This was installed on 6/6/15, do not edit here!

6. The Boltzmann factor in Statistical Mechanics

6.1 Statement of The Boltzmann Extremum Problem

Imagine a system which has m **state levels** into which particles can be placed. At a given energy level ε_i, the number of available **states** is g_i, known as the **degeneracy** at energy ε_i. A particular state of the entire system of particles is characterized by the number of particles N_i in each of the states, so system state = $(N_1,N_2...N_m)$ = **N**, a vector. The total number of particles in the system is fixed at $M = \Sigma_i N_i$, and the total energy is $U = \Sigma_i \varepsilon_i N_i$. The system is isolated so its initial energy U does not change, nor can particles be created or destroyed, so U and M are both fixed constants. The system of particles (perhaps a box of gas atoms) is assumed to be at thermal equilibrium so all its macroscopic characteristics (like pressure and temperature) are stable.

Imagine a billion of these particle systems (an "ensemble"). Each system settles into its own partition set $\{N_i\}$ = **N**. If M is very large, one will find that those billion **N** vectors are all rather similar. This is so because certain $\{N_i\}$ partitions are statistically favored over other partitions just because there are a lot more "states" available to a system with those favored $\{N_i\}$ partitions.

So, how many states are available ("accessible") to such a system characterized by vector **N** which represents some partitioning of the particles $\{N_i\}$?

The number of different "ways" of placing N indistinguishable balls into g boxes is a fascinating elementary problem and the answer is $\binom{N+g-1}{N}$, a binomial coefficient.

> Footnote: One arrives at this answer by thinking of N balls and g-1 "partitions" between boxes, all of which have to be laid out in a row, such as | * * * | * | | * for the case N = 5 and g-1 = 4. Here the leftmost bin is empty, as is the second bin from the right. Each layout of the 9 objects is characterized by picking a committee of 5 balls from the 9 positions (or a committee of 4 partitions). So in this case the number of unique layouts is $\binom{9}{5} = \frac{9!}{5!4!} = \binom{9}{4}$. (6.1.1)

The count $\binom{N+g-1}{N}$ allows multiple balls to be in any given box. This count is then associated with so-called **Bose-Einstein statistics** -- multiple bosons are allowed in the same state (for example, photons or any integral-spin particles). If at most one ball is allowed in any given box , the count of ways is just the number of ways to pick a committee of N boxes from the set of g boxes, and this count is $\binom{g}{N}$. This count is then associated with so-called **Fermi-Dirac statistics** -- only one fermion is allowed in a state (for example electrons or any half-integral spin particles). The fact that electrons are fermions is the reason the periodic table of elements exists.

However, we shall assume that $g_i >> N_i$ (known as the Boltzmann limit), and this allows a simplification of the state count. Both the Bose and Fermi state counts reduce to the same count in this limit:

$$\text{ways} = \binom{N+g-1}{N} \approx \binom{g}{N} = \frac{g!}{N!\,(g-N)!} = \frac{g(g-1)..(g-N+1)}{N!} \approx \frac{g^N}{N!} \ . \tag{6.1.2}$$

So in our particle problem, the number of ways of having N_1 particles in states with energy ε_1 is $\frac{g_1{}^N}{N_1!}$. For each of these ways, we have $\frac{g_2{}^N}{N_2!}$ ways of putting N_2 particles in energy level ε_2. And so on. Thus, the total number of available states ("microstates") for a partition $\{N_i\}$ of the M particles is given by

$$\Omega(N_1,N_2...N_m) = \frac{g_1{}^{N_1}}{N_1!}\frac{g_2{}^{N_2}}{N_2!} \ ....\frac{g_m{}^{N_m}}{N_m!} \ . \tag{6.1.3}$$

So here is our initial **extremum problem**:

Find **N** which maximizes $\Omega(\mathbf{N}) = \frac{g_1{}^{N_1}}{N_1!}\frac{g_2{}^{N_2}}{N_2!} \ ....\frac{g_m{}^{N_m}}{N_m!}$ subject to these two constraints:

$$\Sigma_i N_i = M \qquad \text{and} \qquad \Sigma_i \varepsilon_i N_i = U \ . \tag{6.1.4}$$

When M is a very large number (like Avogadro's number), it turns out that $\Omega(\mathbf{N})$ has a very strong and sharp maximum at a certain **N** which is the solution value **N** for this extremum problem. Thus, when one examines an ensemble of such systems, this solution **N** is the overwhelmingly likely partitioning of the particles into $\{N_i\}$. Our task is then to solve this problem for **N**.

This vector $\mathbf{N} = (N_1, N_2...N_m)$ is the vector $\mathbf{r} = (x_1, x_2...x_N)$ of our general Lagrange Multiplier presentation, so in this application the number of components in the vector is N = m, and we don't want to confuse this previous use of symbol N with **N** or N_i of the present problem!

