An Attempt at the Boltzmann Problem PhL 6.2..15

Here is the problem as stated in my blackbody doc (but perhaps this is not really the blackbody problem)

$$\Omega(N_1,N_2...N_m) = \frac{g_1^{N_1}}{N_1!}\frac{g_2^{N_2}}{N_2!} \cdots \frac{g_m^{N_m}}{N_m!}$$ = number of "accessible states" for the full set of particles

$\Sigma_{i=1}^{m} N_i = M$ // number of particles is a constant (so not photons I guess)

$\Sigma_{i=1}^{m} N_i \varepsilon_i = U$ // total energy is a constant (adiabatic I guess)

Here there are a finite set m of discrete energy states called ε_i. Each state has N_i particles in it. Each state has degeneracy g_i which only effects Ω. You can I guess have N_i larger than g_i if you want, though I think typically $N_i < g_i$. The total number of particles is M, so you cannot create or destroy particles. The total energy of the particles is U.

So this is a classic Lagrange Multiplier problem! I will try to solve it using my new LM document. Let's restate the problem using $N_i \rightarrow x_i$ to fit into my format

$$f(x_1,x_2...x_m) = \frac{g_1^{x_1}}{x_1!}\frac{g_2^{x_2}}{x_2!} \cdots \frac{g_m^{x_m}}{x_m!}$$

$\Sigma_{i=1}^{m} x_i - M = 0$ $a(x_1,x_2...x_m) = 0$
$\Sigma_{i=1}^{m} x_i \varepsilon_i - U = 0$ $b(x_1,x_2...x_m) = 0$

Well, right away we have a distinction with the classical LM case. Here the x_i are all positive integers, not real numbers! But I guess I can "fill in" using the gamma function and $x! = \Gamma(x+1)$.

Then

$$\partial_x (x!) = \partial_x \Gamma(x+1) = [\partial_y \Gamma(y)]^{y=x+1} = ?$$

From GR p 902 we know that

8.36 The psi function $\psi(x)$

8.360 Definition:

1. $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$

This says

$\psi(x) = \partial_x \ln(\Gamma(x)) = 1/\Gamma * \partial_x \Gamma(x)$ $\partial_x \Gamma(x) = \Gamma(x)\psi(x)$

so this gets us involved with "the psi function" which is also known as the "digamma function".

Start with

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

$$\ln\Omega = (N_1 \ln g_1 - \ln N_1!) + (N_2 \ln g_2 - \ln N_2!) + \ .... + (N_m \ln g_m - \ln N_m!)$$

$$= \Sigma_{i=1}^{m} [N_i \ln g_i] - \Sigma_{i=1}^{m} [\ln N_i!]$$

Perhaps $\ln\Omega$ is an easier function to maximize. Stirling says $\ln x! = x \ln x - x.$, so let's assume that all the N_i are very large numbers and then

$$\ln N_i! \approx N_i \ln N_i - N_i.$$

Then we find

$$f \equiv \ln\Omega \approx \Sigma_{i=1}^{m} [N_i \ln g_i] - \Sigma_{i=1}^{m} [N_i \ln N_i - N_i]$$

$$= \Sigma_{i=1}^{m} [N_i \ln \frac{g_i}{N_i}] + \Sigma_{i=1}^{m} N_i$$

$$= \Sigma_{i=1}^{m} [N_i \ln \frac{g_i}{N_i}] + M \qquad \text{// Zemansky (10-6)}$$

and this certainly is a simpler functional form.

Now when we compute the partials f_i only one term is picked out, so we have

$$f_i = \partial_i [N_i \ln \frac{g_i}{N_i}] = ?$$

Simplify for the moment and write

$$\partial_x [x \ln \frac{g}{x}] = \partial_x [x \ln g - x \ln x] = \ln g - \partial_x [x \ln x] = \ln g - \{ x * 1/x + \ln x \}$$

$$= \ln g - \{1 + \ln x \} = \ln (g/x) - 1$$

which is a very simple result! Then

$$f_i = \partial_i [N_i \ln \frac{g_i}{N_i}] = \ln \frac{g_i}{N_i} - 1$$

This may be wrong because I have used the constraint $\Sigma_{i=1}^{m} N_i = M$, but f_i should be computed without applying any of the constraints. Thus, the correct f_i will be different from what I show here. Let's go start a new doc boltz shot 2.

