# Statistical Mechanifesto

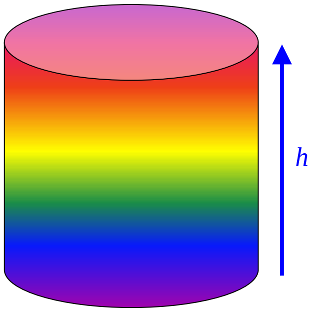
or

## Funky Statistical Mechanics Concepts

The Anti-Textbook*

A Work In Progress.  See elmichelsen.physics.ucsd.edu for the latest versions of the Funky Series.
Please send me comments.

**Eric L. Michelsen**



* Physical, conceptual, geometric, and pictorial physics that didn't fit in your textbook.

Please cite as: Michelsen, Eric L., *Funky Statistical Mechanics Concepts*, elmichelsen.physics.ucsd.edu, 10/10/2019.

2006 values from NIST. For more physical constants, see http://physics.nist.gov/cuu/Constants/ .

| | |
|---|---|
| Speed of light in vacuum | $c \equiv 299\ 792\ 458 \text{ m s}^{-1}$ (exact) |
| Boltzmann constant | $k = 1.380\ 6504(24) \text{ x } 10^{-23} \text{ J K}^{-1} = 8.61734 \text{ x } 10^{-5} \text{ eV/K}$ |
| Stefan-Boltzmann constant | $\sigma = 5.670\ 400(40) \text{ x } 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ |
| Relative standard uncertainty | $\pm 7.0 \text{ x } 10^{-6}$ |
| Avogadro constant | $N_A, L = 6.022\ 141\ 79(30) \text{ x } 10^{23} \text{ mol}^{-1}$ |
| Relative standard uncertainty | $\pm 5.0 \text{ x } 10^{-8}$ |
| Molar gas constant | $R = 8.314\ 472(15) \text{ J mol}^{-1} \text{ K}^{-1} = 0.0820575 \text{ L-atm/(mol-K)}$ |
| calorie | 4.184 J (exact) |
| Electron mass | $m_e = 9.109\ 382\ 15(45) \text{ x } 10^{-31} \text{ kg}$ |
| Proton mass | $m_p = 1.672\ 621\ 637(83) \text{ x } 10^{-27} \text{ kg}$ |
| Proton/electron mass ratio | $m_p/m_e = 1836.152\ 672\ 47(80)$ |
| Atomic mass unit (amu) | $1.660\ 538\ 86 \times 10^{-27} \text{ kg}$ |
| Elementary charge | $e = 1.602\ 176\ 487(40) \text{ x } 10^{-19} \text{ C}$ |
| Electron g-factor | $g_e = -2.002\ 319\ 304\ 3622(15)$ |
| Proton g-factor | $g_p = 5.585\ 694\ 713(46)$ |
| Neutron g-factor | $g_N = -3.826\ 085\ 45(90)$ |
| Muon mass | $m_\mu = 1.883\ 531\ 30(11) \text{ x } 10^{-28} \text{ kg}$ |
| Inverse fine structure constant | $\alpha^{-1} = 137.035\ 999\ 679(94)$ |
| Planck constant | $h = 6.626\ 068\ 96(33) \text{ x } 10^{-34} \text{ J s}$ |
| Planck constant over 2π | $\hbar = 1.054\ 571\ 628(53) \text{ x } 10^{-34} \text{ J s}$ |
| Bohr radius | $a_0 = 0.529\ 177\ 208\ 59(36) \text{ x } 10^{-10} \text{ m}$ |
| Bohr magneton | $\mu_B = 927.400\ 915(23) \text{ x } 10^{-26} \text{ J T}^{-1}$ |

Other values and conversions:

1 eV/particle = 96.485 kJ/mole ≈ 100 kJ/mole

kiloton (of TNT) ≡ 4.184 x $10^{12}$ J = 1 Teracalorie

bar ≡ 100,000 N/m$^2$

atm ≡ 101,325 N/m$^2$ = 1.013 25 bar

torr ≡ 1/760 atm ≈ 133.322 N/m$^2$

Typical chemical bond energy: 1-10 eV

# Contents

# 1 Introduction

## Why Funky?

The purpose of the “Funky” series of documents is to help develop an accurate physical, conceptual, geometric, and pictorial understanding of important physics topics. We focus on areas that don’t seem to be covered well in most texts. The Funky series attempts to clarify those neglected concepts, and others that seem likely to be challenging and unexpected (funky?). The Funky documents are intended for serious students of physics; they are not “popularizations” or oversimplifications.

Physics includes math, and we’re not shy about it, but we also don’t hide behind it.

Without a conceptual understanding, math is gibberish.

This work is one of several aimed at graduate and advanced-undergraduate physics students. Go to http://physics.ucsd.edu/~emichels for the latest versions of the Funky Series, and for contact information. We’re looking for feedback, so please let us know what you think.

## How to Use This Document

This work is not a text book.

There are plenty of those, and they cover most of the topics quite well. This work is meant to be used *with* a standard text, to help emphasize those things that are most confusing for new students. When standard presentations don’t make sense, come here.

You should read all of this introduction to familiarize yourself with the notation and contents. After that, this work is meant to be read in the order that most suits you. Each section stands largely alone, though the sections are ordered logically. Simpler material generally appears before more advanced topics. You may read it from beginning to end, or skip around to whatever topic is most interesting. The “Shorts” chapter is a diverse set of very short topics, meant for quick reading.

If you don’t understand something, read it again *once*, then keep reading.
Don’t get stuck on one thing. Often, the following discussion will clarify things.

The index is not yet developed, so go to the web page on the front cover, and text-search in this document.

### What’s wrong with existing SM expositions?

In an effort to relate to something familiar, many of SM texts jump headlong into complicated thermodynamics before sufficient groundwork has been laid. This leads to lots of logarithms and partial differential equations which obscure the underlying simplicity of SM. They frequently use the same words for different things (e.g. “density of states”) and expect you to instantly understand the distinctions. Also, as in almost all subjects, there are usually just too few simple examples early on to nail down the basics. Finally, there seems to be some kind of exclusion principle disallowing any two books from using the same notation; there’s not much we can do about that, except use the clearest, and then most common, notation we can find.

Also, many expositions expect you to fully understand thermodynamics, and resort to thermodynamic definitions to “derive” mathematical relationships. But that is circular: thermodynamics is empirical; SM is supposed to *derive* thermodynamics from basic physics. Therefore, in this work, we *never* resort to empirical thermodynamic relationships, and derive everything from the fundamentals of SM.

Finally, most works discuss only two specific instances of free energy, rather than the general concept. Here, we described what free energy *is*, so that you understand there are a potenitally infinite variety of free energies (on for each variety of system), you can derive for yourself the appropriate free energy for any given system.

Note that thermodynamics is backwards: thermodynamics usually starts with cycles and engines, which are non-equilibrium systems. It is much easier to start with equilibrium systems, as in statistical mechanics. Also, because thermodynamics was developed empirically, temperature is fundamental, and entropy is derived from it. In SM, entropy is fundamental and temperature is derived from it. The fundamental nature of entropy is why lots of theoretical work is done at constant entropy, instead of constant temperature.

Therefore, temporarily put aside all your thermodynamics before learning SM.

## My Story

The Funky series of notes is the result of my going to graduate school in physics after 20 years out of school. Although I had been an engineer all that time, most of my work involved software and design architectures that are far removed from fundamental science and mathematics. I expected to be a little rusty, but I found that the rust ran deeper than I realized.

There are many things I wish I had understood better while taking my classes (first at San Diego State University, then getting my PhD at University of California, San Diego). The Funky series is my attempt to help other students acquire a deeper understanding of physics.

### Thank You

Thanks to Daniel Arovas for many helpful discussions (arguments). I owe a big thank you to many professors at both SDSU and UCSD, for their generosity, even when I wasn't a real student: Herbert Shore, Peter Salamon, Arlette Baljon, Andrew Cooksy, George Fuller, Tom O'Neil, Terry Hwa, and others.

## Scope

### What This Text Covers

This text covers some of the unusual or challenging concepts in a one-quarter graduate course in Statistical Mechanics (SM). It is also very suitable for undergraduate SM, as well. We expect that you are taking or have taken such a SM course, and have a good text book. *Funky Statistical Mechanics Concepts* supplements those other sources.

### What This Text Doesn't Cover

This text is not a SM course in itself, nor a review of such a course. We do not cover all basic SM concepts; only those that are very important, unusual, or especially challenging (funky?).

### What You Already Know

This text assumes you understand basic integral and differential calculus, partial differential equations, and have some familiarity with probability. You must have a working knowledge of graduate kinematic physics: Hamiltonian mechanics, etc, and especially phase-space. Further, it assumes you have a Statistical Mechanics text for the bulk of your studies, and are using *Funky Statistical Mechanics Concepts* to supplement it.

Statistical Mechanics would simply fall apart without quantum mechanics [Hil p1], but surprisingly, the quantum statistics sections require only a basic understanding of some early quantum mechanics concepts, such as particle in a box, and Pauli exclusion for fermions [Hil p2]. For the most part, if you don't understand the quantum mechanics, you can just take our word for a few simple rules, and run with them.

## Notation

[Brackets indicate an aside, which may be interesting, but also may be omitted without loss of continuity.]

Common misconceptions are sometimes written in dark red dashed-line boxes.

We write the probability of an event as "**Pr(event)**." We write the relative probability as "**RelPr(event)**."

In thermodynamics, upper case variables usually refer to extensive quantities (with the notable historical exception of $T$); lower case variables usually refer to intensive quantities, which are often quantities per unit mass, per mole, per unit volume, or per particle.

| Quantity | Extensive symbol | Intensive symbol |
|---|---|---|
| Temperature | | $T$ |
| Pressure | | $p$ |
| Entropy | $S$ | $s$ |
| Energy | $E$ or $U$ | $u$ |
| Enthalpy | $H$ | $h$ |
| Volume | $V$ | $v$ |
| Mass | $M$ | $m$ (= $M/N$)<br>$\rho$ (= $N/V$ or $M/V$) |
| Number of particles | $N$ | $\rho$ (= $N/V$ or $M/V$) |
| Number of moles | $n$ | |
| Helmholtz free energy | $A$ or $F$ | |
| Gibbs free energy | $G$ | $\mu$ (= $G/N$) |

**MS-Word Bugs:** Word 2002 (10.3416.3501) SP-2 displays the pictures incorrectly on the screen (the arrows are grotesquely large), but prints them properly. On the other hand, the integral signs display correctly on the screen, but print improperly. Petition the Justice Department for the solution to this problem (only monopolies can get away with products so consistently bad).

**TBS** stands for "To Be Supplied," i.e., I'm working on it.

?? For this work in progress, double question marks indicates areas that I hope to further expand in the final work. Reviewers please comment on these areas, and others that may need more expansion.

Keywords are listed in **bold** near their definitions. All keywords also appear in the glossary.

**Formulas:** Evaluation between limits: we use the notation $[\text{function}]_a^b$ to denote the evaluation of the function between a and b, i.e.,

$$[f(x)]_a^b = f(b) - f(a). \qquad \text{For example,} \qquad \int_0^1 3x^2\, dx = [x^3]_0^1 = 1^3 - 0^3 = 1.$$

We write the probability of an event as "Pr(event)."

**Open and closed intervals**: An open interval between $c$ and $d$ is written $(c, d)$. It means the range of numbers from $c$ to $d$ *exclusive* of $c$ and $d$. A closed interval between $c$ and $d$ is written $[c, d]$. It means the range of numbers from $c$ to $d$ *including* $c$ and $d$. A half-open interval $[c, d)$ has the expected meaning of $c$ to $d$ including $c$ but not $d$, and $(c, d]$ means $c$ to $d$ excluding $c$ but including $d$.

**Vector variables**: In some cases, to emphasize that a variable is a vector, it is written in bold; e.g., $V(\mathbf{r})$ is a scalar function of the vector, $\mathbf{r}$. $\mathbf{E(r)}$ is a vector function of the vector, $\mathbf{r}$.

In my word processor, I can't easily make fractions for derivatives, so I sometimes use the notation $d/dx$ and $\partial/\partial x$.

I've never understood the bother about distinguishing between $d/dx$ and $\partial/\partial x$. When the function arguments are independent, both forms of derivative are obviously the same thing; I don't know why there's even two ways to write it. Nonetheless, only as a matter of convention, I use $d/dx$ when a function is clearly a total derivative, and $\partial/\partial x$ when it is clearly a partial derivative. However, in some cases, it's not clear what arguments a function has, and it's not important. In that case, I tend to use $\partial/\partial x$ for generality, but don't worry about it.

And for the record, derivatives *are* fractions, despite what you might have been told in calculus. They are a special case of fraction: the limiting case of differentially small changes. But they are still fractions, with all the rights and privileges thereof. Everyone treats them like fractions, multiplies and divides them like fractions, etc., because they *are* fractions. This is especially relevant to Statistical Mechanics.

# 2 Basic Concepts

## Meet Mr. Mole

We've all heard a lot about Mr. Gaseous Mole, and how "ideal" he is. But what do we *really* know about him? How big is he? How much does he weigh? How stable is he? How does he respond to pressure? Is he dense, or an airhead? How energetic is he? What about his turbulent past? Here we examine the legend, molecule by molecule, to see how he stacks up to common sense. In the process, we gain a familiarity with the magnitudes of gaseous physical properties. We express several of the macroscopic results in imperial units, for those civilizations not yet advanced to the metric system.

We use simple formulas for ideal gasses to compute a whole bunch of physical properties, starting with fairly precise measures, and proceeding to more obscure, and approximate, numbers. All you need to know is the ideal gas law, and some physical constants:

$$pV = nRT \qquad where \quad n = 1\text{ mole} = 6.022 \times 10^{23}\text{ molecules} .$$

NB: "≈" means "approximately equals", and "~" means "is of the order of".

Imagine the archetypal mole of ideal gas, in a box, sitting on your desk. **How big is it?** First, we express $R$ in units of L-atm/K, which is a form of energy per unit temperature, and also a form of size, pressure, and temperature:

$$R = \left(8.314\frac{\text{J}}{\text{K}}\right)\frac{1\text{ atm}}{101{,}325\text{ Pa}}\frac{1000L}{\text{m}^3} = .0821\frac{\text{L-atm}}{\text{K}} \qquad \to \qquad 24.6\text{ L at }300\text{ K for }1\text{ mole}$$

$$1\text{ mole} \to 24.6\text{ L} = .0246\text{ m}^3. \qquad \text{Cubic box edge} = V^{1/3} = 0.291\text{ m}$$

Note that scientific (IUPAC) STP is freezing: 0 C = 273 K. We're more comfortable at room temperature, so we use 300 K, though on the absolute temperature scale, there isn't much difference. The box is a cube a little less than 1 foot on a side, about the size of a portable microwave oven, or a small dog.

**How much does it weigh?** Let's say Mr. Mole is nitrogen, $N_2$, since that is pretty close to air:

$$\text{atomic-weight}(N_2) = 28\text{ amu} \qquad \Rightarrow \qquad 1\,mole\ N_2\ weighs\ 28\text{ g} = 0.028\text{ kg} .$$

This is about an ounce avoirdupois.

**Density**: $$\rho = \frac{m}{V} = \frac{28\text{ g}}{24{,}600\text{ cm}^3} = .0011\text{ g/cm}^3 = 1.1\text{ kg/m}^3 .$$

**What is the specific volume**, i.e. volume per particle, and approximate **particle spacing?**

$$v \equiv \frac{V}{N} = \frac{.0246\ m^3}{6.02\times10^{23}} = 4.1\times10^{-26}\text{m}^3 \qquad \text{particle spacing} \sim v^{1/3} = 3.4\times10^{-9}\text{m or } 3.4\text{ nm} .$$

That's pretty tight! Note that a nitrogen atom is ~0.1 nm across, a $N_2$ molecule ~0.2 nm, so the molecular spacing is only about 15 - 30× the molecular size. The **molecular volume** is about:

$$Volume(N_2) \sim \left(1\times10^{-10}\right)^2\left(2\times10^{-10}\right) = 2\times10^{-30}\text{m}^3 .$$

The specific volume is about 20,000× the molecular volume, which means that at pressures of ~200 atm, the molecular volume is about 1% of the ideal gas specific volume, so the gas starts to become noticeably non-ideal from finite volume. (In reality, inter-molecular forces make it non-ideal at much lower pressures.)

**How much energy does it have?** At room temperature, the rotational modes are excited, so:

$$E = \frac{5}{2}nRT = \frac{5}{2}\left(8.314\ J/K\right)300K = 6{,}236\text{ J}.$$

Mr. Mole is no couch potato. He has the energy of a 100 W light bulb burning for a minute. If you've ever touched a lit 100 W light bulb (which I don't recommend), you can imagine what doing it for a minute means in terms of energy. The kinetic **energy per particle (molecule)** is:

$$\frac{E}{N} = \frac{5}{2}kT = \frac{5}{2}\left(1.38\times10^{-23}\text{ J/K}\right)300K = 1.04\times10^{-20}\text{ J}\,\frac{1\,eV}{1.60\times10^{-19}\,J} = .0647\text{ eV}.$$

Compared to typical chemical bond energies of ~1.5 eV, Mr. Mole is a pretty stable guy.

??Express in kJ/mol.

**What is the molecular RMS (root mean square) speed?** The molecular mass of $N_2$ is 28, so:

$$E = \frac{1}{2}mv_{rms}^{2} \qquad \Rightarrow \qquad v_{rms} = \sqrt{\frac{2E}{m}} = \sqrt{\frac{2\cdot 6.21\times10^{-21}\ J}{28\cdot 1.67\times10^{-27}\,kg}} = 515\text{ m/s}.$$

**Roughly, what is the mean free path?** How far can a molecule go, on average, before it sweeps out a volume equal to that containing one particle (on average)? We approximate the cross sectional area as if the particle were a cube:

$$A \sim v^{2/3} = 1.6\times10^{-20}\text{ m}^2 \qquad A\cdot mfp \sim v \qquad \Rightarrow$$

$$mfp \sim \frac{v}{A} = \frac{4.1\times10^{-26}\text{m}^3}{1.6\times10^{-20}\text{ m}^2} = 2.6\times10^{-6}\text{m} \sim 3\times10^{-6}\text{m} \ \ or\ \ 3\ \mu\text{m}.$$

**Roughly, the mean time to collision?** This is just the mean free path divide by the RMS speed:

$$mttc \sim \frac{mfp}{v_{rms}} \sim \frac{2.6\times10^{-6}m}{515\,m/s} \sim 5\times10^{-9}\text{s} \ \ or\ \ 5\text{ ns}.$$

Our insightful investigation has gleaned the following personal information about the mysterious Mr. Mole:

| **Property** | **Value** |
|---|---|
| size | 24.6 L, or a box 29.1 cm (~1 foot) on a side |
| weight | 28g, or ~1 ounce avoirdupois |
| density | .0011 g/cm$^3$ = 1.1 kg/m$^3$ |
| specific volume | 4.1 × 10$^{-26}$ |
| particle spacing | ~3.4 nm |
| $N_2$ molecule volume | ~2 × 10$^{-30}$ m$^3$ |
| energy | 6,236 J |
| kinetic energy per molecule (including rotation) | 0.0647 eV |
| RMS speed | 515 m/s |
| mean free path | ~3 μm |
| mean time to collision | ~5 ns |

## What Is Everything?

You hear the phrase, "All the thermodynamics follows from ...". What is *all* the thermodynamics? Is it only the equation of state, which relates externally visible parameters such as $P$, $V$, and $T$? No, it's more. Two counter examples:

(1) a block of iron and a block of aluminum sitting on a table. They both have the same equation of state (approximately):

$$V = const \qquad\qquad \text{(block of metal)} .$$

But they differ in their thermodynamics, because they have different heat capacities. The equation of state tells us nothing about *internal energy*, which is related to heat capacity. These quantities (internal energy and heat capacity), too, are part of the thermodynamics of the system.

(2) A more realistic example of the above: a monatomic ideal gas vs. a diatomic ideal gas. They both have the same equation of state: $pV = nRT$. But they have different internal energies, and different heat capacities.

The "thermodynamics" of a system includes the equation of state,
and the heat capacity (as a function of the state).

Other properties of a system that can be included are magnetization, and less commonly, polarization.

## The Fundamental Postulate of Statistical Mechanics

Thermodynamics began as a study of systems, heat, and energy. Every system has a **macro-state**, which comprises observable properties of the system, that could conceivably be measured in a lab. The macro-state includes observables such as temperature, pressure, volume, magnetization, etc. For thermodynamics, that's all there is.

Statistical Mechanics (SM) *derives* the laws of thermodynamics from (1) fundamental
principles of physics (such as mechanics and magnetics), and (2) principles of statistics.

Statistical Mechanics (SM) argues that hidden inside every system's macro-state are billions of micro-states which all appear identical from any realistic laboratory viewpoint. For example, suppose we have a liter of gas at some known temperature, pressure, and volume. These properties are related by the well-known gas laws. However, a liter of gas contains ~$10^{20}$ molecules. Suppose we speed up *one* of those molecules by adding a little energy to it, and at the same time slow down another molecule by removing the same amount of energy. In the sense of micro-states, the modified system is different than the original. However, in the sense of macro-states, the difference is imperceptible, and the system is in the same state as the original. Basic mechanics tells us that the extremely numerous molecules are constantly colliding and exchanging energy in unknowable ways. A tiny perturbation in the energies of two particles is insignificant compared to natural energy exchange processes.

A **micro-state** is a single state of a system, including the exact states of all of the system's constituent particles. Classically, a micro-state includes each particle's position and momentum. Quantum mechanically, a micro-state is an $N$-particle wave function that describes the state of every particle in the system. In either case, a micro-state cannot be observed or determined in a laboratory. However, a knowledge of micro-states from physical theories is essential to applying Statistical Mechanics to predict observable phenomena.

Note that a system is always constrained in some way: an isolated system has fixed energy; a system in a fixed container does no work; the gas under a heavy piston is at fixed pressure; a system in thermal contact with a heat bath is at fixed temperature. Any micro-state of a system must satisfy that system's macroscopic constraints. [Knowledge of the macroscopic constraints often allows us to compute all the micro-states (or all those that matter), and then predict the macro-state of the system.]

The fundamental postulate of Statistical Mechanics is that
"All micro-states of an isolated system are equally likely." [Sch p57m]

Many people have a hard time with that, because we've been (wrongly) trained for so long that systems "seek" certain "preferred" states. The sad truth is that systems don't have "preferences," and don't "seek" anything. Instead, "Probability Rules!"[1] Systems fluctuate randomly, mindlessly, and are usually observed to be in the most likely macro-states (that's really the definition of probability). As we will see, having all micro-states equally likely does *not* mean that all macro-states are equally likely. Far from it. Macro-states are summarizations of micro-states, and hence lose a lot of information. This is what makes "most-likely" macro-states from equally-likely micro-states.

As an example of a "most-likely" macro-state arising from equally likely micro-states, consider a roll of two standard dice. All 36 permutations are equally likely, however the chance of rolling a sum of 7 is 6X higher than the chance of rolling 2. Tossing 100 coins is even sharper. All $2^{100}$ possible outcomes are equally likely, but if we just count the number of heads, getting 50 heads is astronomically more likely than getting 1 head or 100 heads.

We said systems are "usually" observed to be in the most likely macro-states; by "usually" we essentially mean "always." The probability of observing anything else is usually so fantastically small that it would be unlikely to happen in a billion billion lifetimes of the universe.

A big part of SM is "Energy, energy, who's got the energy?" I.e., SM examines how energy is most likely to be distributed throughout the parts of a system.

## The Four Concepts of Statistical Mechanics

In its simplest form, there are only 4 concepts in statistical mechanics. They are critically important, and underlie all other results. In order of dependence:

- Conservation of energy (i.e., the first law). Explicitly note where all the energy comes from, and where it all goes.
- Equilibrium, temperature, and entropy. These are all defined by the same model, so I consider them a single concept.
- The Boltzmann distribution: the likelihood that a subsystem of any size or composition has a given energy.
- Multiplicity defies energy (or, entropy attracts heat).

Everything else is an application of these concepts. We'll talk about all of these. Much of thermodynamics is an accounting of where all the (conserved) energy is.

## What Is ...?

### What is a System?

A **system** is a thing or bunch of things which has, at any instant in time, a definite energy (though you might not know that energy). A system must allow at least 2 different energies (e.g., spin-1/2 particles in a magnetic field). Many systems have an infinite number of allowed energies. Examples:

1. A box of $N$ atoms could be a system.
2. Alternatively, we could consider each atom in the box a system (with excited states, or its kinetic energy, being the multiple energy levels). (Then, a box of $N$ atoms is an "ensemble" of 1-atom systems.)

Choosing a microscopic thing as your system makes some important things very easy. We will see, for example, that such a choice quickly shows that the Maxwell-Boltzmann distribution has essentially the same form as the Boltzmann distribution. Similarly, the Fermi-Dirac and Bose-Einstein distributions are extremely simple with the right choice of "system."

---

[1] Thanks to Laura Tucker for this insightful aphorism.

## What is an Ensemble?

An **ensemble** is a large (possibly hypothetical, possibly real) collection of "systems." Following the examples above:

1. If our system is a box of atoms, we might consider a hypothetical collection of such boxes to be an imaginary ensemble.
2. If each atom in the box is a system, then the box of $N$ atoms is a real ensemble of such systems.

## What is a State?

In general, a **state** is a condition of a system *with a definite energy*. What we mean by "condition" depends on the details of the system, and whether we treat the system classically or quantum mechanically. As is often the case in Statistical Mechanics, the quantum version is both more accurate, and easier to understand [Hil p2], so we start with it.

When our system is obviously quantum in nature, a "state" is an energy eigenstate of that system. Thus, for an atom or molecule, a state is a quantum energy eigenstate, i.e., with some definite energy. For example, a molecule which behaves as a nearly harmonic oscillator has a large number of states, approximately equally spaced by $\hbar\omega$.

Sometimes we use a semi-classical mechanics definition of an energy "state" of a gas of particles, but we still must use the energy density of states (as a function of momentum), derived from quantum mechanics. In the case of a gas, a state is in *reality* also a quantum energy eigenstate, but the quantum nature of the gas is usually too small to measure. Therefore, for a box of $N$ particles: we sometimes say a state is a *classical* $6N$ dimensional phase-space state, i.e. a single point in the $6N$ dimensional phase space of $(x, y, z, p_x, p_y, p_z)$. Such a state must be a quantum energy eigenstate, but for large energies, we *approximate* the possible states as continuous in position, momentum, and energy.

