

Information

The Maxwell, Bose and Fermi statistics

Rastko Vukovic, math prof.

January 20, 2015

We are looking for Shannon's information of the three main energy distributions, classical Maxwell-Boltzmann and the two quantum, Bose-Einstein and Fermi-Dirac. This is the continuation of the article Statistics.

Introduction

The Density functions for the three basic statistics in today's physics can be written in the form:

$$\rho_1(x) = \frac{\alpha_1}{e^x}, \quad (1)$$

$$\rho_2(x) = \frac{\alpha_2}{e^x - 1}, \quad (2)$$

$$\rho_3(x) = \frac{\alpha_3}{e^x + 1}, \quad (3)$$

where the constants alpha yet to be determined. The first should be the Maxwell-Boltzmann classical statistics, the second and third are quantum, the Bose-Einstein and Fermi-Dirac statistics, respectively red, green and blue graph on Figure 1, for the unit values of alpha.

In all three cases $e = 2.71828$ is the base of natural logarithm, and can be substituted $x = E/kT$, where E is the energy state in which is the given particle, $k = 1.38 \times 10^{-23}$ J/K is Boltzmann constant and T temperature in Kelvin.

Generally, let Ω is space of continuum random events. If the density of the probability of the events is $\rho \geq 0$, then the integral equality is true:

$$\int_{\Omega} \rho(\tau) d\tau = 1. \quad (4)$$

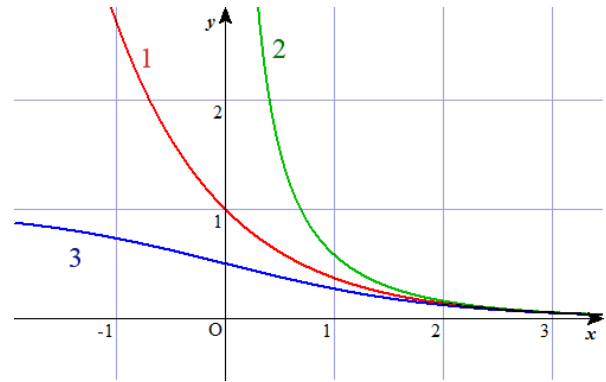


Figure 1: Graphs $\rho_{1,2,3}$ in red, green, blue.

The probability of a random event $S \subset \Omega$ is:

$$\Pr[S] = \int_S \rho(\tau) d\tau. \quad (5)$$

Hartley information is the logarithm

$$\mathcal{I}(\tau) = -\ln \rho(\tau), \quad (6)$$

but Shannon's information is the mean value

$$\bar{\mathcal{I}} = -\int_{\Omega} \rho(\tau) \mathcal{I} d\tau. \quad (7)$$

These are definitions from pure mathematics.

In Quantum Mechanics we use Dirac bra-ket notation. Consistently, the average value (7) should be wroted $\langle \mathcal{I} \rangle$ instead $\bar{\mathcal{I}}$.

The integrals (1-3) are easy to find:

$$\begin{cases} \int \rho_1 dx = -\alpha_1 e^{-x} + C, \\ \int \rho_2 dx = \alpha_2 \ln(1 - e^{-x}) + C, \\ \int \rho_3 dx = -\alpha_3 \ln(1 + e^{-x}) + C. \end{cases} \quad (8)$$

Even easier is the check by direct derivation.

1 Maxwell-Boltzmann

The space Ω of Maxwell-Boltzmann random events is abscissa $x = E/kT$ from 0 to infinity, for

$$\int_0^{\infty} e^{-x} dx = 1. \quad (9)$$

So $\alpha_1 = 1$. The Hartley information (6) is

$$\mathcal{I}_1 = \frac{E}{kT}, \quad (10)$$

for $x = E/kT$.

Hartley uncertainty or the information \mathcal{I}_1 of the state is proportional to the state of the energy E and is inversely proportional to the temperature T of the thermodynamic system.

The Shannon information is:

$$\begin{aligned} \bar{\mathcal{I}}_1 &= - \int_0^{\infty} e^{-x} \ln e^{-x} dx = \\ &= \int_0^{\infty} x e^{-x} dx = - \int_0^{\infty} x d e^{-x} \\ &= -x e^{-x} \Big|_0^{\infty} + \int_0^{\infty} e^{-x} dx = 1, \\ \bar{\mathcal{I}}_1 &= 1. \end{aligned} \quad (11)$$

It is 1 in nats.

For example, the average information 1 in bits has a coin toss:

$$\bar{\mathcal{I}} = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1.$$

Otherwise $\ln x = \ln 2 \log_2 x$, so 1 nat is 1.4427 bit.

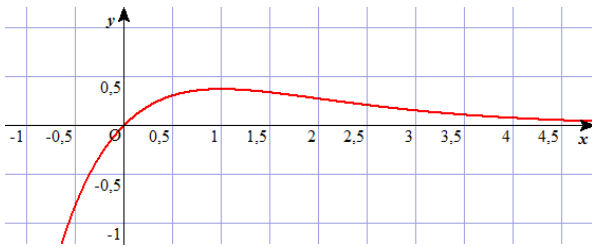


Figure 2: Function $y_1(x) = -\rho_1(x) \ln \rho_1(x)$.

The graph of the function $y_1 = -\rho_1 \ln \rho_1$ is on Figure 2. Shannon information is the area below the abscissa in the first quadrant.