So now we restate our LM problem assuming we are only interested in regions where all N_i are very large:

$$f(N_1,N_2...N_m) = \Sigma_{i=1}^{m} [\, N_i \ln \frac{g_i}{N_i} \,] + M \qquad \text{// maximize f}$$
$$a(N_1,N_2...N_m) = \Sigma_{i=1}^{m} N_i - M = 0$$
$$b(N_1,N_2...N_m) = \Sigma_{i=1}^{m} N_i\, \varepsilon_i - U = 0$$

or in our usual variable names

$$f(x_1,x_2...x_m) = \Sigma_{i=1}^{m} [\, x_i \ln \frac{g_i}{x_i} \,] + M \qquad \text{// maximize f}$$
$$a(x_1,x_2...x_m) = \Sigma_{i=1}^{m} x_i - M = 0$$
$$b(x_1,x_2...x_m) = \Sigma_{i=1}^{m} x_i\, \varepsilon_i - U = 0$$

We know at once that

$$f_i = \ln \frac{g_i}{x_i} - 1 = \ln(g_i/x_i) - 1$$
$$a_i = 1$$
$$b_i = \varepsilon_i$$

That seems to make a relatively simple R matrix:

$$R = \begin{matrix} \ln(g_1/x_1) - 1 & \ln(g_2/x_2) - 1 & ..... & \ln(g_m/x_m) - 1 \\ 1 & 1 & ..... & 1 \\ \varepsilon_1 & \varepsilon_2 & ..... & \varepsilon_m \end{matrix}$$

Now follow the algorithm and see if it leads anywhere useful:

$$H(\mathbf{r},\boldsymbol{\lambda}) \equiv f(\mathbf{r}) + \lambda_1 a(\mathbf{r}) + \lambda_2\, b(\mathbf{r}) + ... + \lambda_{s-1}\, q(\mathbf{r}) \;. \qquad (2.1)$$

or

$$H(\mathbf{r},\boldsymbol{\lambda}) \equiv \Sigma_{i=1}^{m} [\, x_i \ln \frac{g_i}{x_i} \,] + M + \lambda_1[\, \Sigma_{i=1}^{m} x_i - M] + \lambda_2[\, \Sigma_{i=1}^{m} x_i\, \varepsilon_i - U]$$

Next,.

$$H_i = f_i + \lambda_1 a_i + \lambda_2\, b_i + ... + \lambda_{s-1}\, q_i \qquad i = 1,2....N \qquad (2.2)$$

becomes

$$H_i = \ln(g_i/x_i) - 1 + \lambda_1 + \lambda_2 \varepsilon_i$$

Now set all partials to zero, so we then have

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

Here then is the set of equations we have to solve:

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

$$\Sigma_{i=1}^{m} x_i = M$$
$$\Sigma_{i=1}^{m} x_i \varepsilon_i = U$$

There are m+2 equations in m+2 unknowns. Let's try to find candidate solutions for the λ_i as described in my LM doc: [not the right path to take here, but here it is anyway]

$$-\begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_N \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & \dots & q_1 \\ a_2 & b_2 & \dots & q_2 \\ \dots & \dots & \dots & \dots \\ a_N & b_N & \dots & q_N \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_{S-1} \end{pmatrix} \tag{2.7}$$

$$-\begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_m \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ \dots & \dots \\ \dots & \dots \\ a_m & b_m \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_{S-1} \end{pmatrix} \tag{2.7}$$