One might observe that the x_i in the general analysis were reals, whereas the N_i here are integers. We can dispense with this issue by simply writing $N_i! = \Gamma(N_i+1)$ which then serves to interpolate N_i between its integer values. Another approach is to replace N_i everywhere in the above problem statement by $n_i \equiv N_i/M$. Since M and the solution N_i values are very large integers, one can regard this n_i as essentially a continuous real (albeit rational) variable.

It should be noted that the value of U is restricted to a certain range.

$$M\varepsilon_{min} \ \le \ U \ \le \ M\varepsilon_{max} \tag{6.1.5}$$

The limits represent the two extreme possible partitions $\{N_i\}$ of the particles (all in lowest energy state, or all in highest energy state). It is useful to define the average energy of a particle,

$u \equiv U/M$ = average energy of particle in the system (6.1.6)

and then (6.1.5) says,

$$\varepsilon_{min} \le u \le \varepsilon_{max} \quad (6.1.7)$$

6.2 Solution of The Boltzmann Extremum Problem

Maximizing the state count Ω of (6.2) is the same as maximizing $f \equiv \ln\Omega$, and from (6.1.3),

$$f(N_1, N_2...N_m) \equiv \ln\Omega(N_1,N_2...N_m) = \Sigma_i [N_i \ln g_i] - \Sigma_i [\ln N_i!] \quad (6.2.1)$$

where Σ_i means $\Sigma_{i=1}^{m}$. Assuming all the N_i are very large numbers, we may use Stirling's formula,

$$x! \approx \sqrt{2\pi x}\, x^x e^{-x} \quad \Rightarrow \quad \ln x! \approx \ln\sqrt{2\pi} + (1/2)\ln x + x\ln x - x \approx x\ln x - x,\ x >> 1 \ . \quad (6.2.2)$$

Then $\ln N_i! \approx N_i \ln N_i - N_i$ so that,

$$f = \ln\Omega = \Sigma_i [N_i \ln g_i] - \Sigma_i [\ln N_i!] \approx \Sigma_i [N_i \ln g_i] - \Sigma_i [N_i \ln N_i - N_i]$$

$$= \Sigma_i [(1+\ln g_i)N_i - N_i \ln N_i] = \Sigma_i N_i [1 + \ln(g_i/N_i)] \ . \quad (6.2.3)$$

Here then is a restatement of our extremum problem:

Find **N** which maximizes $f(\mathbf{N}) = \Sigma_i [(1+\ln g_i)N_i - N_i \ln N_i]$ subject to these two constraints:

$a(\mathbf{N}) = \Sigma_i N_i - M = 0$

$b(\mathbf{N}) = \Sigma_i \varepsilon_i N_i - U = 0$. (6.2.4)

Following now the prescription of Section 2, we write our "Lagrangian" H as

$$H(\mathbf{N},\lambda) \equiv f(\mathbf{N}) + \lambda_1 a(\mathbf{N}) + \lambda_2 b(\mathbf{N})$$

$$= \Sigma_i [(1+\ln g_i)N_i - N_i \ln N_i] + \lambda_1 [\Sigma_i N_i - M] + \lambda_2 [\Sigma_i \varepsilon_i N_i - U] \ . \quad (6.2.5)$$

The next step (b) is to compute the m derivatives with respect to N_i :

$$f_i = \partial f/\partial N_i = (1+\ln g_i) - N_i (1/N_i) - \ln N_i = \ln g_i - \ln N_i = \ln(g_i/N_i)$$

$$a_i = \partial a/\partial N_i = \partial_i [\Sigma_j N_j - M] = 1$$

$$b_i = \partial b/\partial N_i = \partial_i [\Sigma_j \varepsilon_j N_j - U] = \varepsilon_i \ . \quad (6.2.6)$$

In step (c) we set the partials of H to zero, so

$$0 = H_i(\mathbf{N},\boldsymbol{\lambda}) = f_i + \lambda_1 a_i + \lambda_2 b_i = \ln(g_i/N_i) + \lambda_1 + \lambda_2 \varepsilon_i \quad i = 1,2...m$$

or

$$\ln(g_i/N_i) + \lambda_1 + \lambda_2 \varepsilon_i = 0 \qquad i = 1,2...m$$

or

$$\ln(N_i/g_i) = \lambda_1 + \lambda_2 \varepsilon_i \; . \qquad i = 1,2...m$$

For step (d) we just write down the set of equations we have to solve (6.2.7)

$$\ln(N_i/g_i) = \lambda_1 + \lambda_2 \varepsilon_i \qquad i = 1,2...m$$

$$\Sigma_i N_i = M \qquad // \; a(\mathbf{N}) = 0$$

$$\Sigma_i \varepsilon_i N_i = U \qquad // \; b(\mathbf{N}) = 0 \; . \qquad (6.2.8)$$

