What is Statistical Mechanics?

- **Statistical mechanics** is the application of probability theory, which includes mathematical tools for dealing with large populations, to the field of mechanics, which is concerned with the motion of particles or objects when subjected to a force.
- Statistical mechanics, sometimes called statistical physics, can be viewed as a subfield of physics and chemistry.
- It provides a framework for relating the microscopic properties of individual atoms and molecules to the macroscopic or bulk properties of materials that can be observed in everyday life, therefore explaining thermodynamics as a natural result of statistics and mechanics (classical and quantum) at the microscopic level.
- In particular, it can be used to calculate the thermodynamic properties of bulk materials from the spectroscopic data of individual molecules.

Which distribution to use?

- Fermi-Dirac and Bose-Einstein statistics apply when quantum effects have to be taken into account and the particles are considered "indistinguishable".
- The quantum effects appear if the concentration of particles $(N/V) \geq n_q$ (where n_q is the quantum concentration). The quantum concentration is when the interparticle distance is equal to the thermal de Broglie wavelength i.e. when the wavefunctions of the particles are touching but not overlapping.
- As the quantum concentration depends on temperature; high temperatures will put most systems in the classical limit unless they have a very high density e.g. a White dwarf.
- Fermi-Dirac statistics apply to fermions (particles that obey the Pauli exclusion principle), Bose-Einstein statistics apply to bosons.
- Both Fermi-Dirac and Bose-Einstein become Maxwell-Boltzmann statistics at high temperatures or low concentrations.

- Maxwell-Boltzmann statistics are often described as the statistics of "distinguishable" classical particles.
- In other words the configuration of particle *A* in state 1 and particle B in state 2 is different from the case where particle *B* is in state 1 and particle *A* is in state 2. When this idea is carried out fully, it yields the proper (Boltzmann) distribution of particles in the energy states, but yields non-physical results for the entropy, as embodied in Gibbs paradox.
- These problems disappear when it is realized that all particles are in fact indistinguishable. Both of these distributions approach the Maxwell-Boltzmann distribution in the limit of high temperature and low density, without the need for any ad hoc assumptions.
- Maxwell-Boltzmann statistics are particularly useful for studying gases. Fermi-Dirac statistics are most often used for the study of electrons in solids. As such, they form the basis of semiconductor device theory and electronics.

1. Fermi-Dirac Statistics

- In statistical mechanics, **Fermi-Dirac statistics** is a particular case of particle statistics developed by Enrico Fermi and Paul Dirac that determines the statistical distribution of fermions over the energy states for a system in thermal equilibrium. In other words, it is a probability of a given energy level to be occupied by a fermion.
- More generally, **Fermi-Dirac statistics** means that the total wavefunction of fermions must be antisymmetric under an exchange of every pair of fermions (that is, if one exchanges any fermion with another, the wavefunction gets an overall minus sign).
- Fermions are particles which are indistinguishable and obey the Pauli exclusion principle, i.e., no more than one particle may occupy the same quantum state at the same time. Fermions have half-integral spin. Statistical thermodynamics is used to describe the behaviour of large numbers of particles. A collection of non-interacting fermions is called a Fermi gas.
- F-D statistics was introduced in 1926 by Enrico Fermi and Paul Dirac and applied in 1926 by Ralph Fowler to describe the collapse of a star to a white dwarf and in 1927 by Arnold Sommerfeld to electrons in metals. Pascual Jordan developed in 1925 the same statistics which he called *Pauli statistics*. The problem was that his referee Max Born forgot the paper for six months before finding it again. In the meantime it was independently discovered by Enrico Fermi and Paul Dirac.[1]

For F-D statistics, the expected number of particles in states with energy ε_i is