## What is an Energy Level?

In contrast to a state, which has a definite energy, there are **energy levels**, which are the set of energies of the states. "Energy levels" are also called **allowed energies**. Note that there may be many (quantum) states of the same energy (called "degeneracy" in quantum mechanics); therefore, there are often many more energy *states* than energy *levels*. If we label each allowed energy as $E_i$, then we label the number of states with that energy $g_i$ (also called the **multiplicity** [Wan p84b] or "degeneracy" of the energy level $E_i$). This definition implies that many sums can be written two ways: in terms of energy states, or in terms of energy levels:

$$\sum_{s=1}^{\#\,states} f(E_s) = \sum_{j=1}^{\#\,energy\ levels} g_j f(E_j) \qquad\qquad where\ g_j \equiv multiplicity\ of\ energy\ E_j\,.$$

We will see that having more states of the same energy matters a lot.

## What is Heat?

**Heat** is random energy which can be transferred between different elements of the system. Commonly, heat is random kinetic energy of particles transferred by collisions. However, thermal photons, which are emitted and absorbed, can also be heat.

Bizarrely, even some respectable authorities sometimes say that "heat is infrared radiation." This is so wrong it's hard to believe anyone says it.

Thermally generated infrared radiation can be a form of heat, but heat comes in other forms, and in other wavelengths, as well. IR is in no way what heat "is."

## The Thermodynamic Limit

We frequently speak of the "thermodynamic limit." This refers to a system which is large enough and hot enough that its state variables can be taken as continuous. Therefore, a reasonable definition of the "thermodynamic limit" is that the particle number $N$ is large enough to be taken as continuous, and the energy is high enough that the energy density of states may be taken as continuous. Therefore, we can do things like take derivatives with respect to $N$:

$$\mu \equiv \frac{dA}{dN},$$

and things like integrating the density of states, $g(E)$:

$$Z(T) \equiv \int_{\infty} g(E) e^{-E/kT} \, dE .$$

## That's Intense-ive

Each measurable property of a system is either "intensive" or "extensive." Crudely, "intensive" means the property does not scale with the system size (e.g., $T$), and "extensive" means it is proportional to system size (e.g., mass). More precisely, consider a system, with measurable (macroscopic) parameters such as $P$, $V$, $T$, $M$ (magnetization). Now consider a subset of the system (a subsystem), which has the same set of measurable properties. Those properties of the subsystem which are the same as the whole-system are intensive properties: e.g., $P$, $T$. Those properties of the subsystem which are proportional to the subsystem size are extensive properties: e.g., Mass ($M$), $V$.

I don't know of any properties which are neither the same as the whole-system, nor proportional to the subsystem size, but if there are any, they would be neither intensive nor extensive.

## Don't Be Dense

In SM, the ambiguous phrase "density of states" is often tossed about with inadequate definition. Here, we try to clear up some of the confusion. There are actually several kinds of "density of states," and they are quite different: there are at least three kinds of energy-density of states, and a classical phase space (position-momentum space) density of states:

- Energy density of states per particle, each particle is in exactly one such state.
- Energy density of states for the system, each state possibly being occupied by a particle.
- Energy density of states per unit volume for the system, each state possibly being occupied by a particle.
- Phase-space density of states

We now describe each in turn.

### Energy Density of States Per Particle

**Energy density of states** can be the number of stationary quantum states (energy eigenstates) per unit energy, per particle. For example, for a hypothetical "harmonic oscillator molecule," the energy states are uniformly spaced at $\hbar\omega$, so the density of states for 1 particle is

$$\rho_{particle} = \frac{1}{\hbar\omega} \text{ states per unit energy, per particle} .$$

Energy density of states is usually a function of energy, so we write it $\rho(E)$. Example: For a hydrogen atom, we can find $\rho(E)$ from knowledge of $E_n$, the quantized energy levels. When $n$ is large, the energy levels are very closely spaced, and we can approximate them as continuous:

$$E_n = -\frac{Ry}{n^2}, \qquad \textit{where} \quad Ry \text{ is the Rydberg constant } 13.6 \text{ eV, and all } E_n < 0$$

$$\Rightarrow \qquad E(n) \approx -\frac{Ry}{n^2}, \qquad \textit{where} \quad n \text{ is now considered a continuous variable}$$

$$n(E) = \sqrt{Ry}\left(-E\right)^{-1/2}$$

$$\rho(E) \equiv \frac{dn}{dE} = \frac{\sqrt{Ry}}{2}\left(-E\right)^{-3/2}$$

$\rho(E)$ is sometimes written as $g(E)$.

A bad example: For a quantum particle in a box (which some long, straight molecules approximate for their outer electrons), the energy of successive quantum states, labeled with quantum number $n$, increases as $n^2$. For small ground-state energy ε and large $n$, we try to approximate the energies as a continuous function $E(n)$; hence,

$$E(n) = \varepsilon n^2 \qquad \Rightarrow \qquad n(E) = \left(\frac{E}{\varepsilon}\right)^{1/2} = \frac{1}{\sqrt{\varepsilon}}E^{1/2} \qquad \Rightarrow \qquad \rho(E) = \frac{dn}{dE} = \frac{1}{2\sqrt{\varepsilon}}E^{-1/2} = \frac{1}{2\sqrt{\varepsilon E}} .$$

The energy density of states *decreases* as $E^{-1/2}$. In this case, the continuum approximation fails because the energy states get *farther apart* with increasing energy. The energy states *do not* approximate a continuum. We don't discuss such systems in this work.

## Energy Density of States For a System

Following the example above, for a box with $N$ "harmonic oscillator molecules" that don't interact, the **system energy density of states** is:

$$\rho_{system} = \frac{N}{\hbar\omega} \text{ states per unit energy .}$$

Energy density of states is usually a function of the energy of the system, so we write it $\rho(E)$. If $n(E)$ is the number of states with energy $E$, we define:

$$\rho(E) \equiv \frac{d}{dE}n(E) .$$

Density of states may, in general, depend on temperature. For example, consider a gas at constant pressure. Its volume varies with temperature. The density of states per unit volume (see below) is constant, so the density of states for the system is proportional to volume, and hence a function of temperature.

## Energy Density of States Per Unit Volume For a System

The energy density of states per unit volume for a system is a minor variant on the above energy density of states for a system. This is typically used for a system which is arbitrary in size, and therefore could be made an arbitrary volume. An example is the classical energy density of states for a gas, or the electron density of states near the top of a crystal valence band (or near the bottom of the conduction band). This density of states is determined by the nature of the system (gas of particles, or atomic structure of the crystal). If the particles are non-interacting, the density of states is *independent* of the number of those states that are occupied by particles (to a good approximation). Because this density of states is per unit volume, it is independent of the size of the system.

For an ideal gas, the density of states per unit volume is *not* a function of temperature. But for some systems, the density of states per unit volume may be a function of temperature. For a crystal, it is typically a weak function of temperature (so its temperature dependence may often be ignored) [Wan p85t].

## Phase-Space Density of States

In classical SM, you usually use phase-space density of states, which is often called just "density of states." We assume you are familiar with phase-space from classical mechanics.

**Phase space** for a system of $N$ particles in 3D is a $6N$-dimensional space of the positions and momenta of all $N$ particles: $(x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}, \; x_2, y_2, z_2, p_{x2}, p_{y2}, p_{z2}, \; ... \; x_N, y_N, z_N, p_{xN}, p_{yN}, p_{zN})$. We save lots of writing by defining a point in phase space as:

$$\Gamma \equiv (x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}, \quad x_2, y_2, z_2, p_{x2}, p_{y2}, p_{z2}, \; ... \quad x_N, y_N, z_N, p_{xN}, p_{yN}, p_{zN}) .$$

Analogous to the position vector in real space, $\mathbf{r}$, you may think of $\Gamma$ as a position vector in phase space (but it is not usually written in bold). Just as real space has a differential volume element which may be written in a coordinate independent way as $d^3\mathbf{r}$, phase space has a differential hyper-volume element which we can write as

$$d\Gamma \equiv d^{6N}\Gamma = dx_1 dy_1 dz_1 dp_{x1} dp_{y1} dp_{z1} \; ... \; dx_N dy_N dz_N dp_{xN} dp_{yN} dp_{zN} = d^3\mathbf{x}_1 d^3\mathbf{p}_1 \; ... \; d^3\mathbf{x}_N d^3\mathbf{p}_N ,$$
$$\text{where} \quad \mathbf{x}_i = (x_i, y_i, z_i), \quad \mathbf{p}_i = (p_{xi}, p_{yi}, p_{zi})$$

Each point in phase space specifies the positions and momenta (i.e., the classical or quantum "state") of *all* $N$ particles of a system, at some instant of time. But quantum mechanics says that only some points in phase space are allowed energy eigenstates. Each eigenstate can be represented as a point in phase space. The set of all allowed states is a collection of points in phase space. Then the **phase space density of states** is the number of quantum states per unit (hyper-)volume of phase space:

$$\rho(\Gamma) \equiv \frac{dn}{d\Gamma} \qquad \left(\text{phase space density of states}\right).$$

### Combining Systems

How do phase space densities of states combine when we combine two subsystems? For an $N$-particle system, the dimension of phase space (= $6N$) depends on $N$ (the number of particles), and therefore $\rho(\Gamma)$, the $N$-particle phase-space density of states, must also depend on $N$. If we combine two subsystems into a single system, each subsystem has its own phase-space density of states, and the combined system has a 3rd phase-space density of states. The total number of particles is:

$$N_s = N_1 + N_2 .$$

The combination of all the dynamic variables of subsystem 1 and subsystem 2 are the dynamic variables of the combined system:

$$6N_s = 6N_1 + 6N_2 .$$

We can find $\rho_s(\Gamma_s)$ from $\rho_1(\Gamma_1)$ and $\rho_2(\Gamma_2)$. First note that, from the definition of differential hyper-volume above, the differential hyper-volume of the combined-system phase-space is the product of the differential hyper-volumes of the subsystem phase-spaces:

$$d\Gamma_s = d\Gamma_1 d\Gamma_2 .$$

For each volume element in the combined system, there are $\rho_1(\Gamma_1)d\Gamma_1$ states in subsystem 1, and $\rho_2(\Gamma_2)d\Gamma_2$ states in subsystem 2. Every combination of these subsystem states is a state in the combined system. Therefore,

$$\rho_s(\Gamma_s)d\Gamma_s = \rho_1(\Gamma_1)d\Gamma_1 \rho_2(\Gamma_2)d\Gamma_2, \qquad \text{and} \qquad d\Gamma_s = d\Gamma_1 d\Gamma_2$$
$$\Rightarrow \qquad \rho_s(\Gamma_s) = \rho_1(\Gamma_1)\rho_2(\Gamma_2)$$

That is:

> When combining two systems, phase-space densities of states multiply.

**A word about temperature dependence of density of states:** In many systems, the energy levels of the states, and therefore the density of states, are fairly *in*sensitive to temperature (e.g., gasses, crystals). In most problems, we treat the energies and densities of the system as fixed with temperature. However, the distribution of systems among the states is highly temperature *de*pendent.

> Aside: Each point in phase space can specify the positions and momenta (i.e., the quantum "state") of *all N* particles of a system, at some instant of time. An ensemble of systems, each in a different state, can be represented as a collection of points in phase space, each point representing all the particles of a system in the ensemble. This set of ensemble points is a subset of the allowed states, since each system of the ensemble must always be in an allowed state. If you now forget quantum mechanics, and allow the systems to evolve classically in time, their phase-space points evolve continuously through phase space. Hamilton's equations show that the density of these points (points per unit volume of phase space), seen by an observer "moving" with the points, cannot change through time evolution (Liouville's theorem) [Pat p33]. It is this flow of points ("phase space flow") that is said to be "incompressible."

## μ-Space Particle Density

μ-space particle density is *not* a density of states, it's a density of *particles*, but we describe it here for completeness. (This μ has nothing to do with chemical potential.) What is μ-space? **μ-space** is the phase-space for a single particle. Imagine a gas of particles (in a container). Each particle has a position ($x, y, z$) and momentum ($p_x$, $p_y$, $p_z$). For each particle, we could plot a point in this 6-D space representing its position and momentum. For the gas, we would have $N$ points in the 6-D space. This 6-D space is called μ-space. In any small region of space and momentum, i.e. small region of μ-space, ($dx\ dy\ dz\ dp_x\ dp_y\ dp_z$) around a point ($x, y, z, p_x, p_y, p_z$), we would find some number of particles. The μ-space particle density is the number of particles per unit of μ-space "volume:"

$$\rho(x,y,z,p_x,p_y,p_z)=\frac{\#\ particles\ in\ the\ region\left(dx\ dy\ dz\ dp_x\ dp_y\ dp_z\right)}{dx\ dy\ dz\ dp_x\ dp_y\ dp_z}\ .$$

Each point in μ-space represents the position and momentum of a single particle of the system. The collection of $N$ points in μ-space represents the positions and momenta of all particles of the system.

The Maxwell-Boltzmann distribution of momenta for particles in a gas is a μ-space particle density, where (in the absence of an external potential) the density happens to be uniform over all $x$, $y$, $z$ positions. (Therefore, we could also call it a momentum-space particle density, since it is only a function of ($p_x$, $p_y$, $p_z$), and not a function of position.) However, for a gas in a container, with significant gravity, the μ-space particle density would be a function of $z$, and of ($p_x$, $p_y$, $p_z$).

## Impartial Derivatives and the Constants of Nature

We often write state functions as functions of 3 independent variables, e.g. entropy is $S(N, V, T)$. However, we get to choose what our independent variables are. Typically, for idealized isolated systems, we would specify one of $E$, $T$, or $S$ to indicate how hot something is, and $V$ for its volume (if fixed) or else $p$ for its pressure, and $N$ for number of particles. So we could write:

$$\begin{array}{lllll} A(N,V,T), & E(N,V,T), & and & S(N,V,T) & OR \\ A(N,V,S), & E(N,V,S), & and & T(N,V,S) & OR \\ A(N,V,E), & T(N,V,E), & and & S(N,V,E) & \end{array}$$

The energy of a system is exactly defined only in an idealized system that is *not* in contact with another system. For a system that *is* in thermal equilibrium with another, we can only specify the average energy, and still call it $E$. But this is also true of other state variables: $S$, $p$, etc. Many things can fluctuate in a system in thermal equilibrium with another system (some things, like particle number, may be fixed). Also, we will sometimes consider systems that are not in equilibrium, and consider how they change as they approach equilibrium, *when in contact with a heat bath at fixed T*. In such cases, the energy, entropy, etc. of the system change during the approach to equilibrium. Mathematically then, $T$ is fixed, but one of $E$ or $S$ is still a non-equilibrium state parameter of the system, so we might have:

$$\text{non-equilibrium at fixed } T: \begin{cases} A(N,V,S), & \textit{and} \quad E(N,V,S) \quad \textit{OR} \\ A(N,V,E), & \textit{and} \quad S(N,V,E) \end{cases}$$

When considering parameters of non-equilibrium systems, such as free energy, *A*, we sometimes call the non-equilibrium parameters "variational," such as "variational free energy." This means it is a temporary, or hypothetical, value of free energy at a single instant when the system has not yet reached equilibrium.

In SM, we often write partial derivatives with a qualifier specifying which variables to hold constant:

$$C_V \equiv \left(\frac{\partial E}{\partial T}\right)_V .$$

A common question is "Why do we have to specify which variables to hold constant when we take partial derivatives? Doesn't the *definition* of a partial derivative require holding all the other variables constant?"

Good question. The problem is that, as shown above, a given thermodynamic function can often be written several ways, with different independent variables. If the thermodynamic function is given as a function of the derivative variable, and the variables to hold constant, then the qualified partial derivative is an ordinary partial derivative:

$$\text{If} \qquad E = E(T,V) \qquad \text{then} \qquad C_V \equiv \left(\frac{\partial E}{\partial T}\right)_V = \frac{\partial E(T,V)}{\partial T} .$$

However:

$$\text{If} \qquad E = E(T,P) \qquad \text{then} \qquad C_V \equiv \left(\frac{\partial E}{\partial T}\right)_V \ne \frac{\partial E(T,P)}{\partial T} .$$

In principle, we can write *E* as a function of either (*T*, *V*) or (*T*, *P*) because there is an equation of state which relates *T*, *V*, and *P*. Therefore, we can always eliminate any one of (*T*, *V*, *P*) in favor of the other two.

As a simple abstract example, we do problem 1.45 in [Sch p31], with clarified notation:

$$\text{Let} \qquad w(x,y) = xy \qquad \text{with an equation of state} \quad x = yz.$$

$$\text{Then} \qquad w(x,z) = x^2 / z$$

$$\text{Thus} \qquad \left(\frac{\partial w}{\partial x}\right)_y = y, \qquad \text{but} \qquad \left(\frac{\partial w}{\partial x}\right)_z = \frac{2x}{z} = 2y$$

$$\text{So} \qquad \left(\frac{\partial w}{\partial x}\right)_y \ne \left(\frac{\partial w}{\partial x}\right)_z$$

Similarly, $w(y, z) = y^2 z$. So:

$$\left(\frac{\partial w}{\partial y}\right)_x = x \qquad \text{but} \qquad \left(\frac{\partial w}{\partial y}\right)_z = 2yz = 2x$$

$$\left(\frac{\partial w}{\partial z}\right)_x = -x^2 z^{-2} \qquad \text{but} \qquad \left(\frac{\partial w}{\partial z}\right)_y = y^2 = +x^2 z^{-2} .$$

## Statistics Refresher

Few topics in science are more counter-intuitive than statistics. As should be obvious from the name "Statistical Mechanics," it is essential to have a basic understanding of probability. Surprisingly, though:

For introductory Statistical Mechanics, it is *not* necessary to know advanced statistics.

We quickly outline here the basic concepts needed. See *Funky Mathematical Physics Concepts* [FMPC] for more information on statistics.

The **probability**, or **absolute probability**, of an event is a number in [0,1] (inclusive) that quantifies how likely an event is to occur. Given a large sample of trials, the fraction of occurrences of the event approaches its absolute probability. Zero means the event will never happen; one means the event happens with certainty. We write the probability of an event as "Pr(event)." If an experiment must produce one of *N* mutually exclusive results, each with probability Pr(result *i*), then:

$$\sum_{i=1}^{N} \Pr(result\ i) = 1, \text{ since it is certain that one of the results is produced.}$$

The **relative probability** of an event is a number that quantifies how likely an event is to occur, out of a pool of say, *M* trials. Therefore, the relative probability of an event is proportional to its probability. Given *M*, absolute probability = (relative probability) / *M*. Given a complete set of relative probabilities, that cover all possible outcomes of an experiment, we can find *M*, and therefore all the absolute probabilities:

$$M = \sum_{i}^{all\ results} \mathrm{RelPr}(result\ i), \qquad \text{and} \qquad \Pr(result\ i) = \frac{\mathrm{RelPr}(result\ i)}{M}.$$

*M* (or sometimes 1/*M*) is called the **normalization factor** for the relative probabilities. Even without *M*, you can compare the relative probabilities of two events, to determine the ratio of their absolute probabilities.

The **average** of some random variable (or some probability distribution) is simply the limit of the average of *M* trials, as $M \to \infty$. The term **expected value** is often used for average, but it's a terrible misnomer. In many cases, the so-called "expected" value is very unlikely, or even impossible, to occur. For example, the "expected" value of a roll of a standard 6-sided die is 3.5, but you would never expect to actually roll a 3.5. Therefore, we do not use the term "expected value" in this work.

## The Nature of Nature

Early in our studies of physics, we learn (incorrectly) that systems seek a state of "minimum energy." This is obviously wrong: if I put a glass of cold water on the table, it warms up, and its internal energy increases. The truth is that systems stumble into a state of equilibrium. "Exactly what macroscopic state is equilibrium?" is one of the questions answered by SM.

Of course, systems don't actually "seek" states with given properties. Particles don't have preferences, and nature doesn't favor disorder or shun neatness. Systems fluctuate essentially randomly between states. After a change in external constraints, these fluctuations wander aimlessly between the new sets of allowed micro-states of the system.

The equilibrium macroscopic state is the most likely macro-state for the system.
It has the largest probability-weighted sum of micro-states.

More precisely, SM says that (given enough time) large systems are essentially certain to be found in an equilibrium state, with insignificant probability of deviation from that state.

There's really no physics in this; it's just probabilities. If you flip a coin 100 times, it's almost certain to produce close to 50 heads, even though 100 tosses is a small number. In SM, you may deal with systems of $>>10^{20}$ states. If you flip a coin $10^{20}$ times, you will almost certainly get 50% heads, to nine decimal places.

## Introduction to Entropy

Entropy is almost surely the most misunderstood concept in SM and thermodynamics. We introduce it briefly here, to lay out the big picture of the road leading to a more rigorous definition. M**ultiplicity** is the number of ways a given system can meet some given conditions (constraints). Entropy is the logarithm of multiplicity:

$$S \equiv k \ln \Omega \qquad \textit{where} \quad k \equiv \text{Boltzmann's constant} .$$

We illustrate entropy with the common discussion of the "entropy of mixing" [ref??].

Let us compare two systems with two different sets of constraints; that is the only way to compare entropies. In the first, a box is divided with gas A confined to one half, and gas B confined to the other. The system is isolated with (essentially) fixed total energy. We do not specify any further details. Since there are many discrete quantum states that satisfy these constraints, we can (in principle) count them. That is the multiplicity of our first system with those conditions.

The second system has the same total energy, and the same gas particles, but the divider is removed. Both gasses are free to occupy the entire box. Again, we specify no further details. The number of quantum states satisfying these conditions is more than the first system, therefore, it has higher entropy.

**Note:** we cannot meaningfully discuss the entropy of a single, given microstate of particles. By definition, there is only one such state (multiplicity = 1), and its entropy is therefore zero.

We do not find the imagery of entropy as "chaos" or "disorder" useful.
Multiplicity is what matters, and nothing else.

Chaos and disorder are human aesthetic values, and therefore likely to mislead a scientist.

## Equilibrium, Temperature, and the One True Entropy

[?? I know this can be simplified somehow, but jeepers, it's hard to do.]

The concepts in this section are crucial;
not only do they define the fundamental concepts of temperature and entropy,
but they underlie the Boltzmann distribution, which in turn underlies everything.

This section is similar to section 1.2 of [Pat p11].

### Overview

We will show that **entropy** is a function of the macrostate of a system that defines the difference in heat the system has absorbed from the heat-bath, compared to some other macrostate (state 1) at the same temperature, through the equation (derived later):

$$\Delta Q_{1\to 2} = T\left(S_2 - S_1\right) .$$

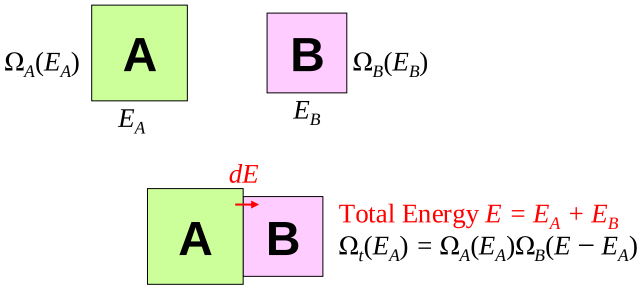
This helps compute thermal equilibrium parameters. Microscopically, entropy is essentially just a measure of **multiplicity**: the number of micro-states with a given macro-state. Many definitions of entropy are possible, but I believe the following one is the most fundamental, and has the fewest approximations (only one: the energy continuum approximation). The **energy continuum approximation** assumes quantized energy states, so we can count them, but also assumes that the quantum energy states are so close together that they compose a virtual continuum of energies (more later). In practice, this is a very realistic approximation. [We do not need the law of large numbers in this derivation.]

For counting states, we define a quantity:

$$\Omega(E) \equiv g(E)\Delta E = \#\text{ of states in range } \Delta E \text{ around the energy } E , \quad \text{where } \Delta E \text{ is very small.}$$

We will call $\Omega$ "density of states," since it is proportional to $g(E)$, with a *constant* factor of $\Delta E$. We will see that the exact value of $\Delta E$ has no bearing on the following derivations; therefore we don't need to specify $\Delta E$ exactly; the only requirement is that it satisfy:

$$\text{quantum spacing} << \Delta E \sim dE << \text{observable energy changes.}$$



Two systems brought into thermal contact.

Consider a total system comprising two dissimilar subsystems, A and B, each with some internal energy, $E_A$ and $E_B$. The total system is isolated, so the total energy of the system, $E$, is fixed:

$$E = E_A + E_B \qquad \Rightarrow \qquad E_B = E - E_A .$$

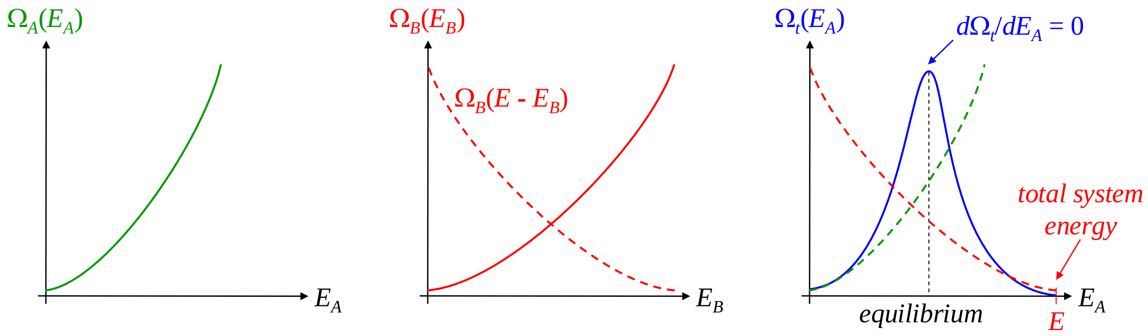
There are no other constraints on the system, i.e., volume, pressure, etc. are arbitrary. We bring the subsystems into contact such that they can freely exchange their internal energy. We wish to find how the total energy divides between systems A and B in equilibrium, i.e. how much energy ends up in A, and how much in B. Equilibrium between systems which can exchange their internal energy is called **thermal equilibrium**. Along the way to equilibrium, we define both temperature and entropy.

We proceed by providing detail to these steps:

1. The total density of states $\Omega_t$ is the product of the subsystem densities of states: $\Omega_t = \Omega_A\Omega_B$.
2. The equilibrium density of states is the maximum total density of states, $\Omega_t$ .
3. The maximum occurs when $d\Omega_t / dE$ is zero, and so that the fractional changes $d\Omega_A / \Omega_A = d\Omega_B / \Omega_B$ , which defines temperature.
4. We recognize that $d\Omega / \Omega = d \ln \Omega$, which defines entropy.

## Details

Each subsystem has its own energy density of states as a function of energy, here called $\Omega_A(E)$ and $\Omega_B(E)$. [See “Don’t Be Dense” above.] The graphs below illustrate typical functions for $\Omega_A(E)$ and $\Omega_B(E)$. Most density of states functions are strictly increasing with $E$ [but see *Negative Temperature* in another section]. The 3rd graph is of $\Omega_t(E_A)$. In equilibrium, system A must have some energy between 0 and $E$, so $\Omega_t(E_A > E) = 0$.



Density of states vs. energy for systems A, B, and the combined system.

