the entropy must return to its original value, entropy must leave the system during part of the cycle to compensate for the entropy gained during the part of the cycle where heat is added. The exhaust heat is

$$Q_X = T_X \Delta S, \tag{4.19}$$

where ΔS is determined by the heat added during the heating portion of the cycle,

$$T_H \Delta S = Q_H. \tag{4.20}$$

Because the net ΔS in a closed cycle will be zero, if one wishes to minimize the exhaust heat and maximize the heat flowing in, one should design a cycle so that the heat flows in while the cycle has its highest temperature and is exhausted at the coldest portion of the cycle. By having the entropy leave when the temperature is as low as possible, Q_X will be as small as possible, and by having the heat enter when T_H is at its highest value, ΔS will be as small as possible, which also helps minimize Q_X. All other parts of the cycle should follow isentropic paths. Such a cycle is known as the Carnot cycle, and has four parts as illustrated in Fig. 4.1.
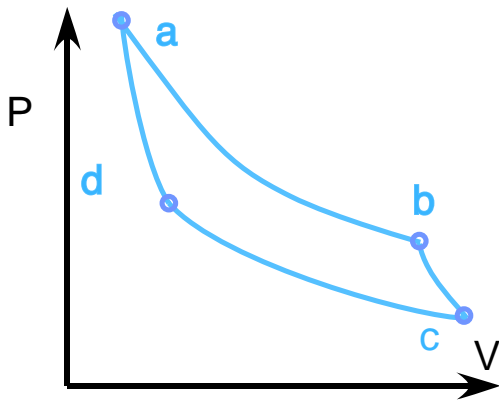
Figure 4.1: Illustration of Carnot Cycle.
(a-b) From a hot and compressed stage characterized by T_H, P_a and V_a, the system expands isothermally to P_b, and V_b. Because the temperature is constant, $P_b V_b = P_a V_a$. Heat is added during this stage.
(b-c) At fixed entropy, the system is cooled to a lower temperature T_X. For this segment, $P_c V_c^\gamma = P_b V_b^\gamma$.
(c-d) The gas is compressed isothermally at temperature T_X until it returns to the original entropy, at which point the pressure and volume are P_d and V_d. Heat is exhausted during this stage.
(d-a) At fixed entropy, the gas is compressed back to its original temperature T_H, pressure and volume. To conserve entropy, $P_d V_d^\gamma = P_a V_a^\gamma$.

The efficiency of a Carnot engine is defined as the work, $W = Q_H - Q_X$, divided by the input energy Q_H,

$$e = \frac{W}{Q_H} = \frac{Q_H - Q_X}{Q_H}. \tag{4.21}$$

Because entropy is only added at the points where heat is added, and because the entropy added at T_H equals the entropy lost at T_X,

$$e = \frac{(T_H - T_X)\Delta S}{T_H \Delta S} = 1 - \frac{T_X}{T_H}. \tag{4.22}$$

The maximum efficiency is thus determined by the temperatures at which heat enters and is exhausted during the cycle. For that reason, diesel engines, which burn hotter than gasoline engines, are more efficient.

Example 4.3:

A Carnot engine uses a gas whose ratio of specific heats is $\gamma = C_P/C_V$. Express the efficiency in terms of γ and the ratio of volumes V_c/V_b from Fig. 4.1.

Because the efficiency is given in terms of the ratio of the temperatures, one must express T_H/T_X in terms of the volumes. For the isentropic expansion from $b - c$,

$$P_b V_b^\gamma = P_c V_c^\gamma,$$

which when combined with the ideal gas law, $PV = NT$,

$$T_H V_b^{\gamma-1} = T_X V_c^{\gamma-1}, \quad \text{or} \quad \frac{T_H}{T_X} = \left(\frac{V_c}{V_b}\right)^{\gamma-1}.$$

This gives

$$e = 1 - \frac{T_X}{T_H} = 1 - \left(\frac{V_b}{V_c}\right)^{\gamma-1}.$$

The efficiency thus improves for a higher compression engine (though the compression that appears here is only that between the b and c point on the cycle). Whereas the compression in a gasoline engine is typically between 7 and 11, diesel engines have compression ratios from 14 to 25. The best gasoline powered engines (gasoline direct injection) approach efficiencies of 35%, whereas the best diesel engines can often reach 45% efficiency. However, due to not running at peak conditions, efficiencies are typically quite a bit less. Power plants burning fossil fuels with large turbines typically maintain efficiencies near 40% while generating electricity. By building plants that operate at higher pressures (essentially higher compressions), the efficiencies increase. Whereas most power plants create steam ~ 200 bar and achieve efficiencies near 40%, super-critical-pressure power plants operate near 300 bar and create power with $\gtrsim 45\%$ efficiency.

4.4   Hydrodynamics

Hydrodynamics explains the dynamics of gases and fluids in the limit that the fluid remains locally equilibrated during the motion. Mathematically, this condition can be stated in terms of the characteristic times associated with the fluid's expansion compared to the characteristic times associated with microscopic equilibration. For instance, the atoms in a gas or fluid will kinetically equilibrate after a few collisions, thus the relevant microscopic time scale is the collision time τ_{coll}. For an isotropic expansion, the expansion rate can be defined as

$$\frac{1}{\tau_{\text{exp}}} \sim \left|\frac{\dot{\rho}}{\rho}\right|, \tag{4.23}$$

where ρ represents any conserved charge or number density. Hydrodynamics is valid in the limit $\tau_{\exp} >> \tau_{\mathrm{coll}}$. The expansion rate can also be related to the divergence of velocity through the equation of continuity,

$$\frac{D}{Dt}\rho = -\rho\nabla\cdot\vec{v}, \quad \frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \vec{v}\cdot\nabla, \tag{4.24}$$

Here $D/Dt = \partial_t + \vec{v}\cdot\nabla$ is the rate of change as observed by an co-mover, i.e., one whose velocity is the same as that of the fluid element. The expansion rate is then equal to the divergence of the velocity,

$$1/\tau_{\exp} = \nabla\cdot\vec{v}. \tag{4.25}$$

For anisotropic expansions, one might have a condition where $\nabla\cdot\vec{v} = 0$ even though the various components $\partial_i v_j \neq 0$, e.g. collapsing along the x axis while expanding along the y axis. The characteristic expansion time would then be defined by the largest components of the velocity gradient, $|\partial_i v_j|$. Ideal hydrodynamics applies when $\tau_{\exp}^{-1}$ is much much less than the collision rate, $\tau_{\mathrm{coll}}^{-1}$. When $\tau_{\exp}^{-1}$ is less than $\tau_{\mathrm{coll}}^{-1}$, but not much much less than the collision rate, one can apply viscous hydrodynamics, which we consider later. Ideal hydrodynamics is based on the assumption that velocity gradients are always sufficiently small that the matter is always locally equilibrated.

Ideal hydrodynamics involves solving three equations simultaneously. The first equation describes the acceleration of the matter,

$$\frac{D\vec{v}}{Dt} = -\frac{1}{\rho_m}\nabla P, \tag{4.26}$$

where ρ_m is the mass density. If one considers a slab of volume $\delta V = A\delta x$, and multiplies both sides of Eq. (4.26) for the xth component by $\rho_m \delta V$, the l.h.s. becomes ma_x and the r.h.s. becomes $P(x)A - P(x+\delta x)A$, which is the net force pushing on the volume element in the x direction. Thus, this expression is simply the differential expression of $\vec{F} = m\vec{a}$, where the force is the pressure multiplied by the area.

The second equation expresses the fact that entropy is conserved, or equivalently that $dU = -PdV$ where U is the thermal energy,

$$\frac{DU}{Dt} = -P\frac{DV}{Dt}. \tag{4.27}$$

If one considers a cubic volume element, $V = L_x L_y L_z$, one can derive that

$$\begin{aligned}
dV &= dL_x L_y L_z + L_x dL_y L_z + L_x L_y dL_z \\
dL_x &= dt[v_x(x=L_x) - v_i(x=0)] = dt L_x \partial_x v_x, \\
\frac{DV}{Dt} &= V\nabla\cdot\vec{v}.
\end{aligned} \tag{4.28}$$

This gives

$$\begin{aligned}
\frac{DU}{Dt} &= -PV\nabla\cdot\vec{v} \\
\frac{D(U/V)}{Dt} &= -(P + U/V)\nabla\cdot\vec{v} \\
\frac{D\epsilon}{Dt} &= -(P + \epsilon)\nabla\cdot\vec{v},
\end{aligned} \tag{4.29}$$

where the final equation simply uses $\epsilon = U/V$ to refer to the energy density as determined in the frame of the matter.

The third equation is simply the equation of continuity for conserved charges. Summarizing, hydrodynamic evolution is determined by three equations:

$$\frac{D\vec{v}}{Dt} = -\frac{1}{\rho_m}\boldsymbol{\nabla}P \tag{4.30}$$

$$\frac{D\epsilon}{Dt} = -(P+\epsilon)\boldsymbol{\nabla}\cdot\vec{v},$$

$$\frac{D\rho}{Dt} = -\rho\boldsymbol{\nabla}\cdot\vec{v},$$

$$\tag{4.31}$$

Respectively, the three equations express: (1) that the net force on an element is the integrated pressure pushing on its surface, (2) entropy conservation and (3) particle number conservation. The co-moving derivative, $D/Dt = \partial_t + \vec{v}\cdot\boldsymbol{\nabla}$, represents the time derivative as viewed by an observer in the frame where $\vec{v} = 0$. When solving the hydrodynamic equations of motion, there are three variables, ρ, ϵ and $\vec{v}$, which one integrates forward with the three equations of motion. The pressure and mass densities are not free variables as P is uniquely determined by ρ and ϵ, and ρ_m is determined by ρ. In many cases, there are multiple conserved currents. In that case one writes conservation laws for a vector of charge densities ρ.

Viscous hydrodynamics incorporates corrections proportional to the expansion rate divided by the equilibration rate. Because the expansion rate is proportional to the derivative of the velocity, the acceleration from Eq. (4.26) is expanded to include terms which are linear in derivatives of the velocity such as:

$$\frac{Dv_i}{Dt} = -\frac{1}{\rho_m}\left\{\partial_i P - \partial_j[\eta\omega_{ij}] - \partial_i[\zeta\boldsymbol{\nabla}\cdot\vec{v}] + \frac{D}{Dt}[\kappa\partial_i T]\right\}, \tag{4.32}$$

$$\omega_{ij} \equiv \partial_j v_i + \partial_i v_j - (2/3)\delta_{ij}(\boldsymbol{\nabla}\cdot\vec{v}).$$

The equation expressing entropy conservation is also altered:

$$\frac{D\epsilon}{Dt} = -(P+\epsilon)\boldsymbol{\nabla}\cdot\vec{v} + \eta\sum_{ij}\omega_{ij}^2 + \zeta(\boldsymbol{\nabla}\cdot\vec{v})^2 + \kappa\boldsymbol{\nabla}^2 T. \tag{4.33}$$

These equations are known as Navier-Stokes hydrodynamics. The coefficient ζ is the referred to as the bulk viscosity. In an expanding system it effectively lowers the pressure proportional to $\boldsymbol{\nabla}\cdot\vec{v}$. The shear viscosity η only manifests itself when the velocity gradients are anisotropic, as $\omega = 0$ for a rotationally symmetric velocity gradient, e.g., $\vec{v} = \vec{v}_0 + a\vec{r}$. The coefficient κ is the heat conductivity and comes from the fact that once the system is not equilibrated, the energy might flow with a different velocity than the particle number. This term appears in variety of forms in the literature depending on whether the velocity $\vec{v}$ is that of the particles (Eckart frame) or that of the energy (Landau frame). For radiatively dominated systems it is more practical to use the energy and momentum density to define the velocity, i.e., the velocity is that required to make the momentum density zero. In the limit where temperatures are very high, this last term can be neglected.

Example 4.4:

Consider matter which initially has a uniform temperature T_0 and a three-dimensional density profile,

$$\rho(r, t = 0) = \rho_0 e^{-r^2/2R_0^2},$$

and expands according to an equation of state $P = \rho T$ according to ideal hydrodynamics.

- Show that if the subsequent evolution of the density and temperature are parameterized by

$$\rho(r, t) = \rho_0 \frac{R_0^3}{R^3(t)} e^{-r^2/2R(t)^2}, \quad T(r, t) = T(t)$$

that entropy will be conserved if

$$R^2(t) T(t) = R_0^2 T_0.$$

The entropy per particle, $\sigma = S/N$, can be calculated by considering a single particle in a volume V/N in the canonical ensemble (see discussion of Gibb's paradox in chapter 2),

$$\sigma = \frac{(\beta E + \ln Z_C)}{N} = (3/2) + \ln\left[\frac{(V/N)}{(2\pi\hbar)^3} \int d^3p e^{-\beta\epsilon_p}\right] = 5/2 + \ln\left[\frac{1}{\rho}\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\right],$$

The total entropy is then

$$
\begin{aligned}
S(t) &= \int d^3r \rho(r)\sigma(r) = \text{constant} + \int d^3r\ \rho(r) \ln\left[T^{3/2}/\rho\right] \\
&= \text{constant}' + \int d^3r\ \rho(r)\left\{\ln(T^{3/2}R^3) + r^2/2R^2\right\} \\
&= \text{constant}' + (N/2)\langle r^2/R^2\rangle + N\ln(T^{3/2}R^3) \\
&= \text{constant}'' + N\ln(T^{3/2}R^3).
\end{aligned}
$$

In the second to the last step, the average $\langle r^2/R^2\rangle = 3$, and was absorbed into the constant. The remaining time dependence comes from the fact that T and R are functions of time. Thus the total entropy will be constant if R^2T is constant.

- Assuming that the velocity profile is linear,

$$\vec{v}(r, t) = A(t)\vec{r},$$

find $A(t)$ and $R(t)$ that satisfy the hydrodynamic equations of motion and current conservation.

Current conservation leads to the expression,

$$
\begin{aligned}
0 &= \frac{\partial\rho}{\partial t} + v\frac{\partial\rho}{\partial r} + \rho\left(\frac{\partial v}{\partial r} + 2\frac{v}{r}\right) \\
&= \rho\left\{-\frac{3\dot{R}}{R} + \frac{r^2}{R^3}\dot{R} - Ar\frac{r}{R^2} + 3A\right\},
\end{aligned}
$$

which fortunately can be satisfied for all r if

$$\dot{R}(t) = A(t)R(t).$$

Next we consider the equations of motion,

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial r} = \frac{-1}{m\rho}\frac{\partial P}{\partial r} = \frac{-T}{m\rho}\frac{\partial \rho}{\partial r}$$

$$\dot{A}r + A^2 r = \frac{Tr}{mR^2}.$$

Again, the equations of motion can satisfied for all r if

$$\dot{A}(t) + A^2(t) = \frac{T(t)}{mR^2(t)}$$

After substituting for T using the fact that $R^2 T = R_0^2 T_0$, we have two equations for the evolution of A and R,

$$\dot{R}(t) = A(t)R(t), \quad \dot{A}(t) + A^2(t) = \frac{T_0}{m}\frac{R_0^2}{R^4(t)}.$$

Using the first equation to eliminate A and $\dot{A}$ in the second equation,

$$m\ddot{R}(t) = \frac{R_0^2 T_0}{R^3(t)}.$$

This can be solved by noting that the term on the right looks like a force affecting a particle of mass m at position R where the potential is $V(R) = R_0^2 T_0/2R^2$. Given the potential one can solve for the trajectory with

$$t = \int_{R_0}^{R}\frac{dx}{v} = \int_{R_0}^{R}\frac{dx}{\sqrt{2(E - V(x))/m}}.$$

The subsequent integration yields:

$$R(t)^2 = R_0^2 + \left(\frac{T_0}{m}\right)t^2, \quad A(t) = \left(\frac{T_0}{m}\right)\frac{t}{R^2(t)}.$$

The temperature can then be found by $T = T_0 R_0^2/R^2$.

4.5   Relativistic Hydrodynamics

Some of the most important applications of hydrodynamics involve relativistic motion. These include astronomical applications as well as some laboratory applications such as relativistic heavy ion collisions. In some cases, such as black holes, general relativity comes into play as

well. Hydrodynamics can incorporate special relativity in a most elegant fashion by considering the stress energy tensor,

$$T^{\alpha\beta} = (P + \epsilon)u^\alpha u^\beta - Pg^{\alpha\beta}, \tag{4.34}$$

where ϵ is the energy density in the frame of the fluid $\epsilon = U/V$, and u^α is the relativistic four velocity. The equations of motion for hydrodynamics and for entropy conservation are expressed as

$$\partial_\alpha T^{\alpha\beta} = 0. \tag{4.35}$$

This is effectively four separate conservation equations, one for each value of β. To see that these represent the equations of motion for hydrodynamics, one can view the system in a frame where the velocities are small, and if they reproduce the non-relativistic equations in that frame, they will be correct in all frames because they have the correct Lorentz form. To see this, we write the 4×4 stress-energy tensor for the limit where $u = (1, v_x, v_y, v_z)$,

$$T^{\alpha\beta} = \begin{pmatrix} \epsilon & (P+\epsilon)v_x & (P+\epsilon)v_y & (P+\epsilon)v_z \\ (P+\epsilon)v_x & -P & 0 & 0 \\ (P+\epsilon)v_y & 0 & -P & 0 \\ (P+\epsilon)v_z & 0 & 0 & -P \end{pmatrix} \tag{4.36}$$

By inspection one can see that the $\partial_\alpha T^{\alpha 0} = 0$ leads to the equation for entropy conservation, Eq. (4.29). The other three equations, $\partial_\alpha T^{\alpha i} = 0$, become,

$$\frac{\partial \vec{v}}{\partial t} = -\frac{1}{P+\epsilon} \nabla P. \tag{4.37}$$

In the non-relativistic limit $P + \epsilon \sim \rho_m$, and these equations reproduce the non-relativisitic expression for acceleration. Finally, after writing the equation for current conservation relativistically, $\partial_\alpha(\rho u^\alpha) = 0$, all the non-relativistic equations are satisfied, and if they are satisfied in one frame, they must be satisfied in all frames given that they have a manifestly consistent Lorentz structure.

One might ask about whether the energy density includes the vacuum energy. In that case there is also a vacuum energy. By "vacuum", one refers to the energy density at zero temperature and particle density. In that case, the pressure is

$$P_{\text{vac}} = -\frac{dE_{\text{vac}}}{dV} = -\frac{d(\epsilon_{\text{vac}}V)}{dV} = -\epsilon_{\text{vac}}, \tag{4.38}$$

where ϵ_{vac} and P_{vac} denote the vacuum energy density and pressure. Thus, the $P + \epsilon$ term in relativistic hydrodynamics is insensitive to the vacuum energy density.

4.6   Hydrodynamics and Sound

To see how a sound wave is a hydrodynamic wave, we linearize the equations of motion and the equation of continuity, keeping terms only of first order in the velocity $\vec{v}$, and first order in the deviation of the mass density, $\delta\rho_m$ ($\rho_m = \rho_m^{(0)} + \delta\rho_m$),

$$\frac{\partial \vec{v}}{\partial t} = -\frac{(dP/d\rho_m)|_S}{\rho_m^{(0)}} \nabla \delta\rho_m, \tag{4.39}$$

$$\frac{\partial \delta\rho_m}{\partial t} = -\rho_m^{(0)} \nabla \cdot \vec{v}.$$

Through substitution, one can eliminate the $\vec{v}$ in the equations by taking the divergence of the upper equation and comparing to the ∂_t of the second equation.

$$\frac{\partial^2}{\partial t^2}\delta\rho_m = \left.\frac{dP}{d\rho_m}\right|_S \nabla^2\delta\rho_m. \tag{4.40}$$

The solution to this equation is a cosine wave,

$$\delta\rho_m(\vec{r}, t) = A\cos(\vec{k}\cdot\vec{r} - \omega t + \phi), \tag{4.41}$$

where ϕ is an arbitrary phase and the speed of the wave is

$$c_s = \omega/k = \sqrt{\left.\frac{dP}{d\rho_m}\right|_S}. \tag{4.42}$$

One could have performed the substitution to eliminate $\delta\rho_m$ rather than v, to obtain the companion expression,

$$\frac{\partial^2}{\partial t^2}\delta\vec{v} = c_s^2\nabla^2\vec{v}, \tag{4.43}$$

for which the corresponding solution for the velocity would be,

$$\vec{v}(\vec{r}, t) = -A\frac{c_s}{\rho_m^{(0)}}\cos(\vec{k}\cdot\vec{r} - \omega t + \phi), \tag{4.44}$$

where again $\omega/k = c_s$.

If matter is in a region where $c_s^2 = dP/d\rho_m|_S < 0$, the speed of sound is imaginary. This means that for a real wavelength $\lambda = 2\pi/k$, the frequency ω is imaginary. Physically, this signals unstable modes which grow exponentially in time. This can occur for matter which suddenly finds itself deep inside a density-temperature region for which it would prefer to separate into two separate phases. See problem 7 of this chapter.

4.7   The Boltzmann Equation

When local kinetic equilibrium is far from being maintained, even viscous hydrodynamics cannot be justified. Fortunately, for many of the examples where hydrodynamic description is unwarranted, Boltzmann descriptions are applicable. The Boltzmann approach involves following the trajectories of individual particles moving through a mean field punctuated by random collisions. The probability of having collisions is assumed to be independent of the particle's past history and correlations with other particles are explicitly ignored. I.e., the chance of encountering a particle depends only on the particles momentum and position, and the average phase space density of the system. Thus, the Boltzmann description is effectively a one-body theory. A Boltzmann description is justified if two criteria are met:

- The particles are sufficiently dilute that classical trajectories do not violate the uncertainty principle. Such violations occur when the collision time falls below the $\hbar/T$. Many examples satisfy this criteria, including most liquids.

- Correlations between particles in space and time are negligible. By definition, liquids are tightly correlated. Applications of the Boltzmann equations is then questionable, however, strongly correlated systems are often ideal candidates for hydrodynamics.

The Boltzmann equation represents a tool for solving for the evolution of the phase space density $f(\vec{p}, \vec{r}, t)$ a.k.a. the phase space occupancy, which is defined by

$$f(\vec{p}, \vec{r}, t) = \frac{1}{(2S+1)} \frac{dN}{d^3p \, d^3r/(2\pi\hbar)^3}. \tag{4.45}$$

The definition is such that $f = 1$ corresponds to one particle per single-particle quantum state. In a kinetically thermalized system, f would be given by the Bose or Fermi distributions, and at each point in coordinate space, a few parameters, $T(x,t)$, $\mu(x,t)$ and the collective velocity $\vec{v}(x,t)$, would determine the entire momentum dependence. In contrast, a Boltzmann approach requires storing the entire structure of $f(\vec{p}, \vec{r}, t)$. Thus, it may involve discretizing both coordinate and momentum space, i.e. storing $f(\vec{p}, \vec{r})$ in a six-dimensional mesh. Because the Boltzmann equation represents the evolution of uncorrelated particles, the Boltzmann equation can also be represented by simulating sample particles colliding through the fluid with specified cross sections. If you over-sample the particles by a factor N_{sample} and reduce the cross sections by the same factor, you wash out correlations and the solution approaches the Boltzmann equation for large N_{sample}.

In the absence of interaction, particles move in uniform trajectories and the Boltzmann equation is simple,

$$\frac{\partial f}{\partial t} + \vec{v}_{\vec{p}} \cdot \nabla f = 0. \tag{4.46}$$

This simply states that the co-moving derivative, $D/Dt = \partial_t + v\partial_x$, of the phase space density of particles with zero velocity is zero. This can be understood by considering a small phase-space element, $\Delta p \delta x$, centered around $p = 0$. If the average number of particles in that element is n_Δ, that number will not change over time because the particles in that element are not moving.

Adding external forces,

$$\begin{aligned} 0 &= \frac{\partial f}{\partial t} + \vec{v}_{\vec{p}} \cdot \nabla f + \frac{d\vec{p}}{dt} \cdot \nabla_p f \\ &= \frac{\partial f}{\partial t} + \vec{v}_{\vec{p}} \cdot \nabla f + \vec{F} \cdot \nabla_p f. \end{aligned} \tag{4.47}$$

Here, the force $\vec{F}$ acts like an external field. It could originate from an external source or it could be the mean field driven by the phase space density itself. After adding this term, the equation now states that after adding an external force, the phase space density will be unchanged if you move and accelerate with the particles. Thus, in the absence of collisions, you can choose any particle with momentum $\vec{p}_0$ at position $\vec{r}_0$, and as you follow the trajectory through time, the phase space density evaluated at that particle's new momentum and position will be the same as it was initially. This can be understood physically by stating that if there are n particles in a phase space cell, that the same n particles will be in that phase space cell if you follow its trajectory.

The fact that the number of particles in a phase space cell is unchanged is not a trivial statement. If one were to choose a phase space cell of dimensions $dx, dy, dz, dp_x, dp_y, dp_z$, then follow the trajectories of the points within the cell, the new phase space cell would be distorted in shape. Despite the fact that $dp'_x \neq dp_x, \cdots$, the product $d^3p'd^3r'$ would equal the original phase space volume d^3pd^3r. This result is known as the Liouville theorem. If one considers the number of effective states within a volume, violating the Liouville theorem would involve either inserting or removing quantum states into the system. As long as the momenta and coordinates satisfy the commutation relation $[p, x] = i\hbar$, the number of states must be fixed. The Liouville theorem can also be applied to other combinations of coordinates and their canonical momentum, such as L_z and ϕ.

To prove the Liouville theorem we consider a rectangular cell of dimensions Δp_i and Δx_i. If the boundaries of the cell move with the velocities corresponding to the fluid velocity at that point of the boundary, the number of particles in the cell will remain fixed. In other words, the fluid element is defined by a fixed number of neighboring particles, and the volume of that element can grow or distort in the presence of a velocity gradient. The phase space density will then remain fixed if the phase-space volume of the cell,

$$\Delta\Omega = \frac{1}{(2\pi\hbar)^3} \prod_i \Delta p_i \Delta x_i, \tag{4.48}$$

remains fixed. In a time step dt the widths Δp_i and Δx_i change by amounts,

$$d\Delta p_i = dt\Delta p_i \frac{\partial F_i(p, x)}{\partial p_i}, \tag{4.49}$$

$$d\Delta x_i = dt\Delta x_i \frac{\partial v_i(p, x)}{\partial x_i}.$$

If v is only a function of p, e.g. $v = p/E(p)$ and if F_i is only a funciton of x, it is clear that the phase space volume is fixed. However, it remains fixed even if there are momentum-dependent forces or position dependent velocities, as can be seen by considering Hamiltonians Eq.s of motion,

$$F_i = -\frac{\partial H}{\partial x_i}, \tag{4.50}$$

$$v_i = \frac{\partial H}{\partial p_i}, \tag{4.51}$$

one finds

$$d(\Delta x_i \Delta p_i) = \Delta x_i d\Delta p_i + \Delta p_i d\Delta x_i = \Delta x_i \Delta p_i dt \left(-\frac{\partial^2 H}{\partial x_i \partial p_i} + \frac{\partial^2 H}{\partial p_i \partial x_i} \right) = 0. \tag{4.52}$$

One of the most notable consequences of the Liouville theorem comes in beam dynamics, where the magnetic forces depend on velocity. Magnets can focus charged particle beams, thus narrowing them down to a small region. However, this focusing is always accompanied by broadening the distribution in momentum space. Conversely, increasing the momentum resolution of a beam always broadens the beam spot in coordinate space.

Collisions ruin the continuity of trajectories, and thus violate Liouville's theorem. They also represent the only means towards achieving thermalization. Collisions are incorporated into the Boltzmann equation through the expression,

$$\frac{\partial f(\vec{p}, \vec{r}, t)}{\partial t} + \vec{v}_{\vec{p}} \cdot \boldsymbol{\nabla} f(\vec{p}, \vec{r}, t) + \vec{F} \cdot \boldsymbol{\nabla}_p f(\vec{p}, \vec{r}, t) \tag{4.53}$$

$$= \int \frac{d^3 p_b'}{(2\pi\hbar)^3} d\Omega \; v_{\rm rel} \frac{d\sigma(\vec{p}_a, \vec{p}_b \to \vec{p}, \vec{p}_b')}{d\Omega} f(\vec{p}_a, \vec{r}, t) f(\vec{p}_b, \vec{r}, t)$$

$$- \int \frac{d^3 p_b}{(2\pi\hbar)^3} d\Omega \; v_{\rm rel} \frac{d\sigma(\vec{p}, \vec{p}_b \to \vec{p}_a', \vec{p}_b')}{d\Omega} f(\vec{p}, \vec{r}, t) f(\vec{p}_b, \vec{r}, t)$$

Here, $v_{\rm rel}$ is the magnitude of the relative velocity between the incoming particles with momenta $\vec{p}_a$ and $\vec{p}_b$, and Ω refers to the angle between the incoming and outgoing relative momentum. For the first term, if one knows the two outgoing momenta $\vec{p}_b'$ and $\vec{p}$, and the scattering angle Ω, the incoming momenta $\vec{p}_a$ and $\vec{p}_b$ are determined by energy and momentum conservation. Similarly, knowing $\vec{p}$ and $\vec{p}_b$ in the second term, along with Ω, determines the two outgoing momenta $\vec{p}_a'$ and $\vec{p}_b'$. The first term on the r.h.s. of Eq. (4.53) is the gain term due to particles with momentum $\vec{p}_a$ and $\vec{p}_b$ scattering into the state $\vec{p}$. The second term is the loss term and represents the scattering of a particle with momentum $\vec{p}$ with momentum $\vec{p}_b$. Given either both outgoing momentum (first term on r.h.s.) or incoming (second term on r.h.s.), momentum and energy conservation along with the angle Ω determine the other two momenta.