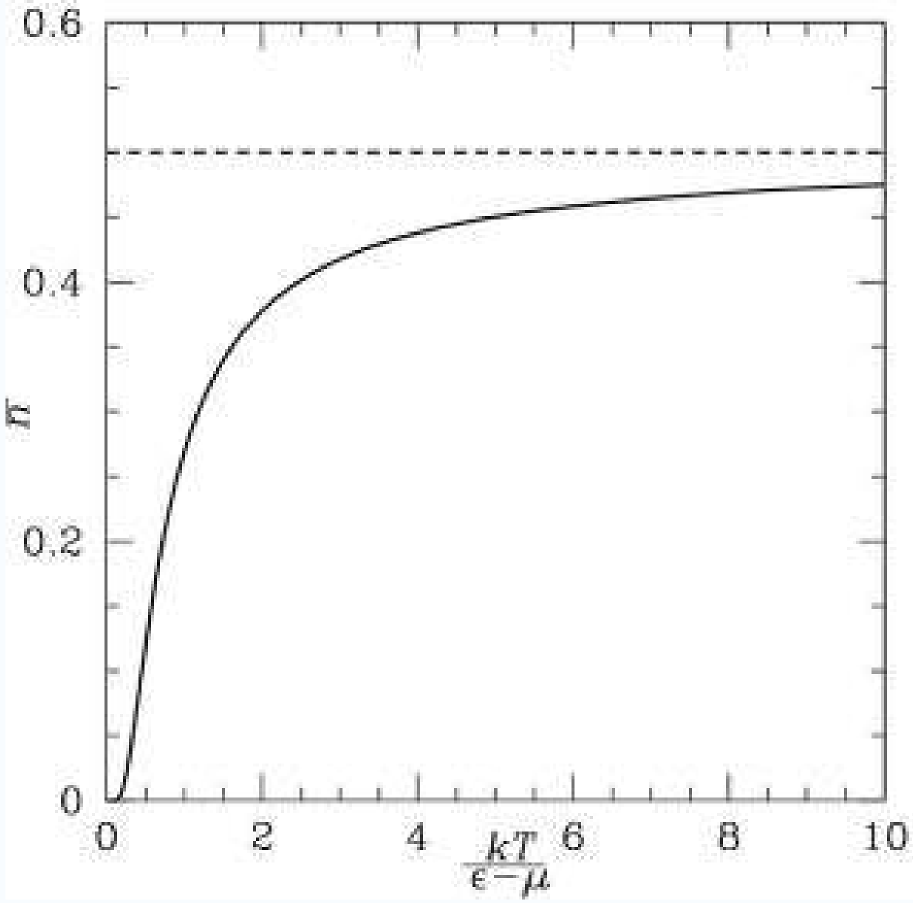
$$n_i = \frac{g_i}{e^{(\varepsilon_i - \mu)/kT} + 1}$$

where:

n_i is the number of particles in state *i*,
ε_i is the energy of state *i*,
g_i is the degeneracy of state(density of states) *i* (the number of states with energy),
μ is the chemical potential (Sometimes the Fermi energy E_F is used instead, as a low-temperature approximation),
k is Boltzmann's constant, and
T is absolute temperature.

In the case where μ is the Fermi energy E_F and g_i=1, the function is called the **Fermi function**:

$$n_i = \frac{1}{e^{(\varepsilon_i - E_F)/kT} + 1}$$



Fermi-Dirac distribution as a function of temperature. More states are occupied at higher temperatures.

Derivation

Suppose there are two fermions placed in a system with four energy levels. There are six possible arrangements of such a system, which are shown in the diagram below.

	ε_1	ε_2	ε_3	ε_4
A	*	*		
B	*		*	
C	*			*
D		*	*	
E		*		*
F			*	*

Each of these arrangements is called a *microstate* of the system. Assume that, at thermal equilibrium, each of these microstates will be equally likely, subject to the constraints that there be a fixed total energy and a fixed number of particles.

- Depending on the values of the energy for each state, it may be that total energy for some of these six combinations is the same as others.
- Indeed, if we assume that the energies are multiples of some fixed value ε, the energies of each of the microstates become:

A: 3ε
B: 4ε
C: 5ε
D: 5ε
E: 6ε
F: 7ε

- So if we know that the system has an energy of 5ε, we can conclude that it will be equally likely that it is in state C or state D.
- Note that if the particles were distinguishable (the classical case), there would be twelve microstates altogether, rather than six.

Now suppose we have a number of energy levels, labeled by index *i*, each level having energy ε_i and containing a total of n_i particles.

Suppose each level contains g_i distinct sublevels, all of which have the same energy, and which are distinguishable. For example, two particles may have different momenta, in which case they are distinguishable from each other, yet they can still have the same energy. The value of g_i associated with level *i* is called the "degeneracy" of that energy level.

The Pauli exclusion principle states that only one fermion can occupy any such sublevel.

Let $w(n, g)$ be the number of ways of distributing *n* particles among the *g* sublevels of an energy level. It's clear that there are *g* ways of putting one particle into a level with *g* sublevels, so that $w(1, g) = g$ which we will write as:

$$w(1,g) = \frac{g!}{1!(g-1)!}$$

We can distribute 2 particles in *g* sublevels by putting one in the first sublevel and then distributing the remaining ($n-1$) particles in the remaining ($g-1$) sublevels, or we could put one in the second sublevel and then distribute the remaining ($n-1$) particles in the remaining ($g-2$) sublevels, etc. so that

$$w'(2, g) = w(1, g-1) + w(1, g-2) + \ldots + w(1, 1)$$

or:

$$w(2,g) = \sum_{k=1}^{g-1} w(1, g-k) = \frac{g!}{2!(g-2)!} = \frac{g(g-1)}{2!}$$

