Chapter 9

Statistical Interpretation of Temperature and Entropy

Temperature and entropy were introduced initially as thermodynamic quantities and interpreted in terms of macroscopic properties of a body in Chapter 2. For example, for every fluid it is possible to find a unique function, $T(P,V)$, of its parameters of state $(P,V)$ which takes the same numerical value when the fluids are in equilibrium with each other. That is, all fluids in equilibrium have the same temperature $T$ and this may be regarded as the thermodynamic definition of temperature. No heat flows between two fluids in equilibrium.

The thermodynamic temperature could also, equivalently, be defined as the integrating factor needed in the equation

$$dS = \frac{dQ}{T}$$

to make the left hand side a perfect differential in a reversible exchange of heat $dQ$. This equation is also the thermodynamic definition of an entropy change, $dS$.

In this chapter we present a statistical interpretation of both temperature and entropy. This will serve to complement the thermodynamic interpretation and heighten the meaning of these two central concepts.

9.1 Temperature

In statistical mechanics the temperature appears fundamentally as a parameter in the Boltzmann factor

$$P_s = e^{-\epsilon_s/kT}/\sum_s e^{-\epsilon_s/kT},$$

the probability of observing a system in energy state $\epsilon_s$. From this we can obtain the function $n(\epsilon_s)$ describing the number distribution of elementary particles in the system over the energy states $\epsilon_s$ available to them. The three examples we considered are the Maxwell-Boltzmann distribution, the Fermi-Dirac distribution and the
Bose-Einstein distribution. Thus if we knew the form and shape of the distribution we could state the temperature of the assembly of elementary particles.

For example, at \( T = 0 \) K (\( \beta = \infty \)) the Fermi-Dirac distribution was such that spin \( \frac{1}{2} \) Fermions occupied the lowest energy states \( \epsilon < \epsilon_F \), two Fermions per state. This remained largely the case for any \( T \ll T_F \). If we observed a Fermi system having few particles in energy states \( \epsilon < \epsilon_F \) we could deduce directly that the temperature was high, \( T \gg T_F \). Similarly for the Bose-Einstein case, at \( T = 0 \) K all the Bosons occupy (condense into) the lowest single particle state. The extent that the higher states are occupied is a direct measure of the temperature. In this sense we have a definite relation between the distribution and \( T \) and we can define a statistical temperature in terms of this distribution. For a system in equilibrium we saw in section 5.1 that this statistical temperature is identical to the thermodynamic temperature.

We have considered already two practical examples of using the statistical distribution to determine the temperature; Doppler Broadening (section 4.3) and Black Body Radiation (section 7.1). In Doppler Broadening we saw that the width of an observed spectral line was a direct measure of the Maxwell-Boltzmann velocity distribution. By measuring the line width we can determine the temperature of the gas. Similarly the spectral distribution of light emitted from a Black Body is directly related to its temperature through Planck’s distribution law. Measurement of the spectral distribution of light emitted from very hot bodies is at present the most practical method for determining their temperature. This is used, for example, in all foundary processes and for measuring the temperature of the sun and stars. What is measured, of course, is the statistical temperature.

9.2 Negative Temperatures

Within the statistical interpretation a state of negative temperature is possible. To explore this we consider a system of independent particles in which the particles have only two levels available to them, \( \epsilon_0 \) and \( \epsilon_1 \). For convenience we can take \( \epsilon_0 = 0 \). This energy level structure is depicted in Fig. 9.1.