Now I pick various choices of 2x2 from this. For example

$$\begin{pmatrix} f_i \\ f_j \end{pmatrix} = \begin{pmatrix} a_i & b_i \\ a_j & b_j \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

or

$$\begin{pmatrix} f_i \\ f_j \end{pmatrix} = \begin{pmatrix} 1 & \varepsilon_i \\ 1 & \varepsilon_j \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & \varepsilon_i \\ 1 & \varepsilon_j \end{pmatrix}^{-1} \begin{pmatrix} f_i \\ f_j \end{pmatrix}$$

Maple computes the inverse matrix:

```
M := matrix(2,2,[1,ei,1,ej]);
```

$$M := \begin{bmatrix} 1 & ei \\ 1 & ej \end{bmatrix}$$

```
inverse(M);
```

$$\begin{bmatrix} -\frac{ej}{-ej+ei} & \frac{ei}{-ej+ei} \\ \frac{1}{-ej+ei} & -\frac{1}{-ej+ei} \end{bmatrix}$$

.

Thus,

$$M^{-1} = \frac{1}{\varepsilon_i - \varepsilon_j} \begin{pmatrix} -\varepsilon_j & \varepsilon_i \\ 1 & -1 \end{pmatrix}$$

and then

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{\varepsilon_i - \varepsilon_j} \begin{pmatrix} -\varepsilon_j & \varepsilon_i \\ 1 & -1 \end{pmatrix} \begin{pmatrix} f_i \\ f_j \end{pmatrix}$$

and this says

$$\lambda_1 = \frac{1}{\varepsilon_i - \varepsilon_j} \left[-\varepsilon_j f_i + \varepsilon_i f_j \right]$$

$$\lambda_2 = \frac{1}{\varepsilon_i - \varepsilon_j} \left[f_i - f_j \right]$$

Not bad. Now

$$f_i - f_j = \left[\ln(g_i/x_i) - 1 \right] - \left[\ln(g_j/x_j) - 1 \right] = \left[\ln(g_i/x_i) - \ln(g_j/x_j) \right]$$

And we can write all out as

$$\lambda_1 = \frac{1}{\varepsilon_i - \varepsilon_j} \left\{ -\varepsilon_j \left[\ln(g_i/x_i) - 1 \right] + \varepsilon_i \left[\ln(g_j/x_j) - 1 \right] \right\}$$

$$= \frac{1}{\varepsilon_i - \varepsilon_j} \left\{ (\varepsilon_j - \varepsilon_i) - \varepsilon_j \ln(g_i/x_i) + \varepsilon_i \ln(g_j/x_j) \right\}$$

$$= -1 + \frac{1}{\varepsilon_i - \varepsilon_j} \left[\varepsilon_i \ln(g_j/x_j) - \varepsilon_j \ln(g_i/x_i) \right]$$

$$\lambda_2 = \frac{1}{\varepsilon_i - \varepsilon_j} \left[\ln(g_i/x_i) - \ln(g_j/x_j) \right]$$

I have no idea whether this will lead anywhere, but it is a great test of my LM doc.

Now back to the equation set. I will mark the choice i,j values as I,J, so

$$\ln(g_i/x_i) - 1 + \lambda_1 + \lambda_2 \varepsilon_i = 0 \qquad i = 1,2....m$$
$$\Sigma_{i=1}^{m} x_i = M$$
$$\Sigma_{i=1}^{m} x_i \varepsilon_i = U$$

becomes

$$\ln(g_i/x_i) - 2 + \frac{1}{\varepsilon_I - \varepsilon_J} \{\varepsilon_I \ln(g_J/x_J) - \varepsilon_J \ln(g_J/x_I)\} + \frac{1}{\varepsilon_I - \varepsilon_J} [\ln(g_I/x_I) - \ln(g_J/x_J)] \varepsilon_i = 0$$
$$i = 1,2,....m$$

$$\Sigma_{i=1}^{m} x_i = M$$
$$\Sigma_{i=1}^{m} x_i \varepsilon_i = U$$

So I have eliminated the λ_1 and λ_2. ALL candidate solutions must solve these equations where you make a selection of I≠J from 1...m.