Continuing this process, we can see that $w(n, g)$ is just a binomial coefficient

$$w(n,g) = \frac{g!}{n!(g-n)!}$$

❖ The number of ways that a set of occupation numbers n_i can be realized is the product of the ways that each individual energy level can be populated:

$$W = \prod_i w(n_i, g_i)$$

❖ Next we wish to find the set of n_i for which W is maximized, subject to the constraint that there be a fixed number of particles, and a fixed energy. We constrain our solution using Lagrange multipliers forming the function:

$$f(n_i) = \ln(W) + \alpha(N - \sum n_i) + \beta(E - \sum n_i E_i)$$

❖ Again, using Stirling's approximation for the factorials and taking the derivative with respect to n_i, and setting the result to zero and solving for n_i yields the Fermi-Dirac population numbers:

$$n_i = \frac{g_i}{e^{\alpha + \beta E_i} + 1}$$

- It can be shown thermodynamically that β = 1/*kT* where *k* is Boltzmann's constant and *T* is the temperature, and that α = -μ/*kT* where μ is the chemical potential, so that finally:

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

Lagrange Multipliers

- In mathematical optimization problems, the method of **Lagrange multipliers**, named after Joseph Louis Lagrange, is a method for finding the extrema of a function of several variables subject to one or more constraints; it is the basic tool in nonlinear constrained optimization.

- Simply put, the technique is able to determine where on a particular set of points (such as a circle, sphere, or plane) a particular function is the smallest (or largest). It does this by generalizing and formalizing the idea that the set of *all* points at height h above the earth's surface is tangent to the peak of a mountain of height h.

- More formally, Lagrange multipliers compute the stationary points of the constrained function. By Fermat's theorem, extrema occur either at these points, or on the boundary, or at points where the function is not differentiable.

- It reduces finding stationary points of a constrained function in n variables with k constraints to finding stationary points of an unconstrained function in $n+k$ variables. The method introduces a new unknown scalar variable (called the Lagrange multiplier) for each constraint, and defines a new function (called the Lagrangian) in terms of the original function, the constraints, and the Lagrange multipliers.

Introduction to Lagrange Multipliers

In general

- in order to minimize the function $f(\vec{x})$
- subject to the constraints $\begin{cases} g_1(\vec{x}) = 0 \\ \vdots \\ g_m(\vec{x}) = 0 \end{cases}$
- one defines the function $F(\vec{x}) = f(\vec{x}) + \lambda_1 g_1(\vec{x}) + \cdots + \lambda_m g_m(\vec{x})$
- and searches for the point $\vec{x}$ and the variables $\lambda_1 \cdots \lambda_m$ such that

$$\nabla F(\vec{x}) = 0 \text{ and } \frac{\partial F}{\partial \lambda_j} = 0 \text{ (for } j = 1 \cdots m)$$

Introduction to Lagrange Multipliers

Example

The entropy of a system with N states that have probabilities $p_1, p_2, \cdots, p_N$

is given by $S = -k_B \sum_{j=1}^{N} p_j \ln p_j$

For what values of the probabilities is the entropy maximal?

Introduction to Lagrange Multipliers

Answer

Probabilities are normalized, i.e. $\sum_{j=1}^{N} p_j = 1$

Hence, we construct the function $F = -k_B \sum_{j=1}^{N} p_j \ln p_j - \lambda \left\{ \sum_{j=1}^{N} p_j - 1 \right\}$

and search for the minimum

$$\frac{\partial F}{\partial p_j} = -k_B (\ln p_j + 1) + \lambda = 0 \Rightarrow p_j = p \text{ constant dependent on } \lambda$$

The constant p has to satisfy

$$\sum_{j=1}^{N} p_j = Np = 1 \Rightarrow p = \frac{1}{N}$$

Uniform distribution: maximal entropy

Stirling's Formula

The formula may also be obtained by repeated integration by parts, and the leading term can be found through the method of steepest descent. The general formula (without the $n^{1/2}$ term) may be quickly obtained by approximating the sum

$$\ln N! = \sum_{1}^{N} \ln n$$

with an integral:

$$\sum_{n=1}^{N} \ln n \approx \int_{1}^{N} \ln n \, dn = N \ln N - N + 1.$$

Bose–Einstein statistics

In statistical mechanics, **Bose-Einstein statistics** (or more colloquially **B-E** statistics) determines the statistical distribution of identical indistinguishable bosons over the energy states in thermal equilibrium.

Bosons, unlike fermions, are not subject to the Pauli exclusion principle: an unlimited number of particles may occupy the same state at the same time. This explains why, at low temperatures, bosons can behave very differently from fermions; all the particles will tend to congregate together at the same lowest-energy state, forming what is known as a Bose–Einstein condensate.