Now, let us understand the meaning of the density (1). The Maxwell-Boltzmann formula

$$P_1(\epsilon) = \frac{1}{Z} g(\epsilon) e^{-\frac{\epsilon}{kT}} \quad (12)$$

tells about particles distribution between the energy levels of the system. The energy is temporarily marked by ϵ instead E for ease of understanding the following text.

Particle has energy in the small range ϵ to $\epsilon + d\epsilon$ with corresponding probability energy $P_1(\epsilon)d\epsilon$. The number of energy states between ϵ and $\epsilon+d\epsilon$ is $g(\epsilon)d\epsilon$ and $Z = Z(T)$ is a function of temperature, called *partition function* that normalizes the probability:

$$\int_0^{\infty} P_1(\epsilon) d\epsilon = 1. \quad (13)$$

This distribution is the base of many properties of materials; it helps us to understand thermodynamics with temperature and entropy.

A *macrostate* of a system makes the average over the microscopic constituencies of the system, such are pressure, volume or temperature of a gas. A *microstate* of a system are the properties of each of the constituent particles, such are the position and momentum of each of the molecules in a sample of gas.

Many different microstates can correspond to a single macrostate and the distribution (12) tells us about energies of the particles in the system. Unlike the continuous ranges of energy, let us consider systems with discrete energy levels which is more accurate in case of quantization.

Example 1. Find the energy levels of harmonic operators.

Solution. The allowed energies for a quantum system are the solutions of Schrodingers equation:

$$\hat{H}\psi = \epsilon\psi \quad (14)$$

where ψ is the wave function of the system, ϵ is an allowed energy, and \hat{H} is the Hamiltonian operator. In one dimension, with the potential

$$V(x) = \frac{1}{2} kx^2$$

for an harmonic oscillator the equation (14) becomes:

$$\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \left(\epsilon - \frac{1}{2} kx^2 \right) \psi = 0. \quad (15)$$

Where \hbar is reduced Plank constant, m is mass in rest of the wave-particle. Solving the equation we find the claimed energy levels:

$$\epsilon_j = \left(j + \frac{1}{2} \right) \hbar\omega, \quad (16)$$

where $j = 0, 1, 2, \dots$, and:

$$\omega = \sqrt{\frac{k}{m}}. \quad (17)$$

So the energy levels are quantized. \square

Assume that particles can exchange energy, but interactions between them are very weak. So the sum of the energy levels of a corresponding single-particle system can be treated as the energy level of the system, or states of energy of the system.

For the case of Maxwell-Boltzmann statistics, the particles are considered distinguishable and can occupy the same quantum state.

For example, in the case of the two particles A and B and two states $s = 2$ we can have the following 4 dispositions:

$$\{AB, \emptyset\}, \{A, B\}, \{B, A\}, \{\emptyset, AB\}. \quad (18)$$

In the case of two particles and three states ($s = 3$) we can have:

$$\begin{aligned} &\{AB, \emptyset, \emptyset\}, \{A, B, \emptyset\}, \{A, \emptyset, B\}, \\ &\{B, A, \emptyset\}, \{\emptyset, AB, \emptyset\}, \{\emptyset, A, B\}, \\ &\{B, \emptyset, A\}, \{\emptyset, B, A\}, \{\emptyset, \emptyset, AB\}. \end{aligned} \quad (19)$$

That is 9 dispositions.

In the case of the two particles and $s = 1, 2, 3, \dots$ states, we have s^2 dispositions. Namely, the first particle may be in any of the states, which makes the s possibilities, the same number as for another particle.

Generally, in the case of $N = 1, 2, 3, \dots$ particles and $s = 1, 2, 3, \dots$ states the Maxwell-Boltzmann statistics allows

$$D_1 = s^N = e^{N \ln s} \quad (20)$$

dispositions. For, each of N particles can be in any of s states.

The probability of one disposition is D_1^{-1} . The question is what the probability of the particularly level of energy is, expressed by a function of the energy or by the number of particles in the level? Let us consider this problem from another side.

A microstate is defined by the energies of the particles in the system. Let us group the microstates into distributions, so the same distribution has the same number of particles in each of the allowed energy levels. On the Figure 3 we can see the only four dispositions of the three particles with zero energy and one particle with energy 4ϵ (the second row in the next table).

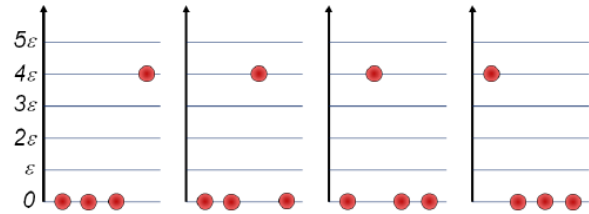


Figure 3: *Distribution of microstates.*

Similarly, to have the total energy 4ϵ we can have two particles with zero energy one particle with energy ϵ and one particle with energy 3ϵ , which can be achieved in 12 dispositions. That is the third row of the table:

	0	ϵ	2ϵ	3ϵ	4ϵ	5ϵ	\rightarrow	No
3	0	0	0	1	0	0	\rightarrow	4
2	1	0	1	0	0	0	\rightarrow	12
2	0	2	0	0	0	0	\rightarrow	6
1	2	1	0	0	0	0	\rightarrow	12
0	4	0	0	0	0	0	\rightarrow	1

Two particle of energy zero and two of energy 2ϵ can be deployed in 6 ways (fourth row). One of energy zero, two with ϵ and one 2ϵ can be in 12 dispositions. Finally, all four particles with energy ϵ have only one disposition. That is total 35 dispositions of 4 particles with the sum energies 4ϵ .