We have a system of m+2 equations in m+2 unknowns which are the N_i, λ_1 and λ_2. To solve these equations, we first solve the first set of m equations for N_i, leaving λ_1 and λ_2 as unknowns:

$$\ln(N_i/g_i) = \lambda_1 + \lambda_2 \varepsilon_i$$

or

$$N_i = g_i \, e^{\lambda_1} \, e^{\lambda_2 \varepsilon_i} \; . \qquad (6.2.9)$$

Already a **major result** has emerged from our Lagrange Multiplier analysis! The population N_i of states at energy level ε_i depends *exponentially* on ε_i. We intuitively expect higher energy levels to be less populated than lower ones, so we expect λ_2 to be negative. For this reason, and following tradition, we set $\lambda_2 = -\beta$ where $\beta > 0$. And while we're at it, we can set $e^{\lambda_1} = A$ to simplify notation. So

$$\beta \equiv -\lambda_2$$

$$A \equiv e^{\lambda_1} \qquad (6.2.10)$$

and then (6.2.9) reads,

$$N_i = A \; g_i \, e^{-\beta \varepsilon_i} \; . \qquad (6.2.11)$$

A study of the statistical mechanics of thermal equilibrium in fact shows that

$$\beta = 1/kT \qquad (6.2.12)$$

where T is **absolute temperature** and k is **Boltzmann's constant**. The factor $e^{-\beta \varepsilon_i}$ in (6.2.11) is referred to as **the Boltzmann factor**. The idea that $\beta = 1/kT$ is really the *definition* of absolute temperature, but these matters are not of immediate interest, and we continue onward.

From (6.2.11) the probability of a particle having energy ε_i is given by

$$p_i = \frac{N_i}{M} = \frac{A}{M}\ g_i\, e^{-\beta\varepsilon_i} \ . \qquad (6.2.13)$$

Since $\Sigma_i p_i = 1$ one finds that $1 = \frac{A}{M}\ [\Sigma_i g_i\, e^{-\beta\varepsilon_i}]$ which then determines the derived Lagrange multiplier constant A (expressed in terms of β),

$$A = \frac{M}{\Sigma_i g_i\, e^{-\beta\varepsilon_i}} \ . \qquad (6.2.14)$$

The denominator $Z \equiv \Sigma_i g_i e^{-\beta\varepsilon_i}$ is called the **partition function** for the system.

Using (6.2.11) for N_i, the two constraint equations shown in (6.2.4) [$\Sigma_i N_i = M$ and $\Sigma_i N_i \varepsilon_i = U$] read

$$\begin{aligned} A\,\Sigma_i g_i e^{-\beta\varepsilon_i} &= M \\ A\,\Sigma_i \varepsilon_i g\, e^{-\beta\varepsilon_i} &= U \ . \end{aligned} \qquad (6.2.15)$$

Dividing these two equations and using (6.1.6) that u = U/M, one gets this self-consistent result,

$$u = \frac{U}{M} = \frac{\Sigma_i g_i \varepsilon_i e^{-\beta\varepsilon_i}}{\Sigma_i g_i e^{-\beta\varepsilon_i}} = \frac{(A/M)\Sigma_i g_i \varepsilon_i e^{-\beta\varepsilon_i}}{(A/M)\Sigma_i g_i e^{-\beta\varepsilon_i}} = \frac{\Sigma_i p_i \varepsilon_i}{1} = \Sigma_i p_i \varepsilon_i = \langle\varepsilon_i\rangle \ . \qquad (6.2.16)$$

To find the derived Lagrange multiplier constant $\beta = -\lambda_2$ we write out the second equation of (6.2.15),

$$\Sigma_i g_i \varepsilon_i\, e^{-\beta\varepsilon_i} = U/A = (U/M)(M/A) = (u)\ (\Sigma_i g_i\, e^{-\beta\varepsilon_i})$$

or

$$\Sigma_i\, [\ \varepsilon_i - u\ ]\ g_i\ e^{-\beta\varepsilon_i} = 0 \qquad (6.2.17)$$

so this is the equation one must solve for β. Recall that ε_i, g_i, and u all known quantities. For the lowest energy states one will have $\varepsilon_i < u$ (recall that $u = \langle\varepsilon_j\rangle$) so $[\varepsilon_i - u]$ will be negative. For high energy states $[\ \varepsilon_i - u\ ]$ will be positive, so that is why (6.2.17) at least stands a chance of having a solution for β. As noted in (6.2.12), the solution for β indicates the temperature of the system of particles. Changing u changes the solution β, so one concludes that the mean particle energy u is a function of temperature.