The expression in Eq. (4.53) is what is usually referred to as the Boltzmann equation, though one might still call it the Boltzmann equation after adding a few options. One such option involves making the scattering kernel non-local. In the kernel above, the a and b particles are assumed to be at the same location. It is straight-forward, though a bit tedious, to add a dependence on $\vec{r}_a - \vec{r}_b$ to the cross section so that the a and b particles scatter at $\vec{r}_a$ and $\vec{r}_b$ respectively. This would also add another layer of integration.

A more common extension to the Boltzmann equation involves incorporating quantum statistics by replacing the products of phase space densities of the outgoing particles,

$$f(\vec{p}_a, \vec{r}) f(\vec{p}_b, \vec{r}) \to f(\vec{p}_a, \vec{r}) f(\vec{p}_b, \vec{r}) [1 \pm f(\vec{p}_a', \vec{r})] [1 \pm f(\vec{p}_b', \vec{r})]. \tag{4.54}$$

With this extension, the equation is sometimes referred to as the Vlasov equation, or the Boltzmann-Uehling-Uhlenbeck equation. The enhancement (Boson) or suppression (Fermion) factors will push the system towards an equilibrated Bose or Fermi distribution. To see how this is accomplished we consider a single momentum mode p, and assume that the rate at which the population changes is:

$$\frac{\partial f_p}{\partial t} = A_p - B_p f_p. \tag{4.55}$$

Here, A_p is the rate at which scatterings fill the level p and B is the rate at which a particle will leave the level if it is already present. If the population equilibrates at $f_p = e^{-\beta(\epsilon_p - \mu)}$, the ratio of the coefficients must be:

$$A_p / B_p = e^{-\beta(\epsilon_p - \mu)}. \tag{4.56}$$

The addition of the enhancement/suppression factors alters the rate equation,

$$\frac{\partial f_p}{\partial t} = A_p (1 \pm f_p) - B_p f_p. \tag{4.57}$$

Solving for the new equilibrated phase space density,

$$f_p = \frac{A_p}{B_p \mp A_p} = \frac{e^{-\beta(\epsilon_p - \mu)}}{1 \mp e^{-\beta(\epsilon_p - \mu)}}. \tag{4.58}$$

For the Bose case the coefficients are sometimes referred to as A and B coefficients, and the enhancement for particles all wanting to be emitted into a state already occupied, provides a crude explanation of lasers.

Example 4.5:

Consider a box of thermalized gas with $f(\vec{p}, \vec{r}, t < 0) = e^{-\beta(\epsilon_p - \mu)}$. The box is infinitely long in the y and z directions but is confined in the x direction by two walls, one at $x = L$ and the other at $x = -L$. At time $t = 0$ the walls magically dissolve. Assuming the gas particles do not interact with one another, solve for $f(\vec{p}, \vec{r}, t > 0)$.

Because f is fixed along a trajectory, The initial phase space density is:

$$f(\vec{p}, \vec{r}, 0) = e^{-\beta(\epsilon_p - \mu)} \Theta(L - x)\Theta(x + L).$$

Using the Liouville theorem, with the particle's velocity being $\vec{v} = \vec{p}/m$,

$$f(\vec{p}, \vec{r}, t) = f(\vec{p}, \vec{r} - \vec{v}t, 0) = e^{-\beta(\epsilon_p - \mu)} \Theta(L - x + v_x t)\Theta(x - v_x t + L).$$

Writing $v_x = p_x/m$, and using the fact that $\Theta(x) = \Theta(Cx)$,

$$f(\vec{p}, \vec{r}, t) = e^{-\beta(\epsilon_p - \mu)} \Theta(p_x - m(x - L)/t)\Theta(m(x + L)/t - p_x).$$

Thus, the distribution at $\vec{r}$ is a thermal one, except for cutting off the tails of the momentum distribution for p_x above $m(x + L)/t$ and below $m(x - L)/t$.

Numerical solutions of the Boltzmann equation tend to fall into two categories. One class of treatments involves storing $f(\vec{p}, \vec{r}, t)$ on a mesh, then explicitly solving for the evolution using the Boltzmann equation. This can be difficult because a three-dimensional Boltzmann equation involves storing six-dimensional phase space information. If each dimension is represented by 40 points, one already begins to challenge the typical amount of memory in a desktop workstation system.

A second approach is to represent the particles as sample particles. For instance, if a nuclear collision has a few hundred nucleons, one might model the scatterings of $100\times$ more particles. If the cross section were scaled down by a factor of 100, the collision rates would be the same as if the sampling ration was 1-to-1. In the limit of a large sample factor, the collisions become local and correlations are washed out. The simulation then approaches the Boltzmann equation.

4.8   Phase Space Density and Entropy

In the first chapter, we derived the fact that entropy is $S = \sum_i -p_i \ln p_i$, where p_i is the probability that a state is occupied. If one looks at a single unit of phase space, the probability of

having the various states $n = 0, 1 \cdots$ is determined by the phase space density (at least in a thermal system). Thus, the entropy in a given cell should be solely determined by the phase space density, and summing over the cells should give the total entropy. Because the phase space density of a given cell is fixed in the absence of collisions, it follows that the entropy is also fixed in the absence of collisions.

If the phase space density is low, there are only two probabilities for a given mode, $n = 0$ and $n = 1$, with probabilities $(1 - f)$ and f respectively. For Fermions, these two probabilities are all that are allowed even if f approaches unity. The entropy for fermions is thus:

$$S = -(1 - f)\ln(1 - f) - f\ln(f) \text{ (for fermions)}, \quad \sim f[1 - \ln(f)] \text{ as } f \to 0. \quad (4.59)$$

Thus, given $f(\vec{p}, \vec{r}, t)$, one can obtain an expression for the entropy that is independent of any assumption about thermalization, and the total entropy is found by summing over phase space cells:

$$S_{\text{total}} = (2J + 1)\frac{1}{(2\pi\hbar)^3} \int d^3p \, d^3r \, [-(1 - f)\ln(1 - f) - f\ln(f)]. \quad (4.60)$$

To see that the total entropy is fixed in a collisionless system, one can consider Eq. (4.60) with the replacement that $f(\vec{p}, \vec{r}, t) = f(\vec{p}, \vec{r} - v_p t, 0)$,

$$S(t) = \frac{(2J + 1)}{(2\pi\hbar)^3} \int d^3p \, d^3r \, \{-(1 - f(\vec{p}, \vec{r} - \vec{v}_p t, 0)\ln(1 - f(\vec{p}, \vec{r} - \vec{v}_p t, 0)) \quad (4.61)$$
$$-f(\vec{p}, \vec{r} - \vec{v}_p t, 0)\ln(f(\vec{p}, \vec{r} - \vec{v}_p t, 0))\}$$

If one makes the substitution $\vec{r}' \equiv \vec{r} - \vec{v}_p t$, the r.h.s. immediately becomes $S(t = 0)$.

Example 4.6:

Assuming that the probability of n bosons being in a specific phase space cell is $p_n \propto e^{-n\beta(\epsilon - \mu)}$, find an expression for the entropy in terms of the phase space density f.

First calculate the normalized value of p_n in terms of $x \equiv e^{-\beta(\epsilon - \mu)}$.

$$p_n = \frac{x^n}{1 + x + x^2 + x^3 \cdots} = x^n(1 - x).$$

The entropy is then

$$S = -\sum_n p_n \ln(p_n) = -(1-x)\sum_n x^n \ln[x^n(1-x)] = -(1-x)\sum_n x^n [n\ln(x) + \ln(1 - x)]$$

Using the fact that $\sum nx^n = x\partial_x \sum x^n$,

$$S = -\frac{x}{1 - x}\ln(x) - \ln(1 - x).$$

Finally, to get an expression in terms of f, one inverts the expression, $f = x/(1 - x)$ to obtain

$$x = \frac{f}{1 + f}, \quad (1 - x) = \frac{1}{1 + f}.$$

66

Inserting this into the expression for the entropy,

$$S = -f \ln(f) + (1 + f) \ln(1 + f).$$

Using the Boson result in the example above, and the Fermion result, Eq. (4.60), the entropy for Bosons/Fermions can be simultaneously expressed,

$$S_{\text{total}} = (2J + 1) \frac{1}{(2\pi\hbar)^3} \int d^3p \, d^3r \, [\pm(1 \pm f) \ln(1 \pm f) - f \ln(f)] . \qquad (4.62)$$

4.9   Hubble Expansion

A Hubble expansion involves the expansion of a system in a translationally invariant manner. Translational invariance implies no pressure gradient (relativistically you must be in the reference frame of the matter) and thus no acceleration. Gravitational effects can still provide acceleration, but we leave that for a course in general relativity. If the matter originates from a point at $x = y = z = t = 0$, and if the acceleration is zero for every fluid element, the collective velocity of the matter at later times must be:

$$\vec{v} = \frac{\vec{r}}{t}. \qquad (4.63)$$

This simply states that an observer moving with a fluid element of velocity $\vec{v}$ will be at position $\vec{r} = \vec{v}t$ if there is no acceleration.

In the absence of collisions, a particle with momentum $\vec{p}$ will slowly traverse matter until it ultimately reaches matter whose velocity is that of its own, $\vec{v}_p = \vec{p}/E$. Thus, if an observer moving with the matter measures a particle's momentum at some time, and if the same particle does not collide and reaches a second observer, who is also co-moving with the local matter, the second observer should see a smaller momentum. If a particle of momentum $\vec{p}_0$ and energy $E_0 = \sqrt{p^2 + m^2}$ at time t_0 at the origin, moves without colliding, at time t, its momentum as measured by a local observer (that moves with the matter) will be:

$$\vec{p} = \gamma \vec{p}_0 - \gamma \vec{v} E_0 = \frac{1}{\sqrt{1 - (r/t)^2}} \vec{p}_0 - \frac{\vec{r}/t}{\sqrt{1 - (r/t)^2}} E_0. \qquad (4.64)$$

Given that the position is

$$\vec{r} = \vec{v}_p(t - t_0) = (\vec{p}_0/E_0)(t - t_0), \qquad (4.65)$$

the second term that is proportional to E_0 can also be written as piece proportional to momentum as measured by a co-mover at the later time is

$$\vec{p} = \frac{1}{\sqrt{1 - (r/t)^2}} \vec{p}_0 - \frac{(t - t_0)/t}{\sqrt{1 - (r/t)^2}} \vec{p}_0 = \frac{\vec{p}_0 t_0}{\sqrt{t^2 - r^2}}. \qquad (4.66)$$

67

The denominator can be identified with the proper time, i.e., the time a co-moving observer would have measured because the matter was initially at the origin:

$$\tau \equiv \frac{t}{\sqrt{1 - v^2}} = \sqrt{t^2 - r^2}. \tag{4.67}$$

Rewriting the expression for the locally-measured momentum,

$$\vec{p} = \vec{p}_0 \frac{\tau_0}{\tau}. \tag{4.68}$$

Here, $\vec{p}_0$ and $\vec{p}$ are the momenta of the particle as measured by observers who are at the same position as the particle, but moving with the local matter. Thus, as long as the particle does not suffer a collision, its locally measured momentum falls inversely with the proper time. If it never collides, it eventually ends up with matter whose velocity matches that of its own, and the locally measured momentum approaches zero.

If none of the particles collide between τ_0 and τ, the distribution of momenta at any point, as measured by a co-mover, must be the same as the distribution at τ_0 after all the momenta are scaled by the factor τ_0/τ. This is true for both the relativistic and non-relativistic ($T \ll m$) limits.

Example 4.7:

Consider a gas of thermalized photons at temperature $T_0 = 4000$ K at proper time $\tau_0 = 10^7$ years. If the photons do not collide (freeze-out) until the current time, 14×10^9 years, what is the effective photon temperature?

Because the phase space density does not change with a moving particle,

$$f(p, \tau) = \frac{e^{-E_0/T_0}}{1 - e^{-E_0/T_0}}.$$

Furthermore, because $p = p_0 \tau_0/\tau$,

$$f(p, \tau) = \frac{e^{-E/(T_0 \tau_0/\tau)}}{1 - e^{-E/(T_0 \tau_0/\tau)}}.$$

Thus, the distribution is identical to a thermal distribution with local temperature,

$$T = T_0 \frac{\tau_0}{\tau} = 2.85K.$$

For hydrodynamic expansion (or collisionless expansions) entropy is conserved. If one is at moving with the fluid the collective velocity is zero and the equation of continuity gives:

$$\frac{\partial s}{\partial \tau} = -s \nabla \cdot \vec{v}. \tag{4.69}$$

Because the collective velocity is $\vec{v} = \vec{r}/\tau$, this gives:

$$\frac{\partial s}{\partial \tau} = -3\frac{s}{\tau}. \tag{4.70}$$

This gives

$$s(\tau) = s(\tau_0)\frac{\tau_0^3}{\tau^3}. \tag{4.71}$$

For a massless gas, the entropy density is proprotional to T^3,

$$\frac{s(\tau)}{s(\tau_0)} = \frac{T^3}{T_0^3} = \frac{\tau_0^3}{\tau^3}, \tag{4.72}$$

and $T = T(\tau_0)\tau_0/\tau$, which is the same expression given in the example before for a collisionless expansion. This equivalence owes itself to Eq. (4.9), which shows that the collisionless system appears locally equilibrated. Because the system is already locally equilibrated, collisions have no effect. For that reason, photon-photon collisions would have no effect on the thermal background radiation spectrum.

For a non-relativistic gas of particles with conserved particle number, the entropy per particle is determined by the factor $\ln[(mT)^{3/2}/\rho]$. The same equation for continuity used for the entropy above could be applied to ρ, and would give,

$$\rho(\tau) = \rho(\tau_0)\frac{\tau_0^3}{\tau^3}. \tag{4.73}$$

Fixing the entropy per particle,

$$\frac{T(\tau)^{3/2}}{\rho(\tau)} = \frac{T(\tau_0)^{3/2}}{\rho(\tau_0)}, \tag{4.74}$$

or

$$T(\tau) = T(\tau_0)\frac{\tau_0^2}{\tau^2}. \tag{4.75}$$

Thus, a massive gas cools much more quickly than a gas of massless particles. In cosmology, this difference in cooling inspired the terminology "radiation dominated" for the massless case vs "matter dominated" for the non-relativistic case. When the universe was approximately a hundred thousand years old, the temperature of the universe left the radiation-dominated era and moved into the matter-dominated behavior era. It should be emphasized that the simple $1/\tau^2$ cooling for the matter (mostly hydrodgen) is not valid due to the heating associated with galaxy and star formation, but that for the photons, the $1/\tau$ cooling is valid as they decouple from other matter when atoms formed, as opposed to plasma, because the neutral atoms are largely transparent to long-wavelength photons.

4.10   Evaporation, Black-Body Emission and the Compound Nucleus

Evaporation plays a crucial role in the cooling of hot stars, hot nuclei, and hot people. For large systems, where the outer surface is cooler than the inside, accurate estimates of the evaporation rate involves integrating over the emission probability over the volume of the star, convoluting with the escape probability. Detailed microscopic approaches are needed for problems of neutrino emission from supernovas, or for calculating details of spectra from stars where the outer gas might absorb specific wave lengths.

Such emission is often approximated by assuming the particles are thermalized at a distance R, and that all outgoing particles at this distance escape. The number emitted per unit area per unit time is:

$$\frac{dN}{dA\ d^3p\ dt} = \vec{v}_{\mathrm{p}} \cdot \hat{n} \frac{dN_{\mathrm{therm}}}{d^3p\ d^3r} \Theta(\vec{v}_{\mathrm{p}} \cdot \hat{n}), \qquad (4.76)$$

where $\hat{n}$ is the unit normal vector of the surface. For an equilibrated gas,

$$\frac{dN_{\mathrm{therm}}}{d^3p\ d^3r} = (2J+1)\frac{f(\vec{p})}{(2\pi\hbar)^3} = \frac{(2J+1)}{(2\pi\hbar)^3} \frac{e^{-(\epsilon_p-\mu)/T}}{1 \mp e^{-(\epsilon_p-\mu)/T}}. \qquad (4.77)$$

One can now calculate the emission rate from Eq.s (4.76) and (4.77) from a spherical surface of radius R,

$$\begin{aligned}
\Gamma_{\mathrm{emm}} &= 4\pi R^2 \int \frac{d^3p}{(2\pi\hbar)^3}\ \vec{v}_{\mathrm{p}} \cdot \hat{n} f(\vec{p})\Theta(\vec{v}_{\mathrm{p}} \cdot \hat{n}) && (4.78) \\
&= 4\pi R^2 \langle \cos\theta \Theta(\cos\theta) \rangle \int \frac{d^3p}{(2\pi\hbar)^3}\ |\vec{v}_{\mathrm{p}}|\ f(\vec{p}), \\
\langle \cos\theta\Theta(\cos\theta) \rangle &= \frac{\int_0^1 d\cos\theta\ \cos\theta}{\int_{-1}^1 d\cos\theta} = \frac{1}{4}. && (4.79)
\end{aligned}$$

Here, the average over emission directions yields a factor of $1/4$. Thus the emission is

$$\frac{d\Gamma_{\mathrm{emm}}}{d^3p} = 4\pi R^2 \frac{|v_p|}{4}(2J+1)\frac{f(\vec{p})}{(2\pi\hbar)^3}. \qquad (4.80)$$

This rate per area looks like a Boltzmann distribution aside from the factor of $|v_p|/4$, which accounts for faster particles hitting the surface more often. This is known as a Maxwell-Boltzmann distribution.

Example 4.8:

The surface of a star has a temperature T, which is much less than the mass of an electron m. If the electrons have chemical potential μ, what is the ratio of emitted photons to emitted electrons. (Assume the electrons are non-degenerate.)

Because both particles have two polarizations, the ratio will be determined by the ratio of velocities and Boltzmann factors:

$$\begin{aligned}
\frac{N_e}{N_\gamma} &= \frac{\int d^3p\ (p/m)e^{-\beta(p^2/2m-\mu)}}{\int d^3p\ e^{-\beta p}/(1-e^{-\beta p})} \\
&= \frac{(mT)^2 8\pi e^{\beta\mu} \int du\ u e^{-u}}{8\pi m T^3 \zeta(3)} \\
&= e^{\beta\mu}\frac{m}{T\zeta(3)}
\end{aligned}$$

For small systems with modest excitation energy, the emitted particle might carry a significant fraction of the total excitation energy and the assumption of an equilibrated temperature and

chemical potential becomes unwarranted. For such systems, the temperature before and after the emission is no longer identical. To address such systems, one can consider two sets of states. In the first set, the drop has energy between E and $E + \delta E$. In the second, the drop has energy between $E - \epsilon_p$ and $E - \epsilon_p + \delta E$, and a particle in an individual quantum state of momentum $\vec{p}$ runs freely. Again, the rates from one set to the other should be equal if the system is at equilibrium, with the absorption rate being,

$$\Gamma_{\text{abs}} = \frac{\sigma_{\text{abs}}(\vec{p})v_p}{V}, \tag{4.81}$$

where V is the volume of the box. Because the emission from a given state in set 1 to a given state in set 2 equals the rate from the same state in set 2 to the specific state in set 1, the ratio of the rates for all states emitted from set 1 to set 2 equals the net rate of the inverse process multiplied by the ratio of states.

$$\Gamma_{\text{emm}}(\vec{p}) = \Gamma_{\text{abs}} \frac{N(E - \epsilon_p)}{N(E)} \tag{4.82}$$

$$= \frac{\sigma_{\text{abs}}(\vec{p})v_p}{V} \frac{N(E - \epsilon_p)}{N(E)}, \tag{4.83}$$

where N is the number of states of the compound nucleus per unit of excitation energy, usually called the level density. This is the rate for emitting a particle into a specific single-particle state in the continuum of volume V. The rate into a range of states in d^3p is found by multiplying by the number of states per d^3p, $(2J + 1)V/(2\pi\hbar^3)$, which gives

$$\frac{d\Gamma_{\text{emm}}}{d^3p} = (2J + 1)\frac{\sigma_{\text{abs}}(\vec{p})v_p}{(2\pi\hbar)^3} \frac{N(E - \epsilon_p)}{N(E)}. \tag{4.84}$$

Using the definition of the entropy, $\Delta S = \ln[N(E)/N(E - \epsilon_p)]$, one finds

$$\frac{d\Gamma_{\text{emm}}}{d^3p} = (2J + 1)\frac{\sigma_{\text{abs}}(\vec{p})v_p}{(2\pi\hbar)^3}e^{S(E-\epsilon_p)-S(E)}. \tag{4.85}$$

This type of reasoning, based on the system that absorbs the particle equilibrating the energy, is known as a Weisskopf argument. If the temperature changes little for such emission, $\Delta S = (\epsilon_p - \mu)/T$, and one obtains the same expression as before (aside from the degeneracy corrections for $f(\vec{p})$, which would be accounted for if one correctly accounted for the probability that the large box might already have particles so that the emission would increase the number of particles in a given momentum state from n to $n + 1$).

Example 4.9:

Consider a star of radius R and uniform temperature T from which spin-1/2 neutrinos are emitted. The rate at which a neutrino is emitted while in the star is γ. Assuming γ is so small that neutrinos typically travel through the star without ever being captured, use Weisskopf arguments to calculate the rate at which neutrinos are emitted from the star. Also, assume that only one species of neutrinos exists, (ν_e).

Here, the absorption rate for a neutrino from a specific quantum level $\vec{p}$ within the large box is

$$\Gamma_{\text{abs}} = \gamma \frac{4\pi R^3}{3V},$$

where V is the volume of the box . Weisskopf arguments state that the rate for emission into the one free quantum state is

$$\Gamma_{\text{emm}}(p) = \gamma \frac{4\pi R^3}{3V} e^{-E_p/T}.$$

Now emmision into a variety of states within d^3p is

$$\frac{d\Gamma_{\text{emm}}}{d^3p} = \frac{(2J+1)}{(2\pi\hbar)^3} \gamma \frac{4\pi R^3}{3} e^{-E_p/T}, \tag{4.86}$$

$$= \gamma V_{\text{star}} \frac{(2J+1)}{(2\pi\hbar)^3} e^{-E_p/T}. \tag{4.87}$$

Here $J = 1/2$. The emmision rate looks like the equilibrated number of neutrinos in the star volume, aside from missing the factor $1/(1 + e^{-E_p/T}))$, multiplied by the absorption rate γ.

4.11   Diffusion

Diffusion is the continuous limit of a statistical random walk. The prototypical example is an impurity in a gas, but there are innumerable examples, both within physical processes or any random process. For instance, the growth and survival of wrapped characters on long paragraphs can be understood with the diffusion equation.

The diffusion equation is based on the drift equation, a.k.a. Fick's law,

$$\vec{j}(x) = -D\nabla\rho. \tag{4.88}$$

This states that a current j will develop in the presence of a density gradient. If one combines this expression with the equation of continuity,

$$\frac{\partial\rho}{\partial t} = -\nabla \cdot \vec{j}, \tag{4.89}$$

one finds the diffusion equation

$$\frac{\partial\rho(x,t)}{\partial t} = D\nabla^2\rho. \tag{4.90}$$

Here D is the diffusion constant.

To see the equivalence with a random walk, we write the equivalent equation for a random walk,

$$\delta N_i = \Gamma\delta t \left\{ (N_{i+1} - N_i) + (N_{i-1} - N_i) \right\}, \tag{4.91}$$

where Γ is the probability of moving in a specific direction per unit time. Thus, in a small time step δt, the number of particles leaving a location is $2\Gamma\delta t$. To compare to the diffusion equation, it is insightful to consider the change in ρ for a small time step δt and to rewrite the diffusion equation with the derivatives being expressed in terms of changes per Δx,

$$\delta\rho(x,t) = \frac{D}{\Delta x^2}\delta t \left\{ [\rho(x+\Delta x,t) - \rho(x,t)] + [\rho(x-\Delta x,t) - \rho(x,t)] \right\} . \tag{4.92}$$

Comparing the diffusion equation in this form to the analogous equation for a random walk in Eq. (4.91), then making the statement that ρ is proportional to N_i, allows one to identify the diffusion constant:

$$D = \Gamma(\Delta x)^2. \tag{4.93}$$

The diffusion constant D can now be identified with microscopic quantities. For instance, if the collision time is τ_{coll} and the mean free path is λ, the size of the random step is $\Delta x \sim \lambda$ and the rate at which one takes some random step in a specific direction is $\Gamma \sim 1/2\tau_{\text{coll}}$. The diffusion constant is:

$$D \sim \lambda^2/\tau = \lambda v_{\text{therm}}. \tag{4.94}$$

Here, we neglect factors of two, because the collision rate depends on both the velocity of the diffusing particles, as well as the velocity of the colliding partners, as well as the fraction of momentum lost per collision. More exact expressions can be evaluated with a detailed microscopic analysis.

A simple solution to the diffusion equation is a Gaussian,

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left\{\frac{-x^2}{4Dt}\right\}. \tag{4.95}$$

The width of the Gaussian increases as $t^{1/2}$, beginning at zero. Thus, if you place particles at a point at $x = 0$, they would diffuse according to Eq. (4.95).

Example 4.10:

Consider molecules that diffuse in a container whose walls are positioned at $x = 0$ and $x = L$. If the molecules collide with the walls, they stick and disappear from the distribution, and the density $\rho(x)$ must satisfy the boundary condition that it becomes zero at $x = 0$ and $x = L$. (If, instead, the boundary was perfectly reflective, the boundary condition would require the current density at the boundary to disappear, or equivalently that $\nabla\rho = 0$.) Consider a solution of the form,

$$\rho(x,t) = A(t)\sin(\pi x/L).$$

Using the diffusion equation find $A(t)$.

Inserting this expression into Eq. (4.90),

$$\dot{A} = -D\left(\frac{\pi}{L}\right)^2 A,$$

$$A(t) = A_0\exp\left\{-D\left(\frac{\pi}{L}\right)^2 t\right\}.$$

One can also check the solution to see that the rate at which the integrated number falls equals the drift into the surfaces. Integrating the density,

$$N(t) = \int dx\, \rho(x,t) = 2A(t)L/\pi,$$

$$\dot{N} = \frac{2\dot{A}L}{\pi} = -\frac{2\pi DA}{L}$$

Because the currents into each wall is $D\partial\rho/\partial x$, one should get the same result from the drift equation,

$$\dot{N} = -2D\frac{\partial\rho}{\partial x} = -2A\frac{D\pi}{L}.$$

Indeed, this agrees.

Example 4.11:

Consider the same problem, but with an initially uniform density ρ_0. Express $\rho(t)$ as a Fourier expansion in terms of $\sin(m\pi x/L)$.

First, recall the properties of a Fourier transform in a finite interval, $0 < x < L$,

$$F(x) = \sum_{m=1,2,3\cdots} F_m \sin(m\pi x/L),$$

$$F_m = \frac{2}{L}\int_0^L dx\, F(x)\sin(m\pi x/L).$$

Now, use the following expression for the density,

$$\rho(x,t) = \sum_m A_m(t)\sin(m\pi x/L).$$

Before proceeding further, find $A_m(t=0)$,

$$A_m(t=0) = \frac{2}{L}\int_0^L dx\, \rho_0 \sin(m\pi x/L) \tag{4.96}$$

$$= \begin{cases} 4\rho_0/m\pi, & \text{odd } m, \\ 0, & \text{even } m. \end{cases} \tag{4.97}$$

Each contribution behaves independently, and can be solved exactly as the previous example, which yields,

$$A_m(t) = A_m(t=0)\exp\left\{-D\left(\frac{m\pi}{L}\right)^2 t\right\}.$$

Thus, the higher order contributions are damped with rapidly increasing exponential damping factors. For large times solutions are independent of the initial conditions and approach the form of the previous example, i.e., only the $m = 1$ term survives.