A system in thermal equilibrium in a given macrostate may be found in any of the allowed microstates with equal probabilities. This is the principle of equal a priori probabilities. Hence the average number of particles with zero energy is:

$$\bar{n}_0 = 3 \cdot \frac{4}{35} + 2 \cdot \frac{12}{35} + 2 \cdot \frac{6}{35} + 1 \cdot \frac{12}{35} + 0 \cdot \frac{1}{35},$$

$$\bar{n}_0 = \frac{12}{7} = 1.71429 \dots$$

Similarly, we find the average numbers \bar{n}_j of the particles with energy $j\epsilon$:

$$\begin{aligned} \bar{n}_0 &\approx 1.71 & \bar{n}_1 &\approx 1.14 & \bar{n}_2 &\approx 0.69 \\ \bar{n}_3 &\approx 0.34 & \bar{n}_4 &\approx 0.11 & \bar{n}_5 &= 0. \end{aligned}$$

When these average numbers presented graphically, with abscissa j and ordinate n_j , it reveals approximately exponential decrease¹.

There are $N!$ ways of arranging N particles into N specified single-particle states. The n_0 particles with zero energy can be arranged in $n_0!$ different ways, the n_1 particles with energy ϵ in $n_1!$ ways, and so

¹see examples on page: <http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/disbol.html>

on. Hence, the number of microstates within a given distribution is:

$$\mathcal{D}_1 = \frac{N!}{n_0!n_1!n_2!\dots n_s!}$$

To find the most probable distribution, we have to find the set of populations $\{n_j\}$ that maximizes \mathcal{D}_1 .

Since N is very large, usually of order 10^{23} , we can use Stirling's approximation:

$$\ln N! = N \ln N - N.$$

Assuming that each of the numbers n_j is also a large, we get:

$$\ln \mathcal{D}_1 = N \ln N - N - \sum_{j=0}^s (n_j \ln n_j - n_j). \quad (21)$$

The number N is constant with variations in n_j , so we have:

$$d(\ln \mathcal{D}_1) = - \sum_{j=0}^s \ln n_j dn_j. \quad (22)$$

The \mathcal{D}_1 has a maximum (with respect to changes in the population of each of the energy levels), when $d(\mathcal{D}_1) = 0$ or

$$d(\ln \mathcal{D}_1) = 0.$$

Thereby the total numbers of particles and the total energy E must remain constant:

$$dN = \sum_{j=0}^s dn_j = 0, \quad dE = \sum_{j=0}^s \epsilon_j dn_j = 0. \quad (23)$$

Therefore we introduce parameters α and β to be determined, and write:

$$- \sum_{j=0}^s (\ln n_j - \alpha - \beta \epsilon_j) dn_j = 0,$$

$$\ln n_j - \alpha - \beta \epsilon_j = 0,$$

$$n_j = e^\alpha e^{\beta \epsilon_j}. \quad (24)$$

This is the Maxwell-Boltzmann distribution.

The constants α, β are determined by (23):

$$\sum_{j=0}^s n_j = N, \quad \sum_{j=0}^s n_j \epsilon_j = E. \quad (25)$$

We have:

$$\sum_{j=0}^s n_j = \sum_{j=0}^s e^\alpha e^{\beta \epsilon_j} = N,$$

$$e^\alpha = \frac{N}{\sum_j e^{\beta \epsilon_j}} = \frac{N}{Z},$$

with Z the partition function. The value of β is determined by the total energy of the system. It comes to zero when the energy increases.

Since the most of the microstates accessible to a system occur in the most probable distribution, the total number of accessible microstates is approximately \mathcal{D}_1 . Using (21) now we find $\ln \mathcal{D}_1$ is

$$\approx N \ln N - N - \sum_j \left(\frac{N}{Z} \ln \left(\frac{N}{Z} \right) e^{\beta \epsilon_j} + \frac{N}{Z} \beta \epsilon_j e^{\beta \epsilon_j} \right),$$

or using

$$\sum_j n_j \epsilon_j = \sum_j \frac{N}{Z} \epsilon_j e^{\beta \epsilon_j} = E,$$

we get

$$\ln \mathcal{D}_1 \approx N \ln Z - \beta E.$$

Because $N = \text{constant}$ and Z is independent of the total energy, we have:

$$\frac{\partial \mathcal{D}_1}{\partial E} \approx -\beta. \quad (26)$$

The β is the rate at which the number of accessible microstates increases with the energy of the system.

In thermodynamics is known the relationship between internal energy E , entropy S and volume V :

$$dE = T dS - p dV, \quad (27)$$

Where T is (thermodynamic) temperature and p is the pressure. If the volume is fixed, $dV = 0$, than we have $dE = T dS$, or

$$\frac{\partial S}{\partial E} = \frac{1}{T}. \quad (28)$$

Routinely here it is combined (20) with (26) to relate entropy with the number of accessible microstates:

$$S = k \ln \mathcal{D}_1, \quad (29)$$

and the distribution parameter β to the thermodynamic temperature:

$$\beta = -\frac{1}{kT}, \quad (30)$$

where k is the Boltzmann constant, with units of joules per kelvin (J/K).