An example is a system of independent, non-interacting spin \( \frac{1}{2} \) particles separated from one another on a lattice in a magnetic field, \( H \). The spins aligned with the field have a lower energy than those aligned against the field. Then \( \epsilon_\uparrow = \epsilon_0 \), \( \epsilon_\downarrow = \epsilon_1 \) and the separation between the energy levels is \( \Delta \epsilon = 2\mu H \). Such a system of free spins is called an Ideal Paramagnet. This ideal paramagnet should not be confused with the Pauli paramagnet discussed in section 8.4. In the Pauli paramagnet the wave functions of the Fermions overlap and the Pauli exclusion principle, which limits state occupation to one Fermion per state,
introduces an effective interaction between the spins. This prevents all the spins from occupying the lowest energy $\epsilon$ state. Ultimately this interaction originates from the overlap of the free Fermion wave functions in the gas. In contrast, we assume here that the spins are completely non-interacting and independent. In practice this requires localizing the spins in space (e.g., on a lattice). This can be done for electrons by localizing electrons spins around atoms in solids. A celebrated example is the salt cerium magnesium nitrate (CMN). Here the cerium has a net spin and each cerium spin is well separated from the others by the other constituents. A second example is the nuclear spins in solid $^3\text{He}$. To a good approximation these spins can be regarded as non-interacting. We shall see that the nearly ideal paramagnet CMN and solid $^3\text{He}$ are used as refrigerator materials for cooling to low temperatures.

For independent spins the occupation of the $\epsilon_1 = \Delta \epsilon$ and $\epsilon_0 = 0$ states is given by the Boltzmann factor,

$$N_1 = \frac{e^{-\beta \Delta \epsilon}}{1 + e^{-\beta \Delta \epsilon}}, \quad N_0 = \frac{1}{1 + e^{-\beta \Delta \epsilon}}$$

and

$$\frac{N_1}{N_0} = e^{-\beta \Delta \epsilon} \quad (9.1)$$

Eq. (9.2) shows that at $T \to 0$ all spins will be in the lowest energy state $\epsilon_0$ and the spin system will be fully aligned in the field ($N_1 = 0$). At $T \to \infty$ there is equal occupation of $\epsilon_1$ and $\epsilon_0$ ($N_1 = N_0$) and zero net magnetization. As noted the spins are either electron spins or nuclear spins and the remainder of the solid is referred to as the lattice. At equilibrium the temperature of the spins (given by $\beta$ in eq. (9.1)) is the same as that of the rest of the lattice and this temperature sets the magnetization in the spin system. However, the thermal contact between the spin system and the lattice can often be very poor so that a distinct “spin” temperature $T_S$ becomes possible. This spin temperature is defined from eq. (9.2) by the ratio of $N_1/N_0$.

To see how $T_S$ can take negative values suppose a field $H$ is applied so that $N_1/N_0$ is given by eq. (9.2). We now reverse the direction of $H$ so that $\epsilon_0 > \epsilon_1$. The spin system takes a finite time to “relax” to the appropriate population of $N_1$ and $N_0$ and immediately after the switch we have

$$\frac{N_0}{N_1} = e^{\beta \Delta \epsilon} \quad (9.3)$$

The spin temperature is then said to be negative. Eventually as the spins “relax” the spin temperature goes over to that of the lattice and becomes positive again. That is, $N_0/N_1$ relaxes to $e^{-\beta \Delta \epsilon}$. The initial negative temperature is defined only in a statistical sense via eq. (9.3) and could not be reached by a slow thermodynamic process from equilibrium state to equilibrium state. For this reason we could not construct a heat engine operating from this negative temperature. Simply, a negative temperature means that higher energy levels have higher occupation than the lower levels so that the sign is reversed in the Boltzmann factor.
Negative temperatures are fundamental to LASER operation (Light Amplification by Stimulated Emission of Radiation). Consider again the two level system depicted in Fig. 9.1 but suppose $\epsilon_0$ and $\epsilon_1$ are electron energy levels. We can stimulate transitions between the two levels by shining light of frequency $h\nu_0 = \Delta\epsilon = \epsilon_1 - \epsilon_0$ on the system. Since the transition probability for absorption and emission of photons is equal the populations $N_1$ and $N_0$ become equal for intense light ($T = \infty$) and there is no net loss or gain of photons passing through the system.