For general values of ε_i (6.2.17) is a transcendental equation which must be solved numerically for β. If it happens that all the ε_i are ratios of integers, the equation can be written as a polynomial equation, but since those integers are likely to be large, the order of this polynomial is large and again one must do a numerical solution.

Notice that the solution to our particle system problem has the same general form regardless of the number m of energy levels ε_i. If we take m very large and make the energy levels ε_i be very closely spaced, they approach a continuum of energy values and the conclusions apply to a classical system. For small finite m, the discrete energy levels indicate a quantum mechanical system.

To summarize, we have used the method of Lagrange multipliers to solve the Boltzmann extremum problem stated in either (6.1.4) or (6.2.4). Since the problem has two constraints, there are two Lagrange multipliers λ_1 and λ_2 which we replaced with the numbers $A \equiv e^{\lambda_1}$ and $\beta \equiv -\lambda_2$. We showed how β is determined by numerically solving (6.2.17), and then A is determined by (6.2.14). Then the solution (extremum) vector $\mathbf{N} = \{N_i\}$ is given by (6.2.11) which says $N_i = A\ g_i\ e^{-\beta\varepsilon_i}$.

What we have not shown is that this extremum solution is a maximum. nor have we shown the fact that $\Omega(\mathbf{N})$ has a very sharp peak at the solution value **N**, which then statistically forces a system to assume a value **N** very close to this solution **N**. We shall deal with these questions in the following section.

6.3 More details of the Boltzmann problem solution

Start with

$$f = \ln\Omega = \Sigma_i\,[(1+\ln g_i)N_i - N_i \ln N_i\,] = \Sigma_i\, N_i\,[1 + \ln(g_i/N_i)]\,. \qquad (6.2.3)$$

The constraints $\Sigma_i N_i = M$ and $\Sigma_i N_i \varepsilon_i = U$ can be regarded as two equation in the two unknowns N_1 and N_2 which are easily solved to obtain,

$$\begin{aligned} N_1 &= +(\varepsilon_2-\varepsilon_1)^{-1}[\,\varepsilon_2 M - U + \Sigma_{i=3}^{m}\,(\varepsilon_i-\varepsilon_2)N_i] \;\; = N_1(N_3,\ldots N_m) \\ N_2 &= -\,(\varepsilon_2-\varepsilon_1)^{-1}[\,\varepsilon_1 M - U + \Sigma_{i=3}^{m}\,(\varepsilon_i-\varepsilon_1)N_i] \;\; = N_2(N_3,\ldots N_m)\;. \end{aligned} \qquad (6.3.1)$$

Notice that, for i = 3,4...m,

$$\begin{aligned} \frac{\partial N_1}{\partial N_i} &= +(\varepsilon_2-\varepsilon_1)^{-1}\,[\,(\varepsilon_i-\varepsilon_2)\,] = \frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} \\ \frac{\partial N_2}{\partial N_i} &= -\,(\varepsilon_2-\varepsilon_1)^{-1}\,[\,(\varepsilon_i-\varepsilon_1)\,] = -\,\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\quad . \end{aligned} \qquad (6.3.2)$$

Now think of f as function of $N_3,,,,N_m$,

$$f(N_1(N_3,\ldots N_m),\, N_2(N_3,\ldots N_m),\, N_3,\, ...N_m)\;. \qquad (6.3.3)$$

Compute the total derivative df/dN_i for i = 3,4...n:

$$\begin{aligned} \frac{df}{dN_i} &= f_1\frac{\partial N_1}{\partial N_i} + f_2\frac{\partial N_2}{\partial N_i} + f_i \qquad\qquad \text{where } f_i \equiv \frac{\partial f}{\partial N_i} \\ &= f_1\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - f_2\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + f_i \qquad\qquad (6.3.2) \\ &= \ln(g_1/N_1)\,\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \ln(g_2/N_2)\,\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + \ln(g_i/N_i) \qquad (6.2.6) \end{aligned} \qquad (6.3.4)$$

Installing the solution values $\ln(N_i/g_i) = \lambda_1 + \lambda_2\varepsilon_i$ shown in (6.2.8) this becomes

$$= (\lambda_1 + \lambda_2\varepsilon_1)\,\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - (\lambda_1 + \lambda_2\varepsilon_2)\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + (\lambda_1 + \lambda_2\varepsilon_i)$$

$$= \lambda_1 \left[\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + 1\right] + \lambda_2 \left[\varepsilon_1 \frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \varepsilon_2 \frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + \varepsilon_i \right]$$

$$= \lambda_1 \left[\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + \frac{\varepsilon_2-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right] + \lambda_2 \left[\varepsilon_1 \frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \varepsilon_2 \frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + \varepsilon_i \frac{\varepsilon_2-\varepsilon_1}{\varepsilon_2-\varepsilon_1} \right]$$

$$= \lambda_1 [0] + \lambda_2[0] = 0 \qquad (6.3.5)$$

This is the expected result since the solution is supposed to be a critical point of f given in (6.3.3).