The last equations are the connection of statistical mechanics and thermodynamic. On the other hand, we can see that the entropy (29) has the same form, and the similar meaning as the Hartley information.

The Maxwell-Boltzmann distribution applies to the classical *ideal gas*, which is composed of many randomly moving point particles that do not interact except when they collide elastically.

If the energy ϵ in Maxwell-Boltzmann distribution (12) is replaced by the kinetic energy $mv^2/2$ of the particle of the gas, than we have probability function of (average) speed v of particles. That distribution is Gaussian.

Eventually, the third form of the Maxwell-Boltzmann distribution is the average number n_j/N of particles per single-particle state, which is:

$$\bar{n}_j = e^{-(\epsilon_j - \mu)/kT}, \quad (31)$$

where ϵ_j is the simple particle energy, μ the chemical potential. The exponent is so big that the ∓ 1 in Bose-Einstein and Fermi-Dirac distributions can be ignored.

2 Bose-Einstein

For the case of Bose-Einstein statistics, the two particles are considered to be indistinguishable but the both can occupy the same quantum state. Denote both particles with A.

If we have the two states ($s = 2$), than we have the following 3 dispositions:

$$\{AA, \emptyset\}, \{A, A\}, \{\emptyset, AA\}. \quad (32)$$

If we have three states ($s = 3$), we have 6 dispositions:

$$\begin{aligned} \{AA, \emptyset, \emptyset\}, \{A, A, \emptyset\}, \{A, \emptyset, A\}, \\ \{\emptyset, AA, \emptyset\}, \{\emptyset, A, A\}, \{\emptyset, \emptyset, AA\}. \end{aligned} \quad (33)$$

If we have $s = 1, 2, 3, \dots$ states, we have

$$\binom{s+1}{2} = \frac{1}{2}s(s+1) \quad (34)$$

dispositions.

Namely, the first particle can be in the 1st state with second particle that can be in any states from 1st to s ; it is s states. Then, the first particle can be in the 2nd state with the second particle in any states from 2nd to s , which is $s - 1$ states more. And so on, the first particle can be in the n -th state with the second in any of n -th to s states, which is $s - n$ additional. Prior 1 to $n - 1$ states were already counted. Finally, we have the sum total of dispositions, the arithmetic sum:

$$s + (s - 1) + \dots + 1 = \frac{1}{2}s(s + 1).$$

This is the number of combinations (34).

Another way to understand the same result (two particles, s states) is as follows. To avoid duplication of particles in the same state, fictitious, imagine that we have one more, the reserve position. If we choose two particles in the same position, then one particle is detached to the reserve position. So we have $s + 1$ positions, each for a maximum one particle. If the reserve position is full, it means that in one of the regular position we have two particles. The number of such situations is exactly s . The total number of dispositions is the number of combinations (34).

This kind of distributions we have in the Bose-Einstein gas, in a collection of photons, or in the vibrations of particles in solid. Generally, it is the distribution of the *bosons*, the integer spin particles.

Let us again, consider an energy level ϵ_j with degeneracy g_j , containing n_j bosons. Imagine we have lined up all the n_j particles in g_j sections separated by $g_j - 1$ borders. The number of distinct orderings of borders and particles is:

$$\mathcal{D}_2(j) = \binom{n_j + g_j - 1}{n_j} = \frac{(n_j + g_j - 1)!}{n_j!(g_j - 1)!}. \quad (35)$$

Hence, the total distribution has

$$\mathcal{D}_2 = \prod_j \frac{(n_j + g_j - 1)!}{n_j!(g_j - 1)!} \quad (36)$$

dispositions.

In the case of the big numbers, $g_j \gg 1$, we can approximate:

$$\mathcal{D}_2 \approx \prod_j \frac{(n_j + g_j)!}{n_j!g_j!}, \quad (37)$$

and follow the same procedure as for the Maxwell-Boltzmann distributions. That is, maximize (37) with constriction (25) on the total number of particles and the total energy:

$$\ln \mathcal{D}_2 \approx \sum_j [(n_j + g_j) \ln(n_j + g_j) - n_j \ln n_j - g_j \ln g_j],$$

$$d \ln \mathcal{D}_2 \approx \sum_j [\ln(n_j + g_j) dn_j - \ln n_j dn_j - \ln g_j dg_j].$$

Using the constraints (23) and Lagrange multipliers α and β we have:

$$d \ln \mathcal{D}_2 \approx \sum \left(\ln \frac{n_j + g_j}{n_j} + \alpha + \beta \epsilon_j \right) dn_j.$$

Hence the solution

$$\ln \frac{n_j + g_j}{n_j} + \alpha + \beta \epsilon_j = 0,$$

$$n_j = \frac{g_j}{e^{-\alpha - \beta \epsilon_j} - 1}. \quad (38)$$

This is the Bose-Einstein distribution.

Formula (38) gives number n_j of particles of an energy level that has energy ϵ_j and degeneracy g_j . The constants α, β can be determined from the constraints on the total number of particles and the total energy of the system.