Suppose now we can, by another process, continually over populate level $\epsilon_1$ so that $N_1 > N_0$ ($T < 0$). This can be done by "pumping" electrons from level $\epsilon_0$ to suitable higher energy states which decay to $\epsilon_1$. This pumping can be achieved using a second light beam of a higher frequency and is referred to as optical pumping. If we now pass the original beam of frequency $h\nu_0$ through this system, having $N_1 > N_0$, it stimulates transitions from $\epsilon_1$ to $\epsilon_0$ and a net emission of photons having identical frequency $\nu_0$ results. The original beam is then amplified and we get Light Amplification by Stimulated Emission of Radiation (photons). This amplification depends entirely on the two level system having a negative temperature.

9.3 Entropy

The statistical interpretation of entropy follows from the Boltzmann's principle

$$S = k \log W_{\text{max}}.$$  \hfill (9.4)

Here $W_{\text{max}}$ is the number of ways that the system may be arranged to produce the same equilibrium thermodynamic state. We showed in section 3.2 that this $S$ and the thermodynamic entropy are identical. In addition we saw there that for a gas $W_{\text{max}}$ could be replaced by $W$ which is the total number of all ways of arranging the gas since this number was dominated by $W_{\text{max}}$—the arrangement leading to the equilibrium state. Thus the complete Boltzmann principle is

$$S = k \log W.$$  \hfill (9.5)

One of the difficulties in using eq. (9.5) is interpreting and evaluating $W$. Indeed Einstein proposed a "reversed" Boltzmann’s principle,

$$W = e^{S/k},$$  \hfill (9.6)

in which $S$ was regarded as known empirically and $W$ was sought. As noted in Chapter 3 Boltzmann’s principle in the form (9.5) was enunciated by Planck including the constant $k$. Boltzmann preferred only to relate the changes in $S$ to changes in the possible number of arrangements of different states of a system. We will effectively use eq. (9.5) in this form. Particularly, as a system goes to equilibrium we will relate the increase in $S$ to transitions from states of artificial order to states of (more probable) disorder.
A formal expression for \( W \) can, however, be readily articulated in terms of the Gibb’s ensemble as

\[
W = \frac{N_A!}{\prod_s N_s!}; \quad \sum_s N_s = N_A
\]

(9.7)

Here \( N_A \) is the total number of systems in the ensemble and \( N_s \) is the number of copies in energy state \( E_s \). The entropy per system is then

\[
S = \frac{1}{N_A} k \log W
\]

(9.8)

\[
= \frac{k}{N_A} \left[ \sum_s N_s \log N_A - \sum_s N_s \log N_s \right]
\]

or

\[
S = -k \sum_s P_s \log P_s
\]

(9.9)

where \( P_s = N_s/N_A \) is the probability of observing a system in state \( S \). If we now substitute \( P_s = e^{-\beta E_s}/Z \) in eq. (9.9) we recover the familiar result for the entropy that can be derived from the canonical partition function,

\[
S = -\left( \frac{\partial F}{\partial T} \right)_V = \frac{\partial}{\partial T} (kT \log Z)_V
\]

\[
S = k \log \sum_s e^{-\epsilon_s/kT} + \frac{1}{T} \sum_s \frac{E_s e^{-\epsilon_s/kT}}{\sum_s e^{-\epsilon_s/kT}}
\]

(9.10)

This last expression for \( S \) is often the most useful for evaluating the entropy if absolute values are desired. Otherwise differences in \( S \) can be readily obtained from eq. (9.5). If the system is a single particle eq. (9.10) holds with \( E_s \) replaced by the single particle energy \( \epsilon_s \).