We know from the fundamental postulate of SM that the equilibrium macrostate is the one with the most microstates having total energy $E_{tot}$. We'll compute the function $\Omega_t(E_A)$, the density of microstates of the *total* system as a function of A's energy, $E_A$. Then we find the maximum of $\Omega_t(E_A)$, and the value of $E_A$ at that maximum. This will be the *total system* equilibrium macrostate. We will see that the equilibrium state is such that a small transfer of energy from A to B causes the fractional *decrease* in subsystem A's density of states to be equal to the fractional *increase* in subsystem B's density of states.

From the discussion of combining density of states, we have:

$$\Omega_t(E_A) = \Omega_A(E_A)\Omega_B(E_{tot} - E_A) .$$

That is, for any given amount of energy in A, $E_A$, there are $\Omega_t(E_A)$ microstates consistent with that (within $\Delta E$ of $E_A$), and $\Omega_t(E_A)$ is the product of the number of states of A with $E_A$ times the number of states of B with the remaining energy. At equilibrium, consider what happens if a small amount of energy, $dE$, moves from A to B. For A, the energy change is $-dE$; for B, the energy change is $+dE$. $\Omega_A$ and $\Omega_B$ change by:

$$(1) \quad \Delta\Omega_A = -\frac{d\Omega_A}{dE} dE < 0, \qquad \Delta\Omega_B = \frac{d\Omega_B}{dE} dE > 0, \qquad \textit{where all } \Omega \textit{ are functions of } E .$$

At the equilibrium point, the slope of $\Omega_t(E_A)$ is zero,

$$\frac{d\Omega_t}{dE} = 0 \qquad \text{at equilibrium} .$$

I.e., the total number of states is stationary (does not change) with small transfers of energy.

Since $\Omega_t(E_A)$ is the product of $\Omega_A(E_A)$ and $\Omega_B(E_{tot} - E_A)$, the tiny *fractional* decrease in $\Omega_A$ must equal the tiny *fractional* increase in $\Omega_B$, so that the product is constant.

In other words, if $\Omega_A$ decreases by 1%, then $\Omega_B$ must *increase* by 1% for the product to remain constant:

$$-\frac{\Delta\Omega_A}{\Omega_A} = \frac{\Delta\Omega_B}{\Omega_B} \quad \Rightarrow \quad \left(\Omega_A + \Delta\Omega_A\right)\left(\Omega_B + \Delta\Omega_B\right) = \Omega_A\Omega_B, \quad \Delta\Omega_A < 0, \quad \left|\frac{\Delta\Omega_A}{\Omega_A}\right| << 1 .$$

Since $\Delta\Omega_A$ and $\Delta\Omega_B$ are proportional to the energy transfer $dE$ (see (1) above), we have:

$$\frac{d\Omega_A}{\Omega_A} = \frac{d\Omega_B}{\Omega_B} .$$

We recognize these fractions as the derivative of the logarithms. Therefore, in equilibrium, for any small energy transfer $dE$:

$$d \ln\Omega_A = d \ln\Omega_B \qquad \Rightarrow \qquad \frac{d}{dE}\ln\Omega_A = \frac{d}{dE}\ln\Omega_B .$$

Conveniently, the left hand side is a function of system A only, and the right hand side is a function of system B only. Therefore, for each system in thermal equilibrium, there exists an equilibrium parameter, $(d/dE)(\ln \Omega)$, which is the same for both systems.

We might have called this parameter "temperature," but historically, it is inverse temperature. For traditional units, we also scale this parameter by a constant scale factor, and define **temperature** as a function of internal energy. This definition applies to *any* system, so we drop the subscripts *A* and *B*:

$$T(E) \equiv \left[k \frac{d}{dE}\ln\Omega(E)\right]^{-1}, \qquad \textit{where} \quad k \quad \text{is Boltzmann's constant} .$$

We will find that the quantity $\ln \Omega$ appears frequently in statistical mechanics. We therefore define it, when scaled by the same factor, $k$, as a system quantity, **entropy**:

$$S(E) \equiv k \ln \Omega(E) \qquad \Rightarrow \qquad T(E) \equiv \left[\frac{dS}{dE}\right]^{-1} = \frac{dE}{dS}. \tag{2.1}$$

From the above, we see that in traditional units, entropy has units of energy / temperature, or J / K.

Entropy is a measure of the multiplicity of a macro-state, on a logarithmic scale.

We see now that the equilibrium state was the one of maximum multiplicity, which is to say:

The equilibrium macrostate of an *isolated* system is that of maximum entropy,
or equivalently, maximum multiplicity.

### What Entropy Is Not

We do not find the imagery of entropy as "chaos" or "disorder" useful.
Multiplicity is what matters, and nothing else.

The fallacy of "disorder" derives from the fact that usually there are many more states that appear (subjectively) "disordered" than ordered. If a gust of wind blows across a pile of paper, the papers will land randomly around the room. They are unlikely to land in a neat pile, because there are many more ways to land that look "disordered" to our sensibilities than ways that look "ordered." But the probability of the papers landing in a neat pile at a given place is the same as them landing in any single given "disordered" set of locations. The incorrect association of entropy with disorder is more a statement of our feelings about neatness than a statement about physics, and it can be misleading.

For example: if we are mixing the sugar with the flour to bake a cake, we dump the two in a bowl. At first, the mixture looks "messy," because the sugar and the flour are not well mixed. We could define such a "not-well-mixed" system by some conditions requiring isolated local groups of sugar particles, and similarly for flour. After we stir for a while, the sugar and flour are mixed very uniformly, and the mixture looks "neat" to us. We could define a "well-mixed" system by requiring all local groups of particles to be smaller than some size. The not-well-mixed system has lower entropy than the well-mixed system. However, to make a nice cake we *want* uniformity; therefore, our perception is that the state of *higher* entropy is "neater" or "more orderly." This is the opposite of the papers, where our perception of neatness is the state of *lower* entropy. The physics hasn't changed, only our aesthetics have.

### Equilibrium More Mathematically

For another perspective, we now do the same thing more mathematically. We find the equilibrium macrostate using basic calculus: we're looking for the equilibrium point, where $\Omega_A$ decreases by the same *factor* that $\Omega_B$ increases, so the product is constant. That is the point where the change in ln $\Omega_A$ is equal and opposite to the change in ln $\Omega_B$. A crucial point is that the absolute magnitudes of $d\Omega_A$ and $d\Omega_B$ don't matter; only the *fractional* change matters, because we are finding the maximum of the product of $\Omega_A$ and $\Omega_B$:

$$\Omega_t(E_A) = \Omega_A(E_A)\Omega_B(E_{tot} - E_A)$$

$$\frac{d\Omega_t}{dE_A} = \Omega'_A(E_A)\Omega_B(E_{tot} - E_A) - \Omega_A(E_A)\Omega'_B(E_{tot} - E_A) = 0 \qquad \text{(at the maximum)}$$

$$\Rightarrow \qquad \Omega'_A(E_A)\Omega_B(E_{tot} - E_A) = \Omega_A(E_A)\Omega'_B(E_{tot} - E_A)$$

This is the condition for thermal equilibrium, however, it is messy, because each side of the equation involves mixing properties of two dissimilar systems. Can we separate the functions of A from the functions of B, so that we can compute something about A from a knowledge of A alone, and something about B from a knowledge of B alone, and know that they are equal in equilibrium? Separating the pieces of the equation above, we find that at equilibrium:

$$\frac{\Omega'_A(E_A)}{\Omega_A(E_A)} = \frac{\Omega'_B(E_{tot}-E_A)}{\Omega_B(E_{tot}-E_A)} \qquad \Rightarrow \qquad \frac{d}{dE}\ln\Omega_A(E_A) = \frac{d}{dE}\ln\Omega_B(E_{tot}-E_A)\,.$$

This is consistent, because conceptually, a change in the logarithm of a number corresponds to a fractional change of that number:

$$\delta + \ln x = \ln\left(e^{\delta}x\right), \qquad \textit{where} \quad e^{\delta} \text{ is the factor by which } x \text{ changes}\,.$$

The rest follows the above reasoning:

$$T(E) \equiv \left[k\frac{d}{dE}\ln\Omega(E)\right]^{-1} \quad \textit{and} \quad \boxed{S(E) \equiv k\ln\Omega(E)} \qquad \Rightarrow \qquad \boxed{T(E) \equiv \left[\frac{dS}{dE}\right]^{-1} = \frac{dE}{dS}}\,.$$

## Sawing Logs

There are at least 3 reasons why it's convenient to define entropy proportional to the logarithm of Ω,

$$S \equiv k\ln\Omega\,.$$

(1) ln Ω comes up naturally when we see that ($d\Omega / \Omega$) is important, because $d\Omega / \Omega = d(\ln \Omega)$.

(2) ln Ω comes up naturally when expressing multiplicity in units of energy to combine it with the Boltzmann factor (more on this later), and we see that:

$$\Omega e^{-U/kT} = e^{-(U-kT\ln\Omega)/kT} = e^{-(U-TS)/kT}\,.$$

(3) When combining systems A and B, the total density of states $\Omega_t = \Omega_A \cdot \Omega_B$. This means that when combining systems, ln Ω adds, so entropy becomes a convenient *extensive* property of a system:

$$S_t = S_A + S_B\,.$$

## The Energy Continuum Approximation

The energy continuum approximation is so good, that it almost doesn't count as an approximation. It is a subtle point about system states. It's a digression from the main reasoning above, but ignoring it risks confusion. The method we used above depends on counting energy states. Therefore, we must avoid constraining our system energy too tightly, or we run into the "granularity" of allowed quantum energies. The standard "trick" is to allow each system some small *constant* latitude in energy, $\Delta E$ [Hua p176b, Kub p7b]. $\Delta E$ must be large compared to quantum energy spacings, i.e. must include a large number of quantum states so they look like a continuum. It may be comparable to our differential energy changes:

$$\text{quantum spacing} << \Delta E \sim dE << \text{observable energy changes}\,.$$

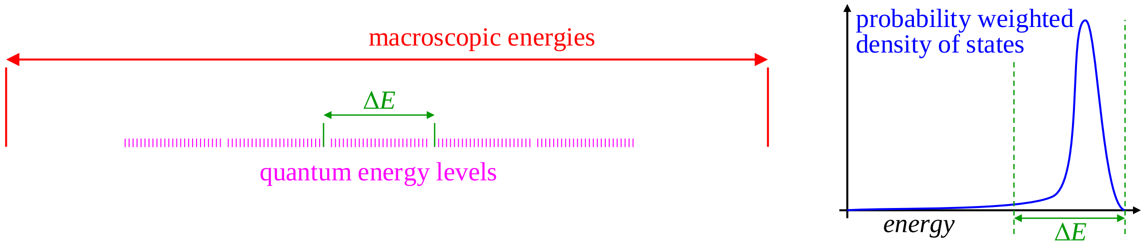


**Figure 2.1** (Left) Qualitative view of the three scales of energy: quantum, mesoscopic, macroscopic. (Right) $\Delta E$ has no practical effect on the number of accessible quantum states.

This allows us to use the energy continuum approximation ($\Delta E$ >> quantum spacing), while still constraining the system to almost fixed energy, at least within measurement accuracy. In practice, this is a very realistic requirement.

For counting states, we define a quantity:

$$\Omega(E) \equiv \rho(E)\Delta E = \#\text{ of states in range } \Delta E \text{ around the energy } E\ .$$

We will call this "density of states," since it is proportional to $\rho(E)$, with a constant factor of $\Delta E$. We will see that the exact value of $\Delta E$ has no bearing on the above derivations; therefore we don't need to specify $\Delta E$ exactly; the only requirement is that it satisfy the above inequality.

### Properties of Temperature and Entropy

The definition of temperature requires a fairly smooth entropy function of energy, $S(E)$, whose derivative is the inverse of temperature. Therefore, small systems, such as atoms, whose energy range is broken up into discrete quantum levels with large spacing (compared to the range) *do not have a temperature*! It's just not defined.

Temperature is only defined for systems with dense energy levels, such that the entropy $S(E)$ has a well-defined derivative over the energies of interest.

TBS: Extensivity of entropy. Arbitrary additive constant; effect of choice of $\Delta E$. Intensivity of temperature. How is this definition of entropy related to the thermodynamic one: $dS = dQ_{\text{reversible}} / T$ ??

## General Equilibrium

The arguments above defining entropy and temperature generalize to any two systems exchanging any conserved thing which affects the multiplicity of a macro-state, i.e. which affects the entropy. It will always be the case that equilibrium is reached when the product of the density of states of the two systems is at a maximum:

$$\Omega_t(X_A) = \Omega_A(X_A)\Omega_B(X - X_A) \qquad \textit{where} \quad X \text{ is the total of some conserved thing}$$

$$X_A \text{ is amount of } X \text{ in } A,\ X_B \text{ is amount of } X \text{ in } B$$

$$\frac{d\Omega_t}{dX_A} = 0 \quad \text{(at the maximum)} \qquad \Rightarrow \qquad \left.\frac{\partial S}{\partial X}\right|_{system-A} = \left.\frac{\partial S}{\partial X}\right|_{system-B} \quad \text{(at equilibrium)}$$

| **Conserved quantity that is exchanged** | **Equilibrium parameter** | **Comments** |
|---|---|---|
| energy | $1/T \sim \partial S/\partial E$ | defines temperature |
| volume | $P \sim \partial S/\partial V$ | say, for A and B separated by movable piston |
| particles | $\mu \sim \partial S/\partial N$ | defines chemical potential $\mu$ |
| any conserved thing, $X$ | $\sim \partial S/\partial X$ | defines some equilibrium parameter |

The chemical potential $\mu$ is special, because it can also be used in the case of non-conserved particles, i.e. when particles can be created or destroyed, e.g. through chemical reactions. In this case, the "conserved" quantity seems somewhat mathematical, but has very real significance. For example, in some reaction:

$$A + B \leftrightarrow C \qquad \text{e.g.} \qquad 2H_2 + O_2 \leftrightarrow 2H_2O$$

The conserved quantity is $(N_{H2} / 2 + N_{O2}) + N_{H2O}$. See later section on $\mu$.

## Spontaneous Processes

There is a fundamental contradiction between the statement that "entropy is only defined for systems in thermal equilibrium," and then comparing the entropies of systems that are clearly not in thermal equilibrium. This is the most egregious failure of most references. Equilibrium is defined as stationary properties (properties don't change with time). Therefore, by definition, any system which spontaneously time evolves was not initially in thermal equilibrium. By our restriction of entropy to equilibrium, it is then meaningless to ask if its entropy has changed. The resolution of this is to define "local thermal equilibrium" (LTE), which allows us to define local entropies in subsystems of the whole system. We then define the total system entropy as the sum of the local entropies, and *only then* can we say the system "entropy" has increased after a spontaneous process.

## Seeing Entropy

We can define that two systems have "significantly" different entropies if a simple tabletop, earthbound experiment at realistic temperatures can demonstrate the difference. I describe some examples shortly. This definition overrides any attempt at quantitative comparisons of absolute entropy vs. entropy differences. A huge baseline entropy has no effect when systems with small entropy *differences* interact. As often stated in thermodynamics, the zero of entropy is almost always irrelevant: it is entropy *differences* that matter. Furthermore, an appropriate quantitative test of entropy differences compares them in light of the height of the free energy barriers between quasi-equilibrium states (which includes system temperature).

I generally think of entropy in the Statistical Mechanics sense: the logarithm of the number of microstates of a system satisfying some conditions. This has the advantage of being inherently path-independent: entropy is a state function, and independent of how the system got into that state. In contrast, the classical thermodynamic definition of integrating reversible heat ($đQ_{rev}/T$) from absolute zero can be misleading, in part because it emphasizes some path of constructing the system, and in part because it is easy to build up to a state which is not in thermal equilibrium.

**An example:** consider a thin, horizontal sheet of oil at room temperature, in which are suspended particles of sugar and coffee. The Brownian motion of those particles is visible under a hand magnifier. For a sheet of a few cm on edge, we could sprinkle sugar on one half, and coffee on the other. Within a few hours or days, the particles would mix. This clearly demonstrates that the system was *not* in thermal equilibrium at first. This initial entropy must be the sum of that of the sugar half of the system plus that of the coffee half. Entropy increases as systems evolve toward equilibrium. Therefore, by my definition, this experiment demonstrates that the entropy is *significantly* lower at first.

Note that if we sprinkled particles directly on a table top, there would be no mixing, even though there is still an entropy difference. However, this experiment introduces large potential barriers between quasi-equilibrium states. The entropy difference between the separated and mixed states is just as large as in the oil-film case, but the experimental setup makes it hard to spontaneously evolve from one state to another, so the difference is not "significant." This has nothing to do with entropy; it has to do with the free-energy barriers. Those are very different things.

A second factor is that the sugar/coffee-on-table experiment conjures the image of the reality of friction on the table top, which our minds must work hard to overcome. If we eliminate friction from the desktop, the particles skitter freely across the surface. They would exhibit Brownian motion, and collisions. It would be a high-speed version of the oil-film experiment, with the same outcome: a significant entropy difference.

In short, the stability of two quasi-equilibrium states says nothing about their entropy differences. Instead, the barriers in the free-energy landscape determine the quasi-equilibrium state lifetimes.

We can visualize which of two systems has higher entropy with the "shake test." Shake up both systems. If one turns into the other, then the latter has higher entropy. One benefit of the "shake test" is that it is independent of the free energy, and intentionally isolates *only* the entropy. Shaking *temporarily* adds a large amount of energy, so that the free-energy landscape presents no barriers. This focuses in on the entropy alone. When we stop shaking, we remove that temporary energy, recalling that final state entropy is independent of path. If the shake test consistently transforms one set of conditions into another, then it must be that the latter state has significantly higher entropy.

We can also see such entropy differences in the classical thermodynamic definition of entropy by integrating from absolute zero. Since entropy is independent of path, there is no need to raise the temperature monotonically. In fact, we can overshoot the final temperature by a large amount, and then come back down. If overshooting results in a different final condition than the monotonic increase, it must be because this new final condition is significantly higher entropy.

In summary, the shake-test isolates entropy, independent of the free-energy landscape, and allows simple, realistic comparisons of entropy differences.

**Summary:** A key point is that entropy is not the whole story, and sometimes not even an important part of the story. Entropy differences are *not the only* driving factor for spontaneous change. In fact, the free-energy landscape, i.e. the availability of paths of evolution that transpire on reasonable time scales, is also a huge factor. In many everyday examples (such as particles on a table), it is the largeness of the energy barriers, not the smallness of the entropy differences, that make thermodynamics inapplicable.

Furthermore, even without large energy barriers, the importance of entropy differences must take into account the temperature. A difference of 1e-20 J/K says nothing directly about how long a spontaneous process might take, or how much energy is absorbed/released. Those depend on the temperature. Smelting ore, carbonizing steel, and diffusing dopants in semiconductors, are all hugely important processes that rely on the tiny entropy difference of mixing, so I must say that the entropy difference is significant. Note that in practice, they are made significant by using high temperatures.

For some applications, it is important to be able to separate entropy differences from energy barriers. My definition for significant entropy differences intentionally isolates an entropy difference: "If I can *easily* demonstrate it, it is significant." But by itself, this says little about practical consequences. A given entropy difference may have no practical consequences in some processes (high barriers or low temperature), but may be clearly significant in others (low barriers or high temperatures).

## Nuts and Boltzmann: The Canonical Distribution

The **Boltzmann distribution** (aka **canonical distribution**) "is the most useful formula in all of statistical mechanics." [Sch p223b]. It underlies a huge amount of SM, and is therefore critically important. In particular, it underlies the "partition function," which we will use heavily later.

The Boltzmann distribution gives the *relative* probability of a system in thermal equilibrium with a heat bath to be in a *single* microstate of given energy;
it is *independent* of the internal structure of the system
(big, small, classical, quantum, fermion, boson, etc.) [Wan p65t]

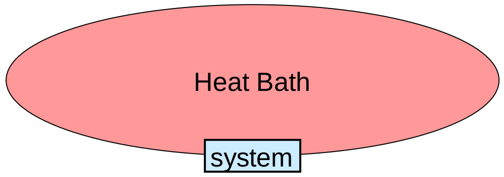
The two constraints for the Boltzmann distribution to apply are:

(1) the system in question must have as its *only significant interaction* the thermal exchange of energy with a large temperature reservoir (aka "heat bath") at fixed temperature, $T$; and

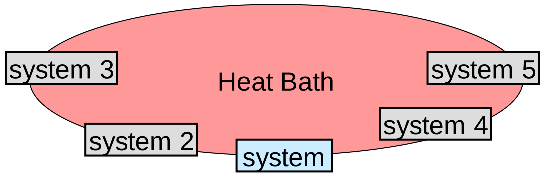
(2) the system must be distinguishable (among an ensemble).

For example, in a classical gas of non-interacting helium atoms, a given atom could be a "system". Its only interactions are through collisions with the other atoms, which compose the heat bath. By definition, "classical" atoms are distinguishable. Therefore the energies of the atoms follow the Boltzmann distribution.

As a counter-example, a single fermion in a fermion gas has significant quantum interactions with the other fermions (via the Pauli exclusion principle), and so such fermions do not meet the requirements of the Boltzmann distribution, and their energies do not follow it. [We consider the microscopic distribution of such systems later.]



System in a large heat bath



Ensemble of distinguishable systems in a large heat bath

We derive the probability of the system having a given energy from *only the properties of the heat bath*! This proves that the construction of the system is irrelevant [Wan p65b]. Many texts use multiple copies of the system (i.e., an ensemble) as the heat bath. This is confusing, and unnecessary. It suggests that there is something special about using identical systems for the heat bath, when there is not. Any large system which can exchange energy with the given system is a heat bath [Wan p64t]. The heat bath's internal structure is not relevant.

Consider a system in thermal equilibrium with a large heat bath (above left). We must work with 3 things: (1) the system, (2) the heat bath, and (3) the combination system/heat-bath. The combination is isolated, therefore all microstates of the combination are equally likely; it has temperature $T$, and it has fixed total energy $E_t$.

What is the probability that the (blue) system is in a microstate $s1$ of zero energy, $E_{s1} = 0$? The probability of the system being in that microstate is proportional to the number of microstates of the heat bath to have energy $E_t$, $\Omega_B(E_t)$:

$$\text{RelPr}(s1) = \Omega_B(E_t)\,.$$

Consider another single microstate $s2$ of the system, of higher energy $E_{s2}$. Its relative probability is

$$\text{RelPr}(s2) = \Omega_B(E_t - E_{s2})\,.$$

Recall that the definition of temperature tells us something about how $\Omega_B$ changes with energy:

$$\frac{1}{T} \equiv k\frac{d\ln\Omega_B}{dE} \qquad \Rightarrow \qquad d\ln\Omega = \frac{dE}{kT}\,.$$

$dE < 0$, since we're considering energy taken from the heat bath and put into the system. Since the heat bath is large and the system small, we can use the linear approximation for $d \ln \Omega_B$:

$$d\ln\Omega_B = \ln\Omega_B(E_t - E_{s2}) - \ln\Omega_B(E_t) = \frac{dE}{kT} \qquad \textit{where} \quad dE = -E_{s2} < 0, \quad E_{s2} > 0$$

$$\ln\frac{\Omega_B(E_t - E_{s2})}{\Omega_B(E_t)} = \frac{-E_{s2}}{kT} \qquad \text{(exponentiate both sides)}$$

$$\frac{\Omega_B(E_t - E_{s2})}{\Omega_B(E_t)} = \frac{\text{RelPr}(s2)}{\text{RelPr}(s1)} = e^{-E_{s2}/kT}$$

The Boltzmann distribution is derived from the properties of only the heat bath, and *not* the system. That is why it is *independent* of the internal structure of the system.

Since relative probabilities have an arbitrary multiplicative factor, we choose RelPr($s$1) = 1. Then we can write:

$$\text{RelPr}(s2) = e^{-E_{s2}/kT}\,.$$

Or, for any single microstate:

$$\Pr(\textit{state} = s \textit{ with energy } E_s) \propto e^{-E_s/kT}\,.$$

We can normalize this to absolute probability by noting that the sum of probabilities of the system being in all possible states (or equivalently, having all possible energies) is 1:

$$\Pr(state\ s) = \frac{e^{-E_s/kT}}{\sum_{r=1}^{\infty} e^{-E_r/kT}} \quad where \quad E_s, E_r \text{ are the quantum energies of the states } s \ \& \ r, \quad OR$$

$$= \frac{e^{-E_s/kT}}{\sum_{j=1}^{\infty} g_j e^{-E_j/kT}} \quad where \quad E_j \text{ are the allowed energies, and } g_j \text{ are their multiplicities}$$

The distribution $g_j$ of allowed energies $E_j$ affects the absolute probabilities of microstates only through the normalization factor. However, the *relative* probabilities of individual microstates are *in*dependent of the allowed energies.

Often, we are interested in the probability of the system having a specified *energy*, rather than being in a specified state. Then we can simply sum the probabilities of all system states with that energy:

$$\Pr(energy = E) = \frac{g_E e^{-E/kT}}{Z(T)} \qquad where \quad g_E \text{ is the multiplicity of energy } E, \text{ and}$$

$$Z(T) \equiv \sum_{r=1}^{\infty} e^{-E_r/kT} = \sum_{j=1}^{\infty} g_j e^{-E_j/kT} \quad where \quad \begin{cases} r, E_r \text{ are states and their energies, } \ and \\ E_j, g_j \text{ are energy levels and their multiplicites (as above)} \end{cases}$$

$Z(T)$ is the famous **partition function**. We'll say more about this later. Notice, though, that the normalization factor, $Z(T)$, is a function of temperature. Therefore:

We can only compare relative probabilities, or Boltzmann factors,
for two things at the same temperature.

For a system of continuous energies (a classical system), we have the continuum equivalent of the above:

$$\Pr(E < energy < E + dE) = \frac{g(E)dE\ e^{-E/kT}}{Z} \quad where \quad g(E) \text{ is the energy density of states, and}$$

$$Z(T) \equiv \int_{E\_\min}^{\infty} dE\ g(E) e^{-E/kT}$$

Recall that systems with large gaps in their quantum energy levels do not have a well-defined temperature.

A system to which the Boltzmann distribution applies may be too small to have a well-defined temperature (e.g., an atom); *T* is the temperature of the heat bath, not the system.

Related to the Boltzmann distribution is the Maxwell-Boltzmann distribution, which is a microscopic distribution of particle states *inside* a single system in thermal equilibrium. For distinguishable particles, the Maxwell-Boltzmann distribution is actually an example of the more-general Boltzmann distribution. More on this later.

Finally, if the Boltzmann distribution seems to suggest that a system "seeks lower energy," remember two things: (1) the Boltzmann distribution describes a *micro*state, not a *macro*state; and (2) the reason the lower-energy microstate is more likely is because the *higher* energy macrostate of the heat bath is more likely, almost as if the heat bath seeks higher energy.

## Multiplicity Defies Energy

This is one of the most important concepts in Statistical Mechanics. We now use the Boltzmann distribution to show:

The equilibrium macrostate of a system is *not* the lowest energy state;
it is the *most likely* state.