As in the previous section, β can be related to the thermodynamic temperature, so the Bose-Einstein distribution takes the form:

$$n_j = \frac{g_j}{B e^{\frac{\epsilon_j}{kT}} - 1}. \quad (39)$$

In the case $n_j \ll g_j$ or $n_j/g_j \ll 1$, we have:

$$\begin{aligned} B e^{\frac{\epsilon_j}{kT}} - 1 &\gg 1, \\ B e^{\frac{\epsilon_j}{kT}} - 1 &\approx B e^{\frac{\epsilon_j}{kT}}, \\ n_j &\approx \frac{g_j}{B} e^{-\frac{\epsilon_j}{kT}}. \end{aligned} \quad (40)$$

This expression can be compared by the Maxwell-Boltzmann probability (12).

In the boundary case (40) the information of the boson gas has the same form as the Maxwell-Boltzmann gas, but in the other cases it is not. Comparing the numbers of disposition (33) and (19) we can see that boson gas has less dispositions, which lead to less information. The fact that the boson not differs minimize their potential accidental deployment and the information they could contain by these possibilities.

The average number of the particles \bar{n}_j of the energy level ϵ_j from (38) lead us to the formula (2).

Shannon's information is the integral of the function

$$y_2(x) = -\rho_2(x) \ln \rho_2(x), \quad (41)$$

whose graph for $\alpha_2 = 1$ is on Figure 4.

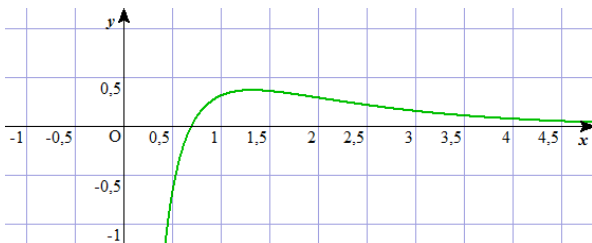


Figure 4: Function $y_2(x) = -\rho_2(x) \ln \rho_2(x)$.

These function intersects x -axis when

$$-\frac{\alpha_2}{e^x - 1} \ln \frac{\alpha_2}{e^x - 1} = 0,$$

$$\ln(e^x - 1) = 0,$$

$$x = \ln 2, \quad (42)$$

or $x_0 = 0.693147$.

So we can find the alpha with:

$$\int_{x_0}^{\infty} \rho_2(x) dx = 1,$$

$$\int_{\ln 2}^{\infty} \frac{\alpha_2}{e^x - 1} dx = 1,$$

$$\alpha_2 \ln(1 - e^{-x}) \Big|_{\ln 2}^{\infty} = 1,$$

$$\alpha_2 = \frac{1}{\ln 2} = \frac{1}{x_0} = 1.4427. \quad (43)$$

On the graph, Figure 4, we can see the function $y_2(x)$ for $\alpha_3 = 1$ has a maximum (x_{max}, y_{max}) . In order to prove it analytically for our parameter $\alpha_3 = 1/\ln 2$, let us find the derivative and equate it with zero. Solutions are stationary points, maxima, minima and saddles of the function.

The first derivative is

$$y_2'(x) = \frac{\alpha_2 e^x}{(e^x - 1)^2} \left(1 - \ln \frac{e^x - 1}{\alpha_2} \right). \quad (44)$$

From $y_2'(x) = 0$ stems:

$$1 - \ln \frac{e^x - 1}{\alpha_2} = 0,$$

$$x_{max} = \ln(1 + e\alpha), \quad y_{max} = \frac{1}{e}, \quad (45)$$

or

$$x_{max} = 1.59364, \quad y_{max} = 0.367879.$$

On Figure 4, for unit α_2 , that abscissa is 1.31 and ordinate 0.33.

Derivation (44) is positive for $x \in (x_0, x_{max})$ and negative for $x > x_{max}$. This means that the point (45) is the maximum, and after the maximum the function is constantly decreasing. The x -axis is a horizontal asymptote of the y_2 because the

$$\lim_{x \rightarrow \infty} y_2(x) = \alpha_2 \lim_{x \rightarrow \infty} \frac{\ln[(e^x - 1)/\alpha_2]}{e^x - 1} = 0,$$

for denominator is the function of higher order than numerator. Therefore, we expect that the integral

$$\bar{I}_2 = \int_{x_0}^{\infty} y_2(x) dx, \quad (46)$$

has a finite value.

The calculation of the integral (46) is a very tricky and lengthy job, so I leave it for another time.

Shortly

$$\bar{\mathcal{I}}_2 = \frac{\pi^2}{\ln(4096)} + \ln \ln 2,$$

or approximately

$$\bar{\mathcal{I}}_2 = 0.820056. \quad (47)$$

This is less than the Shannon information $\bar{\mathcal{I}}_1 = 1$ for Maxwell-Boltzmann distribution (11), as we expected.

3 Fermi-Dirac

The next (3) density is of Fermi-Dirac distribution. First of all, we assume that the

$$\int_{x_0}^{\infty} \frac{\alpha_3}{e^x + 1} dx = 1 \quad (48)$$

where x_0 and α_3 are unknown constants to be determined.

By itself, the function $\rho_3 = \alpha_3 / (\exp[x] + 1)$ allows arbitrary lower integral limit x_0 , but again, we have the quantization and perhaps the restriction on some $x_0 > 0$. On the other hand, if we have quantization, then we also have other restrictions on the integrals.

In the literature there is no a similar solution to this problem, as I know, so we have to continue lonely.

Suppose that $x_0 = 0$. Then the normalization gives:

$$\begin{aligned} \int_0^{\infty} \rho_3(x) dx &= 1, \\ -\alpha_3 \ln(1 + e^{-x}) \Big|_0^{\infty} &= 1, \\ -\alpha_3(0 - \ln 2) &= 1, \\ \alpha_3 &= \frac{1}{\ln 2} = 1.4427, \end{aligned} \quad (49)$$

which is the same constant as α_2 .