For grins, one can use the properties of Fourier transforms to show that $\zeta(2) = \pi^2/6$. If one integrates,

$$\int_0^L dx \rho(x, t = 0) \;=\; \rho_0 L,$$

$$= \sum_{m=1,3\cdots} A_m(t=0) \int_0^L dx \; \sin(m\pi x/L)$$

$$= \sum_{m=1,3\cdots} \frac{8\rho_0 L}{(m\pi)^2}.$$

Some algebra shows that

$$\sum_{m=1,3\cdots} \frac{1}{m^2} = \frac{3}{4} \sum_{m=1,2,3\cdots} \frac{1}{m^2} = \frac{3}{4}\zeta(2).$$

The previous expression thus becomes

$$\rho_0 L \;=\; \frac{8\rho_0 L}{\pi^2} \frac{3}{4}\zeta(2),$$

$$\zeta(2) \;=\; \frac{\pi^2}{6}.$$

For some problems, one may wish to use the method of images, similar to what is used for some boundary-value problems in electromagnetism. A simple example considers placing a charge at x_0, where there is an absorbing plate at the $x = 0$ plane. In the absence of the plate, the density would evolve as

$$\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-(x-x_0)^2/(4Dt)}. \tag{4.98}$$

However, this would not satisfy the boundary condition that $\rho(x = 0, t) = 0$. By considering a second solution centered at $-x_0$, one can subtract the two solutions to find a solution that satisfies both the boundary condition at $x = 0$ and the initial condition, that for $x > 0$ the charge is only at one point at $t = 0$. The overall solution is then

$$\rho(x, t) = \begin{cases} \frac{1}{\sqrt{4\pi Dt}}\left(e^{-(x-x_0)^2/(4Dt)} - e^{-(x+x_0)^2/(4Dt)}\right), & x > 0 \\ 0, & x < 0 \end{cases} \tag{4.99}$$

When x_0 is small, compared to $\sqrt{Dt}$, one can expand the solution in powers of x_0, and find

$$\rho(x, t) \;\approx\; x_0 \left.\frac{d\rho}{dx_0}\right|_{x_0=0} \tag{4.100}$$

$$= \frac{x x_0}{\sqrt{4\pi(Dt)^3}} e^{-x^2/(4Dt)}.$$

For this limiting solution, one can solve for the remaining fraction of original particles,

$$N(t) \;=\; \int_0^\infty dx \; \rho(x, t) \tag{4.101}$$

$$= \frac{x_0}{\sqrt{\pi Dt}}.$$

Further as a check, one can see that $dN/dt = -\frac{x_0}{4\sqrt{\pi D t^3}}$, which indeed matches the flux into the wall, $j = -D\partial_x \rho$ at $x = 0$. The equivalent random walk problem is referred to as the "drunken sailor on a pier".

4.12   Langevin Approaches

The Boltzmann equation handles collisions by convoluting two-particle distributions of momenta p_1 and p_2 to determine the rate they scatter into momenta p_3 and p_4. Each collision conserves both energy and momentum. Langevin approaches model the movement of individual particles, mimicking scattering by combining random forces and drag forces. Langevin approaches can model the motion of a high-energy particle moving through a medium requiring many soft collisions to thermalize, amongst other applications.

The random forces, or more accurately random impulses, have a strength and frequency. Increasing either the strength or frequency of these impulses increases the average kinetic motion and raises the temperature. Drag forces do the opposite, decreasing the kinetic motion and working to lower the temperature. Here, we derive the relation between these to reach a given temperature.

Assume that a particle of mass m feels impulses separated by small times Δt. The impulses have a strength

$$\langle \Delta p_i^2 \rangle = 3\sigma^2, \quad \langle \Delta p_i \rangle = 0. \tag{4.102}$$

The factor of three comes from assuming three dimensions, so that each component has a spread σ_2. Tdrag force is

$$F_d = -\gamma \vec{p}. \tag{4.103}$$

Here, wish to show how the drag coefficient along with the parameters σ and Δt must be adjusted if the particles are to equilibrate at a target temperature T. For a given particle, its momentum at $t = 0$ arises from the previous momentum kicks Δp_n. For the x component,

$$p_x = \sum_n \sigma_n e^{-\gamma n \Delta t}. \tag{4.104}$$

The sum covers previous kicks occurring at times $n\Delta t$. Because these contributions are linear, they decay separately according to the decay constant γ. Each kick σ_n is random, so if one averages over all the kicks,

$$\begin{aligned}
\langle p_x^2 \rangle &= \sum_n \langle \sigma_n^2 \rangle e^{-2\gamma n \Delta t} \tag{4.105} \\
&= \frac{\sigma^2}{\Delta t} \sum_n \Delta t\, e^{-2\gamma n \Delta t} \\
&\approx \frac{\sigma^2}{\Delta t} \int_0^\infty dt\, e^{-2\gamma t}, \\
&= \frac{\sigma^2}{2\gamma \Delta t}.
\end{aligned}$$

From the equipartition theorem, one should have $\langle p_x^2 \rangle = mT$ at equilibrium, so

$$T = \frac{\sigma^2}{2m\gamma\Delta t}.$$ (4.106)

Because adding many small random numbers yields a Gaussian (central limit theorem), the momentum distribution will also have the correct shape for non-relativistic particles. If the drag rate γ is determined by physical arguments, one must adjust the strength of the kicks, σ, and the time step with which they are delivered Δt to satisfy Eq. (4.106). To maintain a given temperature, the strength of the random impulses must decrease if one chooses smaller time steps. As expected, larger σ leads to higher temperatures, and larger damping leads to lower temperatures. Due to the stochastic nature of the implementation, the time step always enters the formalism. This is straight-forward to solve numerically. Typically, the random impulses are chosen with Gaussian distributions, though any short-range even distribution with the variance σ^2 will work.

4.13   Problems

1. A molecule of mass m has two internal states, a spin-zero ground state and a spin-1 excited state which is at energy X above the ground state. Initially, a gas of such molecules is at temperature T_i before expanding and cooling isentropically to a temperature T_f. Neglect quantum degeneracy of the momentum states for the following questions.

   (a) What is the initial energy per particle? Give answer in terms of m, T_i, X and the initial density ρ_i.

   (b) Derive an expression for the initial entropy per particle in terms of the same variables.

   (c) After insentropically cooling to T_f, find the density ρ_f. Give answer in terms of ρ_i, T_i, T_f and X.

2. Repeat problem #1 above, but assume the molecule has the excitation spectrum of a 3-dimensional harmonic oscillator, where the energy levels are separated by amounts $\hbar\omega = X$, with $X \ll T$.

3. A large number of N diatomic molecules of mass m are confined to a region by a harmonic-oscillator potential,

$$V(\vec{r}) = \frac{1}{2}kr^2.$$

   The system is at a sufficient temperature T so that the gas can be considered dilute and the energy levels are practically continuous. The temperature is in the range where rotational modes are routinely excited, but vibrational modes can be neglected.

   (a) What is the energy per particle? Give your answer in terms of m, T, and k.

   (b) Derive an expression for the entropy per particle in terms of the same variables. Begin with the expression,

$$S = \ln Z + \beta E,$$

where

$$Z = \frac{z^N}{N!},$$

and z is the partition function of a single molecule.

(c) If the spring constant is adiabatically changed from k_i to k_f, and if the initial temperature is T_i, find T_f.

Fun facts to know and tell: $\lim_{N \to \infty} \ln(N!) = N \ln N - N$.

4. Consider an ideal gas with $C_p/C_v = \gamma$ going through the Carnot cycle illustrated in Fig. 4.1. The initial volume for N molecules at temperature T_H expands from V_a to $V_b = 2V_a$ and then to $V_c = 2V_b$.

   (a) In terms of NT_H, find the work done while expanding from $a - b$.

   (b) Again, in terms of NT_H, how much heat was added to the gas while expanding from $a - b$.

   (c) In terms of NT_H, find the work done while expanding from $b - c$.

   (d) What is the efficiency of the cycle ($abcd$)?

5. Consider a refrigerator built by an inverse Carnot cycle. What is the efficiency of the refrigerator in terms of the temperatures T_C and T_X?

6. Consider a hydrodynamic slab which has a Gaussian profile along the x direction but is translationally invariant in the y and z directions. Assume the matter behaves as an ideal gas of non-relativistic particles. Initially, the matter is at rest and has a profile,

$$\rho(x, t = 0) = \rho_0 \exp(-x^2/2R_0^2),$$

with an initial uniform temperature T_0. Assume that as it expands it maintains a Gaussian profile with a Gaussian radius $R(t)$.

   (a) Show that entropy conservation requires

$$T(t) = T_0 \left(\frac{R_0}{R} \right)^{2/3}.$$

   (b) Assuming the velocity has the form $v = A(t)x$, show that conservation of particle current gives

$$A = \frac{\dot{R}}{R}.$$

   (c) Show that the hydrodynamic expression for acceleration gives

$$\dot{A} + A^2 = \frac{R_0^{2/3}T_0}{mR^{8/3}}$$

Putting these two expressions together show that $R(t)$ can be found by solving and inverting the integral,

$$t = \sqrt{\frac{m}{3T_0}} \int_{R_0}^{R} \frac{dx}{\sqrt{1 - (R_0/x)^{2/3}}},$$

which, with some work, can be integrated to be

$$\frac{T_0 t^2}{m} = R^2 - 4R_0^2 + 3R_0^{2/3} R^{4/3}.$$

By solving a cubic equation for $R^{2/3}$, one can express $R(t)$ analytically.

7. Consider a gas obeying the Van der Waals equation of state,

$$P = \frac{\rho T}{1 - \rho/\rho_s} - a\rho^2.$$

(a) First, using Maxwell relations, show that the speed of sound for $P(\rho, T)$ is given by

$$mc_s^2 = \left.\frac{\partial P}{\partial \rho}\right|_T + \left(\left.\frac{\partial P}{\partial T}\right|_\rho\right)^2 \frac{T}{\rho^2 C_V}.$$

(b) Consider a gas described by the Van der Waals equation of state, which has a density equal to that of the critical point, $\rho_s/3$. What is the range of temperatures for which the matter has unstable sonic modes? Express your answer in terms of the critical temperature T_c.

8. Consider an initially thermalized three-dimensional Gaussian distribution for the phase space density of non-relativistic particles of mass m,

$$f(\vec{p}, \vec{r}, t = 0) = f_0 \exp\left\{-\frac{r^2}{2R_0^2} - \frac{p^2}{2mT_0}\right\}.$$

Assume the particles move freely for $t > 0$.

(a) At a given $\vec{r}$ and $t > 0$, show that $f(\vec{p}, \vec{r}, t)$ can be expressed in terms of a locally thermalized distribution of the form,

$$f(\vec{p}, \vec{r}, t) = C(t) e^{-r^2/2R^2(t)} \exp\left\{-\frac{(\vec{p} - m\vec{v}(\vec{r}, t))^2}{2mT(\vec{r}, t)}\right\}.$$

Then, find $C(t)$, $R(t)$, $\vec{v}(\vec{r}, t)$ and $T(\vec{r}, t)$. In addition to $\vec{r}$ and t, these parameters should depend on the initial Gaussian size R_0, the initial temperature T_0 and the mass m.

(b) Find the density as a function of $\vec{r}$ and t, then compare your result for the density and temperature to that for a hydrodynamic expansion of the same initial distribution as described in the example given in the lecture notes.

(c) Using the fact that a hydrodynamic expansion assumes infinitely high collision rate, and that the Boltzmann solution was for zero collision rate, make profound remarks about how the hydrodynamic and free-streaming evolutions compare with one another.

(d) Calculate the total entropy as a function of time for the previous problem assuming f_0 is small.

9. Assume that there exists a massive species of neutrinos, $m_\nu = 10$ eV. Further assume that it froze out at the same time as the example of the text, when the Hubble time as 10^7 years and the temperature was 4000 K. If the Hubble expansion was without acceleration after that point, and if the current Hubble time is 14×10^9 years, find:

(a) the current effective temperature of the massive neutrino.

(b) If the neutrino has two polarizations (just like photons) what would be the relative population, N_ν/N_γ, at freezeout? Assume the chemical potential for the neutrino is zero (otherwise there would be more neutrinos than anti-neutrinos) and treat the neutrino non-relativistically,

$$f_\nu(p) = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} \approx e^{-\beta(m+p^2/2m)}.$$

10. Consider a hot nucleus of radius 5 fm at a temperature of 1 MeV. The chemical potential for a cold (or warm) nucleus is approximately the binding energy per particle, $\mu \approx -7$ MeV. Estimate the mean time between emitted neutrons. (Treat the nucleus as if it has fixed temperature and assume a Boltzmann distribution (not Fermi) for $f(\vec{p})$).

11. Consider a hot nucleus of radius R with an electric charge of Z at temperature T. Assuming that protons and neutrons have the same chemical potential, find the ratio of proton spectra to neutron spectra,

$$\frac{dN_{\mathrm{prot}}/d^3pdt}{dN_{\mathrm{neut}}/d^3pdt}$$

as a function of the momentum p. Approximate the two masses as being equal, $m_n = m_p$, and neglect quantum degeneracy. Assume that all incoming nucleons would be captured and thermalized if they reach the position R. HINT: The emission ratio equals the ratio of capture cross sections.

12. A drift detector works by moving the electrons ionized by a track through a gas towards readout plates.

(a) Assuming the mean free path of the electrons is $\lambda = 300$ nm, and assuming their velocity is thermal at room temperature ($v_{\mathrm{therm}} \approx \sqrt{T/m}$), estimate the size R (Gaussian radius) that the diffusion cloud imprints onto the plates after drifting for 200 μs. Use the approximation that $D \approx \lambda v_{\mathrm{therm}}$.

(b) Assume an electric field E is responsible for the drift velocity. If the drift velocity is approximately $a \cdot \tau/2$, where a is the acceleration and $\tau = \lambda/v_{\mathrm{therm}}$ is the collision time, find an analytic expression for the Gaussian size R of the cloud if it travels a distance L. Give R in terms of the temperature T, the electron mass m and charge e, the electric field E, the mean free path λ and L.

13. A cloud of radioactive mosquitos is being blown by the wind parallel to a high-voltage mosquito-zapping plate. At at time $t = 0$, the cloud is at the first edge of the plate ($x = 0$). The probability they start out at a distance y from the plate is:

$$\rho(y, t = 0) = A_0 y e^{-y^2/(2R_0^2)}$$

The speed of the breeze is v_x, and the length of the plate is L. The mosquito's motion in the y direction can be considered a diffusion process with diffusion constant D.

(a) What is the distribution $\rho(y, t)$? Assume that the form for $\rho(t)$ is the same as for ρ_0 only with A and R becoming functions of t. Solve only for $t < L/v_x$.

(b) What fraction of mosquitos survive?

14. For an expanding system, the Diffusion equation is modified by adding an extra term proportional to $\nabla \cdot \vec{v}$,

$$\frac{\partial \rho}{\partial \tau} + \rho(\nabla \cdot \vec{v}) = D\nabla^2 \rho.$$

The second term accounts for the fact that the density would fall due to the fact that the matter is expanding, even if the the particles were not diffusing. For a Hubble expansion, $\vec{v} = \vec{r}/\tau$, and the $\nabla \cdot \vec{v} = 3/\tau$ (would be $2/\tau$ or $1/\tau$ in 2-D or 1-D). Thus in a Hubble expansion, the diffusion equation becomes

$$\frac{\partial \rho}{\partial t} + \frac{3}{\tau}\rho = D\nabla^2 \rho.$$

Instead of the position $\vec{r}$, one can use the variable

$$\vec{\eta} \equiv \frac{\vec{r}}{\tau}.$$

The advantage of using η is that because the velocity gradient is $1/\tau$, the velocity difference between two points separated by $dr = \tau d\eta$ is $dv = d\eta$. Thus if two particles move with the velocity of the local matter, their separation $\eta_1 - \eta_2$ will remain fixed. Next, one can replace the density $\rho = dN/d^3r$ with

$$\rho_\eta \equiv \frac{dN}{d^3\eta} = \tau^3 \rho.$$

Here, we have been a rather sloppy with relativistic effects, but for $|\eta|$ much smaller than the speed of light, they can be ignored. One can now rewrite the diffusion equation for ρ_η,

$$\frac{\partial \rho_\eta}{\partial \tau} = D\nabla^2 \rho_\eta.$$

This looks like the simple diffusion equation without expansion, however because the density is changing D is no longer a constant which invalidates using the simple Gaussian solutions discussed in the chapter. For an ultrarelativistic gas, with perturbative interactions, the scattering cross sections are roughly proportional to $1/T^2$, and the density falls as

$1/T^3$, which after considering the fact that the temperature then falls as $1/\tau$, the diffusion constant would roughly rise linearly with time,

$$D(\tau) = D_0 \frac{\tau}{\tau_0}.$$

This time dependence would be different if the particles had fixed cross sections, or if the gas was not ultra-relativistic. However, we will assume this form for the questions below.

(a) Transform the three-dimensional diffusion equation,

$$\frac{\partial \rho_\eta}{\partial \tau} = D(\tau) \nabla^2 \rho_\eta$$

into an equation where all derivatives w.r.t. r are replaced with derivatives w.r.t. η.

(b) Rewrite the expression so that all mention of τ is replaced by $s \equiv \ln(\tau/\tau_0)$.

(c) If a particle is at $\vec{\eta} = 0$ at τ_0, find $\rho_\eta(\eta, s)$.

15. Consider Fick's law for the number density and the number current, $\vec{j} = -D\nabla\rho$.

(a) Rewrite Fick's law in terms of the gradient of the chemical potential, showing that D is replaced by $D\chi$ where $\chi = \partial\rho/\partial\mu$.

(b) Replacing the gradient of the chemical potential with the gradient of electric potential, assuming the particles have charge e, find an expression for the electric current, $\vec{j}_e = e\vec{j}$, in terms of a gradient of the electric potential energy $e\Phi$.

(c) Express the electric conductivity, σ, in terms of D, χ and e.

16. Consider Eq. (4.106) for a one-dimensional system:

(a) In terms of γ, T, Δt and m, estimate the amount of time required for the variance of the sum of random impulses, Δp, in one direction to reach mT, the thermal variance.

(b) Calculate $\langle v(t=0)v(t)\rangle$, the velocity-velocity correlation in the limit $\Delta t \to 0$.

(c) Calculate the r.m.s. distance traveled by a particle in time t in the large-time limit, $\gamma t >> 1$.
Hint: Use the fact that for a specific particle that had a velocity v_0, the position is $x(t) = \int_0^t dt'\, v(t')$, where $v(t)$ is the contribution from all the previous impulses,

$$
\begin{aligned}
v(t) &= v_0 e^{-\gamma t} + \sum_n \Delta v_n e^{-\gamma(t-t_n)}, \\
\langle \Delta v_m \Delta v_n \rangle &= \frac{\sigma^2}{m^2}\delta_{mn}.
\end{aligned}
$$

(d) For large times, $\gamma t >> 1$, estimate the diffusion constant by comparing $\langle x^2(t)\rangle$ to the result for the diffusion equation.

5   Lattices and Spins

Oh, what tangled webs we weave when first we practice to believe - L.J. Peter

5.1   Statistical Mechanics of Phonons

Phonons are sound waves that move through a crystal or liquid, with the name *phonon* coming from the Greek *phone* for sound, and made to sound like photon, with which phonons have much in common. Phonons play a critical role in determining the specific heat and heat diffusion properties of solids. Additionally, it is the exchange of phonons that provides the attractive interaction necessary for superconductivity.

Sonic excitations in solids have discrete eigen-modes, just as they do for air in organ pipes. Using $\mathbf{k} = (k_x, k_y, k_z)$ to denote the wave number, density waves in a rectangular solid behave as

$$
\begin{aligned}
\delta\rho(\mathbf{r}, t) &= A \sin(\mathbf{k} \cdot \mathbf{r}) \cos(\omega t), \\
k_x &= n_x \frac{\pi}{L_x}, \ k_y = n_y \frac{\pi}{L_y}, \ k_z = n_z \frac{\pi}{L_z}, \ \omega = c_s k,
\end{aligned}
\tag{5.1}
$$

with c_s being the speed of sound. As quantum excitations, the momentum and energy of the excitations are determined by the simple relations,

$$
\mathbf{p} = \hbar \mathbf{k}, \quad \epsilon = \hbar \omega,
\tag{5.2}
$$

which looks exactly the same as that for photons. The difference being that for phonons, $\omega = c_s k$, with c_s being the speed of sound rather than the speed of light. In the macroscopic limit, we state the density of such modes as:

$$
dN = \frac{3V \, d^3 k}{(2\pi)^3} = \frac{3V \, d^3 p}{(2\pi\hbar)^3},
\tag{5.3}
$$

where the factor 3 on the r.h.s. arises because there are 2 transverse modes and one longitudinal mode for any wave number. This is exactly the same as the expression for free particles with $J = 1$. Here, it is assumed that the transverse and longitudinal modes have the same wave velocities, but it is a simple extension to assign different speeds for transverse and longitudinal phonons.

A second difference between phonons and photons is that wavelengths of sonic excitations cannot be arbitrarily short, and are cut off at the scale of the lattice spacing. To determine the wavelength and frequency of the cutoff, the usual argument is based on the number of modes of the sonic excitations as compared to the number of vibrational modes in the individual atoms in the lattice. If one treats the lattice as a collection of coupled three-dimensional harmonic oscillators, each site contributes three modes if the coupling is ignored. Once the coupling between sites is introduced, the new eigen-modes become linear combinations of the previous $3N$ modes. The cutoff in wave number can be determined by equating the number of modes,

$$
3N = \frac{3V}{(2\pi)^3} \int_0^{k_D} d^3 k = \frac{3V}{(2\pi)^3} \frac{4\pi}{3} k_D^3,
\tag{5.4}
$$

which gives the cutoff wave number as

$$k_D = \left(\frac{6\pi^2 N}{V}\right)^{1/3}. \tag{5.5}$$

The subscript D refers to Debye, who first considered the cutoff in 1912. It is more common to refer to the frequency of the cutoff rather than the wave number. The Debye frequency is

$$\omega_D = k_D c_s = c_s \left(\frac{6\pi^2 N}{V}\right)^{1/3}. \tag{5.6}$$

In the real world, transverse and longitudinal modes have different speeds of sound. One can make corrections for this and assign different values of ω_D and k_D for longitudinal and transverse modes, but as the physics of the high-frequency modes is described by the simple cutoff prescription at the $\sim 10\%$ level, we will simply assume that both modes have the same cutoff in frequency.
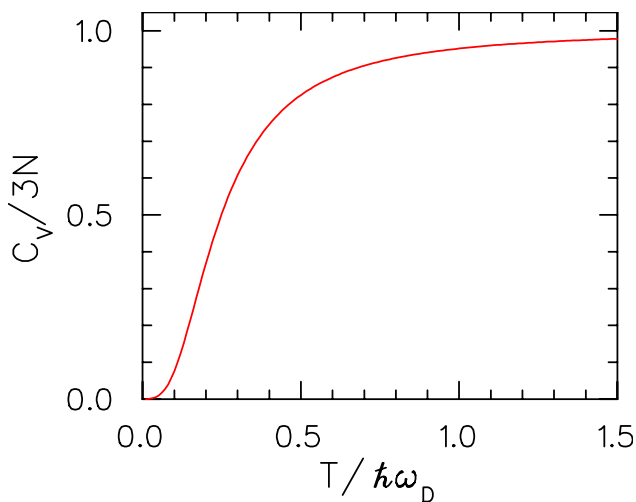
Figure 5.1: The specific heat from phonons rises as T^3 at low temperature, then approaches a constant of $3N$ at large T, where the thermal properties revert to those of the harmonic oscillator states upon which the sonic modes are built.

The occupancy of a given mode is given by

$$f(k) = \frac{e^{-\hbar c_s k/T}}{1 - e^{-\hbar c_s k/T}}, \tag{5.7}$$

which is the expression for Bose modes. The modes are bosonic because they are built on har-

monic oscillator states which allow multiple excitations. The energy and specific heat are:

$$E = \frac{3V}{(2\pi\hbar)^3} \int_{p<\hbar\omega_D/c} d^3p \, c_s p \, \frac{e^{-\beta p c_s}}{1 - e^{-\beta p c_s}}, \tag{5.8}$$

$$
\begin{aligned}
C_V &= \left.\frac{dE}{dT}\right|_V = \frac{3V}{(2\pi\hbar)^3 T^2} \int_{p<\hbar\omega_D/c} d^3p \, (c_s p)^2 \frac{e^{-\beta p c_s}}{(1 - e^{-\beta p c_s})^2} \\
&= \frac{12\pi V T^3}{(2\pi\hbar c_s)^3} \int_0^{x_D = \beta\hbar\omega_D} dx \, x^4 \frac{e^{-x}}{(1 - e^{-x})^2} \\
&= \frac{V\omega_D^3}{2\pi^2 c_s^3} D(x_D) \\
&= 3ND(x_D),
\end{aligned}
$$

$$D(x_D) \equiv \frac{3}{x_D^3} \int_0^{x_D} dx \, x^4 \frac{e^{-x}}{(1 - e^{-x})^2} . \tag{5.9}$$

The last step used the definition of the Debye frequency, Eq. (5.6). All the temperature dependence in carried by $x_D \equiv \beta\hbar\omega_D$.

At high temperatures, the *Debye function* $D(x)$ is found by setting x_D to zero which is accomplished by approximating $e^{-x} \sim 1 - x$ in Eq. (5.9),

$$D(x_D \to 0) \approx 1 \tag{5.10}$$

$$C_V(T \to \infty) \approx 3N. \tag{5.11}$$

This is the expected result for N three-dimensional harmonic oscillators. We could have predicted this result by simply knowing that at high temperature the coupling between oscillators should become irrelevant and the specific heat should be determined by the number of degrees of freedom.

To study the behavior of C_V at low temperature we first integrate by parts, using the fact that $e^{-x}/(1 - e^{-x})^2 = -(d/dx)e^{-x}/(1 - e^{-x})$. This yields,

$$D(x_D) = -3x_D \frac{e^{-x_D}}{(1 - e^{-x_D})} + \frac{12}{x_D^3} \int_0^{x_D} dx \, x^3 \frac{e^{-x}}{1 - e^{-x}}. \tag{5.12}$$

At low temperatures, $x_D \to \infty$, which allows one to ignore the first term while setting the limit on the second term to ∞.

$$D(x_D \to \infty) = \frac{72}{x_D^3} \zeta(4), \tag{5.13}$$

$$C_V(T \to 0) = 216 N \zeta(4) \left(\frac{T}{\hbar\omega_D}\right)^3$$

$$= 36 \frac{V T^3}{\pi^2 (\hbar c_s)^3} \zeta(4),$$

where $\zeta(4) = \pi^4/90$ is the Riemann-Zeta function. Aside from the different spin factor, this is the identical result one would obtain for photons if c_s were replaced by the speed of light. FYI:

the second term in Eq. (5.12) is sometimes expressed in terms of the functions $D_n(x)$, where D_n with the subscript n is the more general definition of a Debye function,

$$D_n(x) \equiv \frac{1}{x^3} \int_0^x dt \, \frac{t^n e^{-t}}{1 - e^{-t}}. \tag{5.14}$$

5.2   Feromagnetism and the Ising Model in the Mean Field Approximation

The term "ferromagnetism" refers to the spontaneous alignment of magnetic spins along a particular direction. Typically, this happens at low temperature and abates when a material is heated. The term "anti-ferromagnetism" describes a system where spins are highly ordered, but in such a way that spins often point in different directions and the net magnetic field is zero. There is actually at term "ferrimagnetism" describing a state where the ordering of spins has a structure where some reduce the magnetization, but do not cancel the magnetization, while still having spontaneous magnetization. For other materials, the terms "paramagnetic" and "diamagnetic" describe whether the induced magnetization enhances or opposes an external magnetic field. Only a few elements are ferromagnetic in their pure form, such as iron, nickel and cobalt.

In the next section we consider the Ising model, which is based on a simple Hamiltonian for spins on a lattice,

$$H = -J \sum_{\text{n.n.}} \sigma_i \sigma_j - \mu B \sum_i \sigma_i. \tag{5.15}$$

Here, the notation on the sum, n.n., refers to the fact that one sums over all pairs of nearest neighbors, and the values of the spins σ_i are limited to ± 1 (not $\pm 1/2$). One can include a higher number of spins, but those models go by a different name, Pott's model. In a two-dimensional Pott's model the spins can point in a finite number n of directions which are separated by $\Delta\theta = 2\pi/n$, and the interactions between neighboring spins behave as $-J\cos(\theta_i - \theta_j)$. In the limit of $n \to \infty$, it is called the XY model, and for $n = 2$ it becomes the Ising model. Despite the apparent simplicity of the Ising model interaction, the solution is remarkably rich, and can only be solved numerically unless one is in one dimension.