Our main interest is in the second derivative d^2f/dN_i^2 . If this comes out negative at the solution point, we know we have a maximum. Applying d/dN_j to the second line in (6.3.5),

$$\frac{d^2f}{d^2N_i} = \frac{df_1}{dN_i}\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - \frac{df_2}{dN_i}\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + \frac{df_i}{dN_i}$$

$$= \left[f_{11}\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - f_{12}\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + f_{1i}\right]\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} \qquad f_{ij} \equiv \frac{\partial^2 f}{\partial N_i \partial N_j}$$

$$- \left[f_{21}\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - f_{22}\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + f_{2i}\right]\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1}$$

$$+ \left[f_{i1}\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1} - f_{i2}\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1} + f_{ii}\right] \qquad i = 3,4....m \qquad (6.3.6)$$

Recall now that

$$f = \ln\Omega = \Sigma_i\,[(1+\ln g_i)N_i - N_i \ln N_i\,] \qquad \textit{(6.2.3)}$$

$$f_i = \ln g_i - \ln N_i \qquad \textit{(6.2.6)}$$

$$f_{ij} = -\,\delta_{ij}(1/N_i) \qquad (6.3.7)$$

Therefore (6.3.6) simplifies to

$$\frac{d^2f}{d^2N_i} = f_{11}\left[\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1}\right]^2 + f_{22}\left[\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right]^2 + f_{ii}$$

$$= -\left\{\frac{1}{N_1}\left[\frac{\varepsilon_i-\varepsilon_2}{\varepsilon_2-\varepsilon_1}\right]^2 + \frac{1}{N_2}\left[\frac{\varepsilon_i-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right]^2 + \frac{1}{N_i}\right\} \qquad (6.3.8)$$

This says that the curvature of f in all independent directions N_i for i = 3,4...m is negative and this is true for *all* vectors **N** = {N_i}. This is of course then true for the solution vector **N**. But for the solution vector **N** we also have $df/dN_i = 0$ from (6.3.5) and therefore the solution is a *maximum* of f and therefore of Ω.

Finally, we may *estimate* the width of the peak of f and then of Ω. The energy fractions in (6.3.8) are on the ballpark order of unity and the N_i are on the order of M, the total number of particles. If we do Taylor expansion of f(**N**) in the i direction only (i = 3,4...m) about the solution **N**, then

$$f(N_i + dN_i) \approx f(N_i) + (df/dN_i)dN_i + (1/2)(d^2f/d^2N_i)\,(dN_i)^2 \tag{6.3.9}$$

But at the solution point $(df/dN_i) = 0$ so we find

$$df \approx (1/2)(d^2f/d^2N_i)\,(dN_i)^2 \tag{6.3.10}$$

so

$$(dN_i)^2 \approx \frac{2\,|df|}{|d^2f/d^2N_i|} \tag{6.3.11}$$

Suppose we are interested in finding dN_i such that the Ω(**N**) drops to half its peak value. Then

$$\Omega_{half}/\Omega_{peak} = 1/2 \quad \Rightarrow \quad \ln \Omega_{half} - \ln \Omega_{peak} = 1/2 \quad \Rightarrow$$

$$f_{half} - f_{peak} = \ln(1/2) \quad \Rightarrow \quad |df| = f_{peak} - f_{half} = \ln(2) = 0.7 \tag{6.3.12}$$

In this case, the half width of the Ω plot is

$$dN_i = \sqrt{\frac{1.4}{|d^2f/d^2N_i|}} \tag{6.3.13}$$

For a ballpark estimate, we can interpret (6.3.8) as saying

$$\frac{d^2f}{d^2N_i} \sim -3/M \quad \Rightarrow \quad dN_i \approx \sqrt{\frac{1.4\,M}{3}} \sim \sqrt{M} \tag{6.3.14}$$

so the fractional half width is roughly.

$$\frac{dN_i}{M} \sim \sqrt{\frac{1}{M}} \tag{6.3.15}$$

For a macroscopic box of gas particles, one might have $M = 10^{20}$ and then $\frac{dN_i}{M} \sim 10^{-10}$. Since this is the fractional half-width of the peak in Ω, the peak is *extremely* sharp. This then is why, in an ensemble of macroscopic particle systems, all systems will have almost the same **N** value.

