Appendix 1: Normal Modes, Phase Space and Statistical Physics

The last line of the introduction to the first edition states that 'it is the wide validity of relatively few principles which this book seeks to demonstrate'. Here we apply that concept to the relationship between normal modes which feature in most of the book, phase space of the final chapter, and statistical physics.

Firstly, we wish to show that the expression for the number of normal modes *per unit volume* in the frequency range ν to $\nu + \mathrm{d}\nu$ given on p. 253 as

$$\mathrm{d}n = \frac{4\pi\nu^2 \mathrm{d}\nu}{c^3}$$

is nothing more than the number of 'cells' of phase space *per unit volume* in the same range ν to $\nu + \mathrm{d}\nu$ available to particles in a statistical distribution.

Moreover, we can easily convert this expression in the frequency ν to one in the velocity v, the momentum $p = mv$ or the energy E.

The particle may be a molecule in the classical Maxwell–Boltzmann distribution (M–B), a fermion of half integral spin in the quantum Fermi–Dirac distribution (F–D) or a boson or any other particle of integral spin in the quantum Bose–Einstein distribution (B–E). Bosons are the messengers of the force fields in physics, e.g. the photon in the electromagnetic field.

We shall see that each of these distributions is nothing more than the statement that

$$n_i = g_i \times \text{probable occupation of the phase space cell.}$$

Here n_i is a number of particles in the distribution and g_i is our expression $4\pi\nu^2\,\mathrm{d}\nu/c^3$ (or its equivalent).

The expression for g_i is common to all three types of distribution but the occupation factor or relative probability of occupation depends on the way in which the particles are allowed to distribute themselves.

Firstly, let us examine the various equivalent forms of g_i. We write

$$g_i = g_i(\nu)\,\mathrm{d}\nu = 4\pi\nu^2\,\mathrm{d}\nu/c^3$$

The Physics of Vibrations and Waves, 6th Edition H. J. Pain

as the number of phase space cells per unit volume in the frequency range ν to $\nu + \mathrm{d}\nu$. For a quantum particle (p. 415) the momentum $p = \hbar k = h\nu/c$ where h is Planck's constant, k is the particle wave number $= 2\pi/\lambda$ and c is the velocity of light, so

$$g_i = g_i(p)\,\mathrm{d}p = 4\pi p^2 \mathrm{d}p/h^3$$

is the number of phase space cells per unit volume in the momentum range p to $p + \mathrm{d}p$. Note that $4\pi p^2 \mathrm{d}p$ is the volume of the shell in momentum space between spheres of radius p and $p + \mathrm{d}p$.

All particles in statistical distributions are required to be free particles, that is having only kinetic energy with no potential energy interaction terms.

Thus, the energy of a particle $E = \frac{1}{2}mv^2 = p^2/2m$ where $p = mv$, m is the particle mass and v is its velocity. Now

$$p^2\,\mathrm{d}p = (2m^3)^{1/2}E^{1/2}\,\mathrm{d}E = m^2v^2m\,\mathrm{d}v = m^3v^2\,\mathrm{d}v$$

so

$$g_i = g_i(E)\,\mathrm{d}E = 4\pi(2m^3)^{1/2}E^{1/2}\,\mathrm{d}E/h^3$$

is the number of phase space cells *per unit volume* in the energy range E to $E + \mathrm{d}E$ and

$$g_i = g_i(v)\,\mathrm{d}v = 4\pi m^3v^2\,\mathrm{d}v/h^3$$

is the number of phase space cells *per unit volume* in the velocity range v to $v + \mathrm{d}v$.

Although we used the phase space of $\dot{x}$ or v with x in our discussion of chaos, the phase space of mv or p with x is much more commonly used in physics. The phase space of (p, x) reveals the significance of h^3 in the denominators of g_i. Consider the expression

$$4\pi p^2\,\mathrm{d}pV/h^3$$

where V is the total volume (not the unit volume) so that the numerator expresses the phase space over the momentum range p to $p + \mathrm{d}p$ and the volume $V = xyz$ of the system.

Heisenberg's Uncertainty Principle, p. 416, tells us that $\Delta x\Delta p \sim h$, so we may write $(\Delta x\Delta p_x)(\Delta y\Delta p_y)(\Delta z\Delta p_z)$ as h^3; that is, the 'volume' of a cell in (p, V) phase space. This volume is the smallest acceptable volume which a particle may occupy for it defines the volume associated with a particle as

$$\left(\frac{h}{\Delta p}\right)^3 \approx \lambda_{\mathrm{DB}}^3$$

where λ_{DB} is the de Broglie wavelength of the particle (p. 412).

So g_i measures the number of phase space cells each of 'volume' h^3 *per unit volume* in the range p to $p + \mathrm{d}p$. Each of these cells may or may not be occupied by a particle.

We now examine what we mean by a statistical distribution in order to find the probable occupation of a cell. This occupation factor is different for each of the three distributions M–B, F–D and B–E.

We consider a system, say a gas, of N particles occupying a volume V and having a total internal energy E. The macroscopic parameters E, V, N define a *macrostate*. The energy E may be partitioned in many different ways among the N particles subject only to the restrictions that $E = \sum n_i \varepsilon_i$ and $N = \sum n_i$ remain constant where ε_i represents the energy levels available to the particles. The probability of a system being found in a particular partition is proportional to W the number of ways of distributing the energy among the particles to achieve that partition.