In this section, we consider the mean-field limit of the Ising model. In that limit, the spins are treated independently, with neighboring spins being treated as if they possessed the average spin, unrelated to the spin of the immediate neighbor. Thus, the interaction for a single spin is

$$H_i = -qJ\langle\sigma\rangle\sigma_i - \mu B\sigma_i, \tag{5.16}$$

where q is a phenomenological parameter, roughly representing the number of nearest neighbors, and $\langle\sigma\rangle$ is the average spin. First, we consider the case with zero external field B. One can then calculate the average of σ,

$$\langle\sigma_i\rangle = \frac{e^{\beta qJ\langle\sigma\rangle} - e^{-\beta qJ\langle\sigma\rangle}}{e^{\beta qJ\langle\sigma\rangle} + e^{-\beta qJ\langle\sigma\rangle}} = \tanh(\beta qJ\langle\sigma\rangle). \tag{5.17}$$

Because the average of a particular spin equals the average of all spins, $\langle\sigma_i\rangle = \langle\sigma\rangle$, and one has a simple equation to solve for $\langle\sigma\rangle$,

$$\langle\sigma\rangle = \tanh(\beta qJ\langle\sigma\rangle). \tag{5.18}$$

The solution can be seen graphically in Fig. 5.2 by plotting the l.h.s. and the r.h.s. of Eq. (5.18) as a function of $\langle\sigma\rangle$ and finding where the two graphs intersect. The $\langle\sigma\rangle = 0$ solution always exists. A second, non-trivial, solution exists only for $\beta qJ > 1$. The solution can be found graphically (see Fig. 5.2) by finding the intersection of the two curves. If $\beta qJ < 1$ the $\boldsymbol{tanh}$ function starts out with a smaller slope, then bends over and the two functions never intersect, whereas for $\beta qJ > 1$ the $\boldsymbol{tanh}$ function starts out above the linear function, and, because it bends over and ultimately becomes horizontal, intersects the linear function at some point. This criteria of having the second solution provides the critical temperature for the transition, $\beta qJ = 1$ or $T_c = qJ$. Some times this temperature is referred to as the Curie temperature. More neighbors or a higher coupling increases T_c. For $T < T_c$, the $\langle\sigma\rangle = 0$ solution also satisfies Eq. (5.18) but yields a higher free energy. Fig. 5.3 shows the mean magnetization as a function of temperature from solving Eq. (5.18).
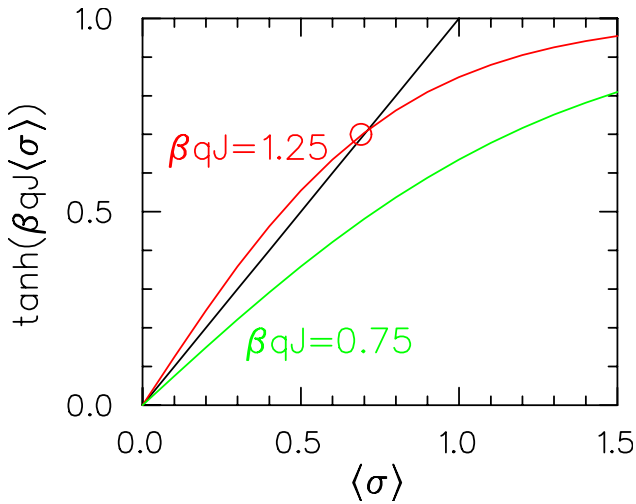
Figure 5.2: Graphical solution to Eq. (5.18). For $\beta qJ = 0.75$, there is no solution aside from $\langle\sigma\rangle = 0$, whereas for $\beta qJ = 1.25$, there is an additional solution as marked where the lines intersect. The solution exists for $\beta qJ > 1$, or for $T < T_c = qJ$.
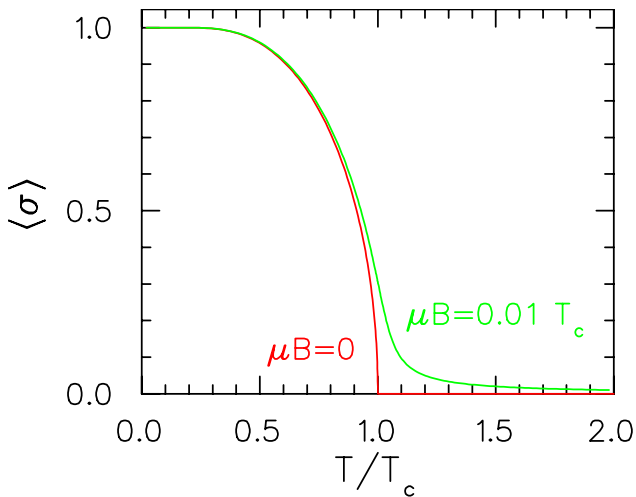
Figure 5.3: The mean magnetization as a function of temperature in the mean field approximation. Spontaneous magnetization ensues for $T < T_c = qJ$ for zero field, whereas for finite the transition is smoothed over and some magnetization ensues for $T > T_c$.
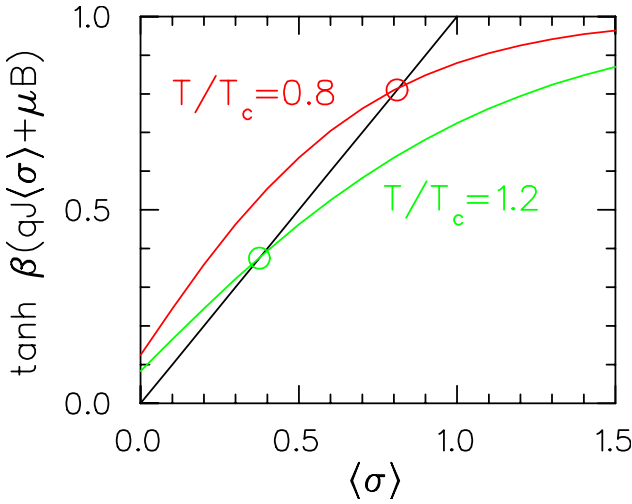
Example 5.1:

Find an expression for $\langle\sigma\rangle$ in the limit of small deviation $\delta T = T_c - T$.

In this limit $\langle \sigma \rangle$ will be small, so the **tanh** function in Eq. (5.18) can be expanded for small σ,

$$\langle \sigma \rangle = \beta q J \langle \sigma \rangle - \frac{1}{3} \left(\beta q J \langle \sigma \rangle \right)^3 + \cdots .$$

Solving for $\langle \sigma \rangle$,

$$\langle \sigma \rangle \approx \sqrt{\frac{3(\beta q J - 1)}{(\beta q J)^3}} \approx \sqrt{\frac{3(T_c - T)}{T_c}}. \tag{5.19}$$

Figure 5.4: Graphical solution to Eq. (5.21), magnetization with external field. The magnetic field raises the **tanh** curves from the origin so that solutions exist for any temperature, not just $T < T_c$.

One can add an external magnetic field to the mean field equation. This adds a term $\mu B \sigma$ to the energy for each spin, which leads to the following expression for the average spin,

$$\langle \sigma_i \rangle = \frac{e^{\beta(qJ\langle\sigma\rangle + \mu B)} - e^{-\beta(qJ\langle\sigma\rangle + \mu B)}}{e^{\beta(qJ\langle\sigma\rangle + \mu B)} + e^{-\beta(qJ\langle\sigma\rangle + \mu B)}}, \tag{5.20}$$

or

$$\langle \sigma \rangle = \tanh \beta \left(qJ\langle\sigma\rangle + \mu B \right). \tag{5.21}$$

Unlike the solution with $B = 0$, the $\langle \sigma \rangle \neq 0$ solution exists for all temperatures. The discontinuity of the slope in the plot of $\langle \sigma \rangle$ vs. T is then smoothed out as shown in Fig. 5.3.

The results described above can alternatively be found by minimizing the free-energy as a function of $\langle \sigma \rangle$. This alternative method is a bit longer, but has the advantage of showing how the $\langle \sigma \rangle = 0$ solution is less favorable. Because the free energy is given by $F = E - TS$, we first find an expression for the entropy in terms of $\langle \sigma \rangle$, the temperature, the field and the coupling qJ. Assuming the probability that a spin in the spin-up state is p,

$$\frac{S}{N} = -\frac{1}{N} \sum_i p_i \ln p_i = -p \ln p - (1 - p) \ln(1 - p). \tag{5.22}$$

To express this in terms of $\langle \sigma \rangle$, one inverts the expression

$$\langle \sigma \rangle = p - (1 - p) = 2p - 1, \tag{5.23}$$

to find

$$p = \frac{1 + \langle\sigma\rangle}{2}, \quad (1 - p) = \frac{1 - \langle\sigma\rangle}{2}. \tag{5.24}$$

The entropy is then

$$\frac{S}{N} = -\frac{1 + \langle\sigma\rangle}{2}\ln\left(\frac{1 + \langle\sigma\rangle}{2}\right) - \frac{1 - \langle\sigma\rangle}{2}\ln\left(\frac{1 - \langle\sigma\rangle}{2}\right) \tag{5.25}$$

The energy per spin is

$$\frac{E}{N} = -\frac{qJ}{2}\langle\sigma\rangle^2 - \mu B\langle\sigma\rangle, \tag{5.26}$$

where the factor of $1/2$ ensures that the interaction between neighboring spins is not double counted. Now, one must simply minimize $F = E - TS$ with respect to $\langle\sigma\rangle$,

$$\frac{\partial}{\partial\langle\sigma\rangle}\frac{F}{N} = 0 \tag{5.27}$$

$$= \frac{\partial}{\partial\langle\sigma\rangle}\left\{-\frac{qJ}{2}\langle\sigma\rangle^2 - \mu B\langle\sigma\rangle + T\frac{1 + \langle\sigma\rangle}{2}\ln\left(\frac{1 + \langle\sigma\rangle}{2}\right) + T\frac{1 - \langle\sigma\rangle}{2}\ln\left(\frac{1 - \langle\sigma\rangle}{2}\right)\right\}$$

$$0 = -qJ\langle\sigma\rangle - \mu B + \frac{T}{2}\ln\left(\frac{1 + \langle\sigma\rangle}{1 - \langle\sigma\rangle}\right).$$

Rearranging the terms so that the logarithm is isolated on one side, then taking the exponential of both sides,

$$e^{2\beta qJ\langle\sigma\rangle + 2\beta\mu B} = \frac{1 + \langle\sigma\rangle}{1 - \langle\sigma\rangle}, \tag{5.28}$$

$$\langle\sigma\rangle = \frac{e^{2\beta qJ\langle\sigma\rangle + 2\beta\mu B} - 1}{e^{2\beta qJ\langle\sigma\rangle + 2\beta\mu B} + 1}$$

$$= \tanh(\beta qJ\langle\sigma\rangle + \beta\mu B).$$

This is precisely the earlier result, Eq. (5.21)

5.3   One-Dimensional Ising Model

The Ising model can be written in any number of dimensions, but the one-dimensional situation is of particular interest because it can be solved analytically. The method for solving the model is of interest in its own right, but it also represents a striking lesson of the folly of applying the mean field approximation in certain situations.

It is intuitively clear that a one-dimensional system cannot undergo spontaneous magnetization. For any temperature above zero, there is always the possibility that any two consecutive spins might not point in the same direction. The penalty for a long chain to break into two equal regions, one spin-up the other spin-down, is thus finite. The same is not true for a two- or three-dimensional system. In those cases, one must flip spins along a boundary whose size grows to infinity as the size of the system goes to infinity. This rather manifest fact is sometimes

stated as a principle: "There is no long-range order in one-dimension" (unless there is a long-range interaction). This principal is manifestly violated in the mean-field approximation which results in spontaneous magnetization, independent of the dimensionality. This underscores the importance of understanding the role of correlations, rather than averaged behavior, in driving the behavior of the system.

Even though the average magnetization for the $B = 0$ case is trivially zero in a one-dimensional system, the $B \neq 0$ behavior seems non-trivial at first glance. However, it can be solved analytically. To see this we consider the partition function,

$$Z = \sum_{\sigma_1 \cdots \sigma_n} \exp \beta \left\{ J(\vec{\sigma}_1 \cdot \vec{\sigma}_2 + \vec{\sigma}_2 \cdot \vec{\sigma}_3 + \cdots \vec{\sigma}_n \cdot \vec{\sigma}_1) + \mu B(\sigma_1 + \sigma_2 \cdots + \sigma_n) \right\}, \quad (5.29)$$

where each spin is coupled through an energy $-J\vec{\sigma}_i \cdot \vec{\sigma}_{i+1}$ to the adjacent spin, and the last spin is coupled back to the first spin, as if the lattice is bent into a ring. The energy can be expressed as a sum over the spin values,

$$H = \sum_i -J\sigma_i\sigma_{i+1} - \frac{1}{2}\mu B(\sigma_i + \sigma_{i+1}), \quad (5.30)$$

and the partition function can then be written as a product of matrix elements,

$$Z = \sum_{\sigma_1 \cdots \sigma_n} \langle \sigma_1 | \mathcal{P} | \sigma_2 \rangle \langle \sigma_2 | \mathcal{P} | \sigma_3 \rangle \cdots \langle \sigma_n | \mathcal{P} | \sigma_1 \rangle, \quad (5.31)$$

$$\mathcal{P}_{i,i+1} = \exp \beta \left\{ J\sigma_i\sigma_{i+1} + \frac{1}{2}\mu B(\sigma_i + \sigma_{i+1}) \right\},$$

$$= \begin{pmatrix} e^{\beta(J+\mu B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu B)} \end{pmatrix}.$$

Here, $\mathcal{P}$ is known as the *transfer matrix*, where the states with $\sigma_i = \pm 1$ are represented by a two-component vector with elements either up or down respectively. The partition function can then be expressed as

$$Z = \text{Tr } P^n, \quad (5.32)$$

which due to the cyclic property of the trace can also be written as

$$Z = \text{Tr } (U\mathcal{P}U^\dagger)^n, \quad (5.33)$$

where U is a unitary transformation ($U^\dagger = U^{-1}$). We will use the transformation that diagonalizes $\mathcal{P}$, $U\mathcal{P}U^\dagger = P_\lambda$, so

$$Z = \text{Tr } (P_\lambda)^n. \quad (5.34)$$

The elements of P_λ are the eigenvalues of $\mathcal{P}$,

$$\mathcal{P}_\lambda = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \quad (5.35)$$

$$\lambda_\pm = e^{\beta J} \cosh(\beta \mu B) \pm \sqrt{e^{2\beta J} \sinh^2(\beta \mu B) + e^{-2\beta J}}.$$

One can now write the partition function as

$$Z = \text{Tr } \mathcal{P}_\lambda^n = \text{Tr } \begin{pmatrix} (\lambda_+)^n & 0 \\ 0 & (\lambda_-)^n \end{pmatrix} = (\lambda_+)^n + (\lambda_-)^n. \tag{5.36}$$

As $n \to \infty$, the second term becomes unimportant as can be seen from expanding $\ln Z$,

$$\ln Z = \ln \left\{ (\lambda_+)^n \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^n \right] \right\} = n \ln(\lambda_+) + \ln \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^n \right], \tag{5.37}$$

and the second term vanishes because $\lambda_- < \lambda_+$. This gives the final result for the partition function,

$$\ln Z = n \ln(\lambda), \ \lambda = e^{\beta J} \cosh(\beta \mu B) + \sqrt{e^{2\beta J} \sinh^2(\beta \mu B) + e^{-2\beta J}}. \tag{5.38}$$

Example 5.2:

In terms of J, μB and T, find the average magnetization (per spin) for the 1-d Ising model.

Because the magnetic field enters the Hamiltonian as $-\mu \vec{B} \cdot \vec{\sigma}$, the average spin is

$$\langle \sigma \rangle = \frac{1}{n} \frac{\partial \ln Z}{\partial(\beta \mu B)}.$$

This gives,

$$\langle \sigma \rangle = \frac{\sinh(\beta \mu B) + \frac{\cosh(\beta \mu B) \sinh(\beta \mu B)}{\sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}}}}{\cosh(\beta \mu B) + \sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}}}$$

$$= \frac{\sinh(\beta \mu B)}{\sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}}}.$$

Although the exact analytic solution only exists for one-dimension, the technique of applying a transfer matrix can be applied to more dimensions. For a finite two-dimensional lattice of size $L_x \times L_y$, one can treat the configuration of each row as a state, with 2^{L_x} possible values. One can then construct the ($2^{L_x} \times 2^{L_x}$) transfer matrix between neighboring columns. This is discussed in Chapter 15 of Huang.

5.4   Lattice Gas, Binary Alloys, Percolation...

Aside from the multitude of theoretical treatments of spins on lattices, lattices are also used to model liquid-gas phase transitions, percolation and fragmentation. Mainly to provide an introduction to the vocabulary, we present a list of several such models here. The first model we describe is the lattice gas. Here, a lattice of sites are assigned either zero (no particle) or 1 (occupied) with a set probability. The configuration is weighted according to the probability

$\exp\{-N_{nn}\epsilon_0/T\}$, where N_{nn} is the number of nearest neighbor sites. Effectively, this models a gas with a short-range attractive attraction, represented by the positive weight (assuming $\epsilon_0 <$ 0). Because the occupancy is never greater than unity, it also mimics a repulsive interaction at short distances. This can be used to model the shapes of drops, which can become very irregular near the critical point.

Binary alloys are more physically represented by lattices. Each site is assigned a specific species. The lattice is first constructed according to the physical structure of the modeled system, e.g., it might be body-centered cubic, before $N_a + N_b = N$ atoms are assigned to the N lattice sites. The entire configuration is weighted according to $e^{-\beta H}$ where

$$H = \epsilon_{aa}N_{aa} + \epsilon_{bb}N_{bb} + \epsilon_{ab}N_{ab}, \tag{5.39}$$

where N_{ij} are the number of neighboring combinations where one species is i and the other is j. For the case of two equally populated species on a BCC lattice, the lowest energy state is chemically divided into alternate planes. Above some critical temperature, the long-range order disappears. Like the lattice gas, the binary model and the Ising model all involve assigning one of two numbers (0 or 1 for a lattice gas, '1' or '2' for a binary alloy, and '-1' or '1' for the Ising model) to lattice sites, then adding a weight which depends only nearest neighbor occupancies. Thus, all three approaches can at times be mapped onto one another.

One numerical strategy for treating a lattice gas or a binary alloy is referred to as *Markov-Chain Monte Carlo* (MCMC). One begins with some configuration, which could be arbitrary, then considers small changes, e.g. two random sites in the binary alloy switch a with b species. The two configurations, which we will call $\mathcal{C}_1$ and $\mathcal{C}_2$, have different energies separated by ΔE. One generates a chain of configurations by following the following algorithm, assuming you already have configuration C_i

1. Choose a configuration C_{i+1} by making a small change to C_i, and calculate $\Delta E = E_{i+1} - E_i$.

2. If $\Delta E < 0$, the configuration is more favorable. You add the new configuration to your list.

3. If $\Delta E > 0$, you throw a random number r from a uniform distribution, $0 < r < 1$. If $r < e^{-\beta|\Delta E|}$, add the new configuration to your list.

4. If $r > e^{-\beta|\Delta E|}$, keep the old configuration, i.e. $C_{i+1} = C_i$.

As long as the chance of choosing C_{i+1} given one is in C_i is identical to the chance of choosing C_i given one is in C_{i+1}, the system obeys time reversal and is effectively ergodic. If one imagines a very large ensemble of chains where the probability of being in a configuration i is $p_i \approx e^{-\beta E_i}$, the probability will stay that way if the rate at which one goes from i to j is the same as the rate one goes from j to i. Given the rate of choosing j given i is the same as the rate of choosing i given j the probability ratio p_i/p_j will only be maintained if the rate of acceptances is proportional to $e^{-\beta\Delta E}$. As long as the chain eventually reaches all configurations, a sampling over time of a given chain should represent a proper statistical sampling. This particular algorithm for the chain is known as Metropolis-Hastings. In some cases, there are barriers, e.g. nucleation barriers for the lattice gas, which require a long chain before it can reach certain configurations. It can also take a long time to build up long-range correlations, such as those that exist near a phase transition. For such problems much more sophisticated strategies need to be developed.

Another class of calculations are called "percolation models". These models usually address the question of how long a given chain might extend across a lattice, given that some of the links might be broken. Percolation models can be divided into bond percolation or site percolation. In bond percolation, a lattice is constructed with potential bonds between all nearest neighbors. Only a fraction p of bonds are unbroken, with the choice of which bonds to break chosen randomly. For a two-dimensional $(L \times L)$ or three-dimensional $(L \times L \times L)$ lattice, one can ask the question of whether an unbroken chain can be found that extends completely across the lattice. As $L \rightarrow \infty$, an infinite chain will exist if and only if $p > p_c$. The critical value p_c depends on dimensionality and the type of lattice, e.g., cubic or BCC, and whether one is doing bond percolation of site percolation. For site percolation, the sites are populated with probability p, and bonds are assumed to exist if neighboring sites are filled. For bond percolation of a simple three-dimensional cubic lattice, $p_c \approx 0.2488$.

(See http://mathworld.wolfram.com/PercolationThreshold.html).

Percolation is used to study electric properties and fragmentation. Percolation models have been applied to fragmentation of nuclei from nuclear collisions. The same approaches have been compared to experiments where a glass is dropped and shattered and analyzed for the probability of a given shard having a specific mass.

The method of linked lists provide a good computational strategy for percolation models. One defines an object for each site, that stores a pointer to the cluster to which it belongs, and a pointer to the next site in the cluster. One also defines an object for each cluster, with each cluster representing a linked list, i.e. the cluster need only be a pointer to the first site in the list. One begins with all the bonds broken and each site is in its own list. As bonds are mended, one checks to see whether the sites are in the same cluster, and if not, the clusters are merged.

5.5   Problems

1. The speed of sound in copper is 3400 m/s, and the number density is $\rho_{Cu} = 8.34 \times 10^{28}$ m^{-3}.

   (a) Assuming there are two free electrons per atom, find an expression for C_V/N_a (where N_a is the number of atoms) from the free electrons in copper by assuming a free gas of electrons. Use the expression from chapter 2 for a low-T Fermi gas,

   $$\delta E = T^2 \frac{\pi^2}{6} D(\epsilon),$$

   where $D(\epsilon)$ is the density of single particle electron states at the Fermi surface. Give answer in terms of T and the Fermi energy ϵ_F.

   (b) In terms of ϵ_F and $\hbar\omega_D$, find an expression for the temperature at which the specific heat from electronic excitations equals that from phonons. Use the low T expression for the specific heat of phonons.

   (c) What is $\hbar\omega_D$ in eV? in K?

   (d) What is ϵ_F in eV? in K?

   (e) What is the numerical value for the answer in (b) in K?

2. Consider the mean-field solution to the Ising model of Eq. (5.18).

   (a) For small temperatures, show that the variation $\delta\langle\sigma\rangle = 1 - \langle\sigma\rangle$ is:

   $$\delta\langle\sigma\rangle \approx 2e^{-2qJ/T}.$$

   (b) Consider a function of the form, $y(T) = e^{-1/T}$. Find dy/dT, d^2y/dT^2, and d^ny/dT^n evaluated at $T = 0^+$. What does this tell you about doing a Taylor expansion of $y(T)$ about $T = 0^+$?

3. Show that in the mean field approximation of the Ising model the susceptibility,

   $$\chi \equiv \frac{d\langle\sigma\rangle}{dB},$$

   becomes

   $$\chi = \frac{(1 - \langle\sigma\rangle^2)\mu}{T - T_c + \langle\sigma\rangle^2 T_c}.$$

4. The total energy for the Ising model in the mean field approximation from summing over all the sites in Eq. (5.16) is

   $$H = -\frac{N}{2}qJ\langle\sigma\rangle^2 - N\mu B\langle\sigma\rangle,$$

   where N is the number of sites, and the factor of $1/2$ is a correction for double counting. In terms of $T, T_c, \mu B$ and $\langle\sigma\rangle$, find an expression for

   $$C_v = \frac{1}{N}\frac{dE}{dT}.$$

5. Consider the one-dimensional Ising model. For the following, give analytic answers in terms of T, μB and J.

   (a) What are the high B and low B limits for $\langle\sigma\rangle$?
   (b) What are the high T and low T limits for $\langle\sigma\rangle$?
   (c) Find an exact expression for the specific heat (per spin) in terms of T, μB and J.
   (d) What are the high B and low $B = 0$ limits for the specific heat?
   (e) What are the high T and low T limits for the specific heat?
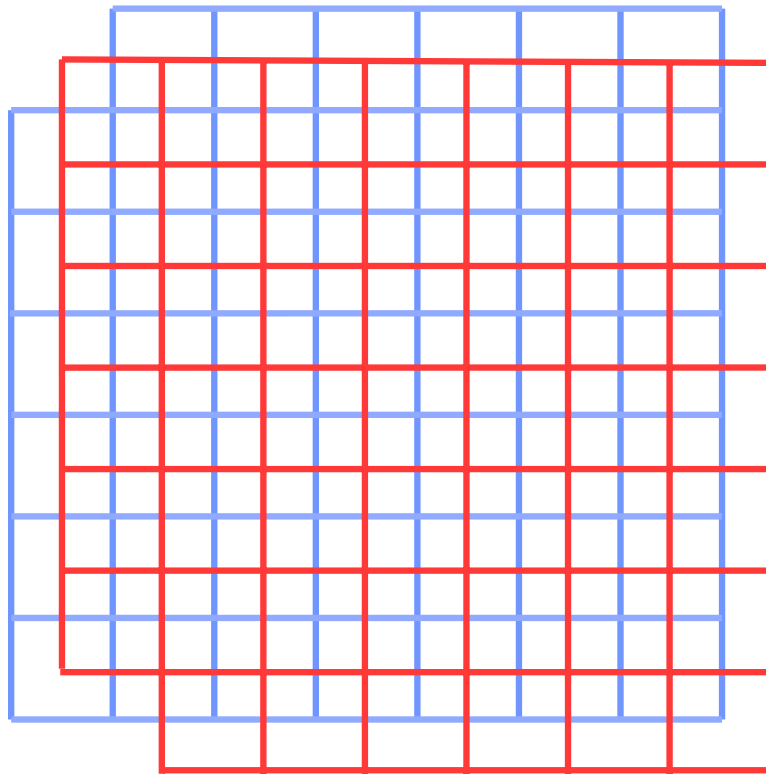
6. Consider two-dimensional bond percolation of a large $L \times L$ simple square lattice. First consider the red lattice, where we will remove a percentage p of the red bonds. Next, we will remove all blue bonds that intersect a surviving red bond. Thus, in the limit of a large lattice, the fraction of blue bonds broken will be $(1 - p)$.

   (a) List which combinations of the following are possible.

      a) A connected string of blue bonds extends all the way from the bottom blue row to the top of the lattice.

      b) There is no connected string of blue bonds that extends all the way from the bottom to the top.

      c) A connected string of red bonds extends all the way from the left side to the right side of the red lattice.

      d) There is no connected string of red bonds that extends all the way from the left side to the right side of the red lattice.

   (b) What is p_c for a simple square lattice in bond percolation?

7. Consider a one-dimensional bond percolation model of an infinitely long string, where the probability of any set of neighbors being connected is p.

   (a) In terms of p, find the probability that a fragment will have size A.

   (b) What is the average size of a fragment?

6   Landau Field Theory

Nothing is real and nothing to get hung about - J. Lennon, P. McCartney

6.1   What are fields?

Field theory represents a means by which the state of a system is expressed in terms of a quantity ϕ which is a function of the position $\vec{x}$. The quantity ϕ represents a measure of some quantity averaged over a small volume δV. One example of such a quantity is the average magnetization per unity volume, $\vec{m}(\vec{x})$, while another is the number density $\rho(\vec{x})$. In each case, when viewed microscopically, the magnetization density and the number density are not smooth quantities, but are carried by point-like objects. Thus, field theories are only applicable when looking at averages or correlations on scales larger than the microscopic scale of the matter. As will be seen further ahead, correlation length scales tend to go to infinity near a critical point, a region where field theories are especially applicable.

Finite-temperature field theories are usually expressed in terms of the free-energy density, $f(\vec{x})$,

$$F \equiv \int d^3x \, f(\phi, \nabla\phi), \tag{6.1}$$

$$f = \frac{\kappa}{2}(\nabla\phi)^2 + \mathcal{V}(\phi).$$

The second statement, that derivatives appear only quadratically, is certainly an ansatz, but as we will see below, the approximation is motivated from physical arguments. The choice of the symbol F for the free energy should not suggest that one is always interested in the Helmholtz free energy, as for some applications it might represent the Gibb's free energy or the entropy. If the parameter ϕ changes slowly compared to the time required to come to thermal equilibrium (constant temperature), the use of the Helmholtz free energy is the appropriate choice.