We consider a system which has discrete (i.e. quantized) microstates, each with a definite energy. The energies may be microscopically close together, and indistinguishable at the scale of laboratory measurements, but QM says the states are discrete. (Fortunately, it is usually easier to think of them as discrete states, so this is a rare case where QM simplifies our lives. In most of what follows, the continuum approximations that might replace the discrete analysis are straightforward.) We can label each system microstate with an index, such as *s*. Note that many states may have exactly the same energy. The number of states with a given energy is the **multiplicity** of that energy (sometimes called the "degeneracy" of the energy).

When a system (be it a molecule, or a macroscopic sample) is in thermal equilibrium with a reservoir, the temperature, or equivalently, the parameter $\beta = 1/kT$, with units of inverse energy, characterizes the system (see Equilibrium, Temperature, and the One True Entropy, above). The Boltzmann distribution shows that the relative probability of the system being in a particular quantum-state, or "micro-state," is related to the state's energy, and $\beta$, by the **Boltzmann factor** as follows:

$$\Pr(state\ s) \propto e^{-\beta E_s} \qquad i.e. \qquad \text{RelPr}(state\ s) = e^{-\beta E_s}\ .$$

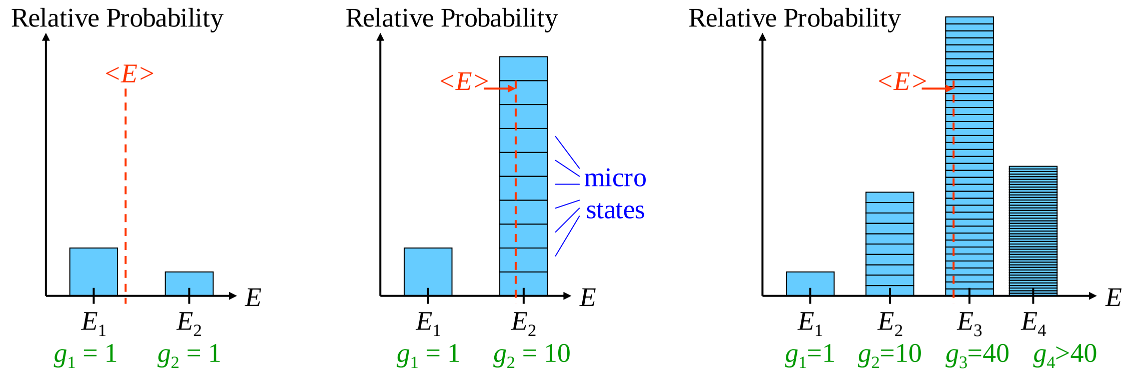
This formula states that for systems which can exchange energy with a large reservoir, a micro-state with a lower energy is more likely to be occupied than another micro-state with higher energy. (We'll assume $\beta > 0$ for now, which is by far the most common case.)

**Example**: It turns out that at room temperature, $\beta \approx 1/(26\text{ meV})$. Consider two states with energies $E_1 = 0$, and $E_2 = 18$ meV. Then:

$$\Pr(E_1) \propto e^{-\beta(0)} = 1, \qquad \Pr(E_2) \propto e^{-\beta(18\,meV)} = 0.5\ .$$

State 1 is twice as likely to be occupied as state 2. If a system has only those two states, then we can normalize the probabilities to find Pr(state 1) = 2/3, and Pr(state 2) = 1/3. Therefore (diagram, below left):

Pr(system energy = $E_1$) = 2/3, and Pr(system energy = $E_2$) = 1/3.



(Left) A system of two microstates. (Center) A system of 11 microstates.
(Right) A system with more energy levels, and still more microstates.

In the diagram, each box is a single microstate. The height of the box is the relative probability (Boltzmann factor) for its energy; all microstates of the same energy have the same relative probability. Higher energies have smaller relative probabilities. The relative probability for an energy is the sum of the relative probabilities of all its microstates, or just the multiplicity, *g*, times the Boltzmann factor.

$$\text{RelPr}(E_j) = g_j e^{-E_j/kT}\ .$$

The average energy of an ensemble of the 2-state systems is:

$$\langle E\rangle = \frac{2}{3}E_1 + \frac{1}{3}E_2 = \frac{2}{3}(0) + \frac{1}{3}(18\text{ meV}) = 6\text{ meV} .$$

which is in the ball park of the minimum energy, but clearly above it. [Almost always, the time average of a single system equals the ensemble average: $<E>_{\text{time}} = <E>_{\text{ensemble}}$. Such a system is called **ergodic**.]

If a similar system had the same two energies, but there were 10 microstates with energy $E_2$, the average energy would be much closer to $E_2$ (above center). With more energies and microstates (above right), we see the average energy is close to the dominant energy level, where dominance is given by the product of the Boltzmann factor times the multiplicity.

Now consider a system with the same energies, but where there are 1,000 states of energy $E_1$, and 1,000,000 states of energy $E_2$. (This is quite plausible.) The relative probability of being in a *particular* micro-state of energy $E_1$ is still 1, and the relative probability of being in a *particular* micro-state of energy $E_2$ is still 0.5. What is the ensemble average energy of this new system? We have 1000 states with relative probability 1, and 1,000,000 states with relative probability 0.5:

$$\text{RelPr}(energy = E_1) = 1000e^{-\beta(0)} = 1000$$

$$\text{RelPr}(energy = E_2) = (1,000,000)e^{-\beta(18\,meV)} = 500,000$$

We're 500 times more likely to have $E = E_2$ than $E_1$. To find <E>, we must convert relative probabilities to absolute probabilities. The probability normalization constant is:

$$N = \sum_s \text{RelPr}(state\ s) = 1,000 + 500,000 = 501,000. \qquad \text{Then}$$

$$\langle E\rangle = \frac{1,000}{501,000}E_1 + \frac{500,000}{501,000}E_2 = (\sim 0)(0) + (\sim 1.0)(18\text{ meV}) \approx 18\text{ meV}$$

Even though the probability of being in any *given* high energy state is only half that of being in a *given* low energy state, the multiplicity of the high energy state vastly outweighs the small energy "advantage" of the lower energy state. Multiplicity defies energy, and raises the average energy to essentially equal the higher energy level. In other words, the system is nearly guaranteed to be in the higher energy level (not what we were taught as freshmen).

We will see in "Free Energy" how much energy a given multiplicity is worth. This quantification of multiplicity leads to the concept of entropy:

Entropy is a measure of multiplicity, on a logarithmic scale, which in turn, is a measure of a macro-state's ability to acquire more energy when in thermal equilibrium.

More on entropy and free energy later.

## Example: The Statistical Chain

Consider a chain of $N$ links, hanging vertically with a constant weight $F$ pulling on the bottom. Each link in the chain has length $l$ , and can be in one of 3 positions: left, right, or down. Note that this system is simplified because it has no kinetic energy (which we will see leads to it having a finite maximum energy, even at arbitrarily large temperature). However, it is realistic, in that some spin-systems have this same property (and even weirder ones; see "Negative Temperature" later on). Questions:

1. What is the low temperature average vertical length of the chain?
2. What is the high temperature vertical length?
3. What is the average length at temperature $T$ ?

**Answers**: 1. At low temperature, there is "no" energy available from the thermal bath. Therefore, all the links are pulled vertically down by the weight. The length is *Nl*. This *is* the state of lowest energy.

2. At high temperature, there is "infinite" energy available from the thermal bath, so the work involved in pulling the chain up, left, or right, is negligible. Therefore, all 3 states of each link are equally likely. Only the down state contributes to vertical length. The average length is simply *Nl* / 3. Notice that this is *not* the state of lowest energy, since some links are Left or Right, and have more energy than Down links.

3. The chain is an ensemble of links; therefore, the links follow the Boltzmann ensemble statistics:

$$\text{RelPr}(E) = e^{-E/kT} = e^{-\beta E} \qquad \textit{where} \quad \beta \equiv 1/kT \text{ is the temperature parameter}$$

In this system, there is no kinetic energy, only potential energy. Consider one link of the ensemble. We could say that being in the left or right position requires doing *Fl* units of work to lift the weight, and being down requires no work. We would then write the energies of the 3 states as $E_L = E_R = Fl, \quad E_D = 0$.

But we can choose the zero-point of energy as we like, and it is equally valid to say that the L and R position are 0 energy, $E_L = E_R = 0$, and the down position has $E_D = -Fl$ (the system gives up potential energy). We choose the latter, because it will make our equations a little simpler if more of the energies are zero:

$$E_L = E_R = 0, \qquad E_D = -Fl$$

We use the Boltzmann distribution to write the relative probabilities of the 3 states:

$$\text{RelPr}(L) = \text{RelPr}(R) = e^{-\beta(0)} = 1, \qquad \text{RelPr}(D) = e^{-\beta(-Fl)} = e^{\beta Fl}$$

We normalize the relative probabilities into absolute probabilities in the usual way: since a link must be in one of the 3 positions, the sum of the absolute probabilities is 1:

$$\Pr(E) = \frac{\text{RelPr}(E)}{\sum_{i=all\ states} \text{RelPr}(E_i)} \quad \Rightarrow \quad \Pr(L) = \Pr(R) = \frac{1}{1+1+e^{\beta Fl}}, \qquad \Pr(D) = \frac{e^{\beta Fl}}{2+e^{\beta Fl}}$$

Each link has the same probability of being down, so the average length is:

$$\langle length \rangle = Nl \Pr(D) = Nl \frac{e^{\beta Fl}}{2+e^{\beta Fl}}. \qquad \text{QED.}$$

Notice that this is *not* the state of lowest energy.

Since this is the general answer for any temperature, we check our first two answers: (1) at high temperature, $\beta \to 0$. <*length*> = *Nl*(1 / 3). (2) at low temperature, $\beta \to \infty$. <*length*> = *Nl*.

Now we can answer more questions:

4. What is the internal energy of the system at low temperature? High temperature?

5. What is the entropy of the system at low temperature? High temperature?

**Answers**: 4. At low temperature, all the links are down (from answer above) and:

$$E = -NFl \qquad (low\, T)$$

At high temperature, 1/3 of the links (the Down ones) have energy –*Fl*, so internal energy is:

$$E = \frac{-NFl}{3} \qquad (high\, T)$$

This is relative to our zero point of energy where every link is Left or Right, and the chain has zero vertical length. This is somewhat unusual that the system has an upper bound on energy. Most systems (e.g. a gas) have kinetic energy which is unbounded, so internal energy is unbounded with increasing *T*.

5. This is a system with discrete energies (rather than a continuum of energies). Therefore, we define Ω(*T*) as the number of states available at temperature *T* (rather than the *density* of states at temperature *T*). At low temperature, there is only one possible state:

$$\Omega(T \to 0) = 1, \qquad S(T) \equiv k \ln \Omega(T) = k \ln 1 = 0 .$$

At high temperature, all states are equally likely, and each link has 3 states:

$$\Omega(T \to \infty) = 3^N, \qquad S(T) \equiv k \ln \Omega(T) = k \ln\left(3^N\right) = Nk \ln 3 .$$

Notice that entropy varies linearly with *N*, so entropy is *extensive*.

We can answer still more questions:

6. At finite temperature, the chain is like a spring: if we pull down on it and let go, it springs back up to its equilibrium length. If we push up the weight, it falls back down to the equilibrium length. Find the effective spring constant in force/length.
7. What is the fractional fluctuation in energy (and length) for a chain of 100 links at temperature *T*? For a chain of $10^{20}$ links?
8. What is the entropy of a long chain at temperature *T*?

**Answers**: 6. The spring constant is *dF*/*dX*, where *X* is length. We already have *X*(*F*) from (3) above. We invert it to find *F*(*X*), and take the derivative:

$$X(F) = Nl \frac{e^{\beta Fl}}{2 + e^{\beta Fl}}, \qquad 2X + Xe^{\beta Fl} = Nle^{\beta Fl}, \qquad 2X = Nle^{\beta Fl} - Xe^{\beta Fl} = e^{\beta Fl}\left(Nl - X\right)$$

$$\beta Fl = \ln \frac{2X}{Nl - X}, \qquad F = \frac{1}{\beta l}\left[\ln 2 + \ln X - \ln\left(Nl - X\right)\right]$$

$$k_{spring} \equiv \left(\frac{dF}{dX}\right)_{\beta} = \frac{1}{\beta l}\left(\frac{1}{X(F)} + \frac{1}{Nl - X(F)}\right)$$

This is called **entropic elasticity**. It results because the chain is *not* in the state of lowest energy, and the number-of-states function Ω(*T*) combined with the Boltzmann distribution makes one particular macrostate (length) more likely than any other. At high temperature ($\beta \to 0$), the spring constant goes to infinity: with infinite thermal energy, the chain can lift any weight.

7. In this system, there is no kinetic energy, only potential energy. Length is proportional to energy:

$$X = -n_D l = \frac{E}{F} \qquad where \quad n_D \equiv \#\text{ of Down links} .$$

The order of magnitude for any fractional fluctuation is almost always $1/\sqrt{N}$. This gives us 10% length (or energy) fluctuation for 100 links, and $10^{-10}$ for $10^{20}$ links. The exact answer uses the binomial distribution: a link is either down, or it's not. The standard deviation of a binomial distribution (see *Funky Mathematical Physics Concepts*) is:

$$\sigma = \sqrt{Np\left(1-p\right)} \qquad \textit{where} \quad N \text{ is the number of "trials" (links)}$$

$$p \equiv \Pr(Down) = \frac{e^{\beta Fl}}{2+e^{\beta Fl}}$$

$$\frac{\sigma}{X} = \frac{\sqrt{Np\left(1-p\right)}}{Nlp} = \frac{1}{\sqrt{N}}\frac{\sqrt{p\left(1-p\right)}}{p} = \frac{1}{\sqrt{N}}\sqrt{\frac{1-p}{p}}$$

At high $T$, $p = 1/3$, and the correction factor to $1/\sqrt{N}$ is only $\sqrt{2}$. As $T$ goes down, the fractional fluctuation increases.

8. The entropy is computed from the number-of-(micro)states of the system that match the macrostate:

$$S \equiv k \ln \Omega(T)\ .$$

For large $N$, the all fluctuations are tiny, so we can use the most likely macrostate, and find $S(T) = S$(most-likely-macrostate). Also for large $N$, the most-likely number of down links is virtually identical to the average number of down links, which we found above:

$$n_D{*} \approx \overline{n}_D = N\frac{e^{\beta Fl}}{2+e^{\beta Fl}}\ .$$

Since this is a discrete system, we can find $\Omega$ as the # of ways to choose $n_D$ down links from $N$ links total:

$$\Omega(T) \textit{ or } \Omega(\beta) = \begin{pmatrix} N \\ n_D{*} \end{pmatrix} \equiv N \text{ choose } n_D{*} = \frac{N!}{n_D{*}!\left(N-n_D{*}\right)!}$$

$$S(T) = k\ln\frac{N!}{n_D{*}!\left(N-n_D{*}\right)!} = k\left[\ln\left(N!\right)-\ln\left(n_D{*}!\right)-\ln\left(\left(N-n_D{*}\right)!\right)\right]$$

We can use some approximations here??

# 3 Classical Statistical Mechanics

So far, we've mostly talked in terms of quantum states. But statistical mechanics is often useful in analysis of classical macroscopic systems, whose detailed quantum behavior is not important. For such cases, we derive some simple rules to obtain SM equations for classical systems. But remember that statistical mechanics cannot exist outside quantum mechanics, because there is no way to count states without quantum mechanics. Our simple rule is based on the fact that, for each dynamic variable of a system, the number of quantum states of a system equals the phase space volume divided by Planck's constant, *h*:

$$n_i = \frac{1}{h}\int dx_i\ dp_i \quad \Rightarrow \qquad n_{system} = \prod_{i=1}^{f} n_i = \frac{1}{h^f}\int dx_1 dp_1\ dx_2 dp_2 \ldots dx_f dp_f = \frac{phase-space-volume}{h^f}$$

$$\text{where} \quad f \equiv \#\text{ dynamic variables (degrees of freedom) of the system}$$

**The continuous energy approximation:** Since the detailed structure of the quantum energy states is not important, we aggregate discrete quantum states into a continuous energy density of states:

$$g_i \to g(E)dE \qquad \text{where} \quad g_i \equiv \text{quantum multiplicity of state } i$$

$$g(E) \equiv \text{continuous energy density of states}$$

This converts formulas with a summation over energies into an integral over energy. For example, the partition function goes to:

$$Z(\beta) = \sum_{j=1}^{\infty} g_j e^{-\beta E_j} \qquad \to \qquad Z(\beta) = \int_0^{\infty} dE\ g(E) e^{-\beta E}\ .$$

This often isn't good enough, since we don't know *g*(*E*). However, we generally know the hamiltonian, which gives *E* as a function of *p* and *x*:

$$E = H(p_1, x_1, p_2, x_2, \ldots, p_f, x_f) \equiv H(p_i, x_i)$$

$$\text{where} \quad f \equiv \#\text{ dynamic variables (degrees of freedom) of the system}$$

*H* is a function of the phase-space state of the system. There are many states with the same energy, but rather than finding *g*(*E*) explicitly, we use the sum-over-states formula for *Z*, and count each $h^f$ units of phase space as one quantum state:

$$Z(\beta) \equiv \sum_{s=1}^{all\ states} e^{-\beta E_s} = \int \frac{dx_1 dp_1 \ldots dx_f dp_f}{h^f} H(p_i, x_i)\ .$$

However, $Z(\beta)$ is arbitrary up to a multiplicative constant, so we can drop the $h^f$:

$$Z(\beta) = \int dx_1 dp_1 \ldots dx_f dp_f\ H(p_i, x_i)\ .$$

Example: One Dimensional Gas: TBS

## Terms of Endearment

You see these same mathematical terms over and over in SM, and they generally have the same meaning everywhere. So here's a simple summary of each common mathematical term. Isolating internal energy on the left hand side, we have:

$$dU = \mu\, dN + T\, dS - P\, dV - F\, dX\ .$$

Alternatively, we can put free energy (described below) on the left hand side:

$$d(\text{free-energy}) = dU + P\, dV + F\, dX - T\, dS\ .$$

| | |
|---|---|
| $dU$ | The change in internal energy. |

| | |
|---|---|
| $\mu\, dN$ | The free-energy it takes to create/add $N$ particles. |
| $T\, dS$ | At constant temperature (i.e., in a heat bath), $T\, dS$ is the heat energy that flows from the heat bath into the system due to an increase in entropy (or out from a decrease). This is the "multiplicity defies energy" principle, also known as the "entropy attracts heat" principle (as shown by the canonical ensemble). |
| $P\, dV$ | At constant pressure, the work done by the system on its surrounding from volume expansion. The surroundings store this work as potential energy, and it could be considered part of the total energy of the system. $P\, dV$ is just a special case of $F\, dx$ work. |
| $F\, dX$ | The generalization of $P\, dV$, where $F$ is any kind of "force" (force, tension, pressure, etc.), and $X$ is the associated "size" of the system (length, area, volume, etc.), such that $F\, dX$ is work done by the system. At "force" $F$, when $X$ increases by $dX$, the work done by the system is $\delta W = F\, dX$. The surroundings store this work as energy, and it can be considered part of the total energy of the system. |

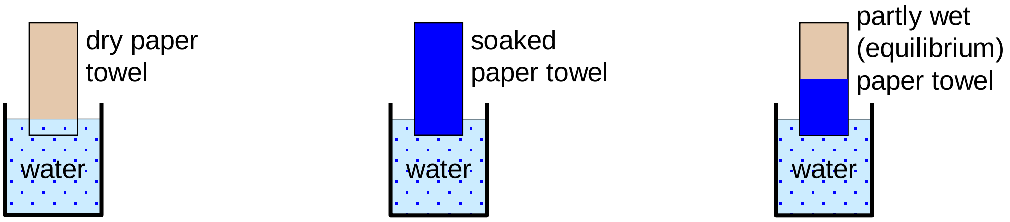
Other work terms include $\sigma\, dA$ (surface tension times area), and $\mathbf{H}\bullet d\mathbf{B}$ (magnetic work).

## Free Energy

Free energy is of fundamental importance to all of statistical mechanics, thermodynamics, and chemistry. However, it can be confusing, so we start with a simple example of free energy, then give several brief high-level views of free energy, and then describe each of these in more detail. This section assumes a thorough understanding of the Boltzmann distribution (described earlier).

### Example of Free Energy: Water Absorbed in a Paper Towel

As an example of free-energy in action, consider a glass of water with a paper towel suspended partially immersed (diagram below).



(Left) System with minimum energy. (Center) System with maximum entropy.
(Right) System with minimum free-energy (in equilibrium).

We find that the equilibrium position has the water soaked part way up the towel (above, right). The system on the left is minimum energy, since all the water is a low as possible (least gravitational potential energy). The system in the middle is maximum entropy, since the water is dispersed among the largest possible area, and consequently has many more ways of rearranging its molecules (maximum number of states). But the equilibrium position is neither minimum energy nor maximum entropy: it is minimum free-energy. The system on the right has made the equilibrium tradeoff between energy and entropy, such that its macro-state is the most likely.

This cannot be completely explained by surface adhesion, as shown by the system with two glasses, and a paper "bridge." TBS.

### Physical View of Free Energy

There are several ways to look at free energy:

1. Free energy is a measure of macrostate probability, on a log scale, and with a minus sign:
   $A = -kT \ln \text{RelPr}(\text{macrostate})$. This implies ...

2. Much like the Boltzmann factor gives the relative probability for a system to be in a *single microstate* of given energy, the Boltzmann factor formula acting on *free energy* gives the relative probability of a system to be in a *macrostate* of given free energy. This implies ...
3. Free energy is minimized within a system in thermal equilibrium (maximum likelihood macrostate).
4. In other words, free energy is a state function of a system at fixed $T$, which when minimized, maximizes the entropy of the universe. Being a state function only of the system in question, it allows us to ignore the rest of the universe, and focus on what we know.
5. Free energy includes the effect of energy "borrowed" from the heat bath due to the system's entropy. This implies that for a system spontaneously evolving from an initial to a final state of equal temperature, the maximum useful work you can get is the difference between the free energies of the initial and final states.
6. Free energy accounts for the "Multiplicity Defies Energy" principle in a single parameter, with units of energy. Essentially, free energy includes a way of measuring multiplicity in energy units, and adding it to the real energy.

Because of a subtlety of potential energy, there is a different "free energy" defined for each set of constraints on (or energy stores in) a system (more later). However:

A given system with constraints has only one kind of free energy relevant to it.
The free energy is always a quantity that is minimized
when a constrained system reaches equilibrium within its constraints.

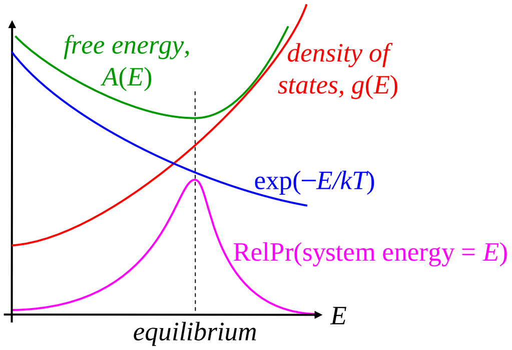
Also, free energy is only relevant for systems at fixed temperature; i.e., for free energy to be meaningful, one of the constraints must be fixed temperature. However, for the "maximum useful work" measure described above, the *process can change temperature in the middle*, but the final temperature must be the same as the initial temperature. We examine this in more detail later.

I would say that it is almost a coincidence that the free energy is both a state-function which is minimized for a system in thermal equilibrium, and also the state-function which gives the maximum useful work obtainable from a spontaneous process. I see no obvious connection between the two, and one might imagine that in some alternate universe, these two phenomena might have been given by two different state-functions.

## Helmholtz Free Energy

The simplest constraint is that of constant temperature: a system held in a heat bath. We now show that for such a system, the free energy is the Helmholtz free energy:

$A(T, V, N) \equiv U - TS$ for a system at given temperature.



Interaction of density of states and the Boltzmann factor to produce the relative probability of the system having a given energy, and the free energy associated with that probability.

This is just an example of the "multiplicity defies energy" concept shown above. First, recall that the relative probability of the system having energy $E_i$ is:

$$\text{RelPr(system-energy is } E_i) = g_i e^{-E_i/kT} \qquad where \quad E_i \text{ are distinct energies of the system,}$$

$$g_i \text{ are \# microstates with energy } E_i \text{ (multiplicities)}$$

The energy with the maximum relative probability is the most likely energy (macro-state). Finding this state involves dealing with two parameters: $g_i$ and $E_i$. It'd be nice if we could combine these two into a single parameter (which we will call **free energy**) which combines the energy and the multiplicity. To do this, we first bring the factor of $g_i$ up into the exponent:

$$\text{RelPr(system-energy} = E_i) = g_i e^{-E_i/kT} = e^{-E_i/kT + \ln g_i} .$$

Next, we measure the multiplicity in energy units, by multiplying $\ln g_i$ by $kT$:

$$\text{RelPr(system-energy} = E_i) = \exp\left(\frac{-\left(E_i - kT \ln g_i\right)}{kT}\right) = \exp\left(\frac{-\left(E_i - TS_i\right)}{kT}\right) \qquad where \quad S_i \equiv k \ln g_i$$

$$= \exp\left(-A_i / kT\right) \qquad where \quad A_i \equiv E_i - TS_i = E_i - kT \ln g_i$$

The energy equivalent of multiplicity depends on temperature: the higher the temperature, the more energy a given multiplicity is "worth".

$$E_{equivalent} = kT \ln g_i = TS .$$

Notice that we could restate the aphorism "Multiplicity defies energy" as "Multiplicity attracts thermal energy (heat)." Higher multiplicity allows for higher energy only for a system in a heat bath, because higher multiplicity allows the system to offset its lower Boltzmann factor for higher energy:

$$\text{RelPr}(E) = g(E) exp(-\beta E) .$$

Where does that higher energy come from? From the heat bath (which keeps the system at constant temperature). For a system with low energy and low entropy, heat flows from the heat bath into the system, because the increase in system entropy more than offsets the decrease in Boltzmann factor for higher energy. When these two effects just balance, the system is in thermal equilibrium. Hence, the multiplicity (entropy) "attracts" heat into the system from the heat bath: Multiplicity attracts heat. The $T\,\Delta S$ term in the Helmholtz free energy is real energy that goes into the system from the heat bath.

The most likely macro-state is that with the highest relative probability, $e^{-A_i/kT}$, which is that when $A_i$ is minimized.

Free energy is a measure of the relative probability of a macro-state. Therefore, random energy transfer causes systems to evolve to the state of minimum free energy.

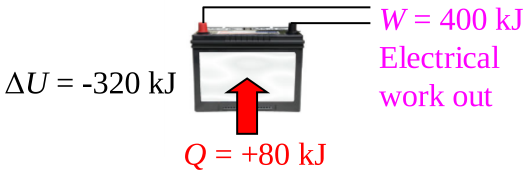
## What's Free About Energy?