Each different way is called a *microstate* and each *microstate* has *a priori* the same probability. Each microstate contributes to the statistical weight of a partition so that the particular partition reached by the greatest number of ways has the greatest statistical weight and is therefore the most probable. The most probable partition with W (maximum) defines the equilibrium of the macrostate and is written Ω (EVN).

It is here that we relate Ω (EVN) to the concept of entropy S. Entropy is a measure of the disorder of a system which increases as the system tends to equilibrium. At constant temperature and volume the internal energy E of the system may be written

$$E = F + TS$$

where T is the temperature, S is the entropy and the product TS is a measure of the energy of the system locked in the disorder amongst the particles and not available for work. F is defined as the Helmholtz free energy and measures the work which can be done by the system at constant temperature. At best, in an ideal reversible thermodynamic process the disorder energy TS remains constant, but in a natural or thermodynamically irreversible process TS increases at the expense of F as E remains constant.

An isolated system in equilibrium with the most probable partition of its energy among its particles represents a maximum of its entropy S and Boltzmann related S and Ω through his expression $S = k \log \Omega$ where k is Boltzmann's constant. Fluctuations from the equilibrium position are very small indeed and $\log \Omega$ is a very sharply defined function.

Calculating the value of W the statistical weight of a partition in order to find W (maximum) $= \Omega$ (EVN) for each of the three distributions is a mathematical exercise which is straightforward and a little tedious but which fails to reveal the underlying physics.

We shall make these calculations at the end of this appendix but we adopt the procedure of quoting the results below together with the forms in which we usually meet them. This will raise questions the answers to which are not evident in the mathematical derivation (Table A1.1).

For all three distributions the particles are identical and indistinguishable, the total energy E and number of particles N are constant. There are no restrictions on the number of particles having a particular energy in the M–B and B–E distributions but in the F–D distribution, Pauli's exclusion principle allows only one fermion per energy level (or two if we include spin).

Note firstly that the occupation factor or relative probability of occupation for each distribution includes the term $\mathrm{e}^{\alpha + \beta \varepsilon_i}$, where α and β arise as multipliers in the mathematical derivation. The index of the exponential requires β to be the inverse of an energy and the relevant term in the normal form of the Fermi–Dirac distribution suggests that α is the ratio of two energies.

Table A1.1 The mathematical derivation for each statistical distribution in the left hand column is compared with its more familiar form on the right

	$n_i = g_i \times$ occupation factor	Normal form
M–B	$n_i = g_i \times \dfrac{1}{\mathrm{e}^{\alpha+\beta\varepsilon_i}}$ $= g_i\,\mathrm{e}^{-\alpha-\beta\varepsilon_i}$	$\dfrac{n}{N} = \dfrac{4\pi p^2 \mathrm{d}p}{(2\pi mkT)^{3/2}}\,\mathrm{e}^{-p^2/2mkT}$ $(p = mv)$
F–D	$n_i = g_i \times \dfrac{1}{\mathrm{e}^{\alpha+\beta\varepsilon_i}+1}$	$n(E)\,\mathrm{d}E = \dfrac{2.4\pi V(2m^3)^{1/2}E^{1/2}}{h^3} \times \dfrac{1}{\mathrm{e}^{(\varepsilon_i-\varepsilon_F)/kT}+1}$
B–E	$n_i = g_i \times \dfrac{1}{\mathrm{e}^{\alpha+\beta\varepsilon_i}-1}$	$n(\nu)\,\mathrm{d}\nu\,h\nu = E(\nu)\,\mathrm{d}\nu$ $= \dfrac{2.4\pi\nu^2\,\mathrm{d}\nu h\nu}{c^3} \times \dfrac{1}{\mathrm{e}^{h\nu/kT}-1}$ Planck's radiation law

In comparing the two columns of the table several questions arise:

1. Is $\beta = 1/kT$?
2. What has happened to the α term in the normal form of M–B?
3. What is the physical significance of the α term?
4. What has happened to the α term in Planck's radiation law?

In question 1 let us integrate by parts the expression

$$\int \mathrm{e}^{-\beta\varepsilon}\,\mathrm{d}p = [p\mathrm{e}^{-\beta\varepsilon}]_{p=-\infty}^{p=+\infty} + \beta \int p\frac{\partial\varepsilon}{\partial p}\,\mathrm{e}^{-\beta\varepsilon}\,\mathrm{d}p$$

where

$$\varepsilon = p^2/2m$$

For $\varepsilon \to \infty$ as $p \to \pm\infty$ the first term on the right hand side equals zero, leaving

$$\frac{1}{\beta} = \frac{\displaystyle\int p\frac{\partial\varepsilon}{\partial p}\,\mathrm{e}^{-\beta\varepsilon}\,\mathrm{d}p}{\displaystyle\int \mathrm{e}^{-\beta\varepsilon}\,\mathrm{d}p} = \overline{p\frac{\partial\varepsilon}{\partial p}}$$

the average value of

$$p\frac{\partial\varepsilon}{\partial p}$$

From the equipartition of energy

$$p\frac{\partial \varepsilon}{\partial p} = \frac{\overline{p^2}}{m} = kT = \frac{1}{\beta}$$

where kT is the average energy per particle.