The gradient term is warranted whenever there is an attractive short-range interaction. As an example, one can look at the two-dimensional $X - Y$ model, where spins on a lattice are denoted by an angle θ, with the following interaction between nearest neighbors,

$$V_{ij} = V_0[1 - \cos(\theta_i - \theta_j)]. \tag{6.2}$$

For perfectly aligned spins $V_{ij} = 0$, whereas $V_{ij} = 2V_0$ when the spins are anti-aligned. For small differences in the angle this can be approximated as:

$$
\begin{aligned}
V &= \frac{V_0}{2} \sum_{i_x, i_y} (\theta_{i_x} - \theta_{i_x-1})^2 + (\theta_{i_y} - \theta_{i_y-1})^2 &(6.3)\\
&= \frac{V_0}{2} \sum_{i_x, i_y} \frac{\left[(\theta_{i_x} - \theta_{i_x-1})^2 + (\theta_{i_y} - \theta_{i_y-1})^2\right]}{(\Delta L)^2}(\Delta L)^2 \\
&= \frac{V_0}{2} \int dx \, dy \, (\nabla\theta)^2,
\end{aligned}
$$

where ΔL is the lattice spacing. Thus, in this example, it is easy to identify κ in Eq. (6.1) with the microscopic quantity V_0. Because the free energy is $F = E - TS$, terms that contribute to the energy density carry through to the free energy density.

A second example is that of a liquid, or gas, where ϕ refers to the density $\rho(\vec{x})$. In that case the attractive interaction is mainly the two-particle potential $v(\vec{r})$,

$$V = \frac{1}{2} \int d^3x \, d^3x' \, \langle \rho(\vec{x})\rho(\vec{x}') \rangle v(\vec{x} - \vec{x}'). \tag{6.4}$$

For a field theory, we are interested in an average density at a given point $\bar{\rho}(\vec{x})$. In terms of the averages, the two-body potential energy is

$$V = \frac{1}{2} \int d^3x \, d^3x' \, \bar{\rho}(\vec{x})\bar{\rho}(\vec{x}')C(\vec{x}, \vec{x}')v(\vec{x} - \vec{x}'), \tag{6.5}$$

$$C(\vec{x}, \vec{x}') \equiv \frac{\langle \rho(\vec{x})\rho(\vec{x}') \rangle}{\bar{\rho}(\vec{x})\bar{\rho}(\vec{x}')}.$$

Here, $C(\vec{r})$ is the correlation (unity if uncorrelated) for two particles to be separate by $\vec{r}$. In the low density limit (particles interacting only two-at-a-time) the correlation is approximately

$$C(\vec{x}, \vec{x}') \approx e^{-v(\vec{r})/T}, \quad \vec{r} \equiv \vec{x} - \vec{x}'. \tag{6.6}$$

Next, one can make a gradient expansion for $\bar{\rho}$,

$$\bar{\rho}(\vec{x}') \approx \bar{\rho}(\vec{x}) + (x_i' - x_i)\partial_i\bar{\rho}(\vec{x}) + \frac{1}{2}\sum_{i,j}(x_i' - x_i)(x_j' - x_j)\partial_i\partial_j\bar{\rho}(\vec{x}), \tag{6.7}$$

and inserting into Eq. (6.5),

$$V = \frac{1}{2} \int d^3x \, d^3r \, \bar{\rho}(\vec{x})\bar{\rho}(\vec{x})e^{-v(\vec{r})/T}v(\vec{r}) + \frac{1}{2} \int d^3x \, d^3r \, \bar{\rho}(\vec{x})\frac{\partial}{\partial x_i}\bar{\rho}(\vec{x})r_i e^{-v(\vec{r})/T}v(\vec{r})$$
$$+ \frac{1}{4} \int d^3x \, d^3r \, \bar{\rho}(\vec{x})\frac{\partial^2}{\partial x_i \partial x_j}\bar{\rho}(\vec{x}) \, r_i r_j e^{-v(\vec{r})/T}v(\vec{r}). \tag{6.8}$$

Because $V(\vec{r})$ is an even function of $\vec{r}$, the odd terms in the expansion can be discarded leaving

$$V = \int d^3x \left\{ \mathcal{V}_{\text{local}}(\bar{\rho}(\vec{x})) - \frac{\kappa}{2}\bar{\rho}(\vec{x})\nabla^2\bar{\rho}(\vec{x}) \right\} \tag{6.9}$$

$$\kappa = -\frac{1}{6} \int d^3r \, r^2 e^{-v(r)/T}v(r),$$

$$\mathcal{V}_{\text{local}}(\bar{\rho}) = \frac{\bar{\rho}^2}{2} \int d^3r \, e^{-v(r)/T}v(r),$$

where the approximation for the correlation, Eq. (6.6), has been used for the correlation function. The extra factor of $1/3$ in the expression for κ comes from assuming that $v(\vec{r})$ is rotationally symmetric, so that the integral with r_x^2 is one third the integral with r^2. The term $\mathcal{V}_{\text{local}}$ is the contribution to the potential energy when density gradients are neglected. We will drop the subscript "local" and simply consider $\mathcal{V}(\bar{\rho})$ to be the free energy density of a system at uniform density. Assuming that other contributions to the energy do not contribute to the ∇^2 term, the free energy becomes:

$$F = \int d^3x \left[\mathcal{V}(\bar{\rho}(\vec{x})) - \frac{\kappa}{2}\bar{\rho}(\vec{x})\nabla^2\bar{\rho}(\vec{x}) \right] \tag{6.10}$$

$$= \int d^3x \left[\mathcal{V}(\bar{\rho}(\vec{x})) + \frac{\kappa}{2}(\nabla\bar{\rho})^2 \right].$$

where the last step involved integrating by parts. Because the definition of density implies an average the "bar" over $\bar{\rho}$ can be dropped .

An obvious question concerns how one might justify dropping higher order gradients in the Taylor expansion of $\rho(\vec{x}')$. Certainly, such terms should exist. Calculating such higher-order terms would involve finding expressions for coefficients like were found for κ in Eq. (6.10) above. However, rather than one extra power of r^2 as in Eq. (6.10), one would find powers of $r^4, r^6 \cdots$. If the potential $v(r)$ falls off exponentially with a sufficiently short range, these terms can be neglected, and even if they fall off with a power-law (such as $1/r^6$ for the Lennard-Jones potential) screening effects often give an exponential fall-off which results in exponential behavior for large r, which keeps the coefficients finite. If one follows through with the algebra for the extra terms, they will fall off as a ratio of the interaction range to the characteristic length for the density changing, $(R_{\text{int}}/L_{\text{char}})^n$. Thus, the higher-order terms can be neglected if one is considering variations of the density on length scales much greater than the range of the interaction.

One common mistake students make when seeing a mean-field theory is to assume that the gradient term originated from kinetic terms in the microscopic equations of motion. Instead, such terms originate from attractive short-range interactions. In fact, if the interactions were repulsive κ would have the wrong sign. The terms do, however, represent kinetic-energy type terms in the free energy. For instance, if V is quadratic in the fields, one can express the free energy as a sum of independent contributions from "momentum" modes where $\delta\rho \sim \sin(kx)$. In that case, the ∇^2 term in the free energy behaves as k^2, just as it would for free particles or sound modes.

We re-emphasize that the application of field theories is based on the assumption that the characteristic length scale of fluctuations is much larger than the range of interaction. It is also assumed that the scale is greater than characteristic quantum length scalses, such as the thermal wavelength $\hbar/sqrtmT$, or the inverse Fermi momentum, $\hbar/p_F$. If one is considering scales of the order of these quantum length scales, one should use a quantum approach such as density-functional theory.

Quantum field theories are often required to study dynamical quantum excitations, such as phonons. These theories significantly differ from the classical, Landau, theories described here. Here, the fields are real quantities like density, whereas quantum fields are often represented by complex numbers and need to be squared to represent an observable quantity.

6.2   Calculating surface energies in field theory

As discussed in Chapter 3, two phases can co-exist in equilibrium. However, there is a penalty associated with the boundary. This is referred to as a surface energy or a surface free energy at finite temperature, and has units of energy per area. To calculate the surface free energy, we consider a large box where the left side ($x \ll 0$) is in a gaseous state and the right side ($x \gg 0$) is in a liquid state. If A is the area of the $y - z$ plane, the total Helmholtz free energy ($F = E - TS$) is

$$F/A = \int dx \left[\mathcal{V}(\rho(x)) + \frac{\kappa}{2}(\partial_x \rho)^2 \right]. \tag{6.11}$$

Our goal is to find the density profile that minimizes the free energy given the constraint that it has the density of the gas at $x = -\infty$ and the density of the liquid at $x = \infty$. We will assume that the temperature is a constant over the profile, which is reasonable if energy moves more quickly than particles. By having the system able to connect to particle baths at $\pm\infty$, one can state that the chemical potential should be chosen so that $\mu(x \to \infty) = \mu(x \to -\infty) = \mu_0$, with μ_0 being chosen so that the two phases are at equilibrium, i.e., they also have the same pressure, $P(x \to \infty) = P(x \to -\infty) = P_0$.

The Helmholtz free energy is the quantity one should minimize if a system has fixed particle number and is connected to a heat bath. If the system can also exchange particles with a bath whose chemical potential is μ_0, one should include the free energy associated with the particles that have been added or subtracted into the bath. In that case, the number is no longer conserved, and the quantity to be minimized is

$$F - \mu_0 N \;=\; A \int dx \left[\mathcal{V}(\rho) - \mu_0 \rho + \frac{\kappa}{2}(\partial_x \rho)^2 \right] \tag{6.12}$$

$$=\; A \int dx \left[-P(\rho) + (\mu - \mu_0)\rho + \frac{\kappa}{2}(\partial_x \rho)^2 \right]. \tag{6.13}$$

The last step used the thermodynamic identity that the free energy density for a gas at uniform density and temperature is $\mathcal{V} = \mu\rho - P$. Figure 6.1 illustrates a plot of $-P + (\mu - \mu_0)\rho$ as a function of the density for fixed temperature. By picking the appropriate value of μ_0 for the specified temperature T, the two minima are at the same level and phase coexistence can occur. In fact, at the minima $\mathcal{V} - \mu_0\rho = -P$, and the statement that both minima are equal is equivalent to saying that for the specified T, μ_0 was chosen so that two densities have the same pressure.
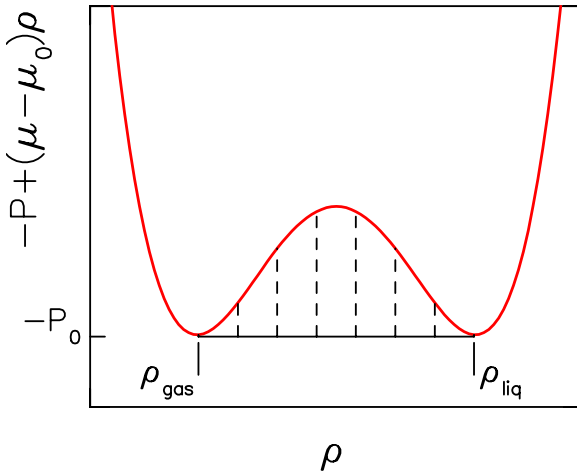
Figure 6.1: Here, $P(\rho, T)$ and $\mu(\rho, T)$ are the pressure and chemical potential of a gas at uniform density ρ and temperature T. If the system is connected to a heat bath with chemical potential μ_0, the system will choose a density to minimize $-P + (\mu - \mu_0)\rho$. If μ_0 equals the chemical potential for phase coexistence, the two minima are degenerate and are located at the coexistence densities, ρ_{gas} and ρ_{liq} for the given T. A boundary between the two regions must traverse the region of less favorable free-energies, hence there is a surface free energy as described in Eq. (6.17). In this expression, the surface energy is found by integrating over the densities between the two coexistence values with the integrand being proportional to the square root of the height, illustrated by the dashed lines.

The cost of the net free energy for a profile $\rho(x)$ that differs from the two coexistence densities

is,

$$\Delta F/A = \int dx \left[-P + (\mu - \mu_0)\rho + \frac{\kappa}{2}(\partial_x \rho)^2 \right] - \int dx\,(-P_0) \tag{6.14}$$

$$= \int dx \left[P_0 - P + (\mu - \mu_0)\rho + \frac{\kappa}{2}(\partial_x \rho)^2 \right],$$

where P_0 is the coexistence pressure for the temperature T. The last term, $\int dx\, P_0$, is a constant for fixed limits, and is subtracted so that the integrand goes to zero at large x, allowing the integration region to be stated as $-\infty < x < \infty$. The quantity $(\mu - \mu_0)\rho + P_0 - P$ is represented by the height of the lines in the shaded region of Fig. 6.1. After replacing $dx(\partial_x \rho)^2 = d\rho(\partial_x \rho)$, the integration can be restated in terms of the densities,

$$\Delta F/A = \int_{\rho_{\text{gas}}}^{\rho_{\text{liq}}} d\rho \left[\frac{P_0 - P + (\mu - \mu_0)\rho}{\partial_x \rho} + \frac{\kappa}{2}\partial_x \rho \right]. \tag{6.15}$$

Next, one chooses $\partial_x \rho$ to minimize the free energy. The first term would be minimized by choosing a step function profile, because there would then be no region in the unfavorable region described in Fig. 6.1. But for a step function $\partial_x \rho = \infty$, and the second term would be minimized by a smooth profile. Minimizing the functional integral yields,

$$\frac{d}{d(\partial_x \rho)} \left[\frac{P_0 - P + (\mu - \mu_0)\rho}{\partial_x \rho} + \frac{\kappa}{2}\partial_x \rho \right] = 0, \tag{6.16}$$

$$\partial_x \rho = \sqrt{2\frac{P_0 - P + (\mu - \mu_0)\rho}{\kappa}},$$

$$\Delta F/A = \sqrt{2\kappa} \int_{\rho_{\text{gas}}}^{\rho_{\text{liq}}} d\rho \sqrt{P_0 - P + (\mu - \mu_0)\rho}.$$

Thus, the coefficient κ is responsible for the surface free energy. If there were no gradient term in the functional for the free energy, the density profile between equilibrated phases would become a step function. Looking back at the expression for κ in the previous section shows that surface energies will be larger for interactions with longer ranges.

Example 6.1:

Let the function $\Delta\Psi \equiv P_0 - P + (\mu - \mu_0)\rho$ be approximated with the following form,

$$\Delta\Psi = \frac{B}{2}\left[(\rho - \rho_c)^2 - \alpha^2 \right]^2.$$

Find the surface energy in terms of B, α and κ.

From Eq. (6.17), the surface energy is

$$F/A = \sqrt{\kappa B} \int_{\rho - \rho_c = -\alpha}^{\rho - \rho_c = \alpha} d\rho \sqrt{[(\rho - \rho_c)^2 - \alpha^2]^2}$$

$$= \sqrt{\kappa B} \int_{-\alpha}^{\alpha} dx\,(\alpha^2 - x^2)$$

$$= \frac{4}{3}\alpha^3 \sqrt{\kappa B}$$

Note that B and α^2 would depend on the temperature, and one could even have κ depend on temperature if one wanted. Because two independent minima no longer exist for $T > T_c$, the coefficient α^2 goes to zero at T_c, which means that the surface energy also becomes zero.

6.3   Correlations and Susceptibilities in the Critical Region

Correlations, e.g.,

$$C(\vec{r} - \vec{r}') \equiv \langle(\phi^*(\vec{r}) - \bar{\phi}^*)(\phi(\vec{r}') - \bar{\phi})\rangle, \tag{6.17}$$
$$\bar{\phi} \equiv \langle\phi\rangle,$$

are central to the study of critical phenomena. Note that here we have assumed a complex field, which has two independent components ϕ_r and ϕ_i. For this case,

$$\langle\phi^*(\vec{r})\phi(\vec{r}')\rangle = \langle\phi_r(\vec{r})\phi_r(\vec{r}')\rangle + \langle\phi_i(\vec{r})\phi_i(\vec{r}')\rangle. \tag{6.18}$$

Thus, if one were to solve for correlations of a one-component (real) field, one would divide the expression derived here for complex fields by a factor of two. We work with complex fields only because the Fourier transforms are somewhat less painful.

Near the critical point the characteristic length of such correlations approaches infinity. To demonstrate this, we consider a purely quadratic functional, as the correlations can be found analytically. The effect of higher-order terms alters the behavior quantitatively (e.g., different values for the critical exponents), but does not alter the qualitative behavior shown here. For quadratic potentials,

$$\mathcal{V}(\phi) = \frac{A}{2}|\phi|^2, \tag{6.19}$$

the problem can be divided into individual modes:

$$F = \int d^3r\, f = \int d^3r\, \left[\frac{A}{2}|\phi|^2 + \frac{\kappa}{2}|\nabla\phi|^2\right] \tag{6.20}$$

$$= \frac{1}{2}\sum_{\vec{k}}\left(A + \kappa k^2\right)\left|\tilde{\phi}_{\vec{k}}\right|^2, \tag{6.21}$$

where the Fourier and inverse-Fourier transforms are defined by

$$\tilde{\phi}_{\vec{k}} \equiv \frac{1}{\sqrt{V}}\int d^3r\, e^{i\vec{k}\cdot\vec{r}}\phi(\vec{r}), \tag{6.22}$$

$$\phi(\vec{r}) = \frac{1}{\sqrt{V}}\sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{r}}\tilde{\phi}_{\vec{k}}.$$

Because the modes are independent, and therefore uncorrelated, the correlation in terms of momentum components is straight-forward,

$$\langle\tilde{\phi}_{\vec{k}}^*\tilde{\phi}_{\vec{k}'}\rangle = \delta_{\vec{k},\vec{k}'}\langle\left|\tilde{\phi}_{\vec{k}}\right|^2\rangle \tag{6.23}$$

$$= \delta_{\vec{k},\vec{k}'}\frac{2T}{(A + \kappa k^2)},$$

where the equipartition theory was used for the last step, i.e. if the free energy behaves as $\alpha|\phi|^2$, then $\alpha\langle|\phi|^2\rangle = T$. The reason it is not $T/2$ comes from the fact ϕ is complex and $|\phi|^2 = \phi_r^2 + \phi_i^2$. One can now calculate the correlation in coordinate space,

$$\langle\phi^*(\vec{r})\phi(\vec{r}')\rangle = \frac{1}{V}\sum_{\vec{k},\vec{k}'} e^{i\vec{k}\cdot\vec{r}-i\vec{k}'\cdot\vec{r}'}\langle\tilde{\phi}_{\vec{k}}^*\tilde{\phi}_{\vec{k}'}\rangle \tag{6.24}$$

$$= \frac{1}{V}\sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}\frac{2T}{(A+\kappa k^2)}.$$

Here, the last term employed the equipartition theorem, Eq (6.23). Now, the sum over modes can be transformed to an integral,

$$\langle\phi^*(\vec{r})\phi(\vec{r}')\rangle = \frac{1}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}\frac{2T}{(A+\kappa k^2)}. \tag{6.25}$$

This integral can be performed by first integrating over the angle between $\vec{k}$ and $(\vec{r}-\vec{r}')$, then performing a contour integration,

$$\langle\phi^*(\vec{r})\phi(\vec{r}'=0)\rangle = \frac{1}{2\pi^2 r}\int_{-\infty}^{\infty} kdk\,\sin(kr)\frac{T}{(A+\kappa k^2)} \tag{6.26}$$

$$= \frac{T}{2\pi\kappa r}e^{-r/\xi},\quad \xi = \sqrt{\kappa/A}.$$

Here, ξ is referred to as the correlation length. Near the critical point, the curvature of the free-energy w.r.t. the order parameter ϕ switches sign, and as A goes to zero, $\xi \to \infty$. The fact that this length becomes much larger than any driving microscopic length legitimizes the scaling arguments used in the discussion of critical phenomena in the next section.

In the case where ϕ refers to the magnetization density $m(\vec{r})$, the correlation function also provides the magnetic susceptibility,

$$\chi \equiv \frac{d}{dB}\langle m(\vec{r}=0)\rangle \tag{6.27}$$

$$= \frac{d}{dB}\frac{\text{Tr}\, e^{-\beta[H_0-\int d^3r\mu Bm(\vec{r})]}m(\vec{r}=0)}{\text{Tr}\, e^{-\beta[H_0-\int d^3r\mu Bm(\vec{r})]}}$$

$$= \frac{\mu\int d^3r\,\text{Tr}\, e^{-\beta H}m(\vec{r})m(\vec{r}=0)}{T}\frac{}{\text{Tr}e^{-\beta H}} - \frac{\mu}{T}\frac{\left[\text{Tr}e^{-\beta H}m(\vec{r}=0)\right]\left[\int d^3r\,\text{Tr}\, e^{-\beta H}m(\vec{r})\right]}{[\text{Tr}\, e^{-\beta H}]^2}$$

$$= \frac{\mu}{T}\int d^3r\,\{\langle m(\vec{r})m(\vec{r}=0)\rangle - \bar{m}^2\} \tag{6.28}$$

$$= \frac{\mu}{T}\int d^3r\,\{\langle(m(\vec{r})-\bar{m})(m(\vec{r}=0)-\bar{m})\rangle\}$$

$$= \frac{\mu}{\kappa}\int rdr\, e^{-r/\xi}$$

$$= \mu\frac{\xi^2}{\kappa},$$

where Eq. (6.26) was inserted into the expression for the correlation in the second-to-last step, and a factor of $1/2$ was added to account for the fact that this is not a complex field. Relations such as this are common in linear response theory. Linear response theory considers the change of an observable $\langle \delta a \rangle$ due to a small field that couples to a second observable b, $H_{\text{int}} = b\delta F$. This dependence can be expressed in terms of a generalized susceptibility, χ, where $\delta a = \chi \delta F$ defines χ. A central theme of linear response theory is that χ can be expressed in terms of correlations of the type $\langle a(0)b(r) \rangle$. More about linear response theory can be found in Chapter 8, or in the first two chapters of Forster[1]. Equation (6.28) will be important in the next sections as it will be used to connect the divergent behavior of the correlation function at the critical point to other quantities such as the susceptibilities and specific heat.

The third line of Eq. (6.27) is actually not completely correct.

$$\partial_\gamma \text{Tr} A e^{-\beta H_0 - \gamma B} \neq -\text{Tr} A B e^{-\beta H_0 - \gamma B},$$

when the operator B does not commute with H_0. This is because $e^{-\beta H_0 - \gamma B} \neq e^{-\beta H_0} e^{-\gamma B}$, unless the operators commute. However, in the classical limit operators commute and any non-commutation of operators typically brings in a power of $\hbar$. Thuys, Eq. (6.27) is correct in the classical limit. This is sometimes described as an expansion in $\hbar$.

6.4   Critical Exponents in Landau Theory

Here, we continue with the assumption that a system's field is fluctuating near the optimized (mininum of $\mathcal{V}(\phi)$), and that in this region one can consider only the quadratic dependence,

$$\mathcal{V}(\phi) \approx V_0 + \frac{A}{2}|\phi|^2. \tag{6.29}$$

The curvature, A, changes sign at the critical point. One can parameterize this as:

$$A(T) = at, \quad t \equiv (T - T_c)/T_c. \tag{6.30}$$

Thus, the correlation length diverges as $t^{-1/2}$ (see Eq. (6.26)) and the susceptibility (in the case of magnetization) diverges as t^{-1} (see Eq. (6.28)),

$$\xi = \sqrt{\frac{\kappa}{A}} = \sqrt{\kappa/a}\, t^{-1/2}, \tag{6.31}$$

$$\chi = \frac{\mu \xi^2}{\kappa} = \frac{\mu}{a} t^{-1}.$$

The exponents describing the divergences here are known as critical exponents. In Landau theory, all three-dimensional systems would have the same *critical exponents*,

$$\xi \sim |t|^{-\nu}, \quad \nu = 1/2, \tag{6.32}$$
$$\chi \sim |t|^{-\gamma}, \quad \gamma = 1.$$

---

[1] D. Forster, Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions

The specific heat is also divergent, but to see that we must first calculate the free energy. Equation (6.20) gives the free energy for a particular set of amplitudes $\phi_1, \phi_2 \cdots$. To calculate the free energy of the entire system, one must sum over all the possible values of ϕ_i,

$$\mathcal{F}_{\text{tot}} = -T \ln \sum_{\text{all configurations } \phi_1, \phi_2 \cdots} e^{-F(\phi_1, \phi_2 \cdots)/T}, \tag{6.33}$$

where F can be written as a sum over individual modes $\vec{k}$,

$$\mathcal{F}_{\text{tot}} = -T \sum_{\vec{k}} \ln \left(\sum_{\phi_k} e^{-F_k(\phi_k)/T} \right). \tag{6.34}$$

As was shown in the path integral discussions of Chapter 3, the sum over states ϕ_k can be written as an integral over the real and imaginary parts of ϕ_k, so that the partition function becomes

$$\begin{aligned}
e^{-F_k/T} &= \frac{1}{2\pi} \int d\phi_r d\phi_i \, e^{-(1/2)(A + \kappa k^2)|\phi_k|^2}, \\
&\sim (A + \kappa k^2)^{-1},
\end{aligned} \tag{6.35}$$

where the last expression ignores additive constants. Finally, the total free energy of the system will be

$$\mathcal{F}_{\text{tot}} = T \sum_{\vec{k}} \ln(A + \kappa k^2). \tag{6.36}$$

This expression applies for any dimensionality, but dimensionality will play a role once the sum over momentum modes is expressed as an integral.

After inserting $A = at$, one can solve for the contribution to the specific heat from the temperature dependence in A. Because $\ln Z = -\mathcal{F}/T$, one can first take a derivative of $\mathcal{F}/T$ w.r.t. t to get an expression for the energy density, then take a second derivative to find the specific heat. The results for three dimensions is

$$C_V/V = \frac{a^2}{(2\pi)^3} \int d^3k \, \frac{1}{(at + \kappa k^2)^2} \sim t^{-1/2}. \tag{6.37}$$

This last term represents another critical exponent,

$$C_V \sim t^{-\alpha}, \quad \alpha = 1/2. \tag{6.38}$$

Another critical exponent describes the rise of the mean magnetization for temperatures just below T_c. In our approximation, we consider only small fluctuations about the free energy minimum, and our choice of the mean magnetization comes from the mean-field approximation from Eq. (5.19) in Chapter 5,

$$\langle m \rangle \sim |t|^\beta, \quad \beta = 1/2. \tag{6.39}$$

Finally, for an external field with $T = T_c$, mean field theory gives (see the homework problem in this chapter)

$$\langle m \rangle \sim (\mu B)^{1/\delta}, \quad \delta = 3. \tag{6.40}$$

The divergence of the specific heat is related to the divergence of the integrand in Eq. (6.37) when t is set to zero. Note that for d dimensions, the integrand goes as $dk\, k^{(d-1)}/k^4$, which will be divergent for $d \leq 3$. Thus, in four dimensions or higher, the specific heat is well defined at $t \to 0$, whereas the behavior is more complicated for the two- and three-dimensional cases.

Summarizing the critical exponents in Landau theory,

$$
\begin{aligned}
\langle \phi(0)\phi(r) \rangle &\sim \xi^{2-d}(r/\xi)^{(1-d)/2}e^{-r/\xi}, \ \ \xi \sim t^{-\nu}, \ \ \nu = 1/2, &(6.41)\\
C_v &\sim t^{-\alpha}, \ \ \alpha = 2 - d/2,\\
<m>(t=0) &\sim H^{1/\delta}, \ \ \delta = 3,\\
<m>(H=0, t<0) &\sim t^{\beta}, \ \ \beta = 1/2,\\
\chi &\sim t^{-\gamma}, \ \ \gamma = 1,
\end{aligned}
$$

where d is the dimensionality. We will see in the next two sections that Landau theory is not valid near T_c for $d \leq 3$. Nonetheless, the features of Landau theory are qualitatively correct, and most importantly, the correct method for visualizing critical behavior requires Landau theory for perspective.