To begin exploring eq. (9.5) consider a gas at \( T = 0 \) K. Both the Fermi and Bose gases are then in their single ground state. In the case of spin \( \frac{1}{2} \) Fermions we have two Fermions per energy state up to \( \epsilon = \epsilon_F \) with all higher energy states empty. In the Bose case, all the Bosons are in the lowest single particle state. Since the particles are indistinguishable, there is only one way to form this state \( (W = 1) \). and from eq. (9.5) \( S = 0 \). That is, \( S = 0 \) at \( T = 0 \) K follows from there being only one single quantum mechanical ground state in the system.

To articulate the last statement we consider \( S \) given by eq. (9.10) and for convenience we take the lowest single state \( E_0 = 0 \) then clearly eq. (9.10)
gives
\[ \lim_{T \to 0} S = k \log 1 = 0 \]  
(9.11)

A more careful analysis of the limit shows we do not have to take \( E_0 = 0 \) to obtain the same result. We need only assume that \( E_0 \) is not degenerate. Also, if the ground state had degeneracy \( g_0 \), \( S = k \log g_0 \) at \( T = 0 \), which is negligible compared to any finite \( T \) entropy.

If \( S \) contained a constant as well as the two terms in eq. (9.10) this constant would be the same for all systems, for the limit (eq. 9.11) is independent of the energy level spectrum and therefore of the type of system. Thus the value of this constant is physically unimportant and we can equally well take it as zero. This is Nerst’s theorem that \( S \) must be zero at \( T = 0 \) in a nutshell.

As \( T \) increases, the particles occupy higher energy states. There then become different ways to arrange the particles to give the same macroscopic state (e.g. the energy) and from eq. (9.10) \( S \) increases. If the energy levels are closely spaced we see from eq. (9.10) that \( S \) increases rapidly with \( T \). The value of \( S(T) \) then depends sensitively on the energy spectrum and becomes a distinct property for distinct systems. From eq. (9.10) \( S \) always increases with temperature and is positive definite. The last statement simply says that \( C_v = (dS/dT)_v \) is always positive. Exceptions are transitions in phase (e.g. melting) at which point the energy state spectrum in eq. (9.10) changes.

The statistical entropy is also additive as required. Suppose we partition a body into two parts. We denote the entropy of each by \( S_1 \) and \( S_2 \). Similarly we denote the configurations on possible arrangements of each by \( W_1 \) and \( W_2 \). If we bring the two parts together the total number of arrangements of the combined body is
\[ W_T = W_1 W_2 \]
The combined entropy from eq. (9.5) is
\[ S_T = k \log W_T = k \log W_1 + k \log W_2 \]
\[ = S_1 + S_2 \]
demonstrating that \( S \) is additive. Similarly, we could divide the body into \( N \) equal parts. An identical argument gives
\[ S_T = S_1 + S_2 + \ldots + S_N = N S_1 \]
which demonstrates that \( S \) in eq. (9.5) is an extensive quantity proportional to the size of the body.

### 9.4 Entropy and Order

Here we explore the relation between the entropy and the order in a system. From thermodynamics we know that an isolated system under-going
spontaneous changes goes from states of lower to higher entropy. The entropy continues to increase until no further changes take place in the system. The final or equilibrium state is then the state having maximum entropy. We also know from our daily lives that systems isolated and left to themselves tend to a state of maximum disorder. A system isolated with no work done on it tends to states of higher and higher disorder. Examples are our clothes, our rooms or our lecture note books. Thus we can associate entropy and order and here we clarify this concept using Boltzmann’s principle.

Gases

Consider a gas of $N$ atoms in the double chamber depicted in Fig. 9.2. Initially, suppose the gas is entirely confined to the left hand chamber while the right hand chamber is evacuated. In this configuration the order of the gas is greater than if the partition were removed and the gas occupied both volumes. In the initial state, we know that all the gas atoms are on the left side only so the order and our information about their location is greater.