6.4 A numerical Boltzmann problem example

For this example we choose three energy levels (m = 3) so the state count is given by

$$\Omega(N_1,N_2,N_3) = \frac{g_1^{N_1}}{N_1!}\frac{g_2^{N_2}}{N_2!}\frac{g_3^{N_3}}{N_3!} \qquad \textit{(6.1.3)} \qquad (6.4.1)$$

and correspondingly (this uses the Stirling approximation),

$$f = \ln\Omega = \Sigma_{i=1}^{3} N_i\,[1 + \ln(g_i/N_i)]\ . \qquad \textit{(6.2.3)} \qquad (6.4.2)$$

Assume the three energy levels are ordered in this manner, where ε_1 is the lowest energy state,

$$\varepsilon_1 < \varepsilon_2 < \varepsilon_3\ . \qquad \text{and recall} \qquad u = U/M \quad \text{and } \varepsilon_1 \le u \le \varepsilon_3 \qquad (6.4.3)$$

We regard both N_1 and N_2 as functions of N_3 where, according to (6.3.1),

$$\begin{aligned} N_1 &= +(\varepsilon_2-\varepsilon_1)^{-1}[\,\varepsilon_2 M - U + (\varepsilon_3-\varepsilon_2)N_3] &&= N_1(N_3) \\ N_2 &= -(\varepsilon_2-\varepsilon_1)^{-1}[\,\varepsilon_1 M - U + (\varepsilon_3-\varepsilon_1)N_3] &&= N_2(N_3)\ . \end{aligned} \qquad \textit{(6.3.1)} \qquad (6.4.4)$$

It seems clear that one must have

$$\begin{aligned} &0 \le N_1 \le M \\ &0 \le N_2 \le M. \end{aligned} \qquad (6.4.5)$$

Using the expressions (6.4.4) in these inequalities and making use of (6.4.3) and the fact that $\varepsilon_1 \le u \le \varepsilon_3$, one finds that all four inequalities can be summarized as just two inequalities which we write as,

$$\begin{aligned} &N_{3min} \le N_3 \le N_{3max} \\ &\quad N_{3min} = \max\left(\frac{u-\varepsilon_2}{\varepsilon_3-\varepsilon_2}M,\ 0\right) \\ &\quad N_{3max} = \min\left(\frac{u-\varepsilon_1}{\varepsilon_3-\varepsilon_1}M,\ M\right) = \frac{u-\varepsilon_1}{\varepsilon_3-\varepsilon_1}M \qquad \text{// since } u \le \varepsilon_3 \end{aligned} \qquad (6.4.6)$$

To find the Lagrange multiplier β we must solve (6.2.17),

$$\Sigma_i\,[\,\varepsilon_i - u\,]\,g_i\ e^{-\beta\varepsilon_i} = 0\ . \qquad \textit{(6.2.17)} \qquad (6.4.7)$$

Then the other derived multiplier is then given by (6.2.14)

$$A = \frac{M}{\Sigma_i g_i\ e^{-\beta\varepsilon_i}} = \frac{M}{Z} \quad \text{where } Z \equiv \Sigma_i g_i e^{-\beta\varepsilon_i}\ . \qquad \textit{(6.2.14)} \qquad (6.4.8)$$

The components N_i of the solution vector are then given by

$$N_i = A\ g_i\ e^{-\beta\varepsilon_i}\ . \qquad (6.2.11) \qquad (6.4.9)$$

The curvature of f at its peak is given by

$$\frac{d^2f}{d^2N_3} = -\left\{ \frac{1}{N_1}\left[\frac{\varepsilon_3-\varepsilon_2}{\varepsilon_2-\varepsilon_1}\right]^2 + \frac{1}{N_2}\left[\frac{\varepsilon_3-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\right]^2 + \frac{1}{N_3} \right\} \qquad (6.3.8) \qquad (6.4.10)$$

The half-width of the Ω peak is then given by

$$dN_i = \sqrt{\frac{1.4}{|d^2f/d^2N_i|}} \qquad (6.3.13) \qquad (6.4.11)$$

For our specific Maple plots, we shall assume

$\varepsilon_3 = 3$ M = 1000 = number of particles
$\varepsilon_2 = 2$ u = 1.5 = average energy
$\varepsilon_1 = 1$ $g_1 = g_2 = g_3 = 5000$ = degeneracy (6.4.12)