In question 2 we note that the term $\mathrm{e}^{-\alpha}$ in M–B has been replaced by $N/(2\pi mkT)^{3/2}$ and that h^3 has been lost from the denominator of $g_i(p)\,\mathrm{d}p$. To explain this and its consequences let us write not *n per unit volume* but n_p in the range p to $p + \mathrm{d}p$ over all $V = xyz$ as

$$n_p = \frac{V 4\pi p^2\,\mathrm{d}p\,\mathrm{e}^{-p^2/2mkT}}{h^3}$$

Then

$$N = \sum n_p = V\int_0^\infty \frac{4\pi p^2\,\mathrm{d}p\,\mathrm{e}^{-p^2/2mkT}}{h^3}$$

where the standard definite integral is well known to have a value of $(2\pi mkT)^{3/2}$.

Thus

$$N = V(2\pi mkT)^{3/2}/h^3$$

Now the average particle momentum $\bar{p} = m\bar{v}$ where $\frac{1}{2}m\bar{v}^2 = kT$ ($\bar{v}$ is the most probable velocity).

Hence

$$(2\pi mkT)^{3/2} \approx \bar{p}^3$$

Thus, $(V/N)\bar{p}^3$ replaces $\mathrm{e}^{\alpha}h^3$ and

$$\mathrm{e}^{\alpha} = \frac{V\bar{p}^3}{N h^3} = \frac{V}{N}\frac{1}{\lambda_{\mathrm{DB}}^3}$$

$$= \frac{\text{Volume available to each particle}}{\text{Volume associated with the thermal de Broglie wavelength of the particle}}$$

The value of $\mathrm{e}^{\alpha} = 0.026\,m^{3/2}T^{5/2}$ at a pressure of one atmosphere, where m is measured in a.m.u. ($\mathrm{O}^{16} = 16$).

For air at STP $\mathrm{e}^{\alpha} \approx 10^6$ so for the Maxwell–Boltzmann distribution

$$\frac{g_i}{n_i} = \mathrm{e}^{\alpha + \beta\varepsilon_i} \approx 10^6\,\mathrm{e}^{\varepsilon_i/kT} \gg 1$$

This states that there are many more states or cells available for occupation than there are particles to fill them, so the probable occupation of each cell is very small. This defines a classical distribution.

For the Bose–Einstein gas He^4 at 4 K and one atmosphere pressure $e^{\alpha} \sim 7.5$ so the gas is not safely classical.

Although it is not strictly applicable, for electrons in a metal at 300 K, $e^{\alpha} \sim 10^{-4}$ so the classical description for the Fermi–Dirac case is totally invalid.

A distribution which is not classical is said to be degenerate. Note that for high enough energies (temperatures) all three distributions become classical.

Before we examine the origin of α and its physical meaning let us note that a factor 2 appears in both the F–D and B–E distributions where each particle has two spin states for each energy level which must be accounted for. In Planck's radiation law these spin states are equivalent to the polarization states of electromagnetic waves. Note also in Planck's law that $E(\nu)\,d\nu$, the energy per unit volume in the frequency range ν to $\nu + d\nu$, is $n(\nu)\,d\nu h\nu$ where $h\nu$ is the photon energy.

Turning to question (iii) on the significance of α we again use the expression $S = k\log\Omega$ or $\Omega = e^{S/k}$. Consider a system in contact with a large reservoir at constant temperature, Figure A1.1, able to exchange both energy and particles with the reservoir. The combination of reservoir and system is isolated and its energy E, volume V and total number of particles N are all fixed and constant.

We ask 'What is the probability of finding the *system* in a particular microstate with n_j particles having *total* energy ε_j?' This will be proportional to the number of microstates in the reservoir after n_j and ε_j are supplied to the system.

The entropy equation with subscript R for reservoir becomes

$$S_R(E - \varepsilon_j, N - n_j) = S_R(E, N) - \varepsilon_j\left(\frac{\partial S}{\partial E}\right)_{NV} - n_j\left(\frac{\partial S}{\partial N}\right)_{EV}$$

where we neglect higher terms in the expansion.

Elementary thermodynamics shows that

$$\left(\frac{\partial S}{\partial E}\right)_{NV} = \frac{1}{T} \quad \text{and} \quad \left(\frac{\partial S}{\partial N}\right)_{EV} = \frac{-\mu}{T}$$

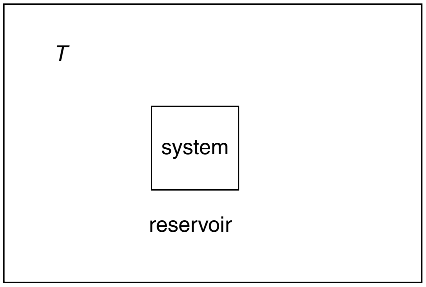

Figure A1.1 When a system, surrounded by a large reservoir with constant N, V and E receives n_j particles and total energy ε_j from the reservoir, the entropy change of the reservoir is $\Delta S = (n_j\mu - \varepsilon_j)/T$ where μ is the chemical potential where μ is called the chemical potential. The chemical potential μ is the free energy per particle lost when the entropy S is increased in the relation $E = F + TS$ where E is constant. Thus, the entropy change may be written

$$\Delta S = S_{\mathrm{R}}(E - \varepsilon_j, N - n_j) - S_{\mathrm{R}}(E, N) = -\frac{\varepsilon_j}{T} + \frac{n_j \mu}{T}$$

Because the statistical weight Ω (EVN) represents the probability of a partition, the probability of the combination of two partitions may be written as the product of their statistical weights so we have

$$\begin{aligned}\Omega(E - \varepsilon_j, N - n_j) &= \Omega(E, N)\,\Omega(\varepsilon_j, n_j)\\ &= \Omega(E, N)\,\mathrm{e}^{\Delta S/k}\\ &= \Omega(E, N)\,\mathrm{e}^{(n_j\mu - \varepsilon_j)/kT}\end{aligned}$$

In order to show the relation between α and $-\mu/kT$, we take as an example a system of four fermions available to occupy any of four single particle energy states ε_1, ε_2, ε_3, ε_4 (Table A1.2). The particles and energies are supplied by the reservoir and each energy level may be filled or empty. The numbers of possible microstates of the system using 0, 1, 2, 3 or 4 particles are shown below together with their relative probabilities.