On the Figure 5 is graph of the function

$$y_3(x) = -\rho_3(x) \ln \rho_3(x). \quad (50)$$

for unit alpha.

The derivation is:

$$y_3'(x) = \frac{\alpha_3 e^x}{e^x + 1} \left(1 - \ln \frac{e^x + 1}{\alpha_3} \right). \quad (51)$$

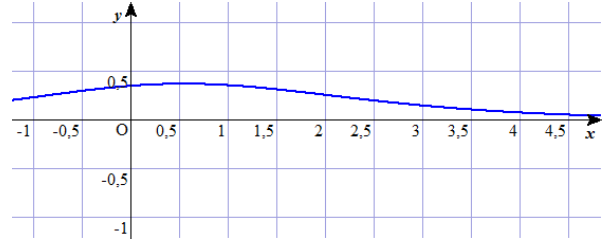


Figure 5: Function $y_3(x) = -\rho_3(x) \ln \rho_3(x)$.

Equating to zero gives the maximum:

$$x_{max} = \ln(1 + \alpha e) = 1.59364, \quad (52)$$

which is exactly equal to the previous case (45). But the ordinate

$$y_3(x_{max}) = 0.344031 \quad (53)$$

is slightly lower than (41).

Let us now compute the Shannon information

$$\bar{\mathcal{I}}_3 = \int_0^{\infty} y_3 dx \quad (54)$$

in a few steps.

Lemma 1. *If*

$$-\int_0^{\infty} \frac{\ln(e^x + 1)}{e^x + 1} dx = b, \quad (55)$$

than

$$-\int_0^{\infty} \frac{\ln[(e^x + 1)/a]}{(e^x + 1)/a} dx = ab - \ln a. \quad (56)$$

where the constant $a > 0$.

Proof. Starting from the assumption, we obtain:

$$\begin{aligned} -\int_0^{\infty} \frac{a}{e^x + 1} \ln \frac{1}{e^x + 1} &= ab, \\ -\int_0^{\infty} \frac{a}{e^x + 1} \ln \frac{1}{e^x + 1} - \ln a &= ab - \ln a. \end{aligned}$$

On the left side of this equality we have:

$$\begin{aligned} -\int_0^{\infty} \frac{a}{e^x + 1} \ln \frac{1}{e^x + 1} - \ln a \int_0^{\infty} \frac{a}{e^x + 1} dx \\ = -\int_0^{\infty} \frac{a}{e^x + 1} \left(\ln \frac{1}{e^x + 1} + \ln a \right) dx \\ = -\int_0^{\infty} \frac{a}{e^x + 1} \ln \frac{a}{e^x + 1} dx. \end{aligned}$$

So

$$-\int_0^{\infty} \frac{a}{e^x + 1} \ln \frac{a}{e^x + 1} dx = ab - \ln a.$$

That is what we are supposed to prove. \square

The next is the corresponding lemma for the indefinite integral.

Lemma 2. *If*

$$\int \frac{\ln(e^x + 1)}{e^x + 1} dx = F(x) + C, \quad (57)$$

than

$$\int \frac{\ln[(e^x + 1)/a]}{(e^x + 1)/a} dx = aF(x) + a \ln a \ln(1 + e^{-x}) + C_1, \quad (58)$$

where C, C_1 and $a > 0$ are constants.

Proof. We start from the left-hand side of equality (58):

$$\begin{aligned} & - \int \frac{a}{e^x + 1} \ln \frac{a}{e^x + 1} dx = \\ & - \int a \frac{\ln a - \ln(e^x + 1)}{e^x + 1} dx = \\ & -a \ln a \int \frac{dx}{e^x + 1} + a \int \frac{\ln(e^x + 1)}{e^x + 1} dx = \\ & = a \ln a \ln(1 + e^{-x}) + aF(x) + C_1, \end{aligned}$$

and this is supposed to prove. \square

So if you know the number $a = \alpha_3$ and the integral (57), or (55), then you know the integral (58), or (56), which are actually Shannon's (54).

On the other side, we have the function:

$$\text{Li}_s(z) = \sum_{j=1}^{\infty} \frac{z^j}{j^s} = z + \frac{z^2}{2^s} + \frac{z^3}{3^s} + \dots \quad (59)$$

that is known as *polylogarithm function*. This definition is valid for arbitrary complex order s and for all complex arguments z with $|z| < 1$. It also can be extended to $|z| \geq 1$ by the process of *analytic continuation*.

In the cases $s = 2$ and $s = 3$ the function (59) are named *dilogarithm* and *trilogarithm* respectively. These names come from the development of a logarithmic function in Taylor series and generalizations.

The starting point is the case $s = 1$ and the natural logarithm function $\ln(1 - z) = -\text{Li}_1(z)$, for $|z| < 1$:

$$\begin{cases} \text{Li}_1(-z) = -z + \frac{z^2}{2} - \frac{z^3}{3} + \dots \\ \ln(1 - z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots \end{cases} \quad (60)$$

Then comes the definition

$$\text{Li}_2(z) = - \int_0^z \frac{\ln(1 - t)}{t} dt, \quad z \in \mathbb{C} \setminus [1, \infty), \quad (61)$$

after follow the equalities:

$$z \frac{\partial}{\partial z} \text{Li}_s(z) = \text{Li}_{s-1}(z). \quad (62)$$

Thus defined polylogarithm functions (59).