Here then is our Maple code. We first enter Ω,

```
restart; with(plots):
Omega := product(g[i]^N[i]/(N[i]!),i=1..3);
```

$$\Omega := \frac{g_1^{N_1}\ g_2^{N_2}\ g_3^{N_3}}{N_1!\ N_2!\ N_3!}$$

The quantities N_1 and N_2 are then replaced as in (6.4.4), [we use e_i in place of ε_i]

```
N[1] := (e[2]*M-U+(e[3]-e[2])*N[3])/(e[2]-e[1]);
```

$$N_1 := \frac{e_2 M - U + (e_3 - e_2) N_3}{e_2 - e_1}$$

```
N[2] := -(e[1]*M-U+(e[3]-e[1])*N[3])/(e[2]-e[1]);
```

$$N_2 := -\frac{e_1 M - U + (e_3 - e_1) N_3}{e_2 - e_1}$$

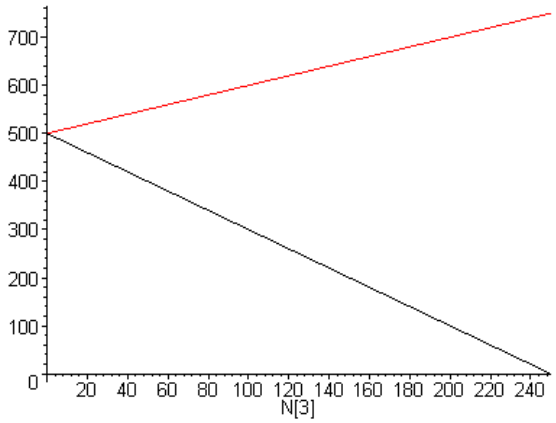
We next enter most of the data shown above and use it to compute N_{3min} and N_{3max}

```
e[1] := 1: e[2] := 2: e[3] := 3: u := 1.5: M := 1000: U := M*u:
N3min := max(0,M*(u-e[2])/(e[3]-e[2]));
N3max := M*(u-e[1])/(e[3]-e[1]);
```

$$N3min := 0$$
$$N3max := 250.0000000$$

so the legal range of N_3 is (0,250). Here is a plot of N_1 (red) and N_2(black) versus N_3 :

```
plot([N[1],N[2]],N[3] = N3min..N3max, color = [red,black]);
```



(6.4.13)

We then enter the degeneracies and take a look at Ω in its numerical form,

```
g[1] := 5000: g[2] := 5000: g[3] := 5000:
Omega;
```

$$\frac{5000^{(500.0+N_3)}\ 5000^{(500.0-2N_3)}\ 5000^{N_3}}{(500.0+N_3)!\,(500.0-2N_3)!\,N_3!} \tag{6.4.14}$$

As is typical in statistical mechanics, at its peak the value of Ω is a *very* large number. Even with our relatively small number of particles M = 1000 and degeneracy g = 5000, the peak value is (see below)

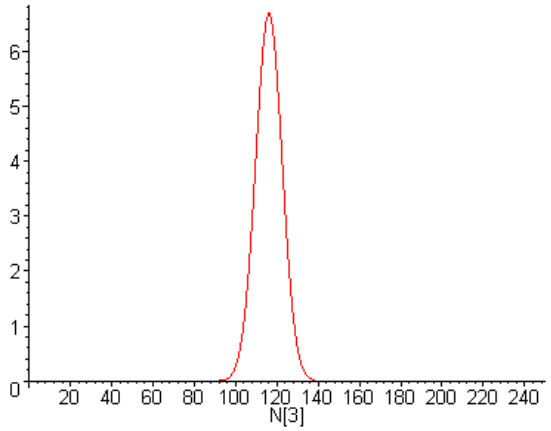
$$\Omega_{peak} = 6.71 \times 10^{1519} \tag{6.4.15}$$

Maple is uncomfortable plotting numbers larger than about 10^{40} so we pre-scale Ω down by 10^{1519} to make the scaled Ω be Maple-digestible. This scaling process in turn creates numbers smaller than 10^{-40} which are also rejected by the Maple plotter, so we use a Heaviside function to pin small numbers to 0 unless they are greater than the arbitrary value .01. Here then is a plot of $\Omega(N_3)$,

```
scale := 1e1519;
```

$$scale := .1\ 10^{1520}$$

```
OmegaScaled := (Omega/scale)*Heaviside(Omega/scale-.01):
xx := plot(  OmegaScaled,N[3] = N3min..N3max,numpoints=100  ):
display(xx);
```

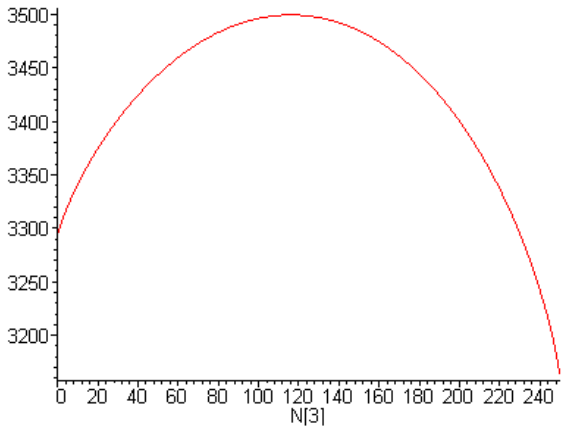


(6.4.16)

We broke the display command into two pieces: first create xx then plot xx. Replacing : with ; after the plot command gives one a list of numbers to be plotted, and from this list one can determine a reasonable scale factor. Once that is found, the plotting coordinates need not be displayed. The curve $\Omega(N_3)$ has the typical Gaussian (normal) shape which no doubt the energetic reader could show results from the central limit theorem.