Now let's back up a little to

$$\ln(g_i/x_i) - 1 + \lambda_1 + \lambda_2 \varepsilon_i = 0 \qquad i = 1,2....m$$
$$\Sigma_{i=1}^{m} x_i = M$$
$$\Sigma_{i=1}^{m} x_i \varepsilon_i = U$$

Remember that the ε_i are givens, not unknowns. So try to solve the first equation for x_i

$$\ln(g_i) - \ln(x_i) - 1 + \lambda_1 + \lambda_2 \varepsilon_i = 0$$

$$\ln(x_i) = \ln(g_i) - 1 + \lambda_1 + \lambda_2 \varepsilon_i$$

$$\ln(x_i/g_i) = -1 + \lambda_1 + \lambda_2 \varepsilon_i$$

$$x_i = g_i \exp(-1 + \lambda_1 + \lambda_2 \varepsilon_i) = g_i e^{-1} e^{-\lambda_1} e^{-\lambda_2 \varepsilon_i}$$

Once again, I have just solved the first m equations to find

$$x_i = g_i e^{-(\lambda_1 + 1)} e^{-\lambda_2 \varepsilon_i}$$

and for the very first time I am seeing something looking like a "Boltzmann energy exponential"!

OK, then we are left with these 2 equations in the two unknowns λ_1 and λ_2 :

$$\Sigma_{i=1}^{m}\ g_i\, e^{-(\lambda_1+1)}e^{-\lambda_2\varepsilon_i} = M$$

$$\Sigma_{i=1}^{m}\ g_i\varepsilon_i\, e^{-(\lambda_1+1)}e^{-\lambda_2\varepsilon_i} = U$$

or

$$\Sigma_{i=1}^{m}\ g_i e^{-\lambda_2\varepsilon_i} = M\, e^{(\lambda_1+1)}$$

$$\Sigma_{i=1}^{m}\ \varepsilon_i g_i e^{-\lambda_2\varepsilon_i} = U e^{(\lambda_1+1)}$$

Divide these equations to get

$$[\Sigma_{i=1}^{m}\ \varepsilon_i g_i e^{-\lambda_2\varepsilon_i}] / [\Sigma_{i=1}^{m}\ g_i e^{-\lambda_2\varepsilon_i}] = U/M = u = \text{average energy per particle}$$

or

$$\frac{\Sigma_{i=1}^{m}\ \varepsilon_i g_i e^{-\lambda_2\varepsilon_i}}{\Sigma_{i=1}^{m}\ g_i e^{-\lambda_2\varepsilon_i}} = u$$

VERY interesting! Let's now define

$p_i \equiv K\, g_i e^{-\lambda_2\varepsilon_i}$ = probability that a particle is in state i. K = not yet known

Then the above states

$$\frac{\Sigma_i \varepsilon_i p_i}{\Sigma_i p_i} = u$$

and this makes total sense with that interpretation. This sort of came out of nowhere!

Now solve for λ_2

$$[\Sigma_{i=1}^{m}\ \varepsilon_i g_i e^{-\lambda_2\varepsilon_i}] / [\Sigma_{i=1}^{m}\ g_i e^{-\lambda_2\varepsilon_i}] = U/M = u$$

$$\Sigma_{i=1}^{m}\ \varepsilon_i g_i e^{-\lambda_2\varepsilon_i} = u\, \Sigma_{i=1}^{m}\ g_i e^{-\lambda_2\varepsilon_i}$$

$$\Sigma_{i=1}^{m} (\varepsilon_i - u) g_i e^{-\lambda_2\varepsilon_i} = 0$$

Here the ε_i are known, u is U/M which is also known, g_i is known. So how solve for λ_2?