Why is it called "free energy"? Perhaps a better term would be "available energy". Given a system with constraints, you can extract energy from it to do useful work, by letting the system evolve spontaneously to a more likely state. In doing so, its state changes from 'initial' to 'final'. How much useful work can you get out? I.e., how much of the system energy is "free" (available) to do useful work? It is the difference between the "free energy" of the initial state, and the "free energy" of the final state. If the initial state has 10 J of free energy, and the final state has 3 J of free energy, then the maximum work I can get done when the system evolves from initial to final state is 7 J.

But here's the really cool part: the difference between the *internal* energy of the system at initial and final states, in some cases, is *less* than 7 J [Sch p154].

The maximum useful work can be *more* than the internal energy lost
from the initial to the final state.

That's possible if the final state has higher entropy than initial, and hence the system absorbs heat energy during the process, some of which can then go out as useful work. This is exactly the case with a car battery: the useful (electrical) work out is *more* than the difference in internal energies of the chemicals before and after delivering electrical work, because the car battery entropy increases while discharging, so it absorbs some heat from the environment.



A car batter converts ambient heat to useful work,
thereby producing *more* work than the decrease in internal energy.

Part of that absorbed heat goes out as useful electrical energy [Sch p154]. This means that your car battery won't work as well if it's thermally insulated, because it can't absorb any heat. The available energy would be less than normal. [Reactions which absorb heat are called **endothermic**.]

Another way the useful work can be more than the change in internal energy is if the size of the system changes. If it shrinks, then the ambient does work on the system, some of which could find its way out as useful work. For a system at constant pressure, the Gibbs free energy accounts for this in the $+PV$ term.

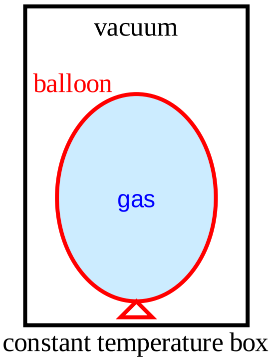
TBS: example of burning hydrogen from [Sch].

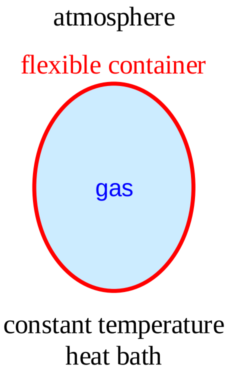
## Gibbs Free Energy and Enthalpy

Gibbs free energy is just Helmholtz free energy in disguise. We will see this by considering a sequence of 3 systems, each slightly different from the previous, starting with a system minimizing Helmholtz free energy, and inevitably ending up at a system minimizing Gibbs free energy. Thus, we will derive the Gibbs free energy from the Helmholtz free energy. Similar arguments apply to **enthalpy**, which is just the sum of internal energy plus the associated potential energy of $PV$ work [Wan p137t]:

$$H \equiv E + PV .$$

Consider a gas, in a rubber balloon, in a vacuum (below left):





Suppose our balloon can stretch, but for simplicity, always exerts constant pressure on the gas. If the gas expands inside the balloon, it does work on the balloon equal to $P\ dV$, which the balloon stores in its stretched rubber. The box is held at constant temperature, and exchanges energy with the balloon/gas system through thermal radiation (e.g., black body). Hence, the gas is also at the same temperature.

**Case 1:** As a first case, take our system to be the balloon and the gas. The internal energy of the system, $E$, then includes the kinetic energy of the gas, and the potential energy stored in the stretched rubber of the balloon, which equals $PV$. There is no other energy, so we have the simple case of Helmholtz free energy being minimized.

**Case 2:** Define the system as the gas inside, but not including the balloon. Changing our definition has no effect on the physics of anything, but now we would say that the $PV$ potential energy is *not* in our system. But clearly, when the volume expands, the gas particles lose kinetic energy through $PV$ work. When the gas contracts, the gas particles acquire energy from $PV$ work done on them. The $PV$ work is still a store of potential energy connected to the gas. Our choice of system boundary which excludes this potential energy store is perhaps, then, misguided, but we can make up for it by including the $PV$ potential energy "by hand":

$$\text{free energy} = E + PV - TS\ .$$

We can think of $(E + PV)$ as a sort of "effective internal energy."

**Case 3:** Replace the balloon with a flexible, but non-stretchy, gas bag, such as a hot-air balloon. Immerse the balloon in a constant pressure environment, such as the atmosphere. Define our system to be the gas in the balloon. (Including or not the balloon itself doesn't matter, since it has no energy.) In this case again, when the gas expands, it loses KE through $PV$ work done on the atmosphere. When it contracts, it acquires KE from the $PV$ work done by the atmosphere on the gas. The gas itself doesn't know or care where this $PV$ energy goes, or where it comes from. Hence, the behavior of our 3rd system is identical to that of both the 1st and 2nd systems. In this case, though, it's even harder to ascribe the $PV$ potential energy to the system itself. In fact, the $PV$ potential energy is stored in the gravitational potential energy of the air in the atmosphere, as the atmosphere is lifted by an expanding gas bag, and the atmosphere falls when the gas bag contracts. But the system behavior is the same as if the $PV$ work were potential energy of the system, so we again include it "by hand" as a kind of "effective internal energy." Since this case arises so often in the real world, we give it a special name, Gibbs free energy:

$$G \equiv E + PV - TS$$

## Generalized Free Energy

As a generalization beyond Gibbs free energy, we note that the *nature* of the "external" state-dependent energy, which is *effectively* internal energy of the system, does not matter. It could be $PV$ work, magnetic work, or any other kind of energy store which is returned to the system upon a change of some observable property [Wan p137m]:

$$free-energy = E + \sum_i F_i X_i - TS \qquad \text{[Reif p??]}$$

$$\Rightarrow \qquad d(free-energy) = dE + \sum_i F_i\, dX_i - T\, dS$$

$$where \quad F_i\, dX_i \text{ is arbitrary work done by the system, and stored as a function of the state}$$

Such a free energy is valid for fixed "forces" $F_i$: For Gibbs free energy, $P$ is fixed. For Helmholtz free energy, there is no state-dependent energy term, so we have the simplest case: just internal energy minus the heat from the bath, $F = U - TS$.

For Gibbs free energy, with fixed $P$, we have the state-dependent energy of $PV$ ($FX$ where $F \to P$, $X \to V$), which could have been considered part of the internal energy of the system, so $G = E + PV - TS$. The case of just $PV$ work is so common in real systems that its free energy has a special name: Gibbs free energy.

For any other free energy, we add internal energy plus all stores of potential energy, minus TS. For example, a magnetically permeable material responds to an applied magnetic field vector **H** with a magnetization vector **M**. The energy stored by the magnetic dipoles in **M** is:

$$\text{magnetic-energy} = -\mathbf{H}\bullet\mathbf{M}\ .$$

Therefore, we would define a free-energy for such a system, valid at fixed **H**, as:

$$F \equiv U + PV + (-\mathbf{H} \bullet \mathbf{M}) - TS$$

Free energies in general are the sum of the internal energy; plus potential energy terms (often confusingly considered "energy out"); minus TS, the contribution to internal energy that comes from the heat bath. When a system changes state, the available work comes from changes in the internal energy, heat energy that comes for "free" from the heat bath, and energy returned to the system from a potential energy store (e.g., *PV*, or magnetic).

More on free energy after the Partition Function, below.

### Free Energy Ladder

Each set of external constraints has only one meaningful "free energy" associated with it, i.e. some thermodynamic function which is minimized at equilibrium. Three common sets of constraints are these:

| **Constraints** | **Free energy** | **Comments** |
|---|---|---|
| Isolated system (no heat exchanged or work done) | $-S$ | Entropy is maximized, so $-S$ is minimized. This isn't really an "energy" (see Note). |
| Fixed temperature | Helmholtz: $A = E - TS$ | Heat is exchanged with heat bath. |
| Fixed temperature and pressure | Gibbs: $G = E - TS + PV$ | Heat is exchanged with heat bath, work can be done and stored in the ambient pressure. |
| Other (general) energy stores | $\text{free-energy} = E + \sum_i F_i X_i - TS$ | Heat is exchanged with heat bath, work can be done and stored in various state-dependent ways. |

**Note** $-S$ isn't really an "energy", but temperature is not relevant for a system not in equilibrium with a heat bath, so we don't define the free energy as $-TS$. We could instead consider $A = E - TS$ as the free energy of an isolated system, since $E$ is fixed, and $T$ doesn't matter.

Notice that entropy always has a negative sign in front of it, because increasing multiplicity (without changing other parameters) will increase likelihood, and therefore be closer to equilibrium. The minus sign means that because higher entropy is more likely, it is *lower* free energy that is more likely.

TBS: Grand free energy??

### How Much Energy Is Multiplicity Worth?

So how much energy is multiplicity worth? In other words, how much energy surplus can multiplicity overcome? We take as our measure of worth the occupation number: given an energy $E_1$ with multiplicity 1, and some other energy $E_2$ of multiplicity $g_2$, how much more energy can the higher state have, and still have the same occupation number as the lower state? Without loss of generality, we take $E_1 = 0$:

$$N_1 = e^{-\beta E_1} \cdot 1 = 1, \qquad N_2 = e^{-\beta E_2} \cdot g_2 = N_1 = 1 \quad \Rightarrow \qquad E_2 = \frac{1}{\beta} \ln g_2 \, .$$

Note that how much energy $g_2$ is worth depends on the equilibrium parameter $\beta$ (which we will see means that it depends on temperature). Also, for any $\beta$, the energy-equivalent of multiplicity goes as ln $g_2$. In other words, at some $\beta$, the occupation number of an energy of multiplicity $g$ is the same as a single multiplicity of energy (ln $g$)/β lower.

The free energy of a system starts with the internal energy, and de-rates it by the energy equivalent of multiplicity. In other words, it gives you an energy equivalent to tell you the occupation number due to a (multiple) energy level, as if it were a single-state energy level ($g$ = 1) of lower energy.

How does free energy become the maximum amount of useful work you can get from a state change?? Because it directly includes the amount of heat "borrowed" from the heat bath.

---

## Enthalpy

Enthalpy uses the same idea as the Gibbs free energy: the energy of a system at fixed pressure can be divided into "internal" energy, *U*, and the energy stored in the "atmosphere" which provides the fixed pressure. Note that the "atmosphere" doesn't have to be the earth's air; it could be a piston, or any other energy storage mechanism that depends on a "size" parameter of the system. In effect, this "external" energy store *must* be included as part of the energy of your system, but we write the "internal" energy and "external" energy as separate terms. For a typical system of varying volume in a fixed pressure, we have:

$$H \equiv U + PV .$$

and this is the true total energy of our system.

Enthalpy is useful for thermally insulated systems, where the heat flow is zero ($Q = 0$).
In this case, all the energy of the system is accounted for in the enthalpy.

An exam question provides a good example of enthalpy in action: Given 1g ($10^{-6}$ m$^3$) of liquid water at 20 C, force it through a thermally insulated porous plug at $10^4$ atm. The plug empties into a room at 1 atm. What happens to the water? (I.e., what is the final state of the water?) Assume that liquid water is incompressible (same volume at $10^4$ atm as at 1 atm).

Since the plug is insulated, there's no heat flow, and enthalpy accounts for all the energy of the system. Even though the pressure of the water changes as a result of the process, the pressure is fixed before, and fixed after, the process. Hence the enthalpy measures the total system energy in each case. We take as our initial state the water at $10^4$ atm. It's total energy is then:

$$H_i = U_i + P_i V_i = U(20\,C) + \left(10^4\,atm\right)\left(10^{-6}\,m^3\right) .$$

Since total energy is conserved, the final enthalpy equals the initial enthalpy:

$$H_f = H_i .$$

The $10^4$ atm does a lot of work on the water to force it through the plug, so we expect the water temperature to rise. How much does it rise? The work done to get it through the plug is ($10^4$ atm)($10^{-6}$ m$^3$), but on the other side, the water does (1 atm)($10^{-6}$ m$^3$) of work to push the atmosphere out of the way. The difference in work heats the water:

$$H_i = H_f \Rightarrow \qquad U_i + P_iV_i = U_f + P_fV_f \qquad \Rightarrow \qquad U_f - U_i = P_iV_i - P_fV_f .$$

But the change in internal energy is just from the temperature rise of the water:

$$U_f - U_i = C(\Delta T) \qquad where \quad C \equiv heat\ capacity = 4.186\ J/(g\text{-}C)$$

$$\Rightarrow \qquad \Delta T = \frac{(10^4\,atm)(10^{-6}\,m^3) - (1\,atm)(10^{-6}\,m^3)}{\left(4.186\ J/(g\text{-}C)\right)(1\,g)} \qquad Use\ 1\ atm \approx 10^5\ N/m^2$$

$$= 239\ C$$

Whoops. The liquid water will boil. So the final state is either a mixture of liquid/gas at 100 C, or pure steam at some higher temperature. We know the "heat of vaporization" of water at 1 atm is 540 kcal/g, or 2,260 J/g. Note that to vaporize water, we must add energy to do 2 things: (1) overcome the attractive potential between liquid water molecules, and (2) push the atmosphere out of the way to make room for the greater volume of steam. In other words,

"Heat of vaporization" is really "enthalpy of vaporization."

We must consider the process in at least 2 steps: first, raise the water temperature to 100 C, then convert some (or all) of the water to steam. As above, the work done is:

$$P_f V_f - P_i V_i = (10^9 \text{ N/m}^2)(10^{-6}\text{m}^3) - (10^5 \text{ N/m}^2)(10^{-6}\text{m}^3) = 1{,}000 \text{ J} .$$

Raising the water temperature requires (100 C - 20 C)(4.186 J/(g-C)) = 335 J, leaving 665 J to boil the water. It would take 2,260 J to boil the whole gram of water, so the process boils:

$$\frac{665 \text{ J}}{2260 \text{ J/g}} = 0.29 \text{ g of water.}$$

Thus the final state is 0.71 g of liquid water and 0.29 g of steam, both at 100 C.

---

## Schrodinger's Cat Says "mu": Chemical Potential

A large number of practical systems involve two or more source particles (reactants) combining or splitting up to produce two or more final (product) particles. For example:

$$A \rightleftarrows B + C, \qquad A + B \rightleftarrows C + D .$$

In chemistry, the "particles" are usually atoms or molecules. In cosmology, the particles may be primordial subatomic particles such as electrons, positrons, neutrinos, photons, etc.

When a system can have a variable number of particles in it, we consider the possibility of adding (or removing) one particle to the system, and ask: Is the system more likely to be found (in a macrostate) *with* this added particle, or without it? And how much more or less likely? Chemical potential answers these questions.

This section assumes a thorough understanding of macrostate probability and free energy (discussed earlier).

**Very brief review of free energy:** A system at constant temperature can potentially be in one of a large number of macrostates. The probability of a given macrostate depends on both its energy (due to the Boltzmann distribution), and its multiplicity (entropy): macrostates with more microstates (i.e., higher multiplicity = more entropy) are more likely. Free energy is a way of adjusting the true energy to account for multiplicity, so that:

Free energy is a direct measure of the macrostate relative probability,
on a logarithmic scale, and with a minus sign:

$$\text{RelPr}(\textit{macro}\text{state with free energy } A) = e^{-\beta A} \quad \textit{where} \quad A \equiv \text{appropriate free energy for the system} .$$

Recall that free energy also includes any energy stored as a byproduct of the macrostate, i.e. "reversible work" such as the $pV$ work in a constant pressure system, or work from surface tension. Because of this, different system constraints require different definitions of free energy. Any given system has only one relevant definition of free energy. The most common examples of free energy are: (1) a system at constant temperature that does no work (e.g., constant volume) is characterized by *Helmholtz free energy* ($F \equiv E - TS$); and (2) a system at constant temperature and constant pressure (which therefore includes the reversible $pV$ work) is characterized by *Gibbs free energy* ($F \equiv E + pV - TS$). NB: Free energies can only be compared between systems at the same temperature.

The chemical potential of a particle in a system is the *free energy cost* of adding the particle:

$$\mu \equiv \frac{\partial A}{\partial N} \quad \textit{where} \quad A \equiv \text{appropriate free energy for the system} .$$

Thus, we can say $\mu$ measures the relative probability of adding a particle to that system. Note that free energy $A$ usually includes an arbitrary additive constant for its energy term, and often *another* arbitrary additive constant for its entropy. However, since $\mu$ is a *difference*, the additive constants cancel, and therefore:

The chemical potential is an absolute quantity, without an additive constant.

(For some applications, however, only differences in chemical potential are important, and such applications would tolerate an additive constant in $\mu$. See below.)

Many systems are a mixture of different kinds of particles, which we label here with the subscript *i*. Each kind of particle has its own chemical potential:

$$\mu_i(T, N, ...) \equiv \frac{dA(T, V, N_i, ...)}{dN_i}, \qquad \textit{where} \qquad A(T, V, N_i, ...) \equiv \text{relevant free energy for the system} .$$

For example, for a system at fixed *T* and *V*:

$$\mu(T, N, V) \equiv \frac{dF(T, N, V)}{dN} = \left(\frac{\partial F}{\partial N}\right)_{T,V}, \quad \textit{where} \qquad \text{[Rei 8.7.10 p313]}$$

$$F(T, N, V) \equiv U(T, N, V) - TS(T, N, V) \qquad \text{is the Helmholtz free energy for the system} .$$

For a system at fixed *T* and *P*:

$$\mu(T, N, p) \equiv \frac{dG(T, N, P)}{dN} = \left(\frac{\partial G}{\partial N}\right)_{T,p}, \quad \textit{where} \qquad \text{[Rei 8.7.12 p313]}$$

$$G(T, N, p) \equiv U(T, N, p) + PV(T, N, p) - TS(T, N, p) \qquad \text{is the Gibbs free energy for the system} .$$

The chemical potential is the free-energy cost of adding a particle to the system.

Therefore, chemical potential includes both real energy and system multiplicity as a function of *N*, the number of particles of a species. The chemical potential can be positive or negative. If $\mu > 0$, the system is more likely to be found *without* the particle, because adding it would *increase* the system free energy, and removing such a particle *decreases* the free energy. If $\mu < 0$, the system is more likely to be found *with* the particle, because adding it would *decrease* the free energy.

The chemical potential for a particle is a function of the macrostate, just like particle density, or pressure.

The chemical potential is *not* an intrinsic property of a particle.

Sometimes, the chemical potential is misleadingly described as the "energy cost" of creating a particle. However, it is the *free* energy cost, which is different.

As previously noted, "probability rules," not energy; equivalently, "free energy" rules (not energy). We can think of the chemical potential of a particle as the "probability cost" of adding a particle, measured by free energy.

For example, consider a monatomic ideal gas at constant temperature (*not* constant entropy), and constant volume. If we add an atom, the system energy increases because the particle must acquire a kinetic energy of $(3/2)kT$. That is the true energy cost. *However*, the entropy of the system *also* increases, which means some heat will flow in from the thermal bath to the system (entropy attracts heat). So the system doesn't lose probability according to the full energy cost of adding the particle: the heat inflow offsets some of the energy cost. I.e., the entropy increase results in a heat inflow:

$$dQ = T\, dS .$$

So the probability cost of creating the particle (in units of energy) is only the *difference* between the true energy increase and the energy inflow from the heat bath:

$$\text{probability cost} \leftarrow dE - TdS \equiv dA \qquad \textit{where} \quad dA \equiv \text{the change in free energy} .$$

Another way to think of this is: chemical potential for a particle is the real energy cost ($dE$) minus the entropy benefit ($T\,dS$). Once again, this includes the "multiplicity defies energy" principle for the particle.

For many systems, especially chemical ones, the pressure is fixed, and the volume varies. We have seen previously that this leads to Gibbs free energy: $G \equiv E + pV - TS$. This is because the work done by the system expanding against the pressure (1) is a *store* of energy, (2) is a function of the system state, and (3) is returned when the system contracts. Therefore, that work can be included as part of the "total energy" of the system [Wan p137t]. The chemical potential is always defined as the change in free energy per particle, using the one-and-only free energy that is appropriate for the constraints of the system. For constant pressure, then, the chemical potential is:

$$\text{probability cost of one particle} = dE + p\,dV - T\,dS \equiv dG \qquad \rightarrow$$

$$\mu \equiv \left(\frac{dG}{dN}\right)_{T,p} \qquad \text{(system at constant } T, p\text{)}.$$

But for constant pressure and temperature, $G$ is extensive, i.e. is proportional to $N$. In this case, the chemical potential is the same for any sized system (i.e., any number of particles), because any sized system has the same particle density, and temperature. Each particle "feels" the same environment, no matter how many other particles are in the system. Therefore, for a system of just one particle type:

$$\mu = G / N \qquad \text{(system at constant } T, p\text{)}. \qquad (3.1)$$

This situation of each particle "feeling" the same environment comes about any time we fix two *intensive* properties (in this case, pressure and temperature).

Note that in this case, $G$ for the system may include an arbitrary additive constant for the energy of each particle, and possibly for the entropy. Therefore, the $\mu$ in (3.1) is *not* absolute, since it includes any per-particle additive constants.

The chemical potential is sometimes described as the "Gibbs free energy per particle." That is misleading on several counts.

For one, it confuses the fact that chemical potential is the free energy cost of adding a particle, using the one-and-only free energy that is relevant for the system at hand. In particular, Gibbs free energy is only meaningful for systems constrained by constant temperature and pressure. For systems with other constraints, the Gibbs free energy is meaningless. For example, systems that include other forms of reversible work, such as surface tension or magnetic work, use a different free energy. And finally, for systems of multiple particle types, each type has its own chemical potential. The definition above says nothing about how to divide up the Gibbs free energy among different particle types, and assign each type some piece of the total Gibbs free energy.

**Maxwell relations for chemical potential:** TBS.

## Comments on the Chemical Potential

At least one reference says $\mu$ is due to inter-molecular forces. The potential energy of such forces are sometimes part of $\mu$, but such forces are *not* required.

In fact, in [Car] problem 9.2, he asks you to compute the chemical potential of an ideal gas. An *ideal* gas, by definition, has no inter-molecular forces. If chemical potential were due to inter-molecular forces, an ideal gas particle could not have $\mu \neq 0$.

As a free energy, the chemical potential includes kinetic energy, potential energy, possibly other energies, and entropy:

$$\mu = kinetic + U - TS \qquad \text{for one particle}.$$

However, recall that the potential energy $U$ and the entropy $S$ both include arbitrary zero points. If we change the zero of either, we also change the zero of $\mu$, at least, for any particle affected by the change. For example, if we add an arbitrary constant to an electrostatic potential energy, then all charged particles have their $\mu$ changed by that amount. However, the $\mu$ for photons and neutral particles such as atoms, molecules, and neutrons, are not affected.

## The Partition Function

The partition function is an odd beast. To my knowledge, it has only one worthwhile use, which is to compute (as a highly accurate approximation) the free energy of a system. But from the free energy, we get all other thermodynamic properties. The partition function is also sometimes used as the normalization constant for the relative probabilities of microstates, and to compute the average energy for a system. These latter two uses are trivial, though, and simply restate basic definitions that don't require the bother of defining any partition function.

The partition function is defined for all systems for which the Boltzmann distribution applies, from single atoms to macroscopic systems. In other words, the partition function is defined for all systems in thermal equilibrium with a heat bath.

This section relies on the Boltzmann distribution. Recall that the Boltzmann distribution gives the relative probability of a system, in thermal equilibrium with a heat bath, to be in a *single* microstate of given energy (see "Nuts and Boltzmann," earlier):

$$\text{RelPr}(\text{microstate of energy } E_s) = \exp(-E_s / kT) .$$

The partition function is also called the "sum of states," which so far as I can tell, means absolutely nothing. It should be called the "sum of relative probabilities of all the microstates," but that doesn't roll off the tongue so well.

We proceed on these lines:

1. Partition function as a normalization factor
2. Using the partition function to compute average energy and such
3. Partition function tells all
4. Partition function and free energy

**The partition function as a normalization factor**: As shown in the section on the Boltzmann distribution, $Z(\beta)$ is the sum of relative probabilities of all the microstates, or equivalently, the sum of the relative probabilities of each energy value. It is defined only for a system in thermal equilibrium (which means it can exchange energy with a temperature bath). For a system with discrete energies:

$$Z(\beta) \equiv \sum_{s=1}^{\textit{all states}} \underbrace{\exp\left(-\beta E_s\right)}_{\text{RelPr(state } s)} = \sum_{j=1}^{\substack{\textit{all energy}\\ \textit{levels}}} \underbrace{g_j \exp\left(-\beta E_j\right)}_{\text{RelPr(energy } E_j)}, \qquad \textit{where} \qquad g_j \equiv \text{multiplicity of energy } E_j .$$

This readily generalizes to a system with continuous energies:

$$Z(\beta) \equiv \int_{E-\min}^{\infty} dE\ \underbrace{g(E)\exp(-\beta E)}_{\text{RelPr(energy } E)} \quad \textit{where} \quad g(E) \equiv \text{energy density of states} . \qquad (3.2)$$

Note that for a system held in a heat bath, the system energy is a variable (*not* a constraint). The system may have more energy than the average, or less, all the way down to zero. That's very unlikely for most systems, but for mathematical expedience, it is still counted as a possibility in $Z(\beta)$.

Since the partition function is a sum of relative probabilities,

The partition function includes an arbitrary multiplicative constant.

This arbitrariness also reflects the arbitrary zero of energy. Changing the zero-energy simply multiplies the partition function by a constant, with no effect on the physics.

The partition function serves as the (inverse of the) normalization factor to convert the relative probability of the system being in some microstate, or having some energy, into an absolute probability:

$$\Pr(\text{state } r) = \frac{\text{RelPr}(E_r)}{\sum_{s=1}^{all\ states} \text{RelPr}(\text{state } s)} = \frac{\exp(-\beta E_r)}{Z(\beta)}$$

$$\Pr\left(\text{energy} = E_k\right) = \frac{\text{RelPr}(E_k)}{\sum_{j=1}^{all\ energy\ levels} \text{RelPr}\left(E_j\right)} = \frac{g_k \exp(-\beta E_k)}{Z(\beta)}$$

But this is trivial, and simply uses the letter $Z$ for something that is already well known.