For any microstate in which a particular energy level is filled we can find another which differs only in having that energy level empty.

Table A1.2 Distribution of four fermions among four single particle energy states with numbers of possible microstates and their relative probabilities

	No particles	One particle	Two particles	Three particles	Four particles
Number of microstates	1	4	6	4	1
Energy level ε_4	0	0	0	0	1
Energy level ε_3	0	0	0	1	1
Energy level ε_2	0	0	1	1	1
Energy level ε_1	0	1	1	1	1
	$n_j = 0$ $\varepsilon_j = 0$	$n_j = 1$ $\varepsilon_j = \varepsilon_1$	$n_j = 2$ $\varepsilon_j = \varepsilon_1 + \varepsilon_2$	$n_j = 3$ $\varepsilon_j = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$	$n_j = 4$ $\varepsilon_j = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4$
Relative probability of microstate	$\mathrm{e}^{(0-0)/kT}$	$\mathrm{e}^{(\mu-\varepsilon_1)/kT}$	$\mathrm{e}^{[2\mu-(\varepsilon_1+\varepsilon_2)]/kT}$	$\mathrm{e}^{[3\mu-(\varepsilon_1+\varepsilon_2+\varepsilon_3)]/kT}$	$\mathrm{e}^{[4\mu-(\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4)]/kT}$

Thus, for example

$$\begin{aligned}\frac{\text{Relative probability of finding } \varepsilon_3 \text{ filled}}{\text{Relative probability of finding } \varepsilon_3 \text{ empty}} &= \frac{p}{1-p}\\ &= \frac{\mathrm{e}^{[3\mu-(\varepsilon_1+\varepsilon_2+\varepsilon_3)]/kT}}{\mathrm{e}^{[2\mu-(\varepsilon_1+\varepsilon_2)]/kT}}\\ &= \mathrm{e}^{(\mu-\varepsilon_3)/kT}\end{aligned}$$

More generally

$$\frac{p}{1-p} = \mathrm{e}^{(\mu-\varepsilon_i)/kT}$$

so

$$p = \frac{1}{\mathrm{e}^{(\varepsilon_i-\mu)/kT}+1} = \overline{n_i}$$

where $n_i = g_i\bar{n}_i$ and $\bar{n}_i$ or the relative probability is the average occupation of a cell.

This is the Fermi–Dirac occupation factor and we can identify $\alpha = -\mu/kT$ (the ratio of two energies) where μ is the chemical potential. For the Fermi–Dirac distribution $\bar{n}_i \leq 1$ and Figure A1.2 shows $\bar{n}_i$ versus ε for electrons in a metal at $T = 0\,\mathrm{K}$.

Each energy level is occupied by one electron until the top energy level ε_{F} the Fermi energy level is reached. At $T = 0\,\mathrm{K}$ the electron with ε_{F} is the only one capable of moving to change the entropy of the system and we identify its free energy with that of the chemical potential μ. Note that, at ε_{F} for $T > 0$, $\bar{n}_i = \frac{1}{2}$ and this is indicated by the dotted curve at ε_{F} in the $\bar{n}_i$ versus ε graph.

We may apply a similar procedure to particles obeying Bose–Einstein statistics where there is no restriction on the number of particles n_i in the energy level ε_i. If n_i can take any value, three identical bosons available to three energy levels $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ can form the

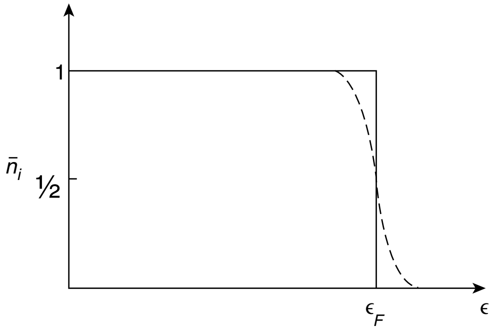

Figure A1.2 Occupation number $\bar{n}_i$ versus energy ε for electrons in a metal at $T = 0\,\mathrm{K}$ (solid line). A slight increase in T permits the electrons near ε_{F} to move to higher energy levels (dotted curve) microstates (3, 0, 0) (0, 3, 0) (0, 0, 3) (2, 1, 0) (0, 2, 1) (1, 0, 2) (0, 1, 2) (2, 0, 1) (1, 2, 0) (1, 1, 1). The energy of each microstate is given by $\varepsilon_j = \sum n_i \varepsilon_i$ with $n_j = \sum n_i$. Suppose, as before, a large reservoir at temperature T surrounds a system to which it can supply particles and energy.

We consider a particular microstate of the system with $n_1, n_2, n_3 \ldots n_i$ particles in the various energy levels to have a *probability* p *when* $n_i = 0$.