$\Sigma_i\ a_i \exp(-x b_i) = 0$ solve for x

Define $z = e^{-x}$ so this then says

$$F(z) = \Sigma_i\ a_i z^{b_i} = 0$$

What kind of equation is this? A polynomial with real instead of integer coefficients. I see the term "exponential polynomials" bandied about and I found one downloadable paper. I think for general b_i this is a transcendental equation. If the b_i are rational, you have a polynomial equation. Go back to

$$\Sigma_i\, a_i \exp(-x b_i) = 0 \qquad \text{solve for x}$$

Write $b_i = N_i/M_i$ and find some large $J = \text{lcm}(M_i)$ so that $J = K_i M_i$. Then

$$\exp(-x b_i) = \exp(-x N_i/M_i) = \exp(-x K_i N_i/J) = \exp(-K_i N_i\,[x/J]\,) = (e^{-x/J})^{K_i N_i}$$

Now let $z \equiv e^{-x/J}$ so you then have

$$\Sigma_i\, a_i\, z^{K_i N_i} = 0$$

and then you have converted your exponential polynomial equation into a normal polynomial equation. In this case, the solution z is an algebraic function according to the wiki definition, since z is then the solution of a polynomial equation with coefficients which are polynomials (here just constants a_i). If the b_i are not rational, like $b_2 = \pi$, then it is a transcendental equation.

In any event, generally higher order polynomials have no analytic solutions for roots. at least no simple ones, and an example is given in this clip:

Algebraic function

From Wikipedia, the free encyclopedia

This article is about algebraic functions in calculus, mathematical analysis, and abstract algebra. For functions in elementary algebra, see function (mathematics).

In mathematics, an **algebraic function** is a function that can be defined as the root of a polynomial equation. Quite often algebraic functions can be expressed using a finite number of terms, involving only the algebraic operations addition, subtraction, multiplication, division, and raising to a fractional power:

$$f(x) = 1/x, f(x) = \sqrt{x}, f(x) = \frac{\sqrt{1+x^3}}{x^{3/7} - \sqrt{7}x^{1/3}}$$

are typical examples.

However, some algebraic functions cannot be expressed by such finite expressions (as proven by Galois and Niels Abel), as it is for example the case of the function defined by

$$f(x)^5 + f(x)^4 + x = 0.$$

In more precise terms, an algebraic function of degree *n* in one variable *x* is a function $y = f(x)$ that satisfies a polynomial equation

$$a_n(x)y^n + a_{n-1}(x)y^{n-1} + \cdots + a_0(x) = 0$$

where the coefficients $a_i(x)$ are polynomial functions of *x*, with coefficients belonging to a set *S*. Quite often, $S = \mathbb{Q}$, and one then talks

So the upshot of this discussion is that we wanted to find λ_2 from

$$\Sigma_{i=1}^{m} (\varepsilon_i - u) g_i e^{-\lambda_2 \varepsilon_i} = 0$$

and in general this needs a numeric solution. For completely general ε_i values, it is a transcendental equation, while for rational ε_i values it is a polynomial equation. So I won't be able to analytically solve for the value λ_2. There might not even BE a solution, but I think in my case there is.

Next question: Assuming we have a solution for λ_2, can we solve for λ_1 ? Our equations were

$$(1/M)\,\Sigma_{i=1}^{m}\, g_i e^{-\lambda_2 \varepsilon_i} = e^{(\lambda_1+1)} = (1/M)\,\Sigma_i p_i$$
$$(1/U)\,\Sigma_{i=1}^{m}\, \varepsilon_i g_i e^{-\lambda_2 \varepsilon_i} = e^{(\lambda_1+1)} = (1/U)\,\Sigma_i p_i \varepsilon_i$$

The first equation says

$$\Sigma_i p_i = M\, e^{(\lambda_1+1)}$$

and this then gives a solution value for λ_1 if you know λ_2. The second equation they just says

$$\frac{\Sigma_i \varepsilon_i p_i}{\Sigma_i p_i} = u$$

and so we can in fact satisfy both.