**Partition function for average energy, and such:** Another use for $Z(\beta)$ is to compute average energy. (This is also trivial, and provides no new physical insight.) Recall from basic statistics that for any discrete random variable, say $E$:

$$\langle E \rangle = \sum_i \Pr(E_i)\, E_i \ . \qquad (3.3)$$

Therefore, for the energy of a system, we have:

$$\langle E \rangle = \sum_{s=1}^{all\ states} \Pr(E_s)\, E_s = \sum_{s=1}^{all\ states} \frac{\text{RelPr}(E_s)}{Z(\beta)} E_s = \frac{1}{Z(\beta)} \sum_{s=1}^{all\ states} e^{-\beta E_s} E_s \ .$$

But purely as a mathematical trick, with no deep physical meaning, we notice that the factor:

$$e^{-\beta E_s} E_s = -\frac{\partial}{\partial \beta} e^{-\beta E_s} \qquad \Rightarrow$$

$$\langle E \rangle = \frac{1}{Z(\beta)} \sum_{s=1}^{all\ states} -\frac{\partial}{\partial \beta} e^{-\beta E_s} = \frac{-1}{Z(\beta)} \frac{\partial}{\partial \beta} \sum_{s=1}^{all\ states} e^{-\beta E_s} = \frac{-1}{Z(\beta)} \frac{\partial}{\partial \beta} Z(\beta)$$

This gives us nothing new, simply using $Z$ for something already well-known. Since the partition function includes all the terms needed for the fundamental formula for average energy, as in Eq (3.3) above, we could have more easily computed the average energy from the fundamental formula. I don't know why anyone bothers tying average energy to the partition function.

As another trick, notice that:

$$\frac{-1}{Z(\beta)} \frac{\partial}{\partial \beta} Z(\beta) = -\frac{\partial}{\partial \beta} \ln Z(\beta) \ .$$

This has the advantage, however, that $\ln Z(\beta) \equiv$ free energy, so there is some physical significance to this. Notice that this leads directly to a thermodynamic identity:

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln Z(\beta) = -\frac{\partial}{\partial \beta} A(\beta) \qquad where \quad A(\beta) \equiv \text{free energy of the system} \ .$$

**Partition function tells all:** In the end, given the partition function $Z(\beta)$ for all values of $\beta$, we can compute every (thermodynamic) thing there is to know about the system. That is, a system's thermodynamics is fully characterized by its partition function. Recall that a system is also fully characterized by its energy density of states. Therefore, knowing the partition function and knowing the density of states are equivalent. How are the two related? Recall the partition function for continuous systems (3.2):

$$Z(\beta) \equiv \int_{E-\min}^{\infty} dE\ g(E) \exp(-\beta E) \quad where \quad g(E) \equiv \text{energy density of states} \ .$$

This is the definition of the Laplace transform of $g(E)$, with transform variable $\beta$. Recall the Laplace transform is *uniquely* related to the original function. Therefore:

The partition function $Z(\beta)$ is the Laplace transform of the density of states $g(E)$,
so knowledge of one is equivalent to knowledge of the other.
Either one fully defines the thermodynamic properties of the system.

However, in practice, you usually need to know $g(E)$ to find the partition function, so this Laplace transform relationship isn't usually very useful. [Pathria's example where it is useful??]

**Example: Thermodynamic zipper**

A horizontal zipper has $N$ links. Each link has two states: closed with energy 0, and open with energy $\varepsilon$. The zipper can only unzip from the left end and the $s^{\text{th}}$ link cannot open unless all of the links to its left $(1, 2, ..., s-1)$ are already open.

(a) Find the partition function for the zipper.

(b) Find the mean number of open links. These kinds of questions always mean "as a function of temperature."

(c) Evaluate your result in both the low and high temperature limits. (The high temperature limit is tricky.) Explain conceptually each result of low, middling, and high temperatures.

**Answer:** (a) We use the variable $s$ as the state variable, defined as number of open links. $s$ runs from 0 to $N$. The system energy is a function of $s$, $E(s) = s\varepsilon$. Then the partition function is:

$$Z(\beta)=\sum_{k}^{\text{all states}}\underbrace{\exp\left(-\beta E_k\right)}_{\substack{\text{Relative Pr}\\ \text{of state } k}}=\sum_{s=0}^{N}\exp\left(-\beta s\varepsilon\right)=\sum_{s=0}^{N}\left[\exp\left(-\beta\varepsilon\right)\right]^{s}. \qquad \text{Use } \sum_{i=0}^{n-1} ar^i=a\frac{1-r^n}{1-r}$$

$$Z(\beta)=\frac{1-\exp\left(-(N+1)\beta\varepsilon\right)}{1-\exp\left(-\beta\varepsilon\right)}$$

This is exact, and the magnitude of $N$ is irrelevant.

(b) $$\langle s\rangle=\frac{\langle E\rangle}{\varepsilon}=\frac{-1}{\varepsilon}\frac{\partial}{\partial\beta}\ln\left(Z(\beta)\right)=\frac{-1}{\varepsilon}\frac{\partial}{\partial\beta}\left[\ln\left(1-\exp\left(-(N+1)\beta\varepsilon\right)\right)-\ln\left(1-\exp\left(-\beta\varepsilon\right)\right)\right]$$

$$=-\frac{1}{\varepsilon}\left[\frac{(N+1)\varepsilon\exp\left(-(N+1)\beta\varepsilon\right)}{1-\exp\left(-(N+1)\beta\varepsilon\right)}-\frac{\varepsilon\exp\left(-\beta\varepsilon\right)}{1-\exp\left(-\beta\varepsilon\right)}\right]$$

$$=-\frac{(N+1)\exp\left(-(N+1)\beta\varepsilon\right)}{1-\exp\left(-(N+1)\beta\varepsilon\right)}+\frac{\exp\left(-\beta\varepsilon\right)}{1-\exp\left(-\beta\varepsilon\right)}$$

(c) We expect at low temperature the zipper is fully closed, since the thermal energy is too low to open even one link. Mathematically, as $T \to 0$, $\beta \to \infty$, $\exp(-\beta\varepsilon) \to 0$:

$$\lim_{T\to 0}\langle s\rangle=-\frac{0}{1}+\frac{0}{1}=0\,,$$

as expected.

The high temperature limit turns out to be more complicated than I expected, and you can skip this if you want. At high temperature, there is unlimited energy, and we might expect every link to open. As $T \to \infty$, $\beta \to 0$, $\exp(-\beta\varepsilon) \to 1$. The denominators go to zero in $\langle s \rangle$, but a 1st order expansion of $(1 - \exp(-x)) = x + O(x^2)$ resolves the problem:

$$\lim_{T\to\infty}\langle s\rangle=-\frac{N+1}{(N+1)\beta\varepsilon}+\frac{1}{\beta\varepsilon}=0\,.$$

Whoa! Can this be right? Yes. Consider the temperature of our zipper, as a function of energy (from (2.1)):

$$T(E) \equiv \left(\frac{dS}{dE}\right)^{-1}, \quad S(E) \equiv k \ln\big(g(E)\Delta E\big) \qquad \textit{where} \qquad \Delta E \equiv \textit{const} << \text{measurable energy change} .$$

Strictly speaking, temperature is only defined for a continuous $g(E)$. If $N$ is large, we can take the states as continuous, and to define a temperature, we take $\Delta E \sim \varepsilon$. The approximate (or "smoothed") density of states is $g(E) = 1/\varepsilon$, since there is only one state of any given energy, and energies are spaced by $\varepsilon$. (We'll analyze the high-temperature limit shortly in a different way.) There is only one microstate for each energy $E$, so:

$$S(E) = k \ln 1 = 0 \quad \text{(constant)}, \qquad \frac{dS}{dE} = 0, \qquad T(E) = \left(\frac{dS}{dE}\right)^{-1} \to \infty .$$

The temperature of the zipper is hotter than any heat bath at finite temperature. It will *always* transfer energy from the zipper to the heat bath, until it has no more energy to give. Thus $<s> = 0$.

A different analysis avoids any zipper temperature by directly considering the entropy of the heat-bath/zipper combined system. It is an isolated system, and so equilibrates in the macrostate of maximum entropy. The entropy of the zipper is fixed at 0, so the maximum entropy of the system occurs at the maximum entropy of the heat bath. The heat bath has finite (and positive) temperature, so its maximum $S(E)$ is at maximum $E$. Thus all the energy leaves the zipper, and goes into the heat bath. The zipper has no open links, as above. But there is still more to the story.

The careful reader must now wonder about $<s>$ at middling temperatures. The entropy argument above suggests that energy will *always* leave the zipper until its energy is zero. How does that square with $<s>$ found in part (b), which is sometimes positive? To illustrate the principle, consider a zipper of only 3 links, at a temperature such that $\exp(-\beta\varepsilon) = ½$. Then:

$$\text{RelPr}(s=0) = 1, \qquad \text{RelPr}(1) = 1/2, \qquad \text{RelPr}(2) = 1/4, \qquad \text{RelPr}(3) = 1/8 .$$

$$Z \equiv \sum_{i=0}^{3} \text{RelPr}(i) = 15/8, \qquad \Pr(s) = \text{RelPr}(s)/Z \qquad \Rightarrow$$

$$\Pr(0) = 8/15, \qquad \Pr(1) = 4/15, \qquad \Pr(2) = 2/15, \qquad \Pr(3) = 1/15 .$$

This shows that the *most likely* state is that of zero energy, but there are always thermal fluctuations. In this case, the probability of an excited state is nearly ½. For $N >> 1$, the probability of an excited state approaches ½. In other words, when $kT \sim \varepsilon$, the probability of an excited zipper is finite. In the high-$T$ limit, though, the entropy of the heat bath is huge, and essentially draws all the energy into it, leaving $<s> \to 0$.

**Summary:** At low $T$, $<s> = 0$, because there is no energy to open a link. At high $T$, $<s> = 0$ because the entropy change of the heat bath for even one link's worth of energy is so huge, it overwhelms the energy fluctuations in the zipper, which approach zero. In the middle, the energy fluctuations of the zipper are substantial, and therefore the average energy (number of open links) is finite.

Perhaps I should find a simpler example for a partition function.

## Free Energy and the Partition Function

When considering the possible macrostates of a system, what really matters is the *most-probable* macrostate of a system, and to be meaningful, it must be the *only* macrostate with any significant chance of occurring. To compute free-energy from the partition function, we use the general conclusion that,

For macroscopic systems,
the partition function is dominated by the one most-probable macrostate.

Why do we bother with a partition function? If we know the most-probable state, shouldn't we just use that? Yes. But often we don't know the most-probable state. And mathematically, it may be easier to compute the complete sum $Z \equiv$ sum-of-(all)-state-relative-probabilities [Rei p219b]. In the thermodynamic limit (macroscopic systems), the most-probable state and the complete sum are equivalent (to 10 orders of magnitude or so), so we do whatever's easier.

The equilibrium macrostate of a system is simply the most-likely macrostate in thermal equilibrium; this, in turn, is that macrostate with the largest number of likely microstates, i.e. largest product of (Boltzmann-factor) × (#-microstates):

$$E_{most-likely} = \max_{j=all\ energies}\left(g_j e^{-\beta E_j}\right).$$

Therefore in principle, we can find the equilibrium macrostate by first identifying all the microstates, their probabilities, and the macrostate for each, and pick that macrostate with the largest number of likely microstates. This is only meaningful if the most likely macrostate has probability near 1, so that all other macrostates are *so unlikely* they can be ignored. For systems with a large number of subsystems, $N$ (e.g., large number of particles), this condition is true: only one macrostate has any reasonable chance of occurring.

In practice, though, finding the "most likely" thing often has no closed-form mathematical expression. But we like formulas, so we use an approximation to the most-likely macrostate. For systems of large $N$, almost all of the likely microstates are crowded very near the most-likely macrostate; therefore, it's usually more convenient to simply sum the relative probabilities of *all* the microstates, as an approximation to the number of likely microstates near the most-likely energy. This often allows closed-form mathematical answers, which would otherwise be mathematically untenable.

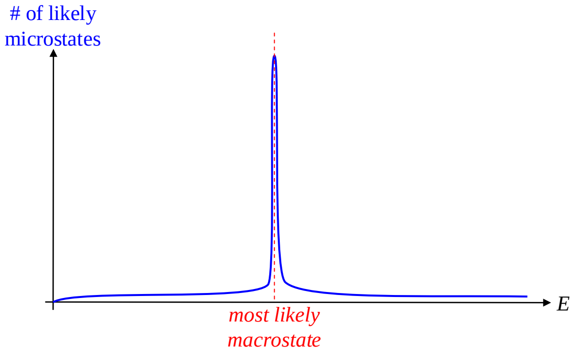


**Figure 3.1** Number of likely microstates of a system as a function of energy

We now show that for systems with a large number of component subsystems (systems of large $N$), there is a simple approximation to the Helmholtz free energy of the system from the partition function:

$$A(\beta) \approx -kT \ln Z(\beta) \qquad \text{(large system)}.$$

Here's why: for large $N$, there will be only a few terms in the partition function which dominate not only all the other terms, but all the other terms *combined*. These few terms all have very nearly the same energy, so they can be lumped together as a single multiplicity for that energy:

$$Z(\beta) = \sum_{j=1}^{\substack{all\ energy\\ levels}} g_j \exp\left(-E_j / kT\right) \approx \sum_{m}^{\substack{states\ near\\ maximum}} g_m \exp\left(-E_{max} / kT\right),$$

$$where \quad E_{max} \text{ is the dominant energy in the sum}$$

This means that the probability of the system being at $E_{max}$ is essentially 1:

$$\Pr(E_{max}) = \frac{\displaystyle\sum_{m}^{\substack{\text{states near}\\ \text{maximum}}} g_m \exp\left(-E_{max}/kT\right)}{Z(\beta)} \approx \frac{Z(\beta)}{Z(\beta)} = 1 .$$

This is as we expect: for systems of large $N$ in thermal equilibrium, there is only 1 realistic macrostate.

Now if $Z(\beta)$ has only one significant energy $E_{max}$, then we can rewrite $Z(\beta)$ or $Z(T)$ to follow the reasoning of free energy. We can replace the sum of the few $g_m$ near $E_{max}$ with a single $g_{eff}$, since they all have essentially the same energy, $E_{max}$.

$$\begin{aligned}
Z(T) &\approx g_{eff} \exp\left(-E_{max}/kT\right) && \textit{where } E_{max} \text{ is the dominant energy level in the sum} \\
&= \exp(\ln g_{eff}) \exp\left(-E_{max}/kT\right) && \text{writing } g_{eff} \text{ as an exponential} \\
&= \exp\left(kT \ln g_{eff} / kT\right) \exp\left(-E_{max}/kT\right) && \textit{where } kT \ln g_{eff} \text{ writes } g_{eff} \text{ in terms of energy} \\
&= \exp\left(-\frac{E_{max} - TS(E_{max})}{kT}\right) && \textit{where} \quad S(E_{max}) \equiv k \ln g_{eff} \\
&= \exp\left(-\frac{A(T)}{kT}\right) && \textit{where} \quad A(T) \equiv \text{free energy of the macrostate at temp } T
\end{aligned}$$

Once again, we're simply quantifying how multiplicity defies energy, as before, and measuring multiplicity in units of energy. Solving for $A(T)$ yields the result given above:

$$A(T) \approx -kT \ln Z(T) \qquad \text{(large system)} .$$

Note that since $Z$ is defined only up to a multiplicative constant, $A(T)$ is defined only up to an additive constant, which is always true for energy (free or otherwise).

Note that for small systems, such as atoms, "temperature" is not defined, and therefore, neither is free energy. The above approximation only works for large systems, where the thermal equilibrium macrostate is sharply defined.

**What kind of free energy?**

The partition function is calculated for a given set of constraints, just like the various kinds of free energies.

For a system constrained as for Helmholtz free energy (fixed $T$, no work done), we compute $Z(T)$ within those constraints, and $A = -kT \ln Z$ gives Helmholtz free energy. If the system had different constraints, such as those of Gibbs free energy (fixed $T$ and $P$), then $Z(T)$ is different, and $G = -kT \ln Z$ gives Gibbs free energy. And so on.

## Big Partition Functions From Little Ones

To effectively use the partition function to compute free energies, we must know how to construct big partition functions from little ones. It is quite common to have a system comprising a large number of simple subsystems, e.g. a gas of particles. We know how to find the partition function for a single subsystem with known energy states. Is there a simple relationship between the 1-particle partition function and the $N$-particle partition function? Of course there is.

We start with a simple case: we define a new system as the combination of two smaller subsystems, A and B. It doesn't matter if they are identical or not, so let's be general, and assume they are different. A microstate of the combined system is defined by specifying the two microstates of the subsystems. If the two subsystems interact, and their energies depend on each other's states, then we can't compute anything. But if the two subsystems are independent, we have a simple result: the probability of being in two given microstates of the subsystems, which is the same as one microstate of the combined system, is just the product of the two probabilities:

$$\begin{aligned}&\text{RelPr}(A\ in\ S_A\ with\ energy\ E_A,\ and\ B\ in\ S_B\ with\ energy\ E_B)=\text{RelPr}(A\ in\ S_A)\text{RelPr}(B\ in\ S_B)\\ &\qquad=e^{-E_A/kT}e^{-E_B/kT}=e^{-(E_A+E_B)/kT}\end{aligned}$$

This is nothing more than the statement that the combined system, with energy $E_A + E_B$, follows the Boltzmann distribution. [It almost seems like a way to *derive* the Boltzmann distribution.]

A few short steps of algebra then shows [Wan p86t]:

Independent distinguishable subsystems combine to give a partition function that is the product of the individual partition functions.

$$\begin{aligned}Z_{AB}(\beta)&\equiv\sum_r^{states\ of\ A}\ \sum_s^{states\ of\ B}\exp\left(-\left(E_r+E_s\right)/kT\right)\\ &=\sum_r^{states\ of\ A}\ \sum_s^{states\ of\ B}\exp\left(-E_r/kT\right)\exp\left(-E_s/kT\right)\\ &=\left(\sum_r^{states\ of\ A}\exp\left(-E_r/kT\right)\right)\left(\sum_s^{states\ of\ B}\exp\left(-E_s/kT\right)\right)=Z_A(\beta)Z_B(\beta)\end{aligned}$$

This immediately gives, for a system of $N$ identical, but distinguishable, subsystems,

$$Z_N(\beta)=\left(Z_1(\beta)\right)^N \qquad \text{(distinguishable subsystems)}.$$

If the subsystems are identical, such as atoms, then strictly speaking, we should divide by $N!$, to account for the $N!$ ways that the $N$-particle quantum state has been overcounted in the above equation (i.e., we should use correct Boltzmann counting):

$$Z_N(\beta)=\frac{\left(Z_1(\beta)\right)^N}{N!} \qquad \text{(identical subsystems)}.$$

This will be critically important later, for the grand partition function. If $N$ is fixed, though, then $N!$ is simply a multiplicative constant. Since the partition function is defined only up to a multiplicative constant, the $N!$ has no effect when $N$ is fixed.

A microscopic individual subsystem, such as an atom, is too small to have a temperature or free energy, but a large number of such subsystems (e.g., gas of atoms) *do* have a temperature and a free energy. This gives the simple result that:

$$A=-kT\ln Z_N(\beta)=-NkT\ln Z_1(\beta).$$

Here again, for identical subsystems, there would be a ($-\ln N!$) term, but that is an additive constant to the (free) energy. Since energy and free energy are only defined up to an additive constant, the ($-\ln N!$) term can be dropped. The important point is that:

$\ln Z_1(\beta)$, for a single small subsystem, doesn't mean anything, because the subsystem is too discrete to have a temperature or free energy. But $N \ln Z_1(\beta)$ *is* meaningful for large $N$, because it describes a system with a nearly continuous energy density of states.

For example, a normal six-sided die has an average roll of 3.5. If I ask you to predict a single roll of a die, knowing the average tells you nothing. However, if I ask you to predict the sum of a billion rolls, knowing the average tells you a lot [Wan p11m].

# 4 The Ensembles

Whole chapter under construction.

There are 3 main classes of systems we consider, distinguished by their constraints. An **ensemble** is a (possibly imaginary) set of systems in identical macro states. Here, we use

$N$ # of particles (or moles, or quantity) of a system

$E$ energy of a system

| | |
|---|---|
| Micro-canonical ensemble (MCE) | $N$, $E$ fixed |
| Canonical ensemble (CE) | $N$ fixed, $E$ variable |
| Grand canonical ensemble (GCE) | $N$, $E$ both variable |

Ergodicity applies to all the ensembles, i.e. the time average of any observable (macroscopic) quantity equals the ensemble average. This is equivalent to saying that given enough time, the system would visit every possible microstate. (However, in practice, it would take inconceivably many multiples of the age of the universe for this to happen.)

## Micro-canonical Ensemble

A **micro-canonical ensemble** is a set of systems, where each is completely isolated from its surroundings: no energy or particle exchange, no work done. The micro-canonical ensemble is an idealization used to develop methods of calculating thermodynamic quantities from microscopic physics. [I'm not sure the "ensemble" idea is a good one.]

## TBS.

System of (nearly) fixed energy [Hua p176b, Kub p7b], volume, particle number. Stop appealing to "classical" thermodynamics; pretend we don't know thermodynamics (stretch). $T$ is defined and it is *intensive* (see Introduction to Entropy, above).

## (Ordinary) Canonical Ensemble

A **canonical ensemble** is a set of systems, each in thermal equilibrium with a large heat bath at a fixed temperature $T$. This means heat can be exchanged with the heat bath, but no other exchanges occur: no particles or work.

Many works have the systems in an ensemble exchanging stuff with each other. This is *not* necessary, and I think confuses the real issues. Such works have a single system in the ensemble use all the other systems in the ensemble as their heat bath. However:

> The nature of the heat bath is irrelevant, and in particular, it need *not* be other identical systems. Only its temperature matters.

TBS. Can't fix energy.

$$Z(\beta,V,N)=\frac{1}{N!h^{3N}}\int_{\infty} d^{3N}p\,d^{3N}q\,e^{-\beta\mathcal{H}(p_i,q_i,N)} \qquad \text{[Hua p149]} \qquad \textit{or}$$

$$Z(\beta,V,N)=\sum_{\{n_i\}}\delta\left(\sum_i n_i-N\right)e^{-\beta E(\{n_i\})}, \quad \textit{where} \quad E(\{n_i\})=\sum_i \varepsilon_i n_i$$

Helmholtz free energy: $A(N,V,T)=U-TS=-kT\ln Z \qquad \mu\equiv\frac{\partial F}{\partial N} \qquad P=-\frac{\partial F}{\partial V}$ .

## Grand Canonical Ensemble

A **grand canonical ensemble** is a set of systems, each in thermal equilibrium with a large heat bath at a fixed temperature *T*, and which can also exchange particles with the heat bath. Thus heat and particles can both be exchanged with the heat bath, but no other exchanges occur: no work.

Many works have the systems in an ensemble exchanging stuff with each other. This is *not* necessary, and I think confuses the real issues. Such works have a single system in the ensemble use all the other systems in the ensemble as their heat and particle bath. However:

> The nature of the heat and particle bath is irrelevant, and in particular, it need *not* be other identical systems. Only its temperature and chemical potential matter.

TBS.

$$\mathcal{Z}(\beta,V,\mu)=\sum_{N=0}^{\infty}\sum_{s=0}^{\infty}e^{-\beta\left(E_s-\mu N\right)} \qquad \textit{where} \qquad s \text{ labels the microstates} \qquad \text{[Sch 7.7 p258]}$$

$$\ln\mathcal{Z}=\beta PV$$

Though I think it's confusing, the grand partition function is sometimes written in terms of "fugacity," *z*, by pulling out the chemical potential factor $z=\exp(\beta\mu)$:

$$\mathcal{Z}(\beta,V,\mu)=\sum_{N=0}^{\infty}e^{\beta\mu N}Z(\beta,V,N)=\sum_{N=0}^{\infty}z^N Z(\beta,V,N) \qquad \textit{where} \qquad z\equiv e^{\beta\mu}\ .$$

This hides the relative nature of the energies with respect to the chemical potential.

Classically:

$$Z(\beta,V,N)=\underbrace{\frac{1}{N!}}_{\text{indistin-guishable}}\ \frac{1}{h^{3N}}\int_{\infty}d^{3N}p\,d^{3N}q\,e^{-\beta\mathcal{H}(q,p)}$$

$$\rho(p_i,q_i,N)=\frac{1}{N!h^{3N}}e^{\beta\mu N}e^{-\beta PV}e^{-\beta\mathcal{H}(q,p)}=\frac{1}{N!h^{3N}}e^{-\beta\left(PV+\mathcal{H}(q,p)-\mu N\right)}$$

If the particles are distinguishable, the factors of *N*! are not present.

Quantum mechanically:

$$\mathcal{Z}(\beta,V,\mu)=\sum_{N=0}^{\infty}e^{\beta\mu N}Z(\beta,V,N)=\sum_{N=0}^{\infty}\sum_{\{n_j\}}\exp\left[\beta\mu\sum_{j}^{\substack{\textit{energy}\\\textit{levels}}}n_j-\beta\sum_{j}^{\substack{\textit{energy}\\\textit{levels}}}\varepsilon_j n_j\right]=\sum_{N=0}^{\infty}\sum_{\{n_j\}}\exp\left[-\beta n_j\left(\varepsilon_j-\mu\right)\right]$$

$$\textit{but} \qquad \mu=\mu(T,V,N) \text{ may vary within the summation}$$

$$\ln\mathcal{Z}=\beta PV \qquad \text{(grand canonical ensemble)}$$

## Equipartition of Energy

> Equipartition of energy is a *classical* result, not strictly true quantum mechanically.

Short version: anything that stores energy is a “system.” Ergo, any dynamic variable, *x* (a position or momentum) whose value determines its energy content (position for potential energy, momentum for kinetic energy), is a “system.” Any ensemble of systems in thermal equilibrium follows the Boltzmann distribution for energy:

$$\text{pdf}(\textit{state with energy } E) \propto e^{-\beta E}, \quad \textit{where } \beta = 1/kT \qquad \text{for continuous states.}$$

Knowing the energy as a function of the dynamic variable, *E*(*x*), and plugging into the Boltzmann distribution, gives the PDF for that variable:

$$\text{pdf}(x) = \frac{e^{-\beta E(x)}}{\int_{all\ x} e^{-\beta E(x)}\, dx} \propto e^{-\beta E(x)} .$$

Integrating energy weighted by pdf(*x*) gives the average energy of a system for the ensemble:

$$E_{avg} = \int_{all\ x} E(x)\, \text{pdf}(x)\, dx = \frac{\int_{all\ x} E(x)\, e^{-\beta E(x)}\, dx}{\int_{all\ x} e^{-\beta E(x)}\, dx} .$$

Therefore, any dynamic variables whose energy functions have the same form (e.g., $E(x) = Cx^2$) will store the same average energy in thermal equilibrium. For this quadratic form, do the math, and $E_{avg} = \frac{1}{2} kT$ .

Even Reif kind of overcomplicates the derivation, and misses the real point. The key is that you can consider any term in the total system energy as a “system” of its own [Sch ??]. For example, in the boring example of a particle in 3D space, you can consider its *x*-motion as a system in thermal equilibrium with the “systems” of *y*- and *z*- motion, and other terms (say, potential energy). A gas of such particles is an ensemble of these systems at a fixed temperature, and hence follows the canonical (i.e. Boltzmann) distribution. The rest is just math (following the short version above):

$$\text{pdf}(\textit{state with energy } E) \propto e^{-\beta E}, \quad \textit{where } \beta = 1/kT \qquad \text{for continuous states.}$$

Now $E(p) = Cp^2 \propto p^2$, (or $E(x) = Cx^2$, then substitute *x* for *p* in the following), so

$$E_{avg} = \int_{-\infty}^{\infty} E(p)\, \text{pdf}(p)\, dp = \frac{\int_{-\infty}^{\infty} E(p)\, e^{-\beta E(p)}\, dp}{\int_{-\infty}^{\infty} e^{-\beta E(p)}\, dp} .$$

Again, the key is recognizing that:

> Each independent dynamic variable (coordinate or momentum) is a “system” in the Statistical Mechanics sense, in thermal equilibrium with its surroundings. Hence the energy of that dynamic variable follows the Boltzmann distribution.