When we open the partition the gas occupies both volumes, a volume of $2V$, and there is a loss of order. We can relate this loss of order to an increase in entropy using eq. (9.5). Since the volume available to the gas has increased from $V$ to $2V$ the volume of phase space available has also increased by two. Hence the number of ways of arranging the atoms among the available states (each of volume $\hbar^3$) has increased by two. The ratio of the initial to final number of possible arrangements of the $N$ atoms is

$$\frac{W_f}{W_i} = 2^N$$  \hspace{1cm} (9.12)

and using the Boltzmann relation the increase in entropy is

$$\delta S = S_f - S_i = k \log \frac{W_f}{W_i} = k N \log 2$$  \hspace{1cm} (9.13)

Hence we obtain a direct relation between an increase in disorder and an increase in $S$.

The instant after the partition is removed the gas is still in the left hand chamber. This can be regarded as an artificial, non-equilibrium state of the gas. The increase in entropy (eq. 9.13) can then be associated with a transition from an artificially prepared state to the most disordered, equilibrium state.
Spin Systems

A second example identical to the gas above is the system of localized, non-interacting spins discussed in section 9.2. In the absence of a magnetic field $H$, the spins have two states, spin up and spin down. The two states have equal energy and are equally available to the spins. In this case the possible arrangements of the spin system is

$$W = 2^N.$$ 

In a strong magnetic field ($\sim 10^4$ Gauss) the spins line up in the direction of $H$. Each spin is then in a single state, spin up, and

$$W = 1^N = 1$$

The entropy of this highly ordered state is

$$S = k \log 1 = 0$$

When $H$ is switched off the spins go from the ordered back to the disordered state. The entropy increase associated with this order to disorder transition is

$$\delta S = k \log 2^N = kN \log 2$$

exactly as in the gas case.

9.5 Adiabatic Demagnetization

The process of aligning and demagnetizing spins is used in low temperature physics as a refrigerator. Temperatures down to $10^{-4}$ K have been reached using adiabatic demagnetization. The sample to be cooled is thermally connected to a paramagnetic salt such as CMN discussed in section 9.2. Both the salt and the sample are cooled to as low temperature as possible in liquid helium, often to less than 30 mK. This cooling is done in a strong magnetic field $H$. The electron spins in the CMN paramagnet are then nearly fully aligned and the spin entropy is effectively zero. The salt and the sample are then isolated from the liquid helium. The magnetic field is switched off and the spins are “demagnetized adiabatically”. The entropy is increased, in the ideal limit by

$$\delta S = kN \log 2.$$ 

This increase in entropy associated with the spin disordering requires a corresponding influx of heat into the spin system

$$\delta Q = T \delta S.$$ 

This heat must come from the sample and the sample is cooled. This cooling is a one shot process and not cyclic as in most refrigerator systems.

Cooling by adiabatic demagnetization may be analysed in more detail using the full entropy expression (eq. 9.10). For a two level spin system having
states $\epsilon_\uparrow = 0$ and $\epsilon_\downarrow = \epsilon_\uparrow + 2\mu H$, this entropy is

\[
S(T, H) = kN \left[ \log \left( 1 + e^{-\frac{2\mu H}{kT}} \right) + \frac{2\mu H}{e^{2\mu H/kT}} \frac{e^{-2\mu H/kT}}{(e^{-2\mu H/kT} + 1)} \right]
\] (9.14)

When $H$ is large the energy level $\epsilon_\downarrow = 2\mu H$ becomes inaccessible and $S \to 0$. Immediately after $H$ is switched off ($H \approx 0$) the spins remain aligned. This aligned state could also be achieved in eq. (9.14) by setting $T = 0$. Thus when $H$ is switched off the effective spin temperature is zero and the sample can be regarded as being cooled through thermal contact with a spin system at very low temperature. Clearly $S$ is a function of both $T$ and $H$ in eq. (9.14). Changing $H$ can be regarded as changing the energy level structure with $S$ depending on the level structure. In practice the detailed analysis of adiabatic demagnetization is complicated but the principle is simple.