Particularly, for $s = 2$ and $z = -e^x$, we have:

$$\begin{aligned} \frac{\partial}{\partial x} \text{Li}_2(-e^x) &= \frac{\partial}{\partial x} \left(-e^x + \frac{e^{2x}}{2^2} - \frac{e^{3x}}{3^2} + \dots \right) \\ &= -e^x + \frac{e^{2x}}{2} - \frac{e^{3x}}{3} + \dots = \text{Li}_1(-e^x). \end{aligned}$$

In the case of a large numbers, $x \rightarrow \infty$, it is valid the approximation:

$$\text{Li}_2(-e^x) = -\frac{1}{2}x^2. \quad (63)$$

as you can see in article [2], §22, page 17. The *Dirichlet eta function*² is:

$$\eta(2) = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \dots = \frac{\pi^2}{12},$$

so

$$\text{Li}_2(-1) = -1 + \frac{1}{2^2} - \frac{1}{3^2} + \dots = -\frac{\pi^2}{12}. \quad (64)$$

Lemma 3. *The derivation of*

$$F(x) = -\text{Li}_2(-e^x) - \frac{1}{2} \ln^2(e^x + 1) \quad (65)$$

is

$$f(x) = -\frac{1}{e^x + 1} \ln \frac{1}{e^x + 1} \quad (66)$$

Proof. We get successively: $F'(x) =$

$$\begin{aligned} &= -\text{Li}_1(-e^x) - \frac{e^x}{e^x + 1} \ln(e^x + 1) \\ &= \ln(e^x + 1) - \frac{e^x}{e^x + 1} \ln(e^x + 1) \\ &= \left(1 - \frac{e^x}{e^x + 1} \right) \ln(e^x + 1) \\ &= \frac{1}{e^x + 1} \ln(e^x + 1) \\ &= -\frac{1}{e^x + 1} \ln \frac{1}{e^x + 1} = f(x). \end{aligned}$$

So $F'(x) = f(x)$, and this is supposed to prove. \square

²Wikipedia: http://en.wikipedia.org/wiki/Dirichlet_eta_function

This is the well-known relation for indefinite integrals

$$\int f(x) dx = F(x) + C \quad (67)$$

where C is a constant of integration.

Using Lemma 2 and Lemma 3 we found easily:

$$\int y_3 dx = \frac{\ln \ln 2}{\ln 2} \ln(1 + e^{-x}) - \frac{1}{2 \ln 2} \ln^2(1 + e^x) - \frac{1}{\ln 2} \text{Li}_2(-e^x) + C_1. \quad (68)$$

This is Shannon's indefinite integral (54).

By this procedure, not so easily, we should also get the result for Shannon's definite integral:

$$\int_0^\infty y_3(x) dx = \frac{\pi^2}{12 \ln 2} + \frac{1}{2} \ln 2 + \ln \ln 2,$$

Approximately

$$\bar{I}_3 = \int_0^\infty y_3(x) dx = 1.16663. \quad (69)$$

In Figure 6 are shown all three of the integrands y_1 , y_2 and y_3 .

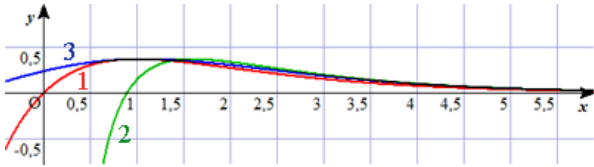


Figure 6: Shannon's integrand $y_{1,2,3}$.

Shannon's information are the definite integrals, which are areas below the corresponding graph to the x -axis for $x > x_0$. As you can see, the graphs confirm the analytical results:

$$\bar{I}_2 < \bar{I}_1 < \bar{I}_3. \quad (70)$$

However, the integral method is not correct for some quantization.

The transition from the smooth curve $y_3(x)$ on quantization, the previous image is slightly changed. In the case of Fermi-Dirac statistic, the two particles cannot be in the same state and they are considered to be indistinguishable. This is the consequence of the Pauli Exclusion Principle which holds for the fermions (half-integer spin), but not for boson (integer spin) particles.

Therefore, for two particles in the two states, there is only one arrangement of

$$\{A, A\}. \quad (71)$$

This is less than the three arrangements (32) for analogous situation for bosons. If we have three states ($s = 3$), the fermions have only 3 dispositions:

$$\{A, A, \emptyset\}, \quad \{A, \emptyset, A\}, \quad \{\emptyset, A, A\}, \quad (72)$$

versus 6 bosons allocations (33).

In the case of $s = 2, 3, \dots$ states, we can choose two on

$$\binom{s}{2} = \frac{s(s-1)}{2} \quad (73)$$

ways.

Example 2. Two particles with the probability p_2 can be in two states, or with the probability p_3 in the three states ($p_2 + p_3 = 1$). Find the information if these particles are:

- 1) classical, 2) bosons, 3) fermions.