If you know the energy as a function of the dynamic variable (coordinate or momentum), that gives you the PDF for the dynamic variable. Then you use the standard formula for computing the average energy given the probability density function for the dynamic variable. This process allows you to answer questions about partition of energy if *E* is not quadratic in *x*.

Note that we don’t aim to separate potential energy from kinetic energy; instead, we separate individual dynamic variables (Hamiltonian coordinates or momenta). In the classic 1D ball-on-a-fixed-spring example, there are two dynamic variables: position of the ball, and momentum of the ball. In a set of *N* such balls-and-springs, energy can be exchanged between any of the *2N* dynamic variables. **Thermal equilibrium** is by definition the stable state of a system when energy can be exchanged between its component subsystems. In this example of $N$ springs/balls, a single ball on a spring is a subsystem, but the position of the ball is a smaller subsystem, and its momentum is another subsystem. The set of $N$ springs/balls is a true, physical ensemble (as opposed to the oft-used "imaginary" ensemble) of $2N$ little systems. Since each of the $2N$ little systems is in thermal equilibrium with all the others, the ensemble statistics follow those of Boltzmann. Hence, we can compute the average energy of each of the $2N$ little systems using Boltzmann statistics. Ergo, (with some math): equipartition of energy.

Once the concepts of equipartition are clear, Reif's derivation provides a solid mathematical foundation for it.

The increasing heat capacity of diatomic gases with temperature, from 3/2 $R$ to 5/2 $R$, and sometimes to 7/2 $R$, demonstrates that equipartition does not apply to quantum systems.

[Aside: Reif cleverly points out that we can find $E_{avg}$ above without evaluating the integral, by using a math trick. Note that the denominator is just a normalization constant for fixed $\beta$, but is a function of $\beta$. Rewrite the numerator in terms of the denominator (we do this all the time in Stat Mech):

$$E_{avg} = \frac{\int_{-\infty}^{\infty} -\frac{d}{d\beta} e^{-\beta E(p)}\, dp}{\int_{-\infty}^{\infty} e^{-\beta E(p)}\, dp}$$

Now interchange the order of integration and differentiation, since they operate on different variables:

$$E_{avg} = \frac{-\frac{d}{d\beta}\int_{-\infty}^{\infty} e^{-\beta E(p)}\, dp}{\int_{-\infty}^{\infty} e^{-\beta E(p)}\, dp}$$

But this is of the form $\frac{1}{U}\frac{dU}{d\beta} = \frac{d}{d\beta}\ln U$ . Hence,

$$E_{avg} = -\frac{d}{d\beta}\ln\int_{-\infty}^{\infty} e^{-\beta E(p)}\, dp . \qquad \text{Now use } E(p) = Cp^2 ,$$

$$E_{avg} = -\frac{d}{d\beta}\ln\int_{-\infty}^{\infty} e^{-\beta Cp^2}\, dp \qquad \text{(Eq. 1)}$$

Focus on the integral for a while: Since we only care about a derivative with respect to $\beta$, we can remove $\beta$ from the integral with a variable substitution:

$$Let \quad y^2 = \beta p^2, \quad \Rightarrow \quad y = \beta^{1/2} p, \quad dy = \beta^{1/2} dp, \quad and \quad dp = \beta^{-1/2} dy. \quad Then$$

$$\int_{-\infty}^{\infty} e^{-\beta Cp^2}\, dp = \beta^{-1/2}\int_{-\infty}^{\infty} e^{-Cy^2} \quad and$$

$$\ln \beta^{-1/2}\int_{-\infty}^{\infty} e^{-Cy^2} = \ln \beta^{-1/2} + const = -\frac{1}{2}\ln\beta + const$$

$$E_{avg} = -\frac{d}{d\beta}\left(-\frac{1}{2}\ln\beta + const\right) = \frac{1}{2}\left(\frac{1}{\beta}\right) = \frac{1}{2}kT$$

Wasn't that fun? ]

# 5 Microscopic Distributions and Quantum Statistics

Given a system of $N$ particles ($N$ large), our purpose here is to derive the average occupation number of a single-particle quantum state (of known energy) within the system. As usual, we consider the particles as "weakly-interacting," which means their interaction energy is negligible, except at very close range where they collide elastically. However, fermions still "interact," not through a potential energy between them, but through the exclusion principle. Similarly, bosons' indistinguishability causes their behavior to deviate from classical behavior. This will have enormous consequences.

We consider the single quantum state in question to be a "subsystem" within a bath, which is the rest of the system. Since we are concerned with how many particles are in the quantum state, we must be concerned with particle exchange between that state and the bath. Therefore, we must include the effects of particle gain or loss by the bath, i.e. we must include the chemical potential of particles in the bath.

> [Some references start the quantum distribution discussion with a bunch of stuff about multi-particle wave-functions and density matrices. They then proceed to derive the microscopic distributions without using any of the introductory quantum physics review material. Here, we skip the multi-particle wave-functions and density matrices, since they aren't needed for a proper understanding.]

Recall that any system in equilibrium has only one relevant state function called the "free energy." The particular form of the free energy depends on the constraints of the system. (Free energy is discussed in detail elsewhere). The chemical potential of a particle in a system is the free energy "cost" to add a particle to the system in equilibrium. Mathematically:

$$\mu \equiv \frac{\partial A}{\partial N} \qquad \textit{where} \quad A \equiv \text{relevant free energy for the system}.$$

Since systems time-evolve toward lower free energy, a system favors particles with lower chemical potential over those with higher chemical potential.

The term "Maxwell-Boltzmann distribution" is ambiguous, because it has at least 3 different meanings. We use it in this chapter as an approximation to *either* the Fermi-Dirac distribution, *or* the Bose-Einstein distribution, at higher energy. We call this the Maxwell-Boltzmann occupation number. [Another commonly used meaning is "the speed distribution of atoms in a gas." What is the other meaning?? Need refs.]

## The Interactions of Non-Interacting Quantum Particles

When we speak of non-interacting bosons or non-interacting fermions, we mean that there is no *energetic* interaction, i.e. no inter-molecular forces, which means the inter-particle potential energy is negligible. However:

> So-called "non-interacting" quantum particles actually *do* interact: though there may be negligible potential energy, fermions still interact through the exclusion principle, and bosons interact in a way that allows them to occupy the same quantum state.

It is in this sense that we speak of three things: (1) "ideal classical gases" (of non-interacting classical particles), (2) "ideal fermion gases," and (3) "ideal boson gases." "Ideal" means no forces, or equivalently, no potential energy between particles, but there may still be quantum interactions [Wan p179b]. These quantum interactions derive from the symmetry (for bosons) and anti-symmetry (for fermions) of their multi-particle wave-functions. The anti-symmetry of fermions implies the Pauli exclusion principle.

Therefore, systems at low density tend to act as ideal classical gases for two separate reasons:

1. Low density implies the particles are far apart, and so their Coulomb potential is low ("ideal," i.e. "non-interacting").
2. Low density implies low fractional occupancy of quantum states, so the fermion or boson quantum interactions are negligible ("classical").

At low energies, the M-B occupation number is unrealistic, because it ignores the quantum implication of indistinguishability. Real particles are quantum, and often indistinguishable (e.g., a gas of identical molecules). When the particles are *in*distinguishable, then the conditions for the *particles* to obey the Maxwell-Boltzmann occupation number do not strictly apply. Specifically, the Maxwell-Boltzmann occupation number requires that the only interaction is thermal exchange of energy. But identical quantum particles interact through their quantum nature, as noted above. Therefore, the *system*, as a whole at fixed *T*, obeys the Boltzmann (not Maxwell) distribution for energy, because (as we saw earlier) the Boltzmann distribution for system energy is independent of the nature of the system. But the quantum states do *not* follow the Maxwell-Boltzmann occupation number, because of the particles' quantum interactions.

### An Ideal World

Now that we know that ideal (so-called "non-interacting") systems *do*, in fact, interact, it is time to more precisely define an "ideal" system, at a given temperature. [This is out of place??]

---

## The Amazing Fermi-Dirac Distribution

We start with fermions, because they are conceptually simpler for this topic.

The Fermi-Dirac distribution tells us, for a given temperature,
the average number of fermions in a single quantum state of given energy;
or equivalently, what is the probability of occupancy of a single quantum state.

When gas particles are fermions, they interact through the Pauli exclusion principle. We usually think of interactions in terms of energy: an interaction potential. But fermions also interact through exclusion, and this isn't described by an interaction potential [Wan p179b].

For any system of identical fermions in thermal equilibrium, the chemical potential, $\mu$, serves as a natural energy reference level. The chemical potential need not be zero, and it follows the arbitrary additive constant of energy (allowed by any system). This allows the Fermi-Dirac distribution to be written with *absolute* probabilities:

$$\langle \textit{occupation number} \rangle = \Pr(\textit{state } s \textit{ with energy } E_s \textit{ is occupied})$$

$$= \frac{1}{e^{(E_s-\mu)/kT}+1} \qquad \textit{where} \quad \mu \equiv \text{chemical potential}.$$

For fermions, the average occupation number of a state is between 0 and 1,
and therefore equals the probability of occupation of the state.

Note that the probability of occupancy of a state at $E = \mu$ is exactly ½.

$\mu$ can be determined by the density of states, $g(E)$, the number of particles in the system, $N$, and the temperature, $T$. Given fixed conditions, there is a unique $\mu$ which satisfies both the Fermi-Dirac distribution, and adds up to the right number of particles:

$$\text{chemical potential} \equiv \mu(N,T) \quad \text{such that} \quad N = \int_0^\infty dE\, g(E) \frac{1}{e^{(E_s-\mu)/kT}+1}.$$

This equation implicitly determines $\mu$.

Some physicists define the **Fermi energy**, or the "Fermi level," as $\varepsilon_F \equiv \mu(T = 0)$, the chemical potential at absolute zero. Others, such as [Wan]. [Rei], and semiconductor physicists, call $\mu$ the Fermi energy at all temperatures: $\varepsilon_F(T) \equiv \mu(T)$, i.e. the Fermi level is just another name for chemical potential.

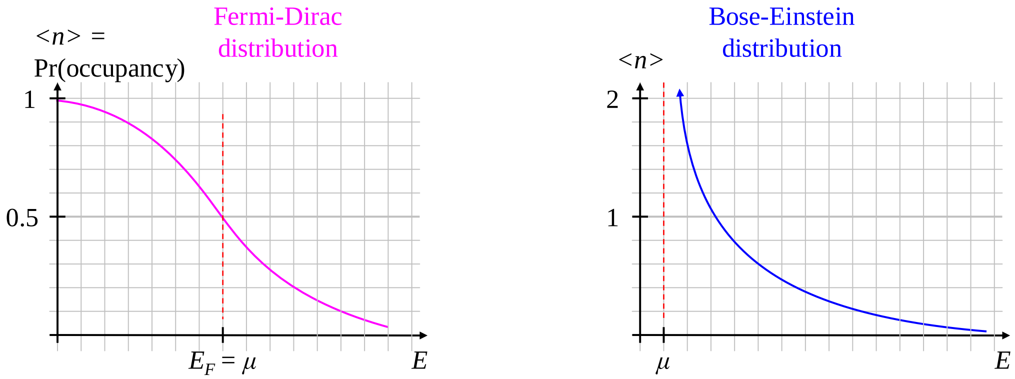


**Figure 5.1** (Left) Fermi-Dirac distribution: <occupation number> = Pr(occupancy). (Right) Bose-Einstein distribution: <occupation number> is unbounded.

We derive the Fermi-Dirac distribution [following Sch p266] by [TBS].

## Summary of the Four Main Statistical Mechanics Distributions

All of these distributions only apply in thermal equilibrium, at temperature *T*.

| Distribution | Formula | What it gives you | What it applies to | Examples |
|---|---|---|---|---|
| **Boltzman** | $e^{-E/kT}$ | *relative* probability of being in a single microstate, of energy *E*. | any system of any size or complexity, quantum or classical. | an atom moving in a gas, electronic excitations of an atom in a gas, a magnetic dipole in a lattice |
| **Maxwell-Boltzmann** | $\dfrac{1}{e^{(E-\mu)/kT}}$<br>$= e^{-(E-\mu)/kT}$ | average occupation number of a given single-particle quantum state of energy *E*; this equals the *absolute* probability that the state is occupied, since <*N*> << 1. | A single-particle quantum state in a classical system at low density, or high temperature, so that <*N*> << 1. | an atom moving in a gas, an electron state in a semiconductor energy band |
| **Fermi-Dirac** | $\dfrac{1}{e^{(E-\mu)/kT}+1}$ | average occupation number of a given single-particle fermion quantum state of energy *E*; since <*N*> < 1, this exactly equals the *absolute* probability that the state is occupied. | A single-particle fermion state in any system; note that 0 < <*N*> < 1. | an electron state in a semiconductor energy band, an electronic energy state in an atom |
| **Bose-Einstein** | $\dfrac{1}{e^{(E-\mu)/kT}-1}$ | average occupation number of a given single-particle boson quantum state of energy *E*. | A single-particle fermion state in any system; note that <*N*> > 0, and unbounded. | a photon state in a photon gas, a $^4$He atom state in a collection of $^4$He atoms |

## The Amazing Bose-Einstein Distribution

The Bose-Einstein distribution tells us the average number of bosons in a single quantum state of given energy; for bosons, this is unbounded.

The Bose-Einstein (B-E) distribution results from the quantum property of bosons which allows any number of bosons to occupy the same quantum state. This quantum constraint is different than fermions. Many books introduce a "fugacity" to describe the B-E distribution, but I think it is easier to start by thinking in analogy to F-D statistics. In F-D statistics, the Fermi energy, $E_F$, served as a "reference point" for energy: the energy that appears in the exponential is the difference between the energy of the quantum state and $E_F$. Similarly, in the B-E distribution, the chemical potential, $\mu$, serves as the "reference point" for energy. An important difference is that in F-D distributions, the average occupation number equals the probability of occupancy. In the B-E distribution, that is not the case [Wan p186]. Also, there is a minus sign in the denominator:

$$\langle occupation\ number \rangle \equiv \langle N \rangle = \frac{1}{e^{(E_s-\mu)/kT}-1} \quad where \quad \mu \equiv \text{chemical potential} .$$

An important thing to notice is the behavior at low energy $E_s$: the occupation number gets very large; similarly, at low $T$, the occupation number gets large. At some point, almost all the particles are in the ground state, which is Bose-Einstein condensation.

The occupation number is valid only if $\mu < E$. It blows up when $\mu = E$, and would be negative for $\mu > E$, which is physically meaningless. Therefore:

For bosons, $\mu$ must be less than the ground state energy.

Photons are a special case of bosons: they have no mass, and cannot exist with zero momentum or zero kinetic energy. This means their particle number is *not* conserved: photons are freely created (radiated) or destroyed (absorbed). This implies that a system in thermal equilibrium has evolved to a point where every photon's chemical potential $\mu = 0$. Thus, there is no free-energy advantage to creating or destroying any photons: they are in equilibrium with the system.

TBS: B-E statistics *favor* multiple occupation of states [Wan p185]. E.g., for 2 bosons distributed among 2 states, the probability of being in the same state is ½ for classical particles, vs. 2/3 for identical bosons:

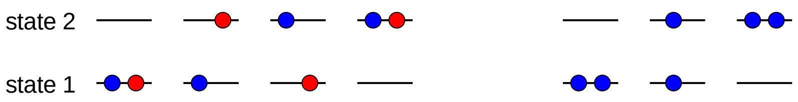


**Figure 5.2** (Left) Classical distribution of two particles. (Right) B-E distribution of two identical particles.

### Black Body Radiation (a Special Case of B-E)

We now derive the black body radiation spectrum from the Boltzmann distribution, and the occupation number formula for bosons, since photons are bosons. We repeat here in brief the derivation of density of quantum states for a 3D system.

**Occupation numbers of boson states:** If we treat each EM radiation mode (**k**, **ε**) as a "system" (e.g., an oscillator with discrete energy states), then we can use the Boltzmann distribution to find the probabilities of every excitation level. (Recall the Boltzmann distribution applies to any system held at fixed temperature [say, by a heat bath], including continuous, discrete, classical, quantum, etc. The Boltzmann distribution applies to all systems because it relies only on the properties of the heat bath, and not at all on properties of the system). Then:

$$\text{RelPr}(s) = e^{-E_s/kT} \qquad \Rightarrow \qquad \text{RelPr}\big(N(\mathbf{k},\boldsymbol{\varepsilon})\big) = e^{-N(\mathbf{k},\boldsymbol{\varepsilon})\hbar\omega/kT}$$

$$\langle N \rangle = \sum_{N=0}^{\infty} N\,\Pr(N) = \frac{\sum_{N=0}^{\infty} N e^{-N\hbar\omega/kT}}{\sum_{N=0}^{\infty} e^{-N\hbar\omega/kT}}, \qquad N \equiv N(\mathbf{k},\boldsymbol{\varepsilon})\,.$$

The numerator of the above sums *relative* probabilities, while the denominator normalizes the result to absolute probabilities. Consider the denominator:

$$\sum_{N=0}^{\infty} e^{-N\hbar\omega/kT} = 1 + x + x^2 + x^3 + \ldots \qquad \textit{where} \quad x \equiv e^{-\hbar\omega/kT}\,.$$

This is a geometric series, with sum:

$$1 + x + x^2 + x^3 + \ldots = \frac{1}{1-x} = \frac{1}{1-e^{-\hbar\omega/kT}}\,.$$

Now consider the numerator:

$$\sum_{N=0}^{\infty} N e^{-N\hbar\omega/kT} = \sum_{N=0}^{\infty} N e^{-Na} \qquad \textit{where} \quad a \equiv \hbar\omega/kT.$$

$$= -\frac{\partial}{\partial a}\sum_{N=0}^{\infty} e^{-Na} = -\frac{\partial}{\partial a}\sum_{N=0}^{\infty}\left(e^{-a}\right)^N = -\frac{\partial}{\partial a}\left(\frac{1}{1-e^{-a}}\right) = +\left(1-e^{-a}\right)^{-2} e^{-a}$$

$$= \frac{e^{-\hbar\omega/kT}}{\left(1-e^{-\hbar\omega/kT}\right)^2}$$

The final average excitation level is:

$$\langle N \rangle = \frac{e^{-\hbar\omega/kT}}{\left(1-e^{-\hbar\omega/kT}\right)^2} \bullet \left(1-e^{-\hbar\omega/kT}\right) = \frac{e^{-\hbar\omega/kT}}{1-e^{-\hbar\omega/kT}} \qquad \text{multiply top \& bottom by } e^{+\hbar\omega/kT}$$

$$= \frac{1}{e^{\hbar\omega/kT}-1}\,.$$

Equivalent to the above, we can think of each mode as a single-photon quantum state. Photons are bosons, so each state has an occupation number, $N$. The average $N$ is that just derived.

The average occupation number of a photon quantum state at temperature $T$
follows directly from the Boltzmann distribution.

Note that $N$ may be greater or less than 1: $N = 0, 1, 2, \ldots$ .

## Black Body Radiation and Einstein Coefficients

We somewhat follow [Bay p276-8]. Einstein used the photon occupation numbers to learn something about the quantized interaction of matter with radiation (which was a mystery at the time). Imagine a box with a photon gas at temperature $T$. The walls have atoms or molecules which can absorb and emit photons; we treat the photons as particles. Consider a wall molecule with two states, separated in energy by $\hbar\omega$, and a photon state (mode) $(\mathbf{k}, \boldsymbol{\varepsilon})$. Let $M \equiv$ # molecules.

We determine the ratio of excited molecules to ground-state molecules by again using the Boltzmann distribution:

$$\frac{P_e}{P_0} = e^{-\hbar\omega/kT} \qquad \textit{where} \quad P_e \equiv \Pr\left(\text{molecule being excited}\right),\ P_0 \equiv \Pr\left(\text{molecule in ground state}\right).$$

On average, the absorption rate (in photons/s) is proportional to the number of photons in the gas, $N(\mathbf{k}, \boldsymbol{\varepsilon}) \equiv N$, and the number of ground-state molecules (excited molecules cannot absorb another photon):

$$w_{abs} = B_{abs} N M P_0 \qquad \text{(photons/s)}. \tag{5.1}$$

Since $M$ (the # molecules in the system) is fixed, we could "absorb" it into the coefficient $B_{abs}$, but we are trying to see how a *single* molecule interacts with radiation, so it helps to include $M$ explicitly.

In thermal equilibrium, the overall absorption rate of photons by the molecules must equal the overall emission rate. The fact that thermal equilibrium between the photons and the molecules can be established at every temperature says that the emission rate must go up with temperature.

From common experience, we expect that an excited molecule will eventually "decay," and radiate a photon: hot bodies radiate and cool, phosphors glow in the dark. Also from macroscopic experience, we might expect to dislodge a photon from an excited molecule if we whack it (a collision). Therefore, we take a leap, and suppose that emission can occur from two causes: spontaneous emission, and **stimulated emission** (aka "induced emission"). Stimulated emission is ambient radiation causing an excited molecule to radiate; in other words, a collision with a photon can stimulate emission much like a collision with a particle can. (At the time, 1917, some physicists thought of this as absorption of a photon followed by emission of two photons, but this "indirect" stimulated emission model is not necessary, and has no supporting evidence.) Einstein guessed that stimulated emission is proportional to the ambient intensity, $N$, which seems reasonable: more collisions induce more radiation. Thus:

$$w_{em,stim} = B_{em} N M P_e \qquad \textit{where} \quad N \equiv \#\ \text{ambient photons}. \tag{5.2}$$

We will see shortly that at high temperature, this must be true. But each molecule has no "idea" of temperature; it knows only that there is an EM field of some intensity impinging on it. (Recall that temperature is an aggregate concept defined by $1/T = dS/dE$; a single molecule does not have a temperature, but it can be in equilibrium with a heat bath that *does* have a temperature.)

For independent molecules, the overall spontaneous emission rate is proportional to the number of excited molecules:

$$w_{em,spon} = A M P_e \qquad \textit{where} \quad A \equiv \text{decay constant of proportionality}. \tag{5.3}$$

In thermal equilibrium, to total emission rate must equal the total absorption rate:

$$w_{abs} = w_{em,stim} + w_{em,spon} \quad \Rightarrow \quad B_{abs} N P_0 = B_{em} N P_e + A P_e = \left(N B_{em} + A\right) P_e. \tag{5.4}$$

As noted, all the coefficients are properties of the molecules, and therefore independent of temperature. We see that for large $N$ (corresponding to high $T$), $A$ becomes insignificant. That is, as $T \to \infty$:

$$\lim_{T\to\infty} \frac{P_e}{P_0} = e^{-\hbar\omega/kT} = 1, \qquad \text{or} \qquad \lim_{T\to\infty} P_e = P_0. \qquad \Rightarrow$$

$$NMB_{abs} = NMB_{em} + \cancel{AM} \to NMB_{em}, \qquad \text{or} \qquad B_{abs} = B_{em}.$$

We can therefore drop the subscript on $B$. Now, at finite temperature, this allows us to solve for $A$ in terms of $B$, recalling that we supposed that $A$ is independent of $T$. Then the equilibrium of rates, (5.4), we have:

$$NMB = \left(NMB + MA\right) e^{-\hbar\omega/kT}$$

$$NB\left(1 - e^{-\hbar\omega/kT}\right) = A e^{-\hbar\omega/kT}.$$

From our earlier analysis of occupation numbers of boson states, we can use for $N$ the average occupation number, and solve for the spontaneous emission coefficient $A$:

$$N = \frac{1}{e^{+\hbar\omega/kT} - 1} \quad \Rightarrow$$

$$\frac{1}{e^{+\hbar\omega/kT} - 1} B\left(1 - e^{-\hbar\omega/kT}\right) e^{+\hbar\omega/kT} = A \qquad \Rightarrow \qquad A = B\,.$$

I other words, the spontaneous emission rate is the same as the stimulated rate from an EM field of a single photon.

Now we return to a single molecule in the presence of *N* ambient photons; thermal equilibrium is no longer a consideration. From the emission and absorption rates (5.1) - (5.3), the emission rate of an excited atom must relate to the absorption rate of a ground-state atom according to:

$$\frac{w_{em}}{w_{abs}} = \frac{BN + A}{BN} = \frac{N+1}{N}\,.$$

How does this square with quantum mechanics? Recall that the dipole radiation matrix element includes a matrix element of the magnetic vector potential (or the electric field with the same result):

$$w_{em} = C_1 \left| \langle N+1 | \hat{\mathbf{A}} | N \rangle \right|^2 = C_2 \left(N+1\right) \qquad \textit{where} \quad N \equiv N(\mathbf{k}, \boldsymbol{\varepsilon}),\ \hat{\mathbf{A}} \equiv A_0(\mathbf{k}) \boldsymbol{\varepsilon}_{\mathbf{k},\lambda} \left(A + A^\dagger\right)$$

$$w_{abs} = C_1 \left| \langle N-1 | \hat{\mathbf{A}} | N \rangle \right|^2 = C_2 N$$

$$\frac{w_{em}}{w_{abs}} = \frac{N+1}{N}\,.$$

Thus, our quantum formulas for the interaction of quantized radiation with matter agree with results demanded by statistical mechanics, and are thoroughly verified by experiment.

## Energy Spectrum of a Photon Gas in Thermal Equilibrium

What is the spectrum of black body radiation? Perhaps the most famous spectrum in physics, this is the spectrum of the Cosmic Microwave Background radiation. We find the energy spectrum from first principles by simply multiplying the photon density of states by the average energy of each state, at a temperature *T*:

$$u(\omega) = \rho_s(\omega) E(T, \omega) \qquad \text{such that} \qquad u(\omega)\, d\omega = \text{energy in } (\omega, \omega + d\omega)\,.$$

The density of states $\rho_s(\omega)$ is the density of EM *modes* per unit angular frequency. We now find $\rho_s(\omega)$ by fitting waves into a finite rectangular box, and then changing variables from **k** to $\omega$ in several steps. (It is not hard to show that for "large" boxes [compared to the significant wavelengths of the photons], neither the actual shape of the box nor the exact boundary conditions significantly matter.) For simplicity, we start with a 1D "particle in a box." We can take the wave-function (or EM wave) to be a standing wave that is zero at the edges:

$$k = n\frac{\pi}{L}, \quad k > 0 \quad \Rightarrow \qquad n(k) = k\left(\frac{L}{\pi}\right), \qquad \rho_{k+}(k) = \frac{dn}{dk} = \frac{L}{\pi}, \qquad n(k) = \int_0^k \rho_{k+}(k')\, dk'\,.$$

Each standing wave is a superposition of left- and right- momenta, but we'd like to count those momenta (*k*-values) separately, so we allow *k* to be negative, and then count those states separately. This halves the *k*-space density of states:

$$\rho_k(k) = \frac{dn}{dk} = \frac{L}{2\pi}, \quad -\infty < k < +\infty; \qquad n(k) = \int_{-k}^k \rho_k(k')\, dk'\,.$$

Generalizing to a 3D box:

$$\rho_{\mathbf{k}}(\mathbf{k}) = \frac{L_x L_y L_z}{(2\pi)^3} = \frac{V}{(2\pi)^3} \quad \text{(a constant)}, \qquad V \equiv \text{volume of the box} .$$

[This implies that each state occupies a volume in phase space of $(2\pi\hbar)^3 = h^3$. TBS.]