Also in practice there is always some interaction between the spins so they are not an ideal paramagnet. This interaction limits the ultimate temperature that can be reached by adiabatic demagnetization. The interaction causes the spins to spontaneously order in zero magnetic field at some low temperature, the Curie Temperature for ferromagnetic ordering or the Néel Temperature for antiferromagnetic ordering. Clearly if the spins reach this ordering temperature then when $H$ is switched off the spins remain ordered and there is no increase in entropy. For electron spins in CMN the exchange interaction causes spontaneous ordering at 10 mK. For nuclear spins the dipole-dipole interaction leads to nuclear spin ordering at about $10^{-6}$ K. The lowest temperatures are therefore reached using nuclear spin demagnetization.

The paramagnets CMN and nuclear spin systems are also used as thermometers in low temperature physics. The extent of ordering and hence the magnetization clearly depends on temperature, as did the entropy in eq. (9.14). The spin temperature may then be determined by measuring the magnetization in the paramagnet.

### 9.6 Pomeranchuk Cooling

An interesting variant of adiabatic demagnetization was proposed by Pomeranchuk\(^1\) in 1950. The spin system here is the nuclear spins of $^3$He atoms in liquid and solid $^3$He. The exchange interaction between the $^3$He nuclear spins is much stronger in the liquid than in the solid phase. This makes sense since the atoms are much less localized in the liquid. The spins therefore order spontaneously at a higher temperature in the liquid than in the solid, about 1 K in the liquid compared to about $\lesssim 10^{-3}$ K in the solid. Thus we may induce a spin order to disorder transition by solidifying helium from the liquid at temperatures between $10^{-3}$ and 1 K. This transition requires an increase in spin order.

entropy and the liquid/solid mixture is cooled since it must supply this entropy.

Fig. 9.3 shows the liquid/solid melting line. Solid helium is formed by increasing the pressure on the liquid. For example, if we begin at point 1 in Fig. 9.3 and apply pressure the liquid solidifies when we cross the liquid/solid transition line. When we reach point 2 all the liquid has solidified. In practice a liquid/solid mixture is used in Pomeranchuk cooling. As pressure is applied more liquid is solidified and the mixture “cools itself” for it must supply entropy (heat) to the spin system to allow the transition to take place.

Pomeranchuk cooling is now widely used for low temperature experiments on liquid and solid helium itself. It was used, for example, in the initial discovery of superfluid liquid $^3$He in 1973. Initially it was not clear whether this transition to the superfluid phase of $^4$He was a transition in the solid or in the liquid since both are present in Pomeranchuk Cooling. The method is not so suitable for cooling other samples since the sample tends to become thermally isolated when solid helium forms around it. The final temperature attainable in Pomeranchuk cooling is set by the antiferromagnetic ordering temperature in the solid—about $10^{-4}$ to $10^{-3}$ K depending upon the pressure.

From Fig. 9.3 we also note that if we cool $^3$He at a pressure $\sim 35$ atmospheres (atm) along the dashed line, the liquid solidifies and then melts again at low temperature. The liquid phase is the low temperature phase since the spin ordering dominates the lattice ordering at that pressure.

### 9.7 Entropy of Mixing

We now evaluate the increase in entropy when two classical, perfect gases are mixed. The mixing is depicted in Fig. 9.4. We begin initially with $N_1$ atoms of type 1 in volume $V_1$ and $N_2$ atoms in $V_2$, separated. The partition is removed and the gases are allowed to mix. In practical cases the pressure