If the system now takes n_i particles each of energy ε_i from the reservoir the probability of the microstate (now with $n_i \neq 0$) is given by

$$p\,\mathrm{e}^{n_i(\mu-\varepsilon_i)/kT} = p\,\mathrm{e}^{n_i x}$$

where $x = (\mu - \varepsilon_i)/kT$.

The total probability for the microstate with $n_i = 0, 1, 2, 3, \ldots$ is

$$1 = \sum_{n_i=0}^{n_i=\infty} p\,\mathrm{e}^{n_i x} = \frac{p}{(1-\mathrm{e}^x)}$$

because $\sum \mathrm{e}^{n_i x}$ is a geometric progression.

Hence

$$p = (1 - \mathrm{e}^x)$$

The average value

$$\bar{n}_i = \sum_{n_i=0}^{n_i=\infty} n_i p\,\mathrm{e}^{n_i x}$$

But

$$\sum n_i \mathrm{e}^{n_i x} = \frac{\mathrm{d}}{\mathrm{d}x}\sum \mathrm{e}^{n_i x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{(1-\mathrm{e}^x)} = \frac{\mathrm{e}^x}{(1-\mathrm{e}^x)^2}$$

Therefore

$$\begin{aligned}\bar{n}_i &= \frac{p\,\mathrm{e}^x}{(1-\mathrm{e}^x)^2} = \frac{(1-\mathrm{e}^x)\mathrm{e}^x}{(1-\mathrm{e}^x)^2} = \frac{\mathrm{e}^x}{(1-\mathrm{e}^x)}\\ &= \frac{1}{\mathrm{e}^{-x}-1} = \frac{1}{\mathrm{e}^{(\varepsilon_i-\mu)/kT}-1}\end{aligned}$$

The general expression for the Bose–Einstein distribution is therefore

$$n_i = g_i \bar{n}_i = g_i \times \frac{1}{\mathrm{e}^{(\varepsilon_i-\mu)/kT}-1}$$

Finally we discuss the absence of α or $-\mu/kT$ in Planck's radiation law, noting that this is a special case and that $-\mu/kT$ is retained in other applications of Bose–Einstein statistics.

Black body radiation is an equilibrium process, so that the system or cavity of a box of photons is in equilibrium with the reservoir at temperature T, the entropy S is a maximum and this process results from the continual emission and absorption of photons by the walls of the cavity. The number of photons in the cavity is not conserved, the energy requirement could be satisfied by a few high energy photons in the γ-ray region or by many photons in the low energy infrared frequencies. This means that the occupation numbers are not subject to the constraint which specifies the total number of particles in the gas.

Since N is not fixed, the entropy S of the reservoir is not affected by the n_j photons in the exponent $n_j\mu$ of the occupation factor for a given microstate; n_j has no role and $n_j\mu = 0$ giving $\mu = 0$.

The graph of the entropy S versus N, the total number of particles, gives low S values, that is few microstates or particle arrangements at low N (γ-rays) and also at high N (infrared) photons.

A typical microstate for γ-rays occupying the energy levels ε_i would read

$$n_1 = 0, \quad n_2 = 0, \quad n_3 = 0 \quad \text{with} \quad n_{\to\infty} \neq 0$$

and for infrared photons a typical microstate would read

$$n_1 \neq 0 \quad n_2 = 0 \quad n_3 = 0$$

Both of these are extremely unlikely and would contribute to partitions of low statistical weight.

At equilibrium the maximum of the S versus N curve occurs at that value of N providing the greatest number of microstates and here

$$\left(\frac{\partial S}{\partial N}\right)_{EV} = \frac{-\mu}{T} = 0$$

again giving $\mu = 0$.

Mathematical Derivation of the Statistical Distributions

The particles are identical but distinguishable by labels. All energy ε states are equally accessible and have the same *a priori* probability of being occupied. The statistical weight or probability of a particular partition is proportional to the number of different ways of distributing particles to obtain that partition.

Maxwell–Boltzmann Statistics

We start by filling the ε_1 states with n_1 particles from the constant total of N particles. We can do this in

$$\frac{N!}{n_1!(N - n_1)!}$$

different and distinguishable ways.

We now fill the ε_2 state with n_2 particles from the $N - n_1$ remaining particles. This gives

$$(N - n_1)!/n_2!(N - n_1 - n_2)!$$

different and distinguishable ways.

Proceeding in this way for all remaining energy states we have

$$W = \frac{N!}{n_1!n_2!n_3!\ldots}$$

as the number of different and distinguishable ways of choosing $n_1, n_2, n_3, \ldots$ from the N particles. Particles with the same ε_i may have g_i differing amounts of angular momentum, etc. This will give g_i cells associated with ε_i in each of which a particle with ε_i may be located. If g_i is the probability of having one particle in the ε_i range of cells then $g_i \times g_i = g_i^2$ is the probability of two particles in that range and $g_i^{n_i}$ is the probability of n_i particles with ε_i being in that range.