6.5   Validity of Landau Theory Near T_c: The Ginzburg Criteria

The exponents above were based on an assumption that one could ignore the quartic piece of the free energy density, which would only be justified if fluctuations of the order parameter were small compared to the mean value. To see whether that is the case, we consider the quartic case

$$
\mathcal{V} = \frac{A}{2}|\phi|^2 + \frac{B}{4}|\phi|^4, \tag{6.42}
$$

where $A = at$. For $t < 0$, A is negative and two distinct minima develop

$$
\langle|\phi|\rangle^2 = -A/B = -at/B. \tag{6.43}
$$

We will consider the case for $t > 0$, which means that the minimum is located at $\phi = 0$, but that as $t \to 0$, the fluctuations will become large. The correlation, see Eq. (6.26), has the form,

$$
\langle\phi(0)\phi(r)\rangle \sim \frac{T}{\kappa}\xi^{2-d}(r/\xi)^{(1-d)/2}e^{-r/\xi}. \tag{6.44}
$$

This expression was derived exactly for three dimensions previously, see Eq. (6.26). The integrals for the three-dimensional integration can be generalized for an arbitrary number of dimensions d in the limit that r is large, resulting in the expression above. If one uses $r = \xi$ as a characteristic length scale for comparison, the typical size for fluctuations are:

$$
\langle\delta\phi\delta\phi\rangle_{\text{typical}} \sim \frac{T}{\kappa}\xi^{2-d}, \tag{6.45}
$$

where dimensionless constants such as $(2\pi)^d$ are ignored. For mean-field theory to be valid, one

must justify ignoring the quartic piece in the free energy above, or

$$\frac{A}{2}\langle\delta\phi\delta\phi\rangle_{\text{typical}} \;\;>>\;\; \frac{B}{4}\langle\delta\phi\delta\phi\rangle^2_{\text{typical}}, \tag{6.46}$$

$$\frac{A}{2} \;\;>>\;\; \frac{B}{4}\langle\delta\phi\delta\phi\rangle_{\text{typical}},$$

$$at \;\;>>\;\; B\frac{T}{\kappa}\xi^{2-d}.$$

Because

$$\xi \sim \sqrt{\kappa/a}t^{-1/2}, \tag{6.47}$$

the criteria for validity becomes

$$t^{2-d/2} >> BT_c\kappa^{-d/2}a^{-2+d/2}. \tag{6.48}$$

As the critical point is approached $t \to 0$, the criteria will be satisfied for $d > 4$. However, Landau field theory is unjustified in the critical region for three dimensions or fewer.

Despite the fact that the theory is unjustified in the three-dimensional world, Landau theory is still enormously useful. First, the theory is still justified away from the critical point, depending on the parameters. Secondly, it is qualitatively correct in many aspects and represents a crucial perspective from which one can understand the more correct approaches based on scaling and renormalization, which are also field theories, but differ in that they consider the effects of the ϕ^4 term.

6.6   Critical Phenomena, Scaling and Exponents

Phase transitions can be classified as either first or second order. First-order transitions are characterized by co-existing phases with the order parameter being different in the two phases. In second-order transitions the derivative (or perhaps n^{th}-order derivative) of the order parameter is discontinuous. The term "order parameter" refers to any measure, not even necessarily experimentally accessible, used to describe a system in a given state. Examples are the magnetization density (ferromagnetic transition), number density (liquid-gas transition), quark-antiquark condensate (chiral transition), Higg's condensate (electro-weak transition), or field phase (superconductivity). Quantities that involve comparing systems with different conditions, such as the specific heat, which involves comparing the entropy at two different temperatures, or the susceptibility, which involves comparing the magnetization at two different external fields, are not order parameters.

Transitions can belong to any of several universality classes, a grouping determined by the symmetry associated with the phase transition. For instance, in the $X - Y$ model the system breaks two-dimensional rotational symmetry by choosing a specific direction for aligning spins. The three-dimensional analog, the Heisenberg model, has a higher degree of symmetry, hence belongs to a different universality class and has different critical exponents, which will be introduced further below.

At T_c Landau's field theory gave the following form for the correlation function for magnetization density in the previous section,

$$\Gamma(r) \equiv \langle m(r = 0)m(\vec{r})\rangle \sim \frac{1}{r}e^{-r/\xi}, \quad \xi = \sqrt{\kappa/A}, \tag{6.49}$$

where the free energy density was treated in the Gaussian approximation,

$$f \approx \frac{A}{2}m^2 + \frac{\kappa}{2}(\nabla m)^2.$$

(6.50)

Here, A is a function of temperature. At T_c, the sign of A will switch, and the correlation length diverges as

$$\xi = \sqrt{\frac{\kappa}{dA/dT}}(T - T_c)^{-1/2} \sim t^{-1/2}, \ \ t \equiv \frac{T - T_c}{T_c}.$$

(6.51)

In interacting theories, interactions change the behavior of ξ near T_c quantitatively, but not qualitatively. Quantum aspects of the interactions lead to screening effects that can be understood in renormalization calculations, but are beyond the scope of this class (See Chapter 18 in Huang).

Scaling arguments are based on the assumption that as $t \to 0$, the correlation functions lose all memory of scale and behave purely as a power.

$$\Gamma(r) \sim r^{2-d-\eta}e^{-r/\xi}, \ \ \xi \sim t^{-\nu}.$$

(6.52)

Here, d is the number of dimensions. As $t \to 0$ only the power-law behavior remains. Rather than assume a power behavior $\sim r^{-p}$, the unknown exponent p is replaced with $2-d-\eta$, so that in the Landau limit $\eta = 0$ for three dimensions. In the three-dimensional Landau calculation, $d = 3$, the two critical exponents are

$$\eta = 0, \ \ \nu = 1/2.$$

(6.53)

If $d \neq 3, \eta = (3 - d)/2$ in the Landau case. However, in the real world they vary from these values, depending on the universality class.

The other four critical exponents used in the last section $(\alpha, \beta, \gamma, \delta)$ are all determined by η and ν. The relations can best be understood from dimensional analysis. These arguments are based on two assumptions:

1. **The non-analytic part of the free energy density scales proportional to $\sim 1/\xi^d$.** From dimensional grounds, if ξ is the only relevant length scale, and because the free-energy density has dimensions of **Energy**$/L^d$, it must behave as ξ^{-d}, or equivalently as $t^{d\nu}$. This is based on a scaling argument that if one looks at a fluctuation of size $X = C\xi$ at temperature t, it will have the same free energy as a fluctuation of size $X' = C\xi'$ at temperature t'. Assuming ξ behaves as $t^{-\nu}$, the free energy scales as $t^{d\nu}$, and the specific heat, which one obtains by taking two derivatives w.r.t. t, behaves as $t^{-(2-d\nu)}$. From Eq. (6.41), one can then see that the critical exponent $\alpha = 2 - d\nu$.

2. **The magnetization and fields all behave as if they have effective dimensionalities, such that changing t, which causes a scale change of ξ, will scale all quantities according to their *effective dimensionality*.** Thus if $\langle m \rangle$ behaves as t^β, Γ must behave as $t^{2\beta}$ if evaluated at the same value of r/ξ, which given the form of the correlation function in Eq. (6.52), requires that

$$\begin{aligned} \Gamma(r/\xi = 1, t) &\sim t^{2\beta}, \\ \xi^{2-d-\eta} &\sim t^{2\beta}, \\ t^{\nu(d+\eta-2)} &\sim t^{2\beta}, \\ \beta &= \frac{1}{2}\nu(d + \eta - 2). \end{aligned}$$

(6.54)

One would have obtained the same result for any choice of r as long as the comparison (before and after the temperature change) was done at the same value of r/ξ. The field H must scale so that the free energy in a fluctuation of characteristic size, $mH\xi^d$, must be independent of t. This means that if $H \sim m^\delta$, and $m \sim t^\beta$, then

$$
\begin{aligned}
m^{1+\delta}\xi^d &\sim t^0, & \text{(6.55)}\\
t^{\beta(1+\delta)}t^{-d\nu} &\sim t^0,\\
\beta(1+\delta) - d\nu &= 0,\\
\delta &= \frac{d-\eta+2}{d+\eta-2}
\end{aligned}
$$

Finally, the susceptibility must have the same scaling as m/H, or

$$
\begin{aligned}
\chi &\sim m^{1-\delta} \sim t^{\beta(1-\delta)} & \text{(6.56)}\\
-\gamma &= \beta(1-\delta) = \nu(2-\eta).
\end{aligned}
$$

Summarizing these arguments, one can express critical exponents as follows:

$$
\begin{aligned}
\langle \phi(0)\phi(r)\rangle &\sim r^{2-d-\eta}e^{-r/\xi}, \ \ \xi \sim t^{-\nu}, & \text{(6.57)}\\
C_v &\sim t^{-\alpha}, \ \ \alpha = 2 - d\nu,\\
<m>(t=0) &\sim H^{1/\delta}, \ \ \delta = \frac{d-\eta+2}{d+\eta-2},\\
<m>(H=0, t<0) &\sim t^\beta, \ \ \beta = \nu(d+\eta-2)/2,\\
\chi &\sim t^{-\gamma}, \ \ \gamma = \nu(2-\eta),
\end{aligned}
$$

where only two of the six exponents are independent. Here, we have used the two exponents used to describe the behavior of the correlation function, η and ν, to define the remaining four. Because the correlation function $\Gamma(r)$ would nominally have dimensions of L^{2-d}, the exponent η is often referred to as the anomalous dimension. It is noteworthy that the scaling relations above are violated in Landau theory, as can be seen by comparing the mean-field value, $\beta = 1/2$, to the expression for β above. This stems from the fact that Landau theory is not based on scaling arguments, and as emphasized by the Ginzburg criteria, should not be applicable for systems of dimensions less than four near T_c.

The relations above were finally rigorously justified with renormalization theory. For this Ken Wilson was awarded the Nobel prize. Renormalization theory is beyond what we will cover in this class, but the interested student should read Chapter 18 of Huang.

6.7   Symmetry Breaking and Universality Classes

In Landau theory, the critical exponents were all determined by the dimensionality alone, and several were even independent of the dimensions, e.g. β in Eq. (6.41). However, in nature they are determined by the symmetry breaking involved in the problem. For instance, if there is a single real field ϕ whose free energy density is of the "mexican hat" form,

$$
f = \frac{A}{2}\phi^2 + \frac{B}{4}\phi^4, \tag{6.58}
$$

the system must choose whether to create a field with $\langle \phi \rangle > 0$ or < 0 when A becomes negative. For a complex field with a free energy density of the same form, $\phi^2 \rightarrow \phi^* \phi$, the system must also choose the complex phase. For a system where the spins are allowed to align in any direction, the system must choose a direction (θ, ϕ) on which to align. Each of these three examples represents a different symmetry breaking and leads to different exponents. The critical exponents are determined by the symmetry breaking and dimensionality, and problems of the same symmetry breaking and dimensionality have identical critical exponents and are referred to as being in the same *universality class*. The breaking of the symmetry at low temperature is called *spontaneous symmetry breaking*, as opposed to the breaking of the symmetry by adding an external field, which is referred to as *explicit symmetry breaking*.

For the broken reflection symmetry above, the system chooses between two discrete minima, whereas for the example of the complex field or for the spin alignment, the set of possible minima form an infinite continuous set. These symmetries are called *continuous symmetries* and have the property that, after being broken, one can make incremental changes in the field while maintaining a minimum in the free energy. In contrast, the real-field case only has a reflection symmetry, which is referred to as a *discrete symmetry*. For instance, consider the case of the complex field. Writing the free energy density in terms of the real and imaginary parts, ϕ_x and ϕ_y,

$$f = \frac{A}{2}(\phi_x^2 + \phi_y^2) + \frac{B}{4}\left(\phi_x^2 + \phi_y^2\right)^2 + \frac{\kappa}{2}(\nabla\phi)^2, \tag{6.59}$$

where we consider the case where $\alpha < 0$ and minima appear for

$$(\phi_x^2 + \phi_y^2) = \phi_0^2 = -\frac{A}{B}. \tag{6.60}$$

If one expands around the minima in the x direction, one can rewrite the fields as

$$\delta\phi_x = \phi_x - \phi_0, \quad \delta\phi_y = \phi_y. \tag{6.61}$$

Rewriting the free energy in terms of $\delta\phi_x$ and $\delta\phi_y$,

$$\begin{aligned}
f &\approx f_{\text{min}} + \frac{1}{2}\delta\phi_x^2\frac{\partial^2 f}{\partial^2\phi_x} + \frac{1}{2}\delta\phi_y^2\frac{\partial^2 f}{\partial^2\phi_y} + \frac{\kappa}{2}(\nabla\phi)^2 \\
&= f_{\text{min}} + \frac{1}{2}A_x\delta\phi_x^2 + \frac{1}{2}A_y\delta\phi_y^2 + \frac{\kappa}{2}(\nabla\delta\phi_x)^2 + \frac{\kappa}{2}(\nabla\delta\phi_y)^2, \\
A_x &= -2A, \quad A_y = 0.
\end{aligned} \tag{6.62}$$

Thus, the y mode contributes to the free energy only through the term proportional to $(\nabla\delta\phi_y)^2$, or in momentum space, the modes contribute to the free energy as k^2. For relativistic field theories, modes for which the contribution vanishes as $k \rightarrow 0$ are called massless modes as they have no energy as $k \rightarrow 0$. These modes are also referred to as Goldstone bosons (Goldstone derived the general idea that massless modes should be associated with breaking a continuous symmetry) or Nambu-Goldstone bosons. Although the most famous example of such symmetry breaking is the electro-weak transition in particle physics, Higgs, Goldstone and Nambu were also concerned with the nuclear physics problem of chiral symmetry breaking, and of course Landau was most concerned with problems in condensed matter physics, the most famous of

which is Ginzburg-Landau theory for superconductivity, which employs precisely this model of complex scalar fields.

In the linear sigma model for nuclear physics, one considers a four-component field, $\vec{\phi}$, where ϕ_0 refers to the σ field, which in QCD is related to the quark-antiquark condensate, and $\phi_{1,2,3}$ are the three pionic fields. The Hamiltonian contains terms of the form,

$$V(\phi) = -\frac{m_0^2}{2}\left(\phi_0^2 + \phi_1^2 + \phi_2^2 + \phi_3^2\right) + \frac{\lambda}{4}\left(\phi_0^2 + \phi_1^2 + \phi_2^2 + \phi_3^2\right)^2. \tag{6.63}$$

Because the first term is negative, spontaneous symmetry breaking ensues. Defining the direction of the condensate as the "zero" direction, the condensate generated is

$$\langle\phi_0\rangle = \sqrt{\frac{m_0}{\lambda}}. \tag{6.64}$$

This would result in one massive mode, from the curvature of the potential in the "zero" direction and three massless modes, because at the minimum,

$$\left.\frac{d^2V}{d\phi_i^2}\right|_{\phi_0=\sqrt{m_0/\lambda}} = 0, \quad i = 1, 2, 3. \tag{6.65}$$

This would predict three Goldstone bosons, which would be massless pions. In nature, pions are indeed very light, but not massless. This is accounted for by adding a small term proportional to ϕ_0 in V, which explicitly breaks the symmetry, and makes the original symmetry approximate. The symmetries of the linear sigma model derive from underlying chiral symmetries in QCD, which is again only an approximate symmetry. The symmetry is broken by the small current quark masses, $\approx 5\,\mathrm{MeV}/c^2$.

In electroweak physics, i.e., the Weinberg-Salam model, the situation is more complicated due to the gauge nature of the theory. Again, there are four scalar fields, or two complex scalar fields, with similar quadratic and quartic couplings. The symmetry breaking is the same, but the three accompanying scalar fields do not become massless due to their gauge coupling with other particles. Remarkably, the three fields play the part of the longitudinal component of the $W^\pm$ and Z massive fields. Whereas the W and Z are massless in the bare Lagrangian, and have only two polarizations, they pick up mass through the symmetry breaking, which requires three polarizations. That extra degree of freedom is provided by the three scalar fields. The lack of Goldstone bosons is related to the fact that in a gauge theory, the local phases of fields is non-observable by construction. Before symmetry breaking, one cannot think of the field as a magnitude and phase, because phase is undefined at $|\phi| = 0$. The two complex fields, thus contribute four degrees of freedom to the specific heat. Once symmetry is broken, there is a magnitude and three non-observable phases. But by introducing $\phi \to \phi_0 + \delta\phi$, the condensate ϕ_0 couples to the $W^\pm$ and Z bosons to provide masses. Whereas massless vector bosons have two polarizations, massive vector bosons have three. This requires one extra degree of freedom for each of the $W^\pm$ and Z bosons. These compensate for the loss of three degrees of freedom from the scalar fields. This transference of degrees of freedom from the scalar to the vector modes during the symmetry breaking, and the associated lack of Goldstone bosons, is known as the Higgs mechanism.

It is no exaggeration to state that the most important problems addressed in the last 40 years in nuclear, particle, and condensed-matter physics concern spontaneously broken symmetries. Furthermore, one of the most exciting astronomical measurements, that of fluctuations of the 3-degree background radiation is related to inflation, which might be associated with the breaking of the electro-weak symmetries or perhaps grand unified theories, under which quarks and leptons appear in a single fermionic multiplet and the gauge bosons all correspond to generators of a single group.

6.8   Problems

1. Consider the example for which the surface energy was calculated, where

$$\Delta\Psi \equiv P_0 - P + (\mu - \mu_0)\rho = \frac{A}{2}\left[(\rho - \rho_c)^2 - \alpha^2\right]^2.$$

   Using the second line of Eq. (6.16), solve for the density profile $\rho(x)$ between the two phases.

2. Consider the one-dimensional Ising model, with the total energy in the mean field approximation being,

$$E = -\sum_i \frac{1}{2}qJ\langle\sigma\rangle\sigma_i.$$

   Note the factor of 1/2 being added relative to the effective energy for one mode to account for double counting.

   (a) Using the definition of entropy, $-\sum_i p_i \ln(p_i)$, show that the entropy per spin is:

$$S/N = -\frac{1 + \langle\sigma\rangle}{2}\ln\left(\frac{1 + \langle\sigma\rangle}{2}\right) - \frac{1 - \langle\sigma\rangle}{2}\ln\left(\frac{1 - \langle\sigma\rangle}{2}\right).$$

   (b) The free-energy, $F = E - TS$, per spin is then

$$F/N = -\frac{1}{2}qJ\langle\sigma\rangle^2 + T\frac{(1 + \langle\sigma\rangle)}{2}\ln\left(\frac{1 + \langle\sigma\rangle}{2}\right) + T\frac{(1 - \langle\sigma\rangle)}{2}\ln\left(\frac{1 - \langle\sigma\rangle}{2}\right).$$

   Show that minimizing the free energy w.r.t. $\langle\sigma\rangle$ gives:

$$2\beta qJ\langle\sigma\rangle = \ln\left(\frac{(1 + \langle\sigma\rangle)}{(1 - \langle\sigma\rangle)}\right).$$

   (c) Compare the expression above to that from the previous expression,

$$\langle\sigma\rangle = \tanh(\beta qJ\langle\sigma\rangle).$$

   (d) If the density of spin sites per unit volume is ρ_0, the free energy density is

$$f(\sigma, T) = \rho_0\left\{\mathcal{V}(\sigma, T) + \frac{\kappa}{2}(\nabla\sigma)^2\right\},$$

$$\mathcal{V}(\sigma, T) = -\frac{1}{2}qJ\sigma^2 + T\frac{(1 + \sigma)}{2}\ln\left(\frac{1 + \sigma}{2}\right) + T\frac{(1 - \sigma)}{2}\ln\left(\frac{1 - \sigma}{2}\right).$$

Derive an expression for the surface free energy that is an integral over σ with limits from $-\sigma_{\text{eq}}$ and $+\sigma_{\text{eq}}$. Write the integrand in terms of ρ_0, κ, T and $\mathcal{R}(\sigma, T) \equiv \mathcal{V}(\sigma, T) - \mathcal{V}(\sigma_{\text{eq}}, T)$, where σ_{eq} is the solution to the transcendental expression for the equilibrated value for $\langle \sigma \rangle$.

(e) Find the surface energy in the limit that $T \to 0$ in terms of κ, ρ_0 and $T_c = qJ$.

3. Repeat the calculation for the correlation function from Sec. 6.3 for the one dimensional case. In this case, the Fourier transforms are defined as:

$$\tilde{\phi}_k \equiv \frac{1}{\sqrt{L}} \int dx\, e^{ikx} \phi(x),$$

$$\phi(x) = \frac{1}{\sqrt{L}} \sum_k e^{-ikx} \phi_k.$$

4. Beginning with the definition of the average density,

$$\langle \rho \rangle = \frac{\text{Tr } e^{-\beta H + \beta \mu \int d^3 r \rho(\vec{r})} \rho(\vec{r} = 0)}{\text{Tr } e^{-\beta H + \beta \mu \int d^3 r \rho(\vec{r})}},$$

prove the following relation between the density-density correlation function and $d\rho/d\mu$,

$$T \frac{d\langle \rho \rangle}{d\mu} = \int d^3 r \, \left\{ \langle \rho(\vec{r}) \rho(\vec{r} = 0) \rangle - <\rho>^2 \right\}.$$

5. Returning to the expression for the mean spin in the Ising model in the mean field approximation,

$$\langle \sigma \rangle = \tanh(\beta qJ \langle \sigma \rangle + \beta \mu B),$$

solve for $\langle \sigma \rangle$ in the limit that μB is small and $T = T_c$. Note this provides the exponent δ in mean field theory. Hint: Expand $\tanh$ and keep only the lowest order non-zero terms for both $\langle \sigma \rangle$ and for μB.

6. Assume the following form for the correlation function,

$$\langle m(0) m(r) \rangle - \langle m \rangle^2 = C r^{2-d-\eta} e^{-r/\xi}, \ \ \xi \sim t^{-\nu}$$

near T_c.

(a) Find the susceptibility using Eq. (8.38) in terms of C, ξ, η and the dimensionality d. When expressing your answer, you can use the shorthand Ω_d for $\int d\Omega$ in d dimensions, i.e.,

$$\int d^d r \dots = \int d\Omega \, r^{d-1} dr \dots$$

For your entertainment, $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$.

(b) If χ diverges as $t^{-\gamma}$ near T_c, find γ in terms of η and ν.

7. Assume that the free energy obeys the following form,

$$F = \int d^3r \left\{ \frac{A}{2}\phi^2 + \frac{C}{2}\phi^6 \right\}.$$

Assuming that near T_c, $A \sim at$, find the critical exponent in mean field theory β where,

$$\langle \phi \rangle \sim t^\beta$$

below T_c.

7   Advanced Methods for Treating Interactions

"Our strength grows out of our weaknesses" – Ralph Waldo Emerson

7.1   Perturbation Theory

The pressure $P(\mu, T)$, density $\rho(\mu, T)$, and energy density $E/V(\mu, T)$, are all affected by interactions. Because all such quantities are derived from the partition function, we will concentrate on calculating the alteration of the partition function. The relations

$$E = -(\partial/\partial\beta)\ln Z_{GC}, \quad \rho = (T/V)(\partial/\partial\mu)\ln Z_{GC}, \quad P/T = (\partial/\partial V)\ln Z_{GC}, \qquad (7.1)$$

hold for any system, interacting or not.

The modern way to include interactions is through perturbation theory. An alternative method is based on "Mayer cluster diagrams", but is applicable only for classical systems with interactions of the type,

$$H_{\text{int}} = \int d^3r_a d^3r_b V(\mathbf{r}_a - \mathbf{r}_b)\rho(\mathbf{r}_a)\rho(\mathbf{r}_b). \qquad (7.2)$$

The cluster technique is well described in Pathria (<u>Statistical Mechanics</u>, R.K. Pathria, Pergammon, or in Huang, Sec. 10.1). More general interactions might involve exchanging off-shell photons or gluons. Because the standard methods of quantum perturbation theory apply to all cases, and naturally incorporate relativity and quantum mechanics, we will focus on these methods. Our discussion will stay brief. Perturbation theory at finite temperature can easily be a subject for an entire course.

Perturbation theory at finite temperature concerns calculation of Z_{GC},

$$Z_{GC} = \sum_i \langle i|e^{-\beta(H-\mu Q)}|i\rangle, \qquad (7.3)$$

where i refers to specific quantum states. Even if the system has 10^{23} particles, and is interacting, eigen-states for the bulk system still exist. We will separate $H - \mu Q$ into two parts, H_0 and V, where H_0 is usually considered to be the part for which the state is expressed simply by listing which single-particle states are occupied, i.e., the many-body wave function is a product of single-particle wave functions. We note that in our notation H_0 includes the contribution from $-\mu Q$. In the language of creation and destruction operators, H_0 is usually the part of the field which can be written as $\sum_i(\epsilon_i - \mu)a_i^\dagger a_i$, while the interaction V represents all other parts of the Hamiltonian where the creation and destruction operators appear in higher powers. The eigenstates are often momentum eigen-states, though they might also be eigen-states of a classical field. In the nuclear shell model, one uses a basis of harmonic-oscillator eigen-states, while in some other variants of nuclear many-body theory one might use eigen-states of a potential $v(r)$, chosen to be consistent with the density. In nuclear physics the interaction V is also often referred to as the "residual" interaction, referring to the fact that it is what is left over after having absorbed as much as possible into the effective potential $v(r)$.

After identifying V, one moves to the interaction representation, with the definitions,

$$
\begin{aligned}
U_{\text{int}}(\beta) &\equiv e^{\beta H_0} e^{-\beta H}, \\
V(\beta) &\equiv e^{\beta H_0} V e^{-\beta H_0},
\end{aligned}
\tag{7.4}
$$

which leads to the relation for the evolution operator U,

$$
\begin{aligned}
\frac{\partial}{\partial \beta} U_{\text{int}}(\beta) &= e^{\beta H_0} H_0 e^{-\beta H} - e^{\beta H_0}(H_0 + V)e^{-\beta H} = -e^{\beta H_0} V e^{-\beta H} \tag{7.5} \\
&= -V(\beta) U_{\text{int}}(\beta).
\end{aligned}
$$

The goal is to calculate the partition function,

$$
Z_{GC} = \text{Tr } e^{-\beta H_0} U_{\text{int}}(\beta).
\tag{7.6}
$$

This is the standard technique for quantum perturbation theory in the interaction representation, only in those cases one is calculating the evolution operator e^{-iHt}. The derivations are equivalent, except that $it/\hbar$ is replaced with β. Thus, the Boltzmann factor is identified with the evolution operator in imaginary time. Furthermore, statistical calculations consider the trace rather than the expectation between specific initial and final states.

The differential eqaution for $U(\beta)$ in Eq. (7.5) can be equivalently stated as an integral equation,

$$
U_{\text{int}}(\beta) = 1 - \int_0^\beta d\beta' V(\beta') U_{\text{int}}(\beta').
\tag{7.7}
$$

One can now replace U in the r.h.s. with the complete expression for U to see the iterative expression,

$$
\begin{aligned}
U_{\text{int}}(\beta) = \quad & 1 - \int_0^\beta d\beta' \, V(\beta') + \int_0^\beta d\beta' \int_0^{\beta'} d\beta'' \, V(\beta')V(\beta'') \tag{7.8} \\
& - \int_0^\beta d\beta' \int_0^{\beta'} d\beta'' \int_0^{\beta''} d\beta''' \, V(\beta')V(\beta'')V(\beta''') + \cdots.
\end{aligned}
$$

Here, $U_{\text{int}}(\beta)$ and $V(\beta)$ are matrices, with the indices referring to specific states. To overcome complications associated with the limits for each variable,

$$
\begin{aligned}
U_{\text{int}}(\beta) = \quad & 1 - \mathcal{T} \int_0^\beta d\beta' \, V(\beta') + \frac{1}{2!} \mathcal{T} \int_0^\beta d\beta' \int_0^\beta d\beta'' \, V(\beta')V(\beta'') \tag{7.9} \\
& - \frac{1}{3!} \mathcal{T} \int_0^\beta d\beta' \int_0^\beta d\beta'' \int_0^\beta d\beta''' \, V(\beta')V(\beta'')V(\beta''') + \cdots,
\end{aligned}
$$

where the limits of the various integrations are now independent from one another, with the fact that the factors $V(\beta)$ at lower values of β are always pushed to the right by the "time-ordering" operator $\mathcal{T}$. Because there are $n!$ ways to order such operators, a factor $1/n!$ is added to correct for the over-counting. If $V(\beta)$ were to commute with $V(\beta')$ (which would be true if $[V, H_0] = 0$), the time-ordering operator would be irrelevant, and U_{int} could be written as $\exp\{-\int d\beta' V(\beta')\}$. This would be the case if V were diagonal in the same basis that diagonalized H_0. This is the case for uniform fields, which might change the energy of a particle of

momentum p, but does not change the fact that p is an eigenstate. The more general case of scattering does not satisfy the commutation condition as particles with momentum p_1 and p_2 are scattered into p'_1 and p'_2 by V. Some classes of interactions represent exchanges of charge or can even change the number of quanta. For instance, in the three-point interaction in electro-weak theory, $V_{ii'}$, i might refer to an electron and i' might refer to a neutrino and W^- boson.