Question: How does this relate do the supposed solution values of λ_1 and λ_2 I found using my LM document method, as outlined above:

$$\lambda_1 = -1 + \frac{1}{\varepsilon_i - \varepsilon_j}\,[\varepsilon_i \ln(g_j/x_j) - \varepsilon_j \ln(g_i/x_i)]$$
$$\lambda_2 = \frac{1}{\varepsilon_i - \varepsilon_j}\,[\ln(g_i/x_i) - \ln(g_j/x_j)]$$

These are candidate solution pairs. But they are expressed in terms of x_i and x_j which are unknowns of the problem, so I guess there is no conflict.

OK, I want now to continue on this general path to see where I have gotten!

Question: How do I know I have found a maximum?

$$f(x_1, x_2 \ldots x_m) = \Sigma_{i=1}^{m}\,[\, x_i \ln \frac{g_i}{x_i}\,] + M$$

I have NO IDEA how to show this is a max. I know that the solution I found is

$$\ln(x_i/g_i) = -1 + \lambda_1 + \lambda_2\, \varepsilon_i$$

I guess I could install that to get

$$f_{\mathbf{extrememum}}(x_1,x_2...x_m) = \Sigma_{i=1}^{m} [- x_i \ \ln(x_i/g_i)] \ + \ M$$

$$= - \Sigma_i \ x_i \ \ln(x_i/g_i) + M$$

$$= - \Sigma_i \ x_i \ [- 1 + \lambda_1 \ + \lambda_2 \ \varepsilon_i] + M$$

$$= - \Sigma_i \ x_i \ [- 1 + \lambda_1 \ + \lambda_2 \ \varepsilon_i] + \Sigma_i x_i$$

$$= - \Sigma_i \ x_i \ [- 2 + \lambda_1 \ + \lambda_2 \ \varepsilon_i]$$

$$= \ \Sigma_i \ x_i \ [2 - \lambda_1 \ - \lambda_2 \ \varepsilon_i]$$

$$= (2-\lambda_1)\Sigma_i \ x_i - \lambda_2 \ \Sigma_i \ x_i\varepsilon_i$$

$$= \ (2-\lambda_1)M - \lambda_2 U$$

Well, maybe this is what you do. Somehow go to a small neighborhood of your solution, construct a legal dr, and see if $\nabla f \bullet d\mathbf{r}$ is positive or negative! A legal d**r** has to respect the two constraints,

$$\Sigma_i x_i \ \ = M$$
$$\Sigma x_i \ \varepsilon_i \ = U$$

If m = 3, we have $x_1 + x_2 + x_3 = M$ and this is a plane in E^3.

Well OK, go back to

$a(x_1,x_2...x_m) = \Sigma_i x_i - M = 0$	$\nabla a = (1,1...1)$	m components
$b(x_1,x_2...x_m) = \Sigma_i x_i \ \varepsilon_i - \ U = 0$	$\nabla b = (\varepsilon_1, \varepsilon_2....\varepsilon_m)$	m components

There are the normals to the constraint planes. Any legal d**r** must be perp to these two normals. So need

$d\mathbf{r} \bullet \nabla a = 0$	or	$\Sigma dx_i = 0$	$dM = 0$
$d\mathbf{r} \bullet \nabla b = 0$	or	$\Sigma \varepsilon_i dx_i = 0$	$dU = 0$

Let's try a Taylor expansion around the solution point

$$f(x_1,x_2...x_m) \ = \Sigma_i \ [\ x_i \ \ln \frac{g_i}{x_i} \ ] \ + \ M$$

$$\partial f/\partial x_i = f_i = \ \ln \frac{g_i}{x_i} - 1 \qquad \text{// first derivatives}$$

$$\partial^2 f/\partial x_i^2 = - 1/x_i$$

$$\partial^2 f/\partial x_i \partial x_j = -(1/x_i)\,\delta_{i,j}$$

The surface is "cupping down" in all directions. This is just like my hemisphere example in LM doc, it just happens. That is encouraging. Is that all I really need to know?