A plot of $f = \ln(\Omega)$ has a much smoother and broader shape,

```
plot(ln(Omega),N[3] = N3min..N3max,numpoints=100 );
```



(6.4.17)

Recall from (6.3.8) that $\frac{d^2 f}{d^2 N_i}$ is everywhere negative, consistent with the cupping down seen in this plot.

Next, we have Maple solve (6.2.17) for the Lagrange multiplier β,

```
eq := sum((e[i]-u)*g[i]*exp(-beta*e[i]),i=1..3) = 0;
```

$$eq := -2500.\, e^{(-1.\,\beta)} + 2500.\, e^{(-2.\,\beta)} + 7500.\, e^{(-3.\,\beta)} = 0$$

```
s := solve(eq,beta);
```

$$s := .2644970943 - 3.141592654\, I,\ .8341151943$$

```
beta := s[2];
```

$$\beta := .8341151943$$

With $w = e^{-\beta}$ the above equation is $3w^2+w-1 = 0$ which has solutions $w = (-1 \pm \sqrt{13})/6$. The minus sign gives the complex β noted, while the plus sign gives the physical solution $\beta = - \ln[(-1 + \sqrt{13})/6] = .834$.

From this β value we compute the other Lagrange parameter A from (6.2.14)

```
Z := sum(g[i]*exp(-beta*e[i]),i=1..3);
```

$$Z := 99083.26914$$

```
A := M/Z;
```

$$A := .01009252126$$

Once these parameters are known, we use (6.2.11) to display the solution vector $\mathbf{N} = \{N_i\}$.

```
for i in [3,2,1] do N[i] := A*g[i]*exp(-beta*e[i]) od;
```

$$N_3 := 116.2040605$$

$$N_2 := 267.5918793$$

$$N_1 := 616.2040607$$

Since we set $u < \varepsilon_2$, we see here a normal distribution where higher energy states have fewer particles.

Our final task is to use (6.3.8) and (6.3.13) to determine the half-width of the Ω peak,

```
curve := (1/N[1])*((e[3]-e[2])/(e[2]-e[1]))^2 + (1/N[2])*((e[3]-e[1])/(e[2]-e[1]))^2 + 1/N[3]
```

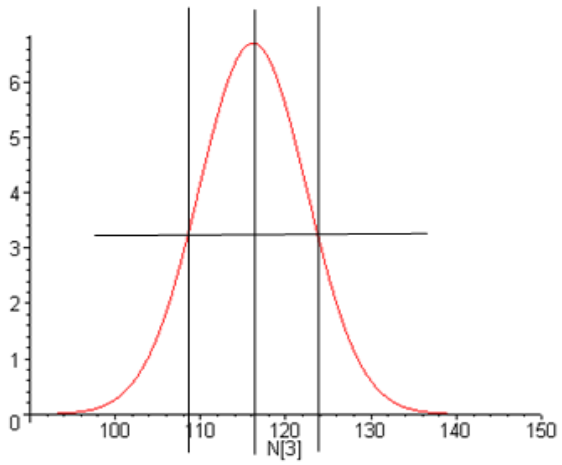
$$curve := .02517652703$$

```
dN3 := sqrt(1.4/curve);
```

$$dN3 := 7.457033725$$

which says $\Delta N \approx 7.5$. This width and the solution $N_3 = 616.2$ can be verified from this blowup plot of Ω,

```
xx := plot(  OmegaScaled,N[3] = 90..150,numpoints=100  ):
display(xx);
```



Notice that this Maple code does not use the Stirling approximation and provides a continuous interpolation of N! as Γ(N+1).

The peak value of Ω is found to be

```
evalf(Omega);
```

$$.6710572393\ 10^{1520}$$

In the above example we used u = 1.5 < ε_2 = 2. If we instead use u = 2.5 > ε_2 = 2, we find that the state populations are as shown above but in reverse order. Such a inverted situation corresponds to a negative absolute temperature T and cannot therefore arise in thermal equilibrium. It could arise in a laser medium which is driven by some power source.