The states i and i' in the matrices $V_{ii'}$ used in perturbation theory are many-body states, with i labeling the configuration of perhaps 10^{23} particles. However, the interaction V rarely involves all the quanta. Instead, it might involve changing the state of two particles, say those with momentum p_1 and p_2 to p'_1 and p'_2. For this two-particle-to-two-particle case, we consider two particles interacting non-relativistically through a potential $U(r_1 - r_2)$. This interaction is diagonal in the r_1, r_2 basis (but clearly not diagonal in the p_1, p_2 basis) because the particles do not move instantaneously and

$$\langle r_1, r_2 | V | r'_1, r'_1 \rangle = U(r_1 - r_2)\delta^3(r_1 - r'_1)\delta^3(r_2 - r'_2). \tag{7.10}$$

The same interaction sandwiched between momentum states describes a change in the relative momentum,

$$q \equiv p_1 - p_2 \rightarrow q' \equiv p'_1 - p'_2, \tag{7.11}$$

with the total momentum not changing. This is expected given the translation invariance of the interaction. Performing the machinations,

$$
\begin{aligned}
\langle p_1, p_2 | V | p'_1, p'_2 \rangle &= \int d^3 r_1 d^3 r_2 d^3 r'_1 d^3 r'_2 \, \langle p_1, p_2 | r_1, r_2 \rangle \langle r_1, r_2 | V | r'_1, r'_2 \rangle \langle r'_1, r'_2 | p'_1, p'_2 \rangle \qquad (7.12)\\
&= \frac{1}{V^2} \int d^3 r_1 d^3 r_2 \, e^{i p_1 \cdot r_1 / \hbar + i p_2 \cdot r_2 / \hbar} U(r_1 - r_2) e^{-i p'_1 \cdot r_1 / \hbar - i p'_2 \cdot r_2 / \hbar}\\
&= \frac{1}{V^2} \int d^3 R \, e^{i(P - P') \cdot R / \hbar} \, d^3 r \, e^{i(q - q') \cdot r / \hbar} U(r)\\
&= \begin{cases} = \frac{(2\pi\hbar)^3}{V^2} \delta^3(P - P') \tilde{U}(k \equiv (q - q')/\hbar), & \text{if } P = P' \text{ not yet constrained.}\\ = \frac{1}{V} \tilde{U}(k \equiv (q - q')/\hbar), & \text{if } P = P' \text{ already constrained.} \end{cases}\\
r &= r_1 - r_2, \quad P = p_1 + p_2, \quad R = \frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}, \quad q = \frac{m_2 p_1 - m_1 p_2}{m_1 + m_2}.
\end{aligned}
$$

The inverse powers of the volume, which originated from having normalized wave functions $e^{ikx}/\sqrt{V}$, will disappear when summing over the density of states, and are often omitted in many text books. Thus, the matrix element between two-particle states typically includes a δ function representing the conservation of momentum and the Fourier transform of the potential $\tilde{U}(k)$, where k is the difference in the relative momenta (divided by $\hbar$) between the incoming and outgoing states.

For many-body theory, it is more physical to think of the interaction in terms of creation and destruction operators. In that case,

$$V = \frac{1}{V} \sum_{p'_1, p'_2, p_1, p_2} \langle p_1, p_2 | V | p'_1, p'_2 \rangle a^\dagger_{p_1} a^\dagger_{p_2} a_{p'_1} a_{p'_2}, \tag{7.13}$$

where $a_p^\dagger / a_p$ are creation/destruction operators that create/remove a particle with momentum p. In the interaction representation, $a_p^\dagger(\beta) = e^{-\beta \epsilon_p} a_p^\dagger$, and

$$V(\beta) = \sum_{p_1', p_2', p_1, p_2} \langle \mathbf{p}_1, \mathbf{p}_2 | V | \mathbf{p}_1', \mathbf{p}_2' \rangle a_{p_1'}^\dagger a_{p_2'}^\dagger a_{p_1} a_{p_2} e^{-\beta(\epsilon_1' + \epsilon_2' - \epsilon_1 - \epsilon_2)}. \tag{7.14}$$

For such a two-particle-to-two-particle matrix element, $V(\beta)$ would destroy particles with momentum p_1 and p_2, then create two new particles with p_1' and p_2'. The initial many-particle state i might undergo several combinations of V at the various "times" before being returned to its original configuration, as demanded by the trace. For some applications, $V(\beta)$ might refer to the annihilation of one particle and the creation of two particles, such as in the decay of a W boson into an electron and a neutrino. Similarly, the merging of the electron and neutrino into a W boson would be represented by a term with two annihilation operators and one creation operator.

Our immediate goal is to show how perturbation theory, whose formalism refers to many-body states, can be rewritten in terms of matrix elements involving only a few particles. We will see that when one calculates the log of the partition function, that indeed, the corrections due to interactions involve the consideration of matrix elements such as those in Eq. (7.12) that involve only a few particles at a time.

To accomplish the reduction mentioned above, we will first show how the nth-order contribution in the perturbative expansion of the grand canonical partition function Z (The GC subscript will be suppressed) can be factorized into a "connected" piece and a lower-order term for the expansion of Z. The nth order term is:

$$Z_n = \frac{(-1)^n}{n!} \int_0^\beta d\beta_1 \cdots d\beta_n \sum_i \mathcal{T} \langle i | e^{-\beta H_0} V(\beta_1) \cdots V(\beta_n) | i \rangle \tag{7.15}$$

We now consider the specific term $V(\beta_1)$ as illustrated a specific path illustrated in Fig. 7.1. Each vertex represents a two-body-to-two-body matrix element, with the two incoming lines (from the left) representing the annihilated momenta, and the two outgoing lines (to the right) representing the new momenta.

The term $V(\beta_1)$ will be "connected" to those other potential terms $V(\beta_i)$ for which one can form an independent chain of initial momenta which are transformed then return to the same independent momenta. In the figure the number of such terms is $m = 5$. By careful inspection one can see that the remainder of the matrix element is independent and factorizes away from those terms connected to $V(\beta_1)$. One can re-express Z_n as a sum over all m,

$$
\begin{aligned}
Z_n &= \frac{(-1)^n}{n!} \sum_{m \leq n} \frac{(n-1)!}{(m-1)!(n-m)!} \left[\int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}} \right] \\
&\qquad \cdot \left[\int d\beta_1' \cdots d\beta_{n-m}' \, \mathcal{T} \langle\langle V(\beta_1') \cdots V(\beta_{n-m}') \rangle\rangle \right] \tag{7.16} \\
&= \frac{1}{n} \sum_{m \leq n} (-1)^m Z_{n-m} \frac{1}{(m-1)!} \int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}}.
\end{aligned}
$$

Figure 7.1: This illustrates how a term $\mathcal{T}\langle V(\beta_1) \cdots V(\beta_n) \rangle$ might look for $n = 11$. The term $V(\beta_1)$ is connected in an $m = 5$ subdiagram which can be factored out of the matrix element. The number of ways that the $V(\beta_1)$ term can be connected to a group of m out of n is $(n-1)!/[(n-m)!(m-1)!]$.

Here, the double brackets denote the thermal trace, $\langle\langle A \rangle\rangle = \sum_i \langle i | e^{-\beta H_0} A | i \rangle$. The preceeding $1/n$ factor prevents one from immediately factorizing Z into Z multiplied by a piece depending only on connected diagrams. To overcome this limitation, we replace V with λV, where $\lambda = 1$. Then, because Z_n will depend on λ as λ^n, we can write

$$\lambda \frac{d}{d\lambda} Z = \sum_n \lambda \frac{d}{d\lambda} Z_n \tag{7.17}$$

$$= \sum_m \sum_{n>m} Z_{n-m} \frac{(-1)^m}{(m-1)!} \int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}}$$

$$= Z \cdot \sum_m \frac{(-1)^m}{(m-1)!} \int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}} \cdot$$

After dividing both sides by Z one finds an expression for $\ln Z$,

$$\frac{d}{d\lambda} \ln Z = \frac{1}{\lambda} \sum_m \frac{(-1)^m}{(m-1)!} \int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}} \cdot \tag{7.18}$$

The final expression then comes from noting the the mth-order connected piece behaves as λ^m. Thus, one can integrate both sides of the equation from $\lambda = 0$ to $\lambda = 1$ to obtain our final result:

$$\ln Z = \ln Z_{(V=0)} + \sum_m \frac{(-1)^m}{m!} \int d\beta_1 \cdots d\beta_m \, \mathcal{T} \langle\langle V(\beta_1) \cdots V(\beta_m) \rangle\rangle_{\text{connected}} \cdot \tag{7.19}$$

This profound expression permits one to calculate corrections to the partition function for 10^{23} particles in terms of calculable terms, each of which represents only a few particles.

Example 7.1:

Consider non-relativistic non-degenerate particles of mass m interacting through a potential

$$U(\mathbf{r}) = U_0 e^{-r/a}.$$

Find the correction to the pressure to first-order in perturbation theory for fixed chemical potential. Give answer in terms of $\alpha = (2j+1)e^{\mu/T}$.

First, calculate the matrix element from Eq. (7.12). Because the first order expression has the same momenta in the bra and ket, we are interested in the matrix element $\langle \mathbf{p}_1, \mathbf{p}_2|V|\mathbf{p}_1, \mathbf{p}_2 \rangle$, and the momentum transfer $\mathbf{k}$ is zero,

$$
\begin{aligned}
\langle \mathbf{p}_1, \mathbf{p}_2|V|\mathbf{p}_1, \mathbf{p}_2 \rangle &= \frac{1}{V}\tilde{U}(\mathbf{k}=0), \\
&= \frac{1}{V}\int d^3r \, U(\mathbf{r}).
\end{aligned}
$$

Now, one can calculate the matrix element necessary for perturbation theory,

$$
\begin{aligned}
\langle \mathbf{p}_1, \mathbf{p}_2|e^{-\beta H_0}V(\beta_1)|\mathbf{p}_1, \mathbf{p}_2 \rangle &= \langle \mathbf{p}_1, \mathbf{p}_2|e^{-\beta H_0}e^{-\beta_1 H_0}V e^{\beta_1 H_0}|\mathbf{p}_1, \mathbf{p}_2 \rangle \qquad (7.20) \\
&= e^{-\beta(P^2/2M + q^2/2m_{\mathrm{red}}) + 2\beta\mu}\frac{1}{V}\int d^3r \, U(\mathbf{r}).
\end{aligned}
$$

The exponentials with β_1 canceled because the state was the same on the left and right of V. This cancellation only happens in first-order.

Approaching the home stretch, one now uses Eq. (7.19) to work toward the answer,

$$
\begin{aligned}
\ln Z &= \ln Z_0 - \frac{1}{2}\sum_{P,q}\left(\int_0^\beta d\beta_1 e^{-\beta(P^2/2M + q^2/2m_{\mathrm{red}}) + 2\beta\mu}\frac{1}{V}\right)\int d^3r \, U(\mathbf{r}) \\
&= \ln Z_0 - e^{2\beta\mu}\frac{\beta}{2}\frac{V(2j+1)^2}{(2\pi\hbar)^6}\int d^3P \int d^3q \, e^{-\beta(P^2/2M + q^2/2m_{\mathrm{red}})}\int d^3r \, U(\mathbf{r}) \\
&= \ln Z_0 - \frac{\beta}{2}V\alpha^2\frac{(MT)^{3/2}(m_{\mathrm{red}}T)^{3/2}}{(2\pi)^3\hbar^6}\int d^3r \, U(\mathbf{r}) \\
&= \ln Z_0 - \frac{1}{2}\frac{\alpha^2 V T^2 m^3}{(2\pi)^3\hbar^6}\int d^3r \, U(\mathbf{r}) \\
&= \ln Z_0 - 4\pi U_0 a^3\frac{\alpha^2 V T^2 m^3}{(2\pi)^3\hbar^6}.
\end{aligned}
$$

Here, $M = 2m$ is the total mass of the pair, and $m_{\mathrm{red}} = m/2$ is the reduced mass, and α was defined as $(2j+1)e^{\mu/T}$. The preceeding factors of $1/2$ correct for double-counting. It is good practice to check the dimensions after a calculations such as this. In this case, all the units cancel as $\ln Z$ is dimensionless.

Although this example may have seen somewhat lengthy, it is the simplest example of a calculation that one can do in perturbation theory at finite T. When performing higher-order calculations, or when including Bose and Fermi effects, calculations are inherently more difficult.

Such calculations will not be considered here, and are more appropriate for a course dedicated to many-body theory. However, even if one does not intend to ever perform such calculations, it is very worthwhile to have an understanding of the fundamentals of perturbation theory shown here:

- Perturbation theory is an expansion in the order of interactions.

- Calculations involve matrix elements of the type $\langle p_1, p_2 | V | p_1, p_2 \rangle$ (or perhaps with different numbers of particles if interactions create or destroy quanta), which can be calculated easily.

- Diagrams, such as the one shown in Fig. 7.1 allow one to separate contributions to $\ln Z$ by whether they involve the interaction of 1 particle, 2 particles $\cdots n$ particles in a "connected" manner.

7.2   An Aside: A Brief Review of Creation and Destruction Operators

When phase space densities are low, the only matrix elements that matter are those connecting an empty single-particle level to one with a singly-occupied level. For higher occupations, one must worry about the more general case, $\langle m | V | n \rangle$, where m and n refer to the number of particles in the level. Although the m and n dependence could be nearly any function, in nearly all cases it will follow the systematics of creation and destruction operators. The systematics plays a central role in the study of condensates and lasers, or in lattice-gauge calculations.

Creation and destruction operators are built on the physics of a simple one-dimensional harmonic oscillator:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{7.21}$$

Here, the spring constant is $k = m\omega^2$, where $\omega = \sqrt{k/m}$ is the angular frequency of the oscillator. Creation and destruction operators are defined as:

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}x + i\sqrt{\frac{1}{2\hbar m\omega}}p, \quad a = \sqrt{\frac{m\omega}{2\hbar}}x - i\sqrt{\frac{1}{2\hbar m\omega}}p. \tag{7.22}$$

With this definition,

$$[a, a^\dagger] = 1, \quad H = \hbar\omega(a^\dagger a + 1/2). \tag{7.23}$$

Here, a is the destruction operator while $a^\dagger$ is the creation operator.

Now, we consider some eigenstate $|m\rangle$ which has an energy ϵ_m,

$$H|m\rangle = \epsilon_m|m\rangle. \tag{7.24}$$

The state $a|m\rangle$ must then also be an eigenstate,

$$H(a|m\rangle) = \hbar\omega\left[(a^\dagger a + 1/2)\right]a|m\rangle = \hbar\omega a\left[a^\dagger a - 1/2\right]|m\rangle = (\epsilon_m - \hbar\omega)(a|m\rangle). \tag{7.25}$$

Thus, for any eigenstate, successive operations of the destruction operator will yield more eigenstates with energies lowered by $\hbar\omega$ – unless there is some state $|0\rangle$ for which

$$a|0\rangle = 0. \tag{7.26}$$

If we assume there is such a state, its energy must be

$$H|0\rangle = \hbar\omega(a^\dagger a + 1/2)|0\rangle = (\hbar\omega/2)|0\rangle. \tag{7.27}$$

One could have also considered the state $a^\dagger|m\rangle$. Similarly, application of the commutation relations leads to fact that

$$H(a^\dagger|m\rangle) = (\epsilon_m + \hbar\omega)(a^\dagger|m\rangle). \tag{7.28}$$

Thus, the eigen-states of the Hamiltonian are:

$$\epsilon = (1/2, 3/2, 5/2\cdots)\hbar\omega. \tag{7.29}$$

Thus, we have shown that the basis of eigenstates are the vacuum, plus some states $(a^\dagger)^m|0\rangle$. However, these states are not yet normalized. To normalize them, we calculate the norm by applying the commutation operators

$$\langle 0|a^m(a^\dagger)^m|0\rangle = m\langle 0|a^{m-1}(a^\dagger)^{m-1}|0\rangle = m! \; . \tag{7.30}$$

The normalized eigenstates are thus:

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle \tag{7.31}$$

In quantum field theory, there are field operators for each point in space x. Each of these points might contain $0, 1, 2\cdots$ particles. Whereas the Schrödinger equation involves wave functions, $\psi(x)$ and $\psi^*(x)$, field operators, $\Psi(x)$ and $\Psi^\dagger(x)$, are actually creation and destruction operators,

$$[\Psi(x), \Psi^\dagger(x')] = \delta(x - x'). \tag{7.32}$$

The fact that quantum field theories behave in this manner is one of the most profound aspects of modern quantum theory. It is as if each point in space is its own harmonic oscillator. The kinetic energy term then acts like a term coupling adjacent oscillators, so that the eigenstates of the Hamiltonian are also eigenstates of momentum. One can then associate a harmonic oscillator of frequency $\hbar\omega = \epsilon_p$ for each value of the momentum p.

Example 7.2:

- Calculate $\langle 0|aaa^\dagger aa^\dagger a^\dagger|0\rangle$
  Commute all the creation operators to the right:

$$\langle 0|aaa^\dagger aa^\dagger a^\dagger|0\rangle = \langle 0|a^3(a^\dagger)^3|0\rangle - \langle 0|a^2(a^\dagger)^2|0\rangle = 3! - 2! = 4.$$

- Calculate $\langle n|a^\dagger a^\dagger a^\dagger a|m\rangle$.
  Commute the destruction operator to the left,

$$\begin{aligned}
\langle n|a^\dagger a^\dagger a^\dagger a|m\rangle &= \langle n|a(a^\dagger)^3|m\rangle - 3\langle n|(a^\dagger)^2|m\rangle \\
&= \frac{1}{\sqrt{n!m!}}\left\{\langle 0|a^{(n+1)}(a^\dagger)^{m+3}|0\rangle - 3\langle 0|a^n(a^\dagger)^{m+2}|0\rangle\right\} \\
&= \delta_{n,m+2}\frac{1}{\sqrt{n!m!}}\left\{(n+1)! - 3n!\right\} \\
&= \delta_{n,m+2}(n-2)\sqrt{n(n-1)}.
\end{aligned}$$

- The wave function for the ground state of the harmonic oscillator is:

$$\psi_0(x) = \frac{1}{(2\pi R^2)^{1/2}} e^{-x^2/4R^2}, \quad R = \hbar/(2m\omega).$$

Find $\psi_1(x)$, the wave function of the first excited state by operating $a^\dagger$ on $\psi_0(x)$. Here, simply use the fact that the creation operator is:

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} x - i\sqrt{\frac{1}{2\hbar m\omega}}(-i\partial_x).$$

Differentiating,

$$\psi_1(x) = \sqrt{\frac{2m\omega}{\hbar}} x \psi_0(x) = (x/R)\psi_0.$$

7.3   The Partition Function as a Path Integral

For interactions that defy perturbation theory, several non-perturbative methods are based on the path-integral picture of the partition function:

$$Z = \sum_i \langle i|e^{-\beta H}|i\rangle \;=\; \sum_{i_1, i_2, \cdots, i_n} \langle i_1|e^{-\delta\beta H}|i_2\rangle\langle i_2|e^{-\delta\beta H}|i_3\rangle\langle i_3|\cdots|i_n\rangle\langle i_n|e^{-\delta\beta H}|i_1\rangle,$$

$$e^{-\delta\beta H} \;\approx\; 1 - \delta\beta H \tag{7.33}$$

Here, $\delta\beta = \beta/n$, and completeness was used to insert the intermediate states. If $n \to \infty$, $\delta\beta \to 0$ and the exponential becomes $1 - \delta\beta H$, which is linear in H. The quantity is considered to be a path integral because each set $i_1 \cdot i_n$ can be considered as a quantum path.

In various forms of the nuclear shell model, a subset of all states is considered in the sum over i. The quantity $M = (1 - \delta\beta H)$ is treated as a matrix, and the path integral is simple the trace of this matrix to the nth power. Because states in the trace are many-body states, even the three-particle - three-hole excitations of a few single-particle levels can amount to millions of many-body states. Nonetheless, such brute-force methods are often applied. The $T = 0$ limit is reached by having $\beta \to \infty$ which requires many applications of M. In fact, this is related to as the Lanczos method for finding the ground state of a matrix Hamiltonian.

An even more brute-force method for calculating the path integral is lattice gauge theory. In lattice calculations, the states $|i\rangle$ refer to configurations of the relativistic quantum field operators at each point in space. In order to make the calculations tenable, the continuum of space points is divided up into a lattice. The finite steps in imaginary time, $\delta\beta$, represent a discretization in a fourth dimension.

One possibility for describing the states $|i\rangle$ would be to consider each point on the lattice as it's own harmonic oscillator basis, then the state i would have $n(i_x, i_y, i_z)$ excitations at each lattice point. However, this is not how lattice calculations are performed. Instead, the state i refers to

a set of complex field values at each point, η_{i_x,i_y,i_z}, where η is a complex number describing the state:

$$
\begin{aligned}
|\vec{\eta}\rangle &= \exp\left(\vec{\eta}\cdot\vec{a}^\dagger - \vec{\eta}^*\cdot\vec{a}\right)|0\rangle \\
&= e^{-|\eta|^2/2}\exp\left(\vec{\eta}\cdot\vec{a}^\dagger\right)|0\rangle
\end{aligned}
\tag{7.34}
$$

Here, $\vec{a}$ is short-hand for $a_1, a_2 \cdots$ where the components refer to different single-particle levels. The last step above exploited the Campbell-Baker-Hausdorff lemma,

$$
e^{A+B} = e^A e^B e^{-[A,B]/2}
\tag{7.35}
$$

The states $|\vec{\eta}\rangle$ are convenient because they are normalized eigenstates of the destruction operators. For instance,

$$
\vec{a}|\vec{\eta}\rangle = \vec{\eta}|\vec{\eta}\rangle.
\tag{7.36}
$$

In order to use this basis in the path-integral formulation, the states must satisfy completeness, which in this case is (homework problem),

$$
\frac{1}{\pi}\int d\eta_r d\eta_i |\eta\rangle\langle\eta| = 1,
\tag{7.37}
$$

where η_r and η_i are the real and imaginary parts of η. Thus, each step in imaginary time can be considered a set of values of η, and the path can be considered as a path through these sets of numbers.

Example 7.3:

The state $|\eta\rangle$ is not orthogonal to $|\eta'\rangle$. Find the overlap, $\langle\eta'|\eta\rangle$.

Expanding the exponentials in the bra,

$$
\begin{aligned}
\langle\eta'|\eta\rangle &= e^{-|\eta'|^2/2}\langle 0|e^{-\eta'^* a}|\eta\rangle \\
&= e^{-|\eta'|^2/2}\langle 0|e^{\eta'^*\eta}|\eta\rangle.
\end{aligned}
$$

Expanding the ket and keeping only the first term, because the others don't overlap with $\langle 0|$,

$$
\langle\eta'|\eta\rangle = \exp\{-|\eta'|^2/2 - |\eta|^2/2 + \eta'^*\eta\}
$$

Because the kinetic terms in the Hamiltonians have gradients, the dominant paths are those where η is smooth. Rather than randomly sampling all paths, lattice calculations use Metropolis algorithms which consider small changes of the path, then keep or reject the new path based on the relative statistical weight of the new and old paths.

In order to perform the summations in Eq. (7.33) one considers element of the type,

$$
\langle\vec{\eta}_{n+1}|e^{-\delta\beta H(\vec{a}^\dagger,\vec{a})}|\vec{\eta}_n\rangle = \langle\vec{\eta}_{n+1}|e^{-\delta\beta H(\vec{\eta}_{n+1}^*,\vec{\eta}_n)}|\vec{\eta}_n\rangle,
\tag{7.38}
$$

which assume that the Hamiltonian is normal ordered with all the creation operators pushed to the left. One difficulty with these elements is that the creation and destruction operators are evaluated for different η. To overcome this, one can use Eq. (7.34) to rewrite,

$$
|\vec{\eta}_n\rangle = e^{-[\delta\vec{\eta}^*\vec{\eta} + \vec{\eta}^*\delta\vec{\eta}]/2}e^{\delta\vec{\eta}\cdot\vec{a}^\dagger}|\vec{\eta}_{n+1}\rangle \quad (\text{where } \delta\eta \equiv \eta_n - \eta_{n+1}),
\tag{7.39}
$$

into the the matrix element to obtain

$$\langle \vec{\eta}_{n+1}|e^{-\delta\beta H}|\vec{\eta}_n\rangle = \langle \vec{\eta}_{n+1}|e^{-\delta\beta H}e^{[\delta\vec{\eta}^*\cdot\vec{\eta}-\vec{\eta}^*\cdot\delta\vec{\eta}]/2}|\vec{\eta}_{n+1}\rangle. \tag{7.40}$$

Here, we have assumed $\delta\beta$ and $\delta\eta$ are small. After referring to $\delta\vec{\eta} = \dot{\vec{\eta}}\delta\beta$, one obtains the relation,

$$\langle \vec{\eta}_{n+1}|e^{-\delta\beta H}|\vec{\eta}_n\rangle = \langle \vec{\eta}_{n+1}|e^{\delta\beta[\dot{\vec{\eta}}^*\eta/2-\vec{\eta}^*\dot{\vec{\eta}}/2-H]}|\vec{\eta}_n\rangle. \tag{7.41}$$

One can now rewrite the path integral in Eq. (7.33) as

$$Z = \frac{1}{\pi^N}\prod_i\left\{\int d\vec{p}_i d\vec{q}_i \langle \vec{p}_i, \vec{q}_i|e^{\delta\beta[i\vec{p}_i\cdot\dot{\vec{q}}_i/2-i\dot{\vec{p}}_i\cdot\vec{q}_i/2-H(\vec{p}_i,\vec{q}_i)]}|\vec{p}_i, \vec{q}_i\rangle\right\}, \tag{7.42}$$

where N is the number of points (Dimension of p times number of time steps) over which fields are integrated. Here, $\vec{p}$ and $\vec{q}$ are the real and imaginary part of $\vec{\eta}$ (aside from a factor of $\sqrt{2}$),

$$\eta \equiv \frac{p+iq}{\sqrt{2}}. \tag{7.43}$$

Calculation of Z thus entails dividing β into an infinite number of steps $\delta\beta$, then integrating over all $\vec{p}$ and $\vec{q}$ for each step, with the boundary conditions, $\vec{p}(\beta) = \vec{p}(0), \vec{q}(\beta) = \vec{q}(0)$. This is a path integral and is often written as:

$$Z = \frac{1}{\pi^N}\int\left\{\prod_i d\vec{p}(\beta_i)d\vec{q}(\beta_i)\right\}\exp\left\{\int_0^\beta d\tau \mathcal{L}\left(\dot{\vec{p}}(\tau), \dot{\vec{q}}(\tau), \vec{p}(\tau), \vec{q}(\tau)\right)\right\}. \tag{7.44}$$

Here, $\mathcal{L} = i\vec{p}\cdot\dot{\vec{q}} - H$ is the Lagrangian, with the factor of i coming from the fact that the time step is in imaginary time. We also note that integration by parts and the boundary conditions that the path return to its original starting point made terms of the form $(p\dot{q} - \dot{p}q)/2$ equivalent to $p\dot{q}$.

Lattice gauge theory is useful for relativistic theories where the particle number is not a good quantum number, and the η basis is as reasonable as any, and is especially useful for situations where condensed fields play a dominant role, such as QCD. Whereas the method described above applies only for bosonic fields (because the creation operators follow bosonic rules), the method, after some contortions, can also be applied for fermions.

Quantum chromodynamics is a non-abelian relativistic field theory, which is very strongly interacting and cannot be treated in perturbation theory. It is also characterized by condensed quark-antiquark pairs and by a condensate of gluonic fields. At temperatures in the neighborhood of 170 MeV, the state of the system changes rather abruptly. Quarks no longer belong to well defined color singlets (color deconfinement) and the background condensate of quark-antiquark pairs melts (restoration of chiral symmetry). Lattice gauge calculation provide the only means to "solve" such a strongly interacting theory. A large community is devoted to precisely such calculations. These calculations involve large and sometimes very specialized computational facilities. Calculations typically exceed 10^{20} floating point operations.

7.4   Problems

1. Calculate the second-order virial coefficient to first order in perturbation theory for distinguishable particles of mass m at temperature T interacting through the two-body potential,

$$V(r) = \left\{ \begin{array}{ll} 0\,, & r < a \\ V_0/r^6\,, & r > a \end{array} \right.$$

2. Consider a gas of distinguishable particles of mass m at temperature T and chemical potential μ in a volume V interacting through a mean field,

$$V = V_0.$$

   (a) Using Eq. (7.19) calculate the correction to $\ln Z$ to first order in perturbation theory.

   (b) Calculate the mth order correction.

   (c) Sum all the corrections to see that

$$\ln Z = \frac{V}{(2\pi\hbar)^3} \int d^3p \, e^{-(\epsilon_p + V_0 - \mu)/T}.$$

3. Consider the state

$$|\eta\rangle \equiv \exp\left(\eta a^\dagger - \eta^* a\right)|0\rangle,$$

   where $a^\dagger$ and a are creation/destruction operators, $[a, a^\dagger] = 1$.