Hence the total number of different distinguishable ways is

$$W = \frac{N!g_1^{n_1}g_2^{n_2}g_3^{n_3}\ldots}{n_1!n_2!n_3!\ldots}$$

The particles are distinguished by labels and if we now remove the labels and the condition of distinguishable particles, we cannot recognize the difference in the partition when particles are exchanged. Therefore all $N!$ permutations among the particles occupying the different states give the same partition with the total number of ways

$$W = \frac{g_1^{n_1}g_2^{n_2}g_3^{n_3}\ldots}{n_1!n_2!n_3!\ldots}$$

We now maximize log W with the constraints that

1. The number of particles $N = \sum n_i =$ constant so that $\mathrm{d}N = \sum \mathrm{d}n_i = 0$.
2. The energy $E = \sum n_i\varepsilon_i =$ constant so that $\mathrm{d}E = \sum \varepsilon_i \mathrm{d}n_i = 0$.

$$\log W = \sum_i (n_i \log g_i - \log n_i!)$$

where for large n_i Stirling's formula gives

$$\log n_i! = n_i \log n_i - n_i$$

Hence

$$\log W = \sum n_i \log \frac{g_i}{n_i} + \sum n_i$$

and

$$\begin{aligned}
\mathrm{d}\,(\log W) &= \sum \mathrm{d}n_i \log\left(\frac{g_i}{n_i}\right) + \sum n_i\,\mathrm{d}\log\left(\frac{g_i}{n_i}\right) + \sum \mathrm{d}\,n_i \\
&= \sum \mathrm{d}n_i \log\left(\frac{g_i}{n_i}\right) - \sum n_i \frac{\mathrm{d}n_i}{n_i} \\
&\qquad \text{(because } g_i \text{ is constant and} \textstyle\sum \mathrm{d}n_i = 0\text{)} \\
&= \sum \mathrm{d}n_i \log\left(\frac{g_i}{n_i}\right)
\end{aligned}$$

If $\sum \mathrm{d}n_i = 0$ then $-\alpha \sum \mathrm{d}n_i = 0$ and
if $\sum \varepsilon_i \mathrm{d}n_i = 0$ then $-\beta \sum \varepsilon_i \mathrm{d}n_i = 0$
where α and β are called Lagrange multipliers.

Adding these constraint conditions to d(log W) gives

$$\mathrm{d}\,(\log W) = \sum \mathrm{d}n_i \left(\log\left(\frac{g_i}{n_i}\right) - \alpha - \beta\varepsilon_i\right)$$

Maximizing W gives d(log W) = 0 which, since all the coefficients dn_i are arbitrary and independent, leaves

$$\log\left(\frac{g_i}{n_i}\right) - \alpha - \beta\varepsilon_i = 0$$

for each n_i.

At $W_{\max}$ we have therefore

$$n_i = g_i \times \frac{1}{\mathrm{e}^{\alpha+\beta\varepsilon_i}}$$

Fermi–Dirac Statistics

We begin again with labelled identical particles. Here the Pauli exclusion principle operates and no two particles may occupy the same state. The g_i are quantum states, e.g. spin gives a factor 2 to each g_i. Also g_i gives the maximum number of particles with ε_i so $n_i \leq g_i$.

To fill the ε_i states with n_i particles we put one particle in a g_i cell and the next particle in any of the $(g_i - 1)$ remaining cells. We can do this in $g_i(g_i - 1)$ ways so the total number of ways of filling the states of energy ε_i with n_i particles is

$$\begin{aligned}
&g_i(g_i - 1)\ldots(g_i - n_i + 1) \\
&= \frac{g_i!}{(g_i - n_i)!}
\end{aligned}$$

If now the labels are removed and the particles become indistinguishable we reduce the total of different distinguishable arrangements to $g_i!/n_1!(g_i - n_i)!$.

Applying this to all g_i gives the total number of different distinguishable ways as

$$W = \frac{g_1!}{n_1!(g_1 - n_1)!} \, \frac{g_2!}{n_2!(g_2 - n_2)!} \, \frac{g_3!}{n_3!(g_3 - n_3)!}$$

Maximizing $\log W$ with $\sum \mathrm{d}n_i = \sum \varepsilon_i \, \mathrm{d}n_i = 0$ we proceed as with the Maxwell–Boltzmann example to obtain for W (max) the condition that

$$\log\left(\frac{g_i}{n_i} - 1\right) - \alpha - \beta\varepsilon_i = 0$$

to give

$$n_i = g_i \times \frac{1}{\mathrm{e}^{\alpha + \beta\varepsilon_i} + 1}$$

Bose–Einstein Statistics

Here there is no exclusion principle and we begin again with labelled identical particles.

The number of distinguishable arrangements of n_i particles in the g_i cells of energy ε_i equals the number of ways of putting n_i objects in g_i boxes with any number allowed in a box. This means putting n_i particles in a row separated by $g_i - 1$ walls so that the number of ways is the number of permutations of $(n_i + g_i - 1)$ objects, i.e. particles and walls. This gives $(n_i + g_i - 1)!$ ways. If we now remove the particle labels to make them indistinguishable we reduce the number of ways by a factor of $n!$ to give $(n_i + g_i - 1)!/n_i!$ ways.