---

of the separated gases is equal so there is gradual mixing by diffusion. Equal
and constant pressure throughout is central in determining the mixing entropy
\( \Delta S_m \), defined as
\[
\Delta S_m = S_m - S_i \tag{9.15}
\]
The initial partition function is just the product of the partition function
for the separated gases (4.9)
\[
Z_i = Z_1 \times Z_2 = \frac{1}{N_1!} \left( \frac{V_1}{\lambda_1^3} \right)^{N_1} \times \frac{1}{N_2!} \left( \frac{V_2}{\lambda_2^3} \right)^{N_2} \tag{9.16}
\]
where \( \lambda_T \) is the thermal wavelength. The partition function after mix ing in
which both the \( N_1 \) and \( N_2 \) non-interacting particles can freely occupy the volume
\( V = V_1 + V_2 \) is
\[
Z_m = \frac{1}{N_1!} \left( \frac{V_1 + V_2}{\lambda_1^3} \right)^{N_1} \times \frac{1}{N_2!} \left( \frac{V_1 + V_2}{\lambda_2^3} \right)^{N_2} \tag{9.17}
\]
The \( Z_m \) is the same as \( Z_i \), except for the volume change. This follows since,
with no interaction, the partition function of \( N_1 \) can be evaluated separately as
if the \( N_2 \) particles were not there. The ratio of the partition functions is then
\[
\frac{Z_m}{Z_i} = \left( \frac{V_1 + V_2}{V_1} \right)^{N_1} \left( \frac{V_1 + V_2}{V_2} \right)^{N_2} \tag{9.18}
\]
With \( F = -kT \log Z \) and \( S = - (\partial F/\partial T)_V \), the mixing entropy is
\[
\Delta S_m = S_m - S_i = \frac{\partial}{\partial T} \left[ kT \log \frac{Z_m}{Z_i} \right]
\]
\[
= k \left[ \log \left( \frac{V_1 + V_2}{V_1} \right)^{N_1} + \log \left( \frac{V_1 + V_2}{V_2} \right)^{N_2} \right] \tag{9.19}
\]
Using the perfect gas equation of state (\( pV = N kT \)) to express \( V \) in terms of
\( N \),
\[
\frac{V_1}{V_1 + V_2} = \frac{N_1}{N_1 + N_2} \equiv x_1, \tag{9.20}
\]
where \( x_1 \) is the fraction of type 1 atoms in the mixture, we have
\[
\Delta S_m = -kN \left[ x_1 \log x_1 + x_2 \log x_2 \right]. \tag{9.21}
\]
For a number of gases mixed together at the same initial and final pressures the
total mixing entropy is
\[
\Delta S_m = -kN \sum_i x_i \log x_i . \tag{9.22}
\]
Returning to eq. (9.19) we may express \( \Delta S_m \) in an equally interesting
form. Using eq. (9.20)
\[ \Delta S_m = k \left[ \log \left( \frac{N_1 + N_2}{N_1} \right)^{N_1} + \log \left( \frac{N_1 + N_2}{N_2} \right)^{N_2} \right] \]

\[ = k \log \left[ \frac{(N_1 + N_2)^{N_1 + N_2}}{N_1^{N_1} N_2^{N_2}} \right] \]

or

\[ \Delta S_m = k \log \left[ \frac{(N_1 + N_2)!}{N_1! N_2!} \right] \quad (9.23) \]

where in the last expression we have used Stirling’s approximation.

The last form is especially interesting when applied to mixing of liquids and solids; again with the restriction that the pressure is equal before and after mixing and interactions are ignored (or taken as equal between all atoms). We may interpret \( \Delta S_m \) in eq. (9.23) in the following manner for a crystal. We bring together a crystal of \( N_1 \) type 1 atoms and a crystal of \( N_2 \) type 2 atoms. The crystals completely mix by diffusion. This mixing is depicted in Fig. 9.5. Before mixing the number of distinguishable ways of permuting the \( N_1 \) atoms among the sites in the crystal 1 is just \( W_I = 1 \), as it is for crystal 2. Although the \( N_1 \) atoms may be permuted \( N_1! \) ways the permutations do not lead to new macroscopic states, since the macroscopic state does not change on interchanging two identical particles.