   (a) Show that $|\eta\rangle$ is an eigenstate of the destruction operator,

$$a|\eta\rangle = \eta|\eta\rangle.$$

   (b) Show that $|\eta\rangle$ obey the relations,

$$\langle\eta'|\eta\rangle = \exp\{\eta'^*\eta - [|\eta|^2 + |\eta'|^2]/2\}.$$

   (c) Show that the completeness relation is satisfied,

$$\frac{1}{\pi} \int d\eta_r d\eta_i \, \langle n_1|\eta\rangle\langle\eta|n_2\rangle = \delta_{n_1,n_2},$$

   where $|n\rangle$ is a normalized state with n particles, $|n\rangle = (1/\sqrt{n!})(a^\dagger)^n|0\rangle$.

4. Consider a Hamiltonian,

$$H = H_0 + j(t)(a + a^\dagger), \quad H_0 = \hbar\omega a^\dagger a,$$

   where $j(t)$ is a real function.

   (a) Show that

$$e^{-iH_0t/\hbar}a = ae^{-iH_0t/\hbar}e^{i\omega t}.$$

(b) Show that the state

$$|\psi(t)\rangle \equiv e^{-iH_0t/\hbar} \exp\left\{-(i/\hbar)\int_{-\infty}^{t} dt'\, j(t')\left[e^{-i\omega t'}a + e^{i\omega t'}a^\dagger\right]\right\}$$

is an eigenstate of the Hamiltonian, i.e., show

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi\rangle.$$

(c) Assuming that $j(t) = 0$ for $t > t'$, find an expression for η in terms of j such that $|\psi(t > t')\rangle$ is a coherent state,

$$|\psi(t)\rangle = e^{-iH_0t/\hbar}\exp\left\{\eta a^\dagger - \eta^* a\right\}|0\rangle.$$

8   Linear Response Theory

Life is simple. You get out what you put in. – Anonymous

8.1   Linear Response and Kubo Relations

Here, we consider the response of thermally equilibrated systems to small external forces, such as an electro-magnetic field or velocity gradient. These external fields lead to changes in a variety of quantities. For instance, an external field generates a current. The ratio of the two quantities, in this case the conductivity, represents a fundamental property of matter. Examples are:

Stimulus	Response	Response/Stimulus
E field	Current	Conductivity
B field	Magnetization	Magnetic susceptibility
Velocity Gradient	Stress-Energy tensor	Viscosity
Density Gradient	Particle Flow	Diffusion Constant
Temperature Gradient	Current	Thermal Conductivity

These properties can be expressed in terms of correlations or fluctuations in the medium that occur in the absence of the stimulus. For example, the conductivity is related to the current-current correlation function in a system with no external field.

Deriving the connections between the driving field (the cause) and the response (effect) is the heart of linear-response theory. A variety of relations, such as the *Fluctuation Dissipation* theorem, *Green-Kubo Relations* all play a role, and one lesson to understand is that dissipation plays a critical role in all these relations. This is not surprising – if there were no dissipation the electrons in a metal would accelerate rather reaching a steady current. Our goals here are modest, as we simply wish to derive an expression for the response of an observable $\langle A \rangle$, which is non-zero due to a small external field F. We want to find $\chi = \langle A \rangle / F$. More generally, one could determine the time dependent generalized susceptibility $\chi(t)$,

$$\langle A(t) \rangle = \int_{-\infty}^{t} dt' \, \chi(t - t') F(t'), \tag{8.1}$$

where the brackets $\langle ... \rangle$ denote thermal averages. This would allow one to understand response due to transient fields or to fields driven with particular frequencies. These broader considerations would open the subject up to that of an entire course.

The goal, as stated above, is to express the response in terms of operators, or correlations of operators, in a thermalized system in the absence of the external field. To begin, one must identify how the external field interacts with the system. For small fields, we assume it connects linearly, through a term $V = FB$, where B is some operator. For instance, if F is an electric field, $E(x)$, the operator $B(x) = -x\rho(x)$, where ρ is the charge density. Because F is small, the correction to A can be found in perturbation theory,

$$\langle A(t) \rangle = \left\langle \left(1 - \frac{i}{\hbar} \int_{-\infty}^{t} V_I(t') dt'\right) A \left(1 + \frac{i}{\hbar} \int_{-\infty}^{t} V_I(t') dt'\right) \right\rangle, \tag{8.2}$$

$$V_I(t) = -F(t)B(t),$$

$$B(t) = e^{-iH_0 t} B e^{iH_0 t}, \tag{8.3}$$

and B is the operator in the Heisenberg representation. Because $\langle A \rangle = 0$ (otherwise we could consider δA), we then find

$$\langle A \rangle = \frac{-i}{\hbar} \int_{-\infty}^{0} dt \, \langle [A(0), B(t)] \rangle F(t). \tag{8.4}$$

This expression is often referred to as a Kubo relation, or a Green-Kubo relation. If A is the current density in the x direction, and F is the electric field, the conductivity is

$$\sigma = \frac{-i}{\hbar} \int_{\infty}^{0} dt' \int d^3r' \, \langle [j_x(\mathbf{r} = 0, t = 0), -x\rho(\mathbf{r}', t')] \rangle. \tag{8.5}$$

The expression is typically modified through some Fourier gymnastics to get it in the form of a current-current correlation. First insert a term $(\partial/\partial t')t'$ into the integral, then integrate by parts to find

$$\sigma = \frac{-i}{\hbar} \int_{\infty}^{0} dt' \int d^3r' \, \langle [j_x(\mathbf{r} = 0, t = 0), -x\rho(\mathbf{r}', t')] \rangle \left(\frac{\partial}{\partial t'} t' \right) \tag{8.6}$$

$$= \frac{-i}{\hbar} \int_{\infty}^{0} dt' \, t' \int d^3r' \, \langle [j_x(\mathbf{r} = 0, t = 0), x\frac{\partial}{\partial t'}\rho(\mathbf{r}', t')] \rangle.$$

Next, use the equation of continuity, $\partial_t \rho = -\nabla \cdot \mathbf{j}$, and a second integration by parts to find

$$\sigma = \frac{-i}{\hbar} \int_{\infty}^{0} dt' \, t' \int d^3r' \, x \langle [j_x(\mathbf{r} = 0, t = 0), -\nabla \cdot \mathbf{j}(\mathbf{r}', t')] \rangle$$

$$= -\frac{i}{\hbar} \int_{\infty}^{0} dt' \, t' \int d^3r' \, \langle [j_x(\mathbf{r} = 0, t = 0), j_x(\mathbf{r}', t')] \rangle \tag{8.7}$$

The terms with j_y and j_z were discarded by symmetry arguments.

Eq (8.7) has nice aesthetics, but is difficult to grasp physically. If one were to consider the classical limit, the current operators at different times would commute, but then $\hbar \to 0$, so one would have zero divided by zero. Fortunately, the conductivity can be re-expressed in terms of anti-commutators, and without dividing by $\hbar$. This requires analyticity arguments. For short hand, we consider the expression,

$$\sigma = \frac{-i}{\hbar} \int_{-\infty}^{\infty} dt \, t G(t), \tag{8.8}$$

$$G(t) \equiv \langle J(t) J(0) \rangle,$$

$$J(t) = \frac{1}{\sqrt{V}} \int d^3r \, j_x(\mathbf{r}, t).$$

Here, the volume is V, and we have made use of the fact that the commutator is an odd function in t, which allows us to remove one term of the commutator by extending the integration over all time.

Next, we can make use of properties of the thermal trace to derive symmetry relations for $G(z)$
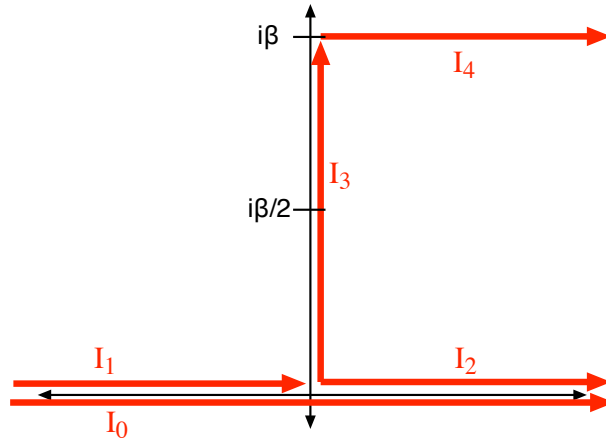
Figure 8.1: The path integrals, $I_0 \cdots I_4$, are illustrated above. Using symmetry and analyticity arguments, relations between the various segments lead to Eq. (8.12) to show that the segment $I_0 = 0$, which then provides the connection between the Kubo commutator and anti-commutator version of the Kubo relations.

in the complex plane,

$$
\begin{aligned}
\langle J(0)J(i\beta\hbar/2 + z)\rangle &= \operatorname{Tr} e^{-\beta H} J e^{iH(i\beta/2 + z/\hbar)} J e^{-iH(i\beta/2 + z/\hbar)}, && (8.9) \\
&= \operatorname{Tr} e^{iH(i\beta/2 + z/\hbar)} J e^{-iH(i\beta/2 + z/\hbar)} e^{-\beta H} J \\
&= \operatorname{Tr} e^{-\beta H} J e^{iH(i\beta/2 - z/\hbar)} J e^{-iH(i\beta/2 - z/\hbar)} \\
&= \langle J(0)J(i\beta/2 - z/\hbar)\rangle, \\
G(i\beta\hbar/2 + z) &= G(i\beta\hbar/2 - z).
\end{aligned}
$$

Here, we have made use of cyclic properties of the trace. This expression allows us to see that $G(z)$ has an even reflection symmetry about the point $i\beta/2$ in the complex plane. We consider the integrals,

$$
\mathcal{I}_a = \int_a dz \, (z - i\beta\hbar/2) G(z). \tag{8.10}
$$

We consider four paths for the integral (illustrated in Fig. 8.1),

- I_0, which goes from $z = -\infty$ to ∞ along the real axis.

- I_1, which goes from $z = -\infty$ to 0 along the real axis.

- I_2, which goes from $z = 0$ to ∞ along the real axis.

- I_3, which goes from $z = 0$ to $z = i\beta\hbar$ along the imaginary axis.

- I_4, which goes from $z = i\beta\hbar$ to $z = i\beta\hbar + \infty$, parallel to the real axis.

From analyticity and the reflection symmetry about $z = i\beta\hbar/2$, one can see the following conditions,

$$
\begin{aligned}
I_3 &= 0, \text{ reflection symmetry}, && (8.11) \\
-I_2 + I_3 + I_4 &= 0, \text{ integrating analytic function about closed loop}, \\
I_4 &= -I_1, \text{ reflection symmetry}.
\end{aligned}
$$

129

One can then see that

$$
\begin{aligned}
I_0 &= I_1 + I_2 \\
&= -I_4 + I_2 \\
&= -I_3 = 0.
\end{aligned}
\tag{8.12}
$$

Thus,

$$
\int_{-\infty}^{\infty} dt\, t G(t) = i\frac{\beta\hbar}{2}\int_{-\infty}^{\infty} dt\, G(t),
\tag{8.13}
$$

$$
\begin{aligned}
\sigma &= \frac{\beta}{2V}\int_{-\infty}^{\infty} dt\, G(t) \\
&= \frac{\beta}{2}\int_{-\infty}^{\infty} dt\int d^3r\, \langle j_x(0)j_x(\mathbf{r},t)\rangle \\
&= \frac{\beta}{2}\int_{0}^{\infty} dt\int d^3r\, \langle \{j_x(0), j_x(\mathbf{r},t)\}\rangle,
\end{aligned}
$$

where the curly brackets denote anti-commutators. The division by $\hbar$ has vanished, and in the classical limit the operators commute.

This last expression, Eq. (8.13), can then be compared to simple arguments for the conductivity. Consider a gas of charged particles ($q = e$), of mass m at temperature T. For $t = 0$, the current-current correlation function only has contributions between a particle and itself,

$$
j_x(\mathbf{r}, t = 0) = \frac{1}{V}\sum_i e\frac{p_{i,x}/m}{},
\tag{8.14}
$$

$$
\begin{aligned}
\int d^3r\, \langle j_x(\mathbf{r} = 0, t = 0)j_x(\mathbf{r}, t = 0)\rangle &= \frac{e^2}{V}\sum_i \frac{p_{i,x}^2}{m^2}, \\
&= e^2 n\frac{T}{m}.
\end{aligned}
\tag{8.15}
$$

Here, the density of particles is n, and the equipartition theorem has been used, $\langle p_x^2/2mT\rangle = T/2$. Now, assume the correlation decays with a relaxation time τ_R, which is likely approximately two mean collision times. This would give

$$
\int_{0}^{\infty} dt\int d^3r\, \langle j_x(\mathbf{r} = 0, t = 0)j_x(\mathbf{r}, t = 0)\rangle = e^2 n\tau_R\frac{T}{m}
\tag{8.16}
$$

$$
\sigma = \frac{e^2 n\tau_R}{m}
$$

The factor of T was canceled by the factor of β in Eq. (8.13). The conductivity estimated from the Kubo relation above can now be compared to that from a simple kinetic model. In that case, the current density is

$$
j_x = en\langle p_x/m\rangle,
\tag{8.17}
$$

where the term in the brackets is the average velocity due to the acceleration during time τ_R,

$$
\langle p_x\rangle = eE\tau_R.
\tag{8.18}
$$

Inserting this into the expression for the current,

$$j_x = \frac{e^2 n \tau_R}{m} E, \tag{8.19}$$

$$\sigma = \frac{e^2 n \tau_R}{m},$$

which matches the previous expression from the Kubo relation.

Kubo relations are powerful. They allow calculation of fundamental properties of matter from microscopic pictures. Calculations can vary from sophisticated quantum treatments based on advanced many-body theory, to the simple kinetic picture above.

8.2   Response at Non-Zero Frequency

In the previous subsection, the response was calculated for a constant field F, whereas often applied fields are sinusoidal with frequency ω. The generated response has the same frequency. The general Kubo relation, Eq. (8.4), shows that the general response of a quantity a to a field $F(t)$ that couples to an operator b through $H_I(t) = -b(t)F(t)$, is

$$\delta\langle a(t=0)\rangle = \int dt' \chi_{ab}(t-t') F(t'), \tag{8.20}$$

$$\chi_{ab}(t) = 2i\chi''_{ab}(t)\Theta(t),$$

$$\chi''_{ab}(t) = \frac{1}{2\hbar}\langle [a(t), b(0)]\rangle.$$

In frequency, the response then becomes

$$\hat{a}(\omega) = \chi_{ab}(\omega) F(\omega), \tag{8.21}$$

$$\hat{\chi}_{ab}(\omega) = \int dt\, e^{i\omega t} \chi_{ab}(t),$$

and the quantity of interest is $\hat{\chi}_{ab}(\omega)$.

One can determine $\hat{\chi}_{ab}(\omega)$ from $\hat{\chi}''_{ab}(\omega)$ using contour integration,

$$e^{i\omega t}\Theta(t) = \frac{1}{2\pi i}\int \frac{e^{i\omega' t} d\omega'}{\omega' - \omega - i\epsilon}, \tag{8.22}$$

$$\hat{\chi}_{ab}(\omega) = \frac{1}{\pi}\int d\omega' dt \frac{e^{i\omega' t} d\omega'}{\omega' - \omega - i\epsilon}\chi''(t)$$

$$= \frac{1}{\pi}\int d\omega' \frac{\hat{\chi}''_{ab}(\omega')}{\omega' - \omega - i\epsilon}. \tag{8.23}$$

Using $1/(x - i\epsilon) = \pi i \delta(x) + \mathcal{P}(1/x)$, one can express the integrals as

$$\hat{\chi}_{ab}(\omega) = \hat{\chi}'_{ab}(\omega) + i\hat{\chi}''_{ab}(\omega), \tag{8.24}$$

$$\hat{\chi}'_{ab}(\omega) \equiv \frac{1}{\pi}\int d\omega'\, \mathcal{P}\frac{1}{\omega' - \omega}\hat{\chi}''_{ab}(\omega').$$

This is a Kramers Kronig relation, [2].

The symmetry properties of $\chi_{ab}''(t)$ and $\hat{\chi}_{ab}''(\omega)$ are important to understand.

$$
\begin{aligned}
\chi_{ab}''(t) &= -\chi_{ba}''(t), & (8.25) \\
&= -[\chi_{ab}''(t)]^*, & (8.26) \\
&= -\epsilon_a \epsilon_b \chi_{ab}''(-t), & (8.27) \\
\hat{\chi}_{ab}''(\omega) &= \epsilon_a \epsilon_b \hat{\chi}_{ba}''(\omega), & (8.28) \\
&= \epsilon_a \epsilon_b [\hat{\chi}_{ab}''(\omega)]^*, & (8.29) \\
&= -\epsilon_a \epsilon_b \hat{\chi}_{ab}''(-\omega), & (8.30)
\end{aligned}
$$

where ϵ_a and ϵ_b are each ± 1 and represent the behavior of the operators under time reversal. For instance, densities are even under time reversal, whereas currents and magnetizations are odd. Depending on whether $\epsilon_a \epsilon_b = \pm 1$, $\hat{\chi}_{ab}''$ is an even or odd function of ω, which means that one need only consider the even or odd part of $\mathcal{P}/(\omega' - \omega)$ in the expression for $\hat{\chi}'(\omega)$ above. Thus,

$$
\hat{\chi}_{ab}'(\omega) = \begin{cases} \frac{1}{\pi} \int d\omega' \, \mathcal{P} \frac{\omega'}{\omega'^2 - \omega^2} \hat{\chi}_{ab}''(\omega'), & \epsilon_a \epsilon_b = 1 \\ \frac{1}{\pi} \int d\omega' \, \mathcal{P} \frac{\omega}{\omega'^2 - \omega^2} \hat{\chi}_{ab}''(\omega'), & \epsilon_a \epsilon_b = -1 \end{cases} \tag{8.31}
$$

Thus $\hat{\chi}_{ab}'(\omega)$ is even when $\hat{\chi}''(\omega)$ is odd, and is odd when $\hat{\chi}''(\omega)$ is even. When $\hat{\chi}''$ is real, so is $\hat{\chi}'$, and when $\hat{\chi}''$ is imaginary, so is $\hat{\chi}'$.

If $\epsilon_a \epsilon_b = 1$, both $\hat{\chi}'(\omega)$ and $\hat{\chi}''(\omega)$ are real, with $\hat{\chi}'(\omega)_{ab}$ being an even function, while $\hat{\chi}_{ab}''(\omega)$ being odd. In this case, $\hat{\chi}'$ represents the response that is in phase with the driving force, and $\hat{\chi}''$ represents the out-of-phase response. If $\epsilon_a \epsilon_b = -1$, $\hat{\chi}''$ is an even function and imaginary, while $\hat{\chi}'$ is an odd function while also imaginary. Thus, when $\epsilon_a \epsilon_b = -1$ $\hat{\chi}'$ represents the out-of-phase response and $\hat{\chi}''$ represents the in-phase response.

Unfortunately, $\chi_{ab}''(t)$ is difficult to grasp physically. This is because it is a commutator, and in the classical limit becomes zero. However, the response does not disappear in the classical limit, because the commutator is divided by $\hbar$, which also goes to zero. Fortunately, one can relate the commutator to either the anti-commutator,

$$
\Phi_{ab}(t) \equiv \frac{1}{2} \langle \{a(t), b(0)\} \rangle, \tag{8.32}
$$

or to the simple correlator

$$
\begin{aligned}
S_{ab}(t) &\equiv \langle a(t) b(0) \rangle, & (8.33) \\
\chi_{ab}''(t) &= \frac{1}{2\hbar} \left(S_{ab}(t) - S_{ba}(-t) \right), \\
\Phi_{ab}(t) &= \frac{1}{2} \left(S_{ab}(t) + S_{ba}(-t) \right),
\end{aligned}
$$

through the fluctuation-dissipation theorem. Before progessing, we rewrite the relations above

---

[2]https://en.wikipedia.org/wiki/Kramers-Kronig_relations

in frequency space,

$$\hat{\chi}''_{ab}(\omega) = \frac{1}{2\hbar}\left(\hat{S}_{ab}(\omega) - S_{ba}(-\omega)\right), \tag{8.34}$$

$$\hat{\Phi}(\omega) = \frac{1}{2}\left(\hat{S}_{ab}(\omega) + S_{ba}(-\omega)\right)$$

The fluctuation-dissipation theorem relates these quantities in frequency space. Deriving the relation begins by exploiting the cyclic property of the trace,

$$\begin{aligned}
S_{ab}(t) &= \mathrm{Tr}\, e^{-\beta H} a(t) b(0) \\
&= \mathrm{Tr}\, e^{-\beta H} e^{-iHt/\hbar} a(0) e^{iHt/\hbar} b(0) \\
&= \mathrm{Tr}\, e^{-\beta H} e^{iHt/\hbar} b(0) e^{-iHt} e^{-\beta H} a(0) e^{\beta H} \\
&= \mathrm{Tr}\, e^{-\beta H} b(-t) a(-i\hbar\beta) \\
&= S_{ba}(-t + i\beta).
\end{aligned} \tag{8.35}$$

One can also use the fact that $e^{\tau \partial_t} f(t) = f(t + \tau)$ to rewrite the last line as

$$S_{ab}(t) = e^{i\hbar\beta\partial_t} S_{ba}(-t). \tag{8.36}$$

In Fourier space, ∂_t becomes $-i\omega$ in the transformation and

$$\hat{S}_{ab}(\omega) = e^{\beta\hbar\omega} \hat{S}_{ba}(-\omega). \tag{8.37}$$

One can now insert these into the expressions for $\hat{\chi}''$ and $\hat{\Phi}$ in Eq.s (8.34),

$$\begin{aligned}
\hat{\chi}''_{ab}(\omega) &= \frac{1}{2\hbar}\hat{S}_{ab}(\omega)(1 - e^{-\beta\hbar\omega}), \\
\hat{\Phi}_{ab}(\omega) &= \frac{1}{2}\hat{S}_{ab}(\omega)(1 + e^{-\beta\hbar\omega}), \\
\hat{\chi}''_{ab}(\omega) &= \frac{1}{\hbar}\hat{\Phi}_{ab}(\omega)\tanh(\beta\hbar\omega/2).
\end{aligned} \tag{8.38}$$

These relations are referred to as the fluctuation dissipation relation. The quantities $\Phi(t)$ and $S(t)$ can be understood in terms of correlations of classical quantities by assuming the operators commute. For χ'', the classical limit is found by taking the lowest order expansion in $\hbar$. In that limit,

$$\begin{aligned}
\hat{\chi}''_{ab}(\omega) &\approx \beta\omega \hat{S}_{ab}(\omega), \tag{8.39} \\
\hat{\Phi}_{ab}(\omega) &\approx \hat{S}_{ab}(\omega). \tag{8.40}
\end{aligned}$$

Example 8.1:

Express the electric conductivity in terms of the current-current correlation, $S_{jj}(\omega)$, for a field oscillating with frequency ω:

Here, to reduce the burden of subscripts, the currents and electric fields are implicitly assumed to be in the x direction. First, we write the interaction,

$$
\begin{aligned}
H_I(t) &= -\int d^3r\; x E(t)\delta\rho(\vec{r},t) \\
&= -\int \frac{1}{2\pi}d\omega'\int d^3r\; x\hat{E}(\omega')\delta\rho(\vec{r},t)
\end{aligned}
$$

The evolution operator is

$$
\begin{aligned}
U(t) &= 1 - \frac{i}{\hbar}\int_{-\infty}^t dt'\; H_I(t') & (8.41)\\
&= 1 - \frac{i}{\hbar}\int d^3r\; x\int \frac{d\omega'}{2\pi}\hat{E}(\omega')\int_{-\infty}^t dt'\delta\rho(x,t')e^{-i\omega't'} & (8.42)\\
&= 1 - \frac{i}{\hbar}\int d^3r\; x\int \frac{d\omega'}{2\pi}\hat{E}(\omega')\int_{-\infty}^t dt'\delta\rho(x,t')\frac{1}{-i\omega'}\partial_{t'}(e^{-i\omega't'}-e^{-i\omega't}) & (8.43)\\
&= 1 + \int d^3r\; x\int \frac{d\omega'}{2\pi\hbar\omega'}\hat{E}(\omega')\int_{-\infty}^t dt'\partial_x j(x,t')(e^{-i\omega't'}-e^{-i\omega't}). & (8.44)
\end{aligned}
$$

The induced current density is then

$$
\begin{aligned}
\langle \delta j(0,t)\rangle &= \int d^3r \int \frac{d\omega'}{2\pi\hbar\omega'}\hat{E}(\omega')dt'(e^{-i\omega't'}-e^{-i\omega't})\Theta(t-t')\langle[\delta j(0,t)\delta j(x,t')]\rangle \\
&= \int \frac{d\omega'}{2\pi\omega'}\hat{E}(\omega')dt'(e^{-i\omega't'}-e^{-i\omega't})\chi''_{JJ}(t-t') \\
\chi''_{JJ}(t) &= \frac{1}{\hbar}\Theta(t)\langle[J(t),J(0)]\rangle, \\
J(t) &= \frac{1}{\sqrt{\Omega}}\int d^3r\; j(\vec{r},t),
\end{aligned}
$$

where Ω is the volume. Performing more Fourier transforms,

$$
\begin{aligned}
\langle \delta\hat{j}(0,t)\rangle &= \int \frac{d\omega'}{(2\pi)^2\omega'}E(\omega')dt'(e^{-i\omega't'}-e^{-i\omega't})\int d\omega'' e^{-i\omega''(t-t')}\hat{\chi}''_{JJ}(\omega''), \\
&= \int \frac{d\omega'}{2\pi\omega'}E(\omega')\int d\omega'' e^{-i\omega''(t-t')}\hat{\chi}''_{JJ}(\omega'')\cdot\left[\delta(\omega'-\omega'')e^{-i\omega't}-\delta(\omega'')e^{-i\omega't}\right] \\
&= \int \frac{d\omega'}{2\pi\omega'}E(\omega')e^{-i\omega't}\left[\hat{\chi}''_{JJ}(\omega')-\chi'_{JJ}(0)\right] \\
\langle \delta\hat{j}(\omega)\rangle &= \frac{E(\omega)}{\omega}\left[\hat{\chi}''_{JJ}(\omega)-\hat{\chi}''_{JJ}(0)\right].
\end{aligned}
$$

The conductivity is then

$$
\begin{aligned}
\hat{\sigma}(\omega) &= \frac{\hat{\chi}''_{JJ}(\omega)-\hat{\chi}''_{JJ}(0)}{\omega} \\
&= \frac{\hat{S}_{JJ}(\omega)}{2\hbar\omega}\left(1-e^{-\beta\hbar\omega}\right) \\
&= \frac{\hat{\Phi}_{JJ}(\omega)}{2\hbar\omega}\tanh(\beta\hbar\omega),
\end{aligned}
$$

134

where the last two expressions used the fluctuation-dissipation theorem, Eq. (8.38).

8.3   Problems

1. Assuming that current is carried by particles of charge e, show that the diffusion constant and conductivity are related by

$$D = \frac{\sigma T}{\chi}.$$

   where χ is the charge fluctuation.

$$\chi = \frac{1}{V}\langle (Q - \bar{Q})^2 \rangle = \frac{\partial \rho}{\partial(\mu/T)}.$$

   The conductivity and diffusivity are defined by

$$\vec{J} = -D\nabla\rho,$$
$$\vec{J} = -\sigma\vec{E}.$$

   Hint: Assume the current responds to the electric potential the same as it would a chemical potential,

$$\Phi \leftrightarrow \mu.$$

2. Consider an interaction of the magnetic field with the magnetization density $M(\mathrm{r})$,

$$H_{\mathrm{int}} = \mu \int d^3 r M(\mathrm{r}) B.$$

   Find an expression for the magnetic susceptibility χ, defined by

$$\langle M \rangle = \chi B,$$

   in terms of correlations of the magnetic field, i.e. $\langle M(0)M(\mathrm{r},t)\rangle$, and the temperature.

3. Show that in the limit where $\omega \to 0$, the conductivity $\sigma(\omega \to 0)$ from the final example matches that found in Eq. (8.13).

4. Consider that the equal-time correlation of the current density is from Eq. (8.14),

$$\int d^3 r \, \langle \delta J(0,t)\delta J(\vec{r},t)\rangle = e^2 n\frac{T}{m}.$$

   Assume the classical limit, and assume that the correlation dies in time as

$$\int d^3 r \, \langle \delta j(0,t)\delta j(\vec{r},0)\rangle = \int d^3 r \, \langle \delta j(0,t)\delta j(\vec{r},0)\rangle e^{-|t|/\tau}.$$

   (a) Derive $\hat{S}_{JJ}(\omega)$.

   (b) What is $\hat{\sigma}(\omega)$?

   (c) As $\omega \to 0$, what is the conductivity?

5. A magnetic field is applied is applied with frequency ω. Find the magnetic susceptibility in the classical limit, $\hat{\chi}_M(\omega)$, as a function of ω assuming that you know the susceptibility in the zero frequency limit, $\hat{\chi}_M(\omega \to 0)$, and the exponential time for which correlations decay, τ.