B-E statistics was introduced for photons in 1920 by Bose and generalized to atoms by Einstein in 1924.

The expected number of particles in an energy state i for B-E statistics is:

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

where:

n_i is the number of particles in state i
g_i is the degeneracy of state i
ε_i is the energy of the i-th state
μ is the chemical potential
k is Boltzmann's constant
T is absolute temperature

This reduces to M-B statistics for energies ($\varepsilon_i - \mu$) >> kT.

History

In the early 1920s Satyendra Nath Bose, a professor of University of Dhaka was intrigued by Einstein's theory of light waves being made of particles called photons.

Bose was interested in deriving Planck's radiation formula, which Planck obtained largely by guessing. In 1900 Max Planck had derived his formula by manipulating the math to fit the empirical evidence.

Using the particle picture of Einstein, Bose was able to derive the radiation formula by systematically developing a statistics of massless particles without the constraint of particle number conservation.

Bose derived Planck's Law of Radiation by proposing different states for the photon. Instead of statistical independence of particles, Bose put particles into cells and described statistical independence of cells of phase space. Such systems allow two polarization states, and exhibit totally symmetric wavefunctions.

He developed a statistical law governing the behaviour pattern of photons quite successfully. However, he was not able to publish his work; no journals in Europe would accept his paper, being unable to understand it. Bose sent his paper to Einstein, who saw the significance of it and used his influence to get it published.

A derivation of the Bose–Einstein distribution

- Suppose we have a number of energy levels, labeled by index i, each level having energy E_i and containing a total n_i of particles. Suppose each level contains g_i distinct sublevels, all of which have the same energy, and which are distinguishable. For example, two particles may have different momenta, in which case they are distinguishable from each other, yet they can still have the same energy. The value of g_i associated with level i is called the "degeneracy" of that energy level. Any number of bosons can occupy the same sublevel.

- Let w(n,g) be the number of ways of distributing n particles among the g sublevels of an energy level. There is only one way of distributing n particles with one sublevel, therefore w(n,1)=1. It is easy to see that there are (n+1) ways of distributing n particles in two sublevels which we will write as:

$$w(n,2) = \frac{(n+1)!}{n!1!}$$

- With a little thought it can be seen that the number of ways of distributing n particles in three sublevels is

$$w(n,3) = w(n,2) + w(n-1,2) + \ldots + w(1,2) + w(0,2)$$

- so that

$$w(n,3) = \frac{(n+2)!}{n!2!}$$

- Continuing this process, we can see that is just a binomial coefficient

$$w(n,g) = \frac{(n+g-1)!}{n!(g-1)!}$$

❖ The number of ways that a set of occupation numbers can be realized is the product of the ways that each individual energy level can be populated:

$$W = \prod_i w(n_i, g_i) = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \approx \prod_i \frac{(n_i + g_i)!}{n_i!(g_i)!}$$

❖ where the approximation assumes that g_i>>1. Following the same procedure used in deriving the Fermi-Dirac Statistics, we wish to find the set of n_i for which W is maximized, subject to the constraint that there be a fixed number of particles, and a fixed energy. The maxima of W and lnW occur at the value of N_i and, since it is easier to accomplish mathematically, we will maximise the latter function instead. We constrain our solution using Lagrange multipliers forming the function:

❖ Using the approximation and using Stirling's approximation for the factorials ln(x!)=xlnx-x gives

$$F(n_i) = \sum_i (n_i + g_i)\ln(n_i + g_i) - n_i \ln(n_i) - g_i \ln(g_i) + \alpha\left(N - \sum_i n_i\right) + \beta\left(E - \sum_i n_i E_i\right)$$

❖ Taking the derivative with respect to n_i, and setting the result to zero and solving for n_i, yields the Bose–Einstein population numbers:

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

❖ It can be shown thermodynamically that β=1/kT, where k is Boltzmann's constant and T is the temperature.

❖ It can also be shown that α=-μ/kT, where μ is the chemical potential, so that finally:

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

References

Annett, James F., "Superconductivity, Superfluids and Condensates", Oxford University Press, 2004, New York.

Carter, Ashley H., "Classical ans Statistical Thermodynamics", Prentice-Hall, Inc., 2001, New Jersey.

Griffiths, David J., "Introduction to Quantum Mechanics", 2nd ed. Pearson Education, Inc., 2005.

Charles Kittel and Herbert Kroemer, *Thermal Physics*, 2nd ed. (W.H. Freeman, 1980)

C. J. Pethick and H. Smith, "Bose-Einstein Condensation in Dilute Gases", University Press, 2002, Cambridge.