However, all permutations of the $g_i - 1$ dividing walls among the n_i particles give the same physical state, so the number of different distinguishable ways is given by $(n_i + g_i - 1)!/n_i!(g_i - 1)!$ and for all particles we have the number of ways

$$W = \frac{(n_1 - g_1 - 1)}{n_1!(g_1 - 1)} \, \frac{(n_2 + g_2 - 1)}{n_2!(g_2 - 1)} \cdots \text{etc.}$$

Maximizing $\log W$ as for the other two distributions gives $\mathrm{d}(\log W) = 0$ when

$$\log\left(\frac{g_i}{n_i} + 1\right) - \alpha - \beta\varepsilon_i = 0$$

that is, when

$$n_i = g_i \frac{1}{\mathrm{e}^{\alpha + \beta\varepsilon_i} - 1}$$

Table A1.3 Waves incident normally on a plane boundary between media of characteristic impedances Z_1 and Z_2

			Amplitude Coefficients			
Wave type	Impedance Z $+$ve for wave in $+$ve direction $-$ for wave in $-$ve direction	Boundary conditions	$\frac{Reflected_r}{Incident_i} = \frac{Z_1 - Z_2}{Z_1 + Z_2}$	$\frac{Transmitted_t}{Incident_i} = \frac{2Z_1}{Z_1 + Z_2}$	$\frac{Reflected_r}{Incident_i} = \frac{Z_2 - Z_1}{Z_1 + Z_2}$	$\frac{Transmitted_t}{Incident_i} = \frac{2Z_2}{Z_1 + Z_2}$
Transverse on string (p. . . .)	$\frac{-T(\partial y/\partial x)}{\dot{y}} = \rho c = (T\rho)^{1/2}$	$y_i + y_r = y_t$ or $\dot{y}_i + \dot{y}_r = \dot{y}_t$ $T\left[\frac{\partial y_i}{\partial x} + \frac{\partial y_r}{\partial x} = \frac{\partial y_t}{\partial x}\right]$	y and $\dot{y}$		$-T\frac{\partial y}{\partial x}$	
Longitudinal acoustic (p. . . .)	$\frac{\rho}{\dot{\eta}} = \rho_0 c = (B_a\rho)^{1/2}$	$\dot{\eta}_i + \dot{\eta}_r = \dot{\eta}_t$ $p_i + p_r = p_t$	η and $\dot{\eta}$		p	
Voltage and current on transmission line (p. . . .)	$\frac{V}{I} = \sqrt{\frac{L}{C}}$	$I_i + I_r = I_t$ $V_i + V_r = V_t$	I		V	
Electro-magnetic (p. . . .)	$\frac{E}{H} = \sqrt{\frac{\mu}{\varepsilon}}$	$H_i + H_r = H_t$ $E_i + E_r = E_t$	H		E	

All waves $\frac{\text{Reflected intensity}}{\text{Incident intensity}} = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2}\right)^2$ $\frac{\text{Transmitted intensity}}{\text{Incident intensity}} = \frac{4Z_1Z_2}{(Z_1 + Z_2)^2}$

Appendix 2: Kirchhoff's Integral Theorem

Kirchhoff's Integral Theorem is valid for any solution E of the scalar time independent Helmholtz equation (3), p. 187, that is

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + k^2 \mathbf{E} = 0$$

For the radial direction r in a spherical coordinate system this becomes

$$\frac{\partial^2 \mathbf{E}}{\partial r^2} + \frac{2}{r}\frac{\partial \mathbf{E}}{\partial r} = 0$$

which is satisfied by

$$\mathbf{E} = \frac{E_0}{r}\,\mathrm{e}^{\mathrm{i}kr}$$

where E_0/r is the amplitude at a distance r from the origin O of a spherical electromagnetic wave. We note that the amplitude of such a wave decays as $1/r$ where r is the distance from O.

Kirchhoff's Theorem states that the complex amplitude $\mathbf{E}_P$ at a point P is related to the complex amplitude $\mathbf{E}$ on a surface S enclosing P by

$$\mathbf{E}_P = \frac{1}{4\pi}\iint_S \left(\mathbf{E}\frac{\partial}{\partial n}\frac{\mathrm{e}^{\mathrm{i}kR}}{R} - \frac{\mathrm{e}^{\mathrm{i}kR}}{R}\frac{\partial \mathbf{E}}{\partial n}\right)\mathrm{d}S$$

where R is the distance from P to the surface element d**S** and **n** is the direction normal to d**S** (Figure A2.1).

If r is the distance from O to dS, then

$$\mathbf{E} = \frac{E_0}{r}\,\mathrm{e}^{\mathrm{i}kr}$$

The Physics of Vibrations and Waves, 6th Edition H. J. Pain

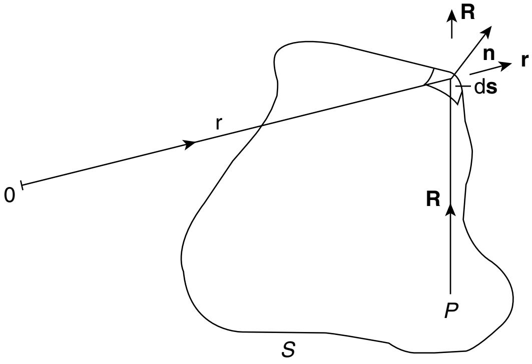

Figure A2.1 0 is the origin of an electromagnetic wave. Kirchoff's Theorem relates its complex amplitude $\mathbf{E}_P$ at a point P to the complex amplitude $\mathbf{E}$ on a surface S enclosing P

and

$$\frac{\partial \mathbf{E}}{\partial n} = \frac{E_0}{r}\,\mathrm{e}^{\mathrm{i}kr}\left(\mathrm{i}k - \frac{1}{r}\right)\cos\left(\mathbf{n}, \mathbf{r}\right)$$

The term $(\mathrm{i}k - 1/r)$ shows that inside S there is a phase shift of $\pi/2$ rad and an amplitude factor $1/r$. However, for $r = m\lambda$, where m is large, then

$$k = \frac{2\pi}{\lambda} \gg \frac{1}{r} = \frac{1}{m\lambda}$$

so that $1/r$ may be neglected for distances much greater than λ.