Solution. If the two particles with the probability $p_s \geq 0$ chose the $s = 2, 3, \dots$ states with n_s possible dispositions ($p_2 + p_3 + \dots = 1$), than the information

$$I = -p_2 n_2 \ln \frac{p_2}{n_2} - p_3 n_3 \ln \frac{p_3}{n_3} - \dots \quad (74)$$

If the particles are:

- 1) classic (20) we have $n_s = s^2$ dispositions,
- 2) bosons (34) have $\frac{1}{2}s(s+1)$ dispositions,
- 3) fermions (73) have $\frac{1}{2}s(s-1)$ dispositions.

Especially, for $s = 2$ or $s = 3$ and $p_2 + p_3 = 1$, the corresponding information are

$$\begin{aligned} 1) \text{ classic: } & I_1 = -4p_2 \ln \frac{p_2}{4} - 9p_3 \ln \frac{p_3}{9}, \\ 2) \text{ bosons: } & I_2 = -3p_2 \ln \frac{p_2}{3} - 6p_3 \ln \frac{p_3}{6}, \\ 3) \text{ fermions: } & I_3 = -p_2 \ln(p_2) - 3p_3 \ln \frac{p_3}{3}. \end{aligned} \quad (75)$$

These are the required solutions.

Eventually if the probabilities for the states are equal, $p_2 = p_3 = 0.5$, we get

$$\begin{cases} 1) \text{ classic: } & I_1 = 17.1656, \\ 2) \text{ bosons: } & I_2 = 10.1424, \\ 3) \text{ fermions: } & I_3 = 3.03421. \end{cases}$$

As we see, these discrete information are much larger than the appropriate continual (11), (47) and (69), and are not in the same order (70). \square

Let us go to the method of the huge amount of particles to find the numbers of dispositions, like (36) for bosons but now for the fermions.

Again, we consider states corresponding to energy levels ϵ_j . In each state can be distinct single-particle states with the same energy. The number of states with energy ϵ_j is g_j , and the number of such particles

is n_j . But now, because of the Pauli Exclusion Principle for fermions, the largest number of particles with energy ϵ_j is g_j .

The number of deployment g_j single-particle states is

$$\mathcal{D}_3(j) = \binom{g_j}{n_j} = \frac{g_j!}{n_j!(g_j - n_j)!}. \quad (76)$$

Hence, the total number of microstates for fermions is

$$\mathcal{D}_3 = \prod_j \frac{g_j}{n_j!(g_j - n_j)!}. \quad (77)$$

For the principle of equal a priori probabilities, the most likely distributions is that with the largest number of microstates. So we are looking the maximum of the function $\ln \mathcal{D}_3$ (or \mathcal{D}_3) with respect to small changes in the energy level.

By Stirling's approximation we have:

$$\begin{aligned} \ln \mathcal{D}_3 &= \sum_j \ln \left(\frac{g_j}{n_j!(g_j - n_j)!} \right) \\ &\approx \sum_j [g_j \ln g_j - n_j \ln n_j - (g_j - n_j) \ln (g_j - n_j)], \\ d(\ln \mathcal{D}_3) &\approx \sum_j \ln \left(\frac{g_j - n_j}{n_j} \right) dn_j, \end{aligned}$$

for $dg_j = 0$.

These changes have its constraints:

$$\begin{aligned} \sum_j n_j &= N, & \sum_j n_j \epsilon_j &= E, \\ \sum_j dn_j &= 0, & \sum_j \epsilon_j dn_j &= 0. \end{aligned}$$

Here again, we have the (new) constants, the total number of particles (N) in the system and the total energy (E).

Introducing the Lagrange multipliers α and β , we have that the

$$d \ln \mathcal{D}_3 \approx \sum_j \left[\ln \left(\frac{g_j - n_j}{n_j} \right) + \alpha + \beta \epsilon_j \right] dn_j = 0,$$

must be true in all the variations dn_j . Hence:

$$\begin{aligned} \ln \left(\frac{g_j - n_j}{n_j} \right) + \alpha + \beta \epsilon_j &= 0, \quad \forall j = 1, 2, 3, \dots, \\ n_j &= \frac{g_j}{e^{-\alpha - \beta \epsilon_j} + 1}. \end{aligned}$$

Here comes $\alpha = \mu/kT$ and $\beta = -1/kT$ and we get the well-known Fermi-Dirac distribution:

$$n_j = \frac{g_j}{e^{\frac{\epsilon_j - \mu}{kT}} + 1}. \quad (78)$$

With the substitutions

$$x_j = \frac{\epsilon_j - \mu}{kT}, \quad \alpha_3(j) = g_j,$$

we have the transition to the previous continuum distribution (3).

However, the transition between the discrete quantum states and continuous distribution is not always correct, as the Plank discovered in 1900.

4 Epilogue

This is a continuation of some of my private and informal research regarding the information in quantum mechanics.

So simple and appealing the idea about information in the physics shows mathematically very complicated and repulsive. Thence, on the one hand, there is always the same question, whether it is naturally something to be so confusing, and on the other hand there is the suspicion that we still have not mastered the right mathematical tools.

Nevertheless, the results presented in this article are mathematically correct. That's at least something to believe.

References

- [1] R. Vukovic: *Statistics*, Distributions and information of Planks law³, January 13, 2015.
- [2] D. Wood: *The computation of polylogarithms*, School of computing, University of Kent, site www.cs.kent.ac.uk/pubs/1992/110/.
- [3] C. Walck: *Hand-book on Statistical Distributions* for experimentalists, Particle Physics Group Fysikum, University of Stockholm, Internal Report SUFPFY/9601, 11 December 1996.

³Statistics: www.academia.edu/10144096/