After mixing the total number of permutations of the atoms among the sites is \( N! = (N_1 + N_2)! \). However permutations of the \( N_1 \) atoms among their sites in the mixture again does not lead to distinguishable states. Hence the number of distinguishable permutations (\( \equiv \) no. of macroscopic states) is the total number \( N! \) divided by those which lead to the same macroscopic state. This is

\[ W_m = \frac{(N_1 + N_2)!}{N_1! N_2!} \quad (9.24) \]

and

\[ \Delta S_m = S_m - S_I = k \log \frac{W_m}{W_I} = k \log \left[ \frac{(N_1 + N_2)!}{N_1! N_2!} \right] \quad (9.25) \]

In this way we see that \( \Delta S_m \) is applicable beyond the perfect gas to liquids and solids mixed at constant pressure and for which the constituents
A particularly interesting application is thermal vacancy creation in crystals at high temperature. At high temperature the free energy of a crystal is lower when a number \( n \) of the sites do not contain an atom. The crystal must then cover \( n + N \) sites with \( n \) of these vacant. The increase in “configuration” entropy when the crystal contains \( n \) vacancies is, from eq. (9.23)

\[
\Delta S_v = k \log \left( \frac{(n + N)!}{n! N!} \right).
\]

There is a local free energy increase \( g \) around each vacancy (assumed isolated and non-interacting). The total free energy change in the crystal required to create \( n \) vacancies is then

\[
\delta G = ng - kT \log \left( \frac{(n + N)!}{n! N!} \right).
\]

Minimizing \( \delta G \) with respect to \( n \) gives the equilibrium fraction of vacant sites expected at temperature \( T \) as

\[
\frac{n}{(N + n)} = e^{-g/kT}.
\]

In most metals at the melting temperature about 1 site in \( 10^4 \) is vacant, \( n/(n + N) \approx 10^{-4} \).

### 9.8 Gibbs’ Paradox

The famous paradox in statistical mechanics pointed out by Gibbs’ in 1905 inspired the corrected Boltzmann counting discussed in section 3.3. The paradox involves the mixing of two gases at constant pressure discussed in section 9.7. The entropy increase on mixing is given by eq. (9.21). The two gases could be isotopes, say \(^3\)He and \(^4\)He which differ little. Provided they differ, however little, we can distinguish the initial separated state of the gases in Fig. 9.4 from the mixed state (after the separating partition has been removed) and \( \Delta S_m \) is given by eq. (9.21). However if the two gases are identical then we cannot distinguish the initial from the final state and \( \Delta S_m = 0 \).

The paradox arose since prior to the corrected Boltzmann counting the relation between \( Z \) and \( Z_0 \) was taken as \( Z = Z_0^N \) rather than the correct expression \( Z = \frac{1}{N!} Z_0^N \), since classical particles were regarded as distinguishable. If we employ the \( Z = Z_0^N \) in eqs. (9.16) and (9.17) we arrive at \( \Delta S_m \) given by eq. (9.21), even for identical gases. However \( \Delta S_m \) must be zero for mixing of identical gases and this is Gibbs’ paradox.

The paradox is immediately removed using the corrected Boltzmann counting. In this case for two identical gases the initial partition function \( Z_1 \) is the same as eq. (9.16) but \( Z_m \) is

\[
Z_m = \frac{1}{N!} \left[ \frac{(V_1 + V_2)^N}{N^N T^N} \right].
\]
with \(N = N_1 + N_2\). The ratio of the two partition function is then,

\[
\frac{Z_m}{Z_I} = \frac{(N_1 + N_2)!}{N_1! N_2!} \frac{(V_1 + V_2)^{N_1 + N_2}}{V_1^{N_1} V_2^{N_2}}
\]

which using eq. (9.20) and Stirling’s approximation is unity for perfect gases and

\[
\Delta S_m = k \log \left( \frac{Z_m}{Z_I} \right) = 0
\]

as required. In this way Gibbs’ paradox is completely resolved using the correct Boltzmann counting of section 3.3.