Similar arguments hold for

$$\frac{\partial}{\partial n}\frac{\mathrm{e}^{\mathrm{i}kR}}{R}$$

Thus, if P and O are many wavelengths from S, Kirchhoff's integral becomes

$$\mathbf{E}_P = \frac{-i}{\lambda}\iint E_0 \frac{\mathrm{e}^{\mathrm{i}k(r+R)}}{rR}\frac{(\cos\mathbf{n},\mathbf{R} - \cos\mathbf{n},\mathbf{r})}{2}\,\mathrm{d}\mathbf{S}$$

where the cosine terms generate an inclination factor $K(\chi)$ and $\cos(\mathbf{n},\mathbf{R}) = \cos\chi$.

The problem of showing that Huygens wavelets on an unobstructed wavefront do not propagate backwards reduces to that of demonstrating that $K(\chi)$ can be zero. This occurs where

$$\cos\left(\mathbf{n}, R\right) = \cos\pi = -1$$

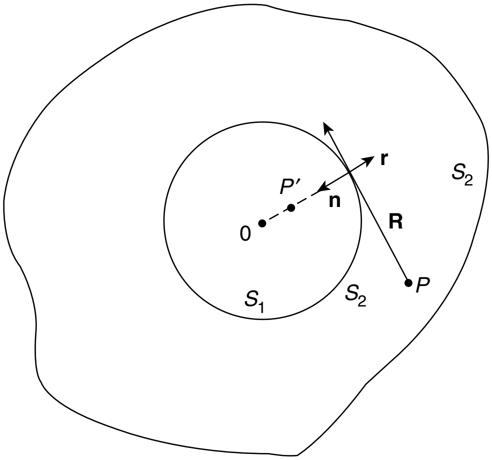

Figure A2.2 When P' is located on **r** within the surface of the spherical wavefront S_1, situated within S, $\mathbf{E}_{P'}$ is reduced to zero proving that Huygens wavelets do not propagate backwards

and

$$\cos(\mathbf{n},\mathbf{r}) = \cos\pi = -1$$

This is achieved in the following way.

The surface S designated S_2 now encloses a spherical wavefront surface S_1 centre O. S_1 and S_2 are said to be doubly connected and the surface integral now includes S_1 and S_2 (Figure A2.2). At S_1 the normal **n** to dS on S_2 now points towards O and if the outer surface of S_2 is allowed to expand to infinity its contribution to the integral becomes zero. This leaves only the integral over the surface where S_1 and S_2 coincide. The singularity E_0/r at O is excluded from the integral.

If P is now located on **r**, at P', that is in the direction of backward propagation of Huygens wavelets, then

$$\cos(\mathbf{n},\mathbf{R}) = \cos\pi = -1$$

and

$$\cos(\mathbf{n},\mathbf{r}) = \cos\pi = -1$$

$K(\chi)$ is then equal to zero. Any other position for P gives

$$K(\chi) = \frac{\cos\chi - \cos(\mathbf{n},\mathbf{r})}{2} = \frac{1+\cos\chi}{2}$$

Appendix 3: Non-Linear Schrödinger Equation

This equation describes phenomena in non-linear media with strong dispersion. It appears in several forms. For optical soliton purposes, Mollenauer *et al.* (1982) derive it from the equation

$$\mathrm{i}\left(\frac{\partial u}{\partial z} + k_1 \frac{\partial u}{\partial t}\right) = \frac{-k_2}{2}\frac{\partial^2 u}{\partial t^2} + \gamma |u|^2 u \tag{A3.1}$$

where

$$k_1 = \frac{\partial k}{\partial \omega}, \qquad k_2 = \frac{\partial^2 k}{\partial \omega^2}, \quad \text{and} \quad \gamma = \frac{1}{2} k_0 \frac{n_2}{n_0}$$

n_2 and n_0 appear in the Kerr Optical Equation $n - n_0 = n_2 I$.

Equation (A3.1) is satisfied by a pulse of the form

$$E(z,t) = u(z,t)\,\mathrm{e}^{\mathrm{i}(\omega_0 t - k_0 z)}$$

Using the transformation of Mollenauer *et al.* (1980), (A3.1) assumes the dimensionless form

$$-\mathrm{i}\frac{\partial v}{\partial \xi} = \frac{1}{2}\frac{\partial^2 v}{\partial s^2} + |v|^2 v \tag{A3.2}$$

which has a soliton solution $u(\xi, s) = \operatorname{sech}(s)\mathrm{e}^{i\xi/2}$ where

$$s = T^{-1}(t - k_1 z) \qquad \xi = |k_2| T^{-2} z$$

and

$$v = T\left(\frac{\gamma}{|k_2|}\right)^{1/2} u$$

where T is a measure of the width of the input optical pulse.

The first term on the right hand of equation (A3.2) describes the effects of dispersion which may be seen as the kinetic energy term in the linear Schrödinger equation, while the second term corresponds to the energy of a self-trapping potential proportional to $|u|^2$ arising from the non-linear refractive index which may be interpreted in probability terms.

The Physics of Vibrations and Waves, 6th Edition H. J. Pain