We are often interested in the *energy* spectrum (not *k* or momentum spectrum), so we must convert this to an *energy* density of states (and then to *frequency* density of states). Since the energy of a photon $E$ depends only on $|\mathbf{k}|$, we change variables to $E$ in two steps: (1) we convert $\rho_{\mathbf{k}}$ to the density per unit *magnitude* of $\mathbf{k}$; and (2) convert from $k \equiv |\mathbf{k}|$ to $E$. First, the number of states in a $k$-space spherical shell of fixed $k$ is simply the surface area of the shell times its thickness:

Define: $k \equiv |\mathbf{k}|$. Then:

$$\rho_k(k)\, dk = 4\pi k^2\, \rho_{\mathbf{k}}(k\hat{\mathbf{n}})\, dk, \qquad \text{or} \qquad \rho_k(k) = 4\pi k^2 \frac{V}{(2\pi)^3} .$$

Second, we use the relation between momentum and energy (i.e., the “dispersion relation”) for a photon:

$$\rho_E(E) = \rho_k(k)\frac{dk}{dE}, \qquad E = pc = \hbar kc, \qquad k = \frac{E}{\hbar c}, \quad \frac{dk}{dE} = \frac{1}{\hbar c} \qquad \Rightarrow$$

$$\rho_E(E) = 4\pi\left(\frac{E}{\hbar c}\right)^2 \frac{V}{(2\pi)^3}\frac{1}{\hbar c} .$$

We must now include a factor of 2, because each $\mathbf{k}$ value has 2 independent polarization modes, yielding:

$$\rho_E(E) = E^2 \frac{V}{\pi^2 (\hbar c)^3} .$$

Finally, converting to energy density per unit $\omega$:

$$u(\omega) = \hbar\omega N(\omega)\rho_E(E)\frac{dE}{d\omega} \qquad \textit{where} \quad E = \hbar\omega, \quad \frac{dE}{d\omega} = \hbar$$

$$= \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}\cdot\frac{(\hbar\omega)^2 V}{\pi^2(\hbar c)^3}\hbar = \frac{\hbar}{\pi^2}\cdot\frac{\omega^3}{c^3}\cdot\frac{1}{\left(e^{\hbar\omega/kT} - 1\right)}V .$$

## Massive Non-interacting Boson Gas: Bose-Einstein Condensation

Many systems exhibit interesting and useful properties when they are very cold. One example of this is Bose-Einstein condensation (BEC), defined as a system of bosons so cold that a “significant” fraction of the particles is in the system’s ground state. There are many variations of BEC, depending on the particle interactions and sometimes the confining potential. For example, liquid $^4$He becomes a superfluid at 2.17 K, and involves significant inter-molecular forces.

We consider here a very simple case: bosons in a cube (3D particles in a box). The ground state wavelength is always about twice the box size, regardless of the size of the box, or the details of the confining potential (Figure 5.3). The ground state kinetic energy is *not* zero; for the idealized particle in a cube, it is:

$$KE_0 = n_x{}^2 + n_y{}^2 + n_z{}^2 = 3 \ \text{(in some units)} \qquad \textit{where} \quad KE \equiv \text{kinetic energy} .$$

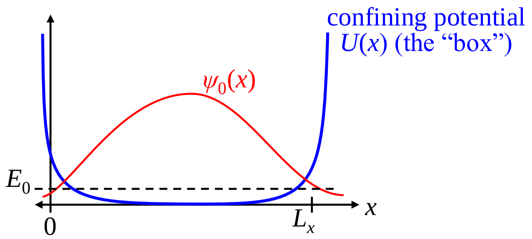


**Figure 5.3** The ground state wave-function is always about a half-wavelength, regardless of the size of the box, or the details of the confining potential.

In a large system, this is tiny, but conceptually, it is finite. The 1st excited state has kinetic energy:

$$KE_1 = 1^2 + 1^2 + 2^2 = 6 \qquad \text{(twice } E_0\text{)}.$$

We choose $U(\mathbf{r}) = 0$ everywhere, so:

$$E_n = KE_n, \text{ and in particular:} \qquad E_0 = KE_0 \ .$$

(Some references choose $U(\mathbf{r}) = -KE_0$, so $E_0 = 0$. In this case, one cannot then claim that the momentum is zero because the energy is zero. Other references invoke periodic boundary conditions, which would allow the ground state energy and momentum to be zero, but periodic boundary conditions are inconsistent with the ground state. Any bound particle has nonzero $KE$, and therefore a nonzero momentum.)

Despite many claims to the contrary, the ground-state momentum is not zero. In an idealized box:

$$\psi_0(x,y,y) = N\left(\sin(k_x x)\sin(k_y y)\sin(k_z z)\right) \quad \textit{where} \quad k_x = \pi / L_x, \text{ etc.}$$

In each dimension ($x$, $y$, and $z$) the momentum is a superposition of a positive-going and negative-going nonzero momentum. This is consistent with the fact that the ground-state kinetic energy is nonzero.

At all temperatures, there must exist a chemical potential $\mu = \mu(T)$ such that the average occupation numbers of all the states of the system add up to $N$, the known number of particles:

$$N = \sum_{\substack{\textit{quantum} \\ \textit{states } s}} \frac{1}{e^{(E_s - \mu)/kT} - 1} \qquad \text{(exact)} \ .$$

At warm temperatures, we can approximate the sum of average occupation numbers with an integral (Figure 5.4, left):

$$N \approx \int_0^\infty g(E) \frac{1}{e^{(E-\mu)/kT} - 1} \, dE \ .$$

At every temperature, the ground state occupation number $N_0$ is finite, but above the critical temperature, it is nearly zero (negligibly small).

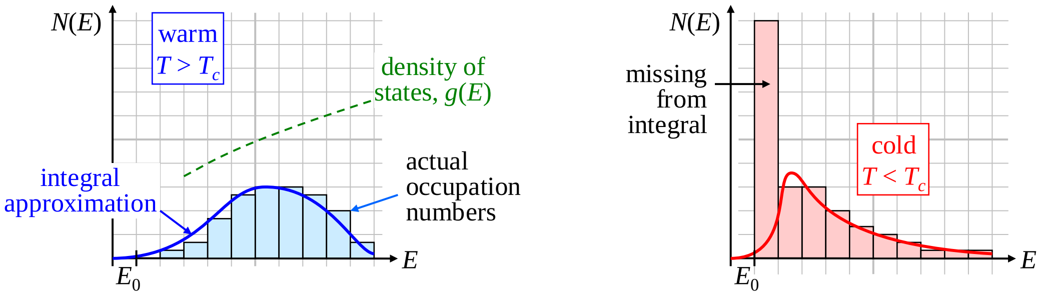


**Figure 5.4** (Left) The integral approximation to the occupation number sum works well at high $T$. (Right) The integral approximation to the occupation number sum fails at low temperatures.

However, at very low temperatures, the integral approximation for $N$ fails because the ground state occupation number is mostly excluded from the integral (Figure 5.4, right). To get a more accurate counting of the occupation numbers, we must separate the ground state occupation number ($N_0$) from all the excited state occupation numbers ($N_{ex}$):

$$N = N_0 + N_{ex} \approx N_0 + \int_0^\infty g(E) \frac{1}{e^{(E-\mu)/kT} - 1} \, dE \, . \tag{5.5}$$

It can be shown that for large $N$, only the ground state needs to be taken outside the integral [Mandl p294b].

The rapid increase in the ground state occupation number $N_0$ as $T$ decreases leads to a "critical behavior" of the system.

To see the behavior of the occupation number sum (5.5), we must find $g(E)$. This is a standard result: we start with our known density of states for $k \equiv |\mathbf{k}|$, and change variables to the energy density of states:

$$\rho_k(k) = 4\pi k^2 \frac{V}{(2\pi)^3}$$

$$g(E) = \rho_k(k) \frac{dk}{dE}, \qquad E = \frac{\hbar^2 k^2}{2m}, \qquad k = \frac{\sqrt{2mE}}{\hbar} \qquad \Rightarrow \qquad \frac{dk}{dE} = \frac{\sqrt{2m}}{\hbar}\left(\frac{1}{2}E^{-1/2}\right)$$

$$g(E) = \underbrace{4\pi\left(\frac{2mE}{\hbar^2}\right)\frac{V}{(2\pi)^3}}_{\rho_k} \underbrace{\frac{\sqrt{2m}}{\hbar}\left(\frac{1}{2}E^{-1/2}\right)}_{dk\,/\,dE} = \underbrace{4\sqrt{2}\pi \frac{m^{3/2}}{(2\pi\hbar)^3} V}_{\text{constant}} E^{1/2} \, . \qquad \text{[Schroeder 7.123 p316]}$$

The important point is that $g(E) \sim E^{1/2}$ (and this depends on our having chosen $U(\mathbf{r}) = 0$, so that $E_0 = KE_0$). Note that if we had *chosen* to make the ground state zero energy by shifting $E_0$ to 0, then in principle, we should shift our energy density of states function by the same amount. In practice, the difference is small, and is often neglected.

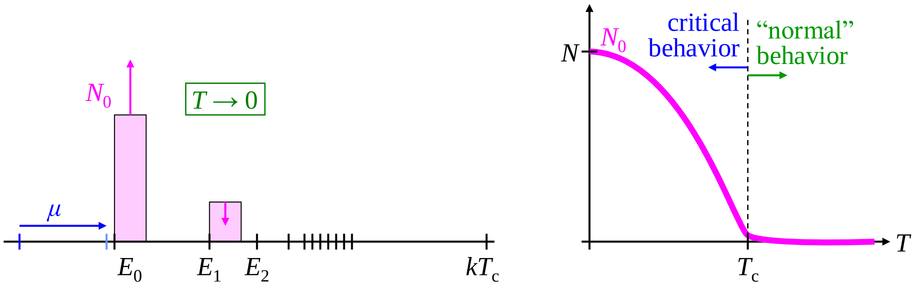


**Figure 5.5** (Left) As $T$ decreases below $T_c$, $\mu \to E_0$, and $N_0 \to N$. (Right) Ground state occupation number vs. temperature, showing the critical region.

At low temperatures, as more particles go into the ground state, $\mu$ must approach $E_0$. $N_{ex}$ can then be upper-bounded (and approximated) by letting $\mu(T) \to E_0$:

$$N_{upper} \approx \int_0^\infty g(E) \frac{1}{e^{(E-E_0)/kT} - 1} \, dE = C \int_0^\infty \frac{E^{1/2}}{e^{(E-E_0)/kT} - 1} \, dE \, .$$

To evaluate the integral, we substitute $x = (E - E_0)/kT$ in the denominator, and $x \approx E/kT$ in the numerator (which inflates $N_{upper}$ a little, but that's OK, because it's an upper bound). Finally, $dE = kT\,dx$, which leaves us with a temperature dependence times a definite integral (a constant):

$$N_{upper} = C\int_0^\infty \frac{(kTx)^{1/2}}{e^x - 1} kT\ dx = C(kT)^{3/2} \underbrace{\int_0^\infty \frac{x^{1/2}}{e^x - 1} dx}_{\text{constant}}$$

$$= 2.6 \frac{(2\pi mkT)^{3/2}}{(2\pi\hbar)^3} V . \qquad (\text{NB: } 2\pi\hbar \equiv h) \qquad [\text{Schroeder 7.125 p316}].$$

We define the critical temperature $T_c$ as that when the upper bound matches the actual number of particles, $N_{upper} = N$. Rearranging the above equation:

$$kT_c = 0.53 \frac{(2\pi\hbar)^2}{2\pi m}\left(\frac{N}{V}\right)^{2/3} \qquad [\text{Schroeder 7.126 p317}] .$$

At the critical temperature, $\mu$ is very close to $E_0$ (Figure 5.5, left), such that:

$$E_0 - \mu << E_1 - E_0 << kT_c \qquad [\text{cf Schroeder p319m}] .$$

$T_c$ is typically very small, millikelvins or less. Using $T_c$, we can simplify the expression for $N_{ex} \approx N_{upper}$ for $T < T_c$:

$$N_{ex} \approx N_{upper} = N\left(\frac{T}{T_c}\right)^{3/2} \qquad T < T_c ,$$

$$N_0 = N - N_{ex} = N\left(1 - (T/T_c)^{3/2}\right), \quad T < T_c .$$

This shows that below $T_c$, a significant fraction of the particles are in the ground state, with minimum momentum (but finite), and minimum KE (also finite).

Note that $\mu(T)$, and also $N_0(T)$, are *smooth* functions of $T$, so there is no *true* "kink" in $N_0(T)$ at $T = T_c$; however, the transition from $N_0 \approx 0$ to $N_0 \sim (T_c - T)^{3/2}$ is a small range: $\Delta T << T_c$, so the transition is considered "abrupt."

When a parameter (such as $N_0$, or an observable consequence of $N_0$) varies as a power law of $(T_c - T)$, for some critical temperature $T_c$, the parameter is said to exhibit **critical behavior** (Figure 5.5, right). Since $N_0$ is continuous, this is sometimes said to be a higher order phase transition. Note that this is distinct from, say, steam condensing to liquid water (a 1st-order phase transition), which involves two distinct phases: gas and liquid. In contrast, the BEC critical behavior goes all the way to absolute zero. There is no phase below the temperature range of critical behavior, i.e. no analog to the liquid phase.

It is often said that BEC occurs when the particles' de Broglie wavelength is of the order of the particle spacing, so multiple particles have to start crowding into a single quantum state. It's not clear this is physically valid, though, because 2D systems experience the same de Broglie wavelength effect, but do not exhibit critical behavior.

**Thermodynamic limit?** A reasonable definition of the "thermodynamic limit" is that the particle number $N$ is large enough to be taken as continuous, and the energy is high enough that the energy density of states may be taken as continuous. By this definition, a BEC is not in the thermodynamic limit, because it specifically distinguishes the ground state from all the excited states. Therefore, $N$ may be large, but the energy is insufficient to allow the energy density of states to be considered continuous.

**Other forms of BEC:** Liquid $^4$He exhibits a superfluid transition at 2.17 K. It can be explained as a kind of BEC, however it is quite different from the ideal gas we analyzed above. After all, it's a liquid and so must have strong intermolecular forces. The critical exponent is different from 3/2, as well.

## The Amazing Maxwell-Boltzmann Occupation Number Distribution

The term "Maxwell-Boltzmann distribution" is ambiguous, because it has at least 3 different meanings. We use it here as an approximation to *either* the Fermi-Dirac distribution, *or* the Bose-Einstein distribution, at higher energy. It therefore gives the average occupation number of a single quantum state of given energy, which must be << 1 for the approximation to hold. When $(E-\mu) > (few)kT$, the Fermi-Dirac and the Bose-Einstein distributions both approach the Maxwell-Boltzmann distribution:

$$(E-\mu) > (few)kT \qquad \Rightarrow \qquad e^{(E-\mu)/kT} >> 1 \quad \Rightarrow$$

$$\underbrace{\frac{1}{e^{(E-\mu)/kT}+1}}_{\text{Fermi-Dirac}} \approx \underbrace{\frac{1}{e^{(E-\mu)/kT}-1}}_{\text{Bose-Einstein}} \approx \underbrace{\frac{1}{e^{(E-\mu)/kT}}}_{\text{Maxwell-Boltzmann}} << 1.$$

Note that at the higher energies where this approximation applies, the occupation number is << 1, and is therefore equal to the probability that the state is occupied, since the chance of double occupancy is negligible.

At higher energies, for any particles, the average occupation number of a state << 1,
and therefore equals the probability of occupation of the state.

**Maxwell-Boltzmann energy distribution:** The Maxwell-Boltzmann energy distribution describes the distribution of energies of molecules in an ideal gas. (This is different than the occupation number distribution, though they can be related through the density of states.) There's no physics here. Just combinatorics. We can use the reasoning of the Boltzmann distribution to show that the most populous way to distribute a fixed number of $N$ *distinguishable* particles throughout arbitrary energy levels, while maintaining a fixed total energy, is to populate each possible quantum state according to these relative probabilities

$$\Pr(E_s) \propto e^{-\beta E_s} \qquad or \qquad \text{RelPr}(E_s) = e^{-\beta E_s}$$

We show this by considering each particle to be a microscopic "system" which is in thermal equilibrium with the "bath" of other particles [Sch??]. Our given macro-system is then an ensemble of little-systems (particles) in thermal equilibrium. These are exactly the conditions of the canonical ensemble, and hence the ensemble (i.e., the little-systems, aka particles) follow the Boltzmann distribution. When applied to the microscopic distribution of states *within* a system, these statistics are called **Maxell-Boltzmann statistics**.

Because this is a Boltzmann distribution, it has the amazing property that the *relative* probability of a particle being in a given state depends only on the state's energy $E$, and is *independent* of the distribution of energy levels, $\rho(E)$! However, the *absolute* probabilities *do* depend on $\rho(E)$ through the normalization factor, just as for any Boltzmann distribution.

I repeat: there is no physics here. Nature doesn't really "seek" this distribution; it doesn't "seek" disorder; it doesn't "favor" lower energy states. The particles stumble about randomly between all possible configurations, *with each configuration as likely as any other*. There are simply more ways to distribute your particles among the energy levels according to the M-B (Maxwell-Boltzmann) distribution than there are any other ways to distribute them (keeping total energy fixed). This means if you examine a system at some random time, it's simply much more likely to be in the M-B distribution than any other. No physics, just combinatorics (unless you consider combinatorics a part of physics. Hint, hint.)

But how likely is a system to be in the M-B distribution? For $N \sim 10^{20}$ systems (particles), it's virtually guaranteed. Deviations from the M-B distribution are exponentially unlikely, i.e. their probabilities fall off rapidly.

Subtle point: in some works, the derivation of the M-B distribution uses phrases like "we seek to maximize" some thing. Actually, we don't *try* to maximize anything. The distribution is fixed by combinatorics, and we have no freedom. There's no knob we can turn to change (or maximize) anything. We seek to *find* the most populous distribution, knowing that our large-$N$ system is exceedingly likely to be in such a state. This is a subtle, but important, point. Nature isn't trying to maximize or minimize anything. Some things just come out maximized or minimized by combinatorics. All we can do is find out what things those are.

**Maxwell-Boltzmann speed distribution:** The Maxwell-Boltzmann speed distribution describes the distribution of speeds of molecules in an ideal gas. You can derive it from the M-B energy distribution by noting that *speed* = $(2E/m)^{1/2}$.

# 6 Thermodynamics

Thermodynamics isn't really Statistical Mechanics, because it's empirically based on observations of macroscopic systems. But SM is supposed to be able to derive all of thermodynamics, so they are related.

In thermodynamics, "temperature" isn't even well defined. Usually, the best one can do is define "temperature" in terms of ideal gasses at low density and/or high temperature, where they experimentally follow the ideal gas law very closely. Lower temperatures are defined with "ladders" of connections to other materials that have some simple, accurate equation of state at low temperatures. Therefore, in thermodynamics, the relation:

$$\frac{1}{T} = k\left(\frac{\partial S}{\partial E}\right)_{N,V} \qquad \text{is empirical in thermodynamics.}$$

But in Statistical Mechanics, the relation:

$$\frac{1}{T} \equiv k\frac{dS}{dE} \qquad \text{is a precise } \textit{definition} \text{ in Statistical Mechanics.}$$

## Reversible Vs. Irreversible, and Entropy

Reversibility is a huge stumbling block for many people, and many (most?) textbook descriptions are weak.

A process is defined as **reversible** if an arbitrarily small change in system (or outside) conditions will make the process spontaneously reverse direction, and restore the original state.

Any process is "reversible" if:

(1) there is only infinitesimal temperature difference causing any heat flow, and

(2) only infinitesimal force differences doing any work, and

(3) there are no dissipative forces (e.g., friction).

But there are many ways to draw system boundaries, and a process that is reversible within a subsystem, may not be reversible within the larger system. Therefore, it is important to distinguish between reversibility of changes that occur *within* a system, from changes that are done *to* a system by outside influences. We talk first about reversibility *within* a system.

### Reversibility *Within* a System

Consider a system comprising two parts: subsystem A and subsystem B. Imagine A is infinitesimally hotter than B; then heat flows slowly from A to B. It's reversible because with only an infinitesimal change of system conditions, we could make B slightly hotter than A, and let the heat flow back. A system with only infinitesimal temperature differences within it is in **quasi-equilibrium**, i.e., it's almost in equilibrium, but not quite. Because the temperature difference is very small, the heat flow rate is very slow.

In general, entropy increases when you equilibrate subsystems at finitely different temperatures. This is intuitive, because once you equilibrate finitely different temperatures, you can't separate them back into hot and cold parts without substantial effort. But a quasi-equilibrium process is (essentially) isentropic, that is, we can make the entropy change arbitrarily small by making the temperature difference small enough. Continuing our example of subsystems A and B, we can see mathematically why entropy does not increase when the system is in quasi-equilibrium. Imagine $đQ$ joules of heat flow from A to B:

$$dS_A \equiv \frac{đQ_{reversible}}{T_A} = \frac{-đQ}{T_A}, \qquad dS_B = \frac{đQ}{T_B}, \qquad dS_{system} = dS_A + dS_B$$

$$\text{But} \qquad T_A = T_B \qquad \Rightarrow \qquad dS_{system} = \frac{-đQ}{T_A} + \frac{đQ}{T_B} = 0\,.$$

Quasi-equilibrium is **isentropic** (constant entropy), and slow. But slow is not necessarily quasi-equilibrium (= isentropic). Consider a thermos bottle filled with hot water, sitting in a box (our system). The heat flows very slowly from the hot water into the box, because the thermos is well insulated, but the system is not quasi-equilibrium. There is a large temperature difference between the hot water and the air in the box. The entropy of the system increases as the temperatures equilibrate, even though it happens slowly.

Some references use the term “quasi-static” to mean “quasi-equilibrium.” This is misleading, because “static” implies motionless, and “quasi-static” suggests “slow.” But we have already seen that being slow doesn’t mean being isentropic. However, being nearly in equilibrium *does* mean isentropic, i.e. reversible.

When considering changes *within* a system, reversible implies slow, but slow does not imply reversible.

Notice that any friction or other dissipative losses (e.g., magnetic hysteresis) converts energy into heat, and cannot be undone by infinitesimal changes within the system.

## Reversible Changes *To* a System

Now let’s consider a different view: we have a system, and outside influences change its state. These changes can be made to happen slowly, with the system continually in quasi-equilibrium, or they may be made quickly, where the system doesn’t have time to equilibrate after transfer of significant heat or work. If changes are made slowly, the system is said to be changed *reversibly*, because infinitesimal changes in the outside influences can cause the system to slowly retrace its path back to its original state. In this case, slow *does* mean reversible! However, if the changes were done quickly, so that the system was knocked out of equilibrium, there is no easy way to make the system go out of equilibrium in the reverse direction. Such changes to the system are said to be *irreversible*.

When considering changes *to* a system, reversible and slow are equivalent.

For example, if a very hot reservoir conducts heat slowly through an insulator to a system, the system has no way of “knowing” how the heat got there, or that it was a large temperature difference. All the system can respond to is how fast heat flows into it. If it’s slow, the system is in quasi-equilibrium. In this case, the formula for the entropy change of the system can be written

$$dS = \frac{đQ_{slow}}{T} \qquad \text{(changes } to \text{ a system)}\,.$$

Notice that isothermal changes *to* a system (system remains at constant temperature) require the system to be in quasi-equilibrium the whole time, are necessarily slow, and therefore $dS = đQ/T$.

The reasons for “slow” are (1) to have well-defined temperatures throughout the system, and (2) to avoid accelerations, which result in kinetic energy of motion, which then is usually lost to the environment when the motion stops, through heat-of-deformation, friction, sound, etc. However, if there are no losses (e.g., if the system is constant volume, and does no work), then the internal energy increases by the heat input:

$$dU = đQ \qquad \text{(no losses)}\,,$$

no matter how fast or slowly the heat is transferred. In that case, since entropy is a function of the system state, we can remove the restriction of “slow,” and:

$$dS = \frac{đQ}{T} = \frac{dU}{T} \qquad \text{(any } fast\ or\ slow \text{ heat flow } to \text{ a system that has no energy losses)}\,.$$

Furthermore, the state of the system after energy input, with no energy losses, is determined only by the new internal energy. This means that the energy input doesn't have to be heat: it can be radiation, mechanical agitation, chemical reaction, etc. Therefore:

$$dS = \frac{dU}{T} \qquad \text{(any } \textit{energy} \text{ flow } \textit{to} \text{ a system that has no energy losses)} .$$

## The Second Law of Thermodynamics, and Touchy, Feely Entropy

From our Statistical Mechanics, we know that entropy is multiplicity, measured on a logarithmic scale, and usually measured in units of energy per unit temperature. Here are some qualitative implications of this.

The second law quantifies the fact that there are 2 kinds of energy: "ordered" energy, which can theoretically be converted without loss, and random or "thermal" energy, which is harder to get ordered work from. It's natural that a bunch of particles flying randomly about in all directions are less useful than an orderly chunk of energy waiting to be put to use, even if the two sources have the same energy. The second law quantifies how much useful work you can get out of thermal (random) energy. Any time you change energy from ordered to random, you decrease its usefulness. Hence, friction is bad. It's not that friction destroys energy; total energy is *always* conserved. But friction converts useful ordered energy into less useful thermal energy. Note that having separate hot and cold reservoirs is some kind of order, and more useful than having a single tepid reservoir. Hence, mixing hot and cold reservoirs wastes potential for useful work (i.e. increases entropy).

There are two kinds of energy: ordered, and thermal (random).
Ordered energy is more useful than thermal energy.

Crudely, entropy is a way of quantifying how random, or how disordered, the energy of a system is. Hence, for a given amount of energy, lower entropy means greater usefulness; higher entropy means less usefulness. As we've seen in Statistical Mechanics, however, it is not really disorder that matters, it is multiplicity of the system state. More accurately, entropy is a measure of the number of micro-states of a given macro-state (set of constraints), on a log scale. Lower multiplicity states evolve into higher multiplicity states; higher multiplicity states are those that often seem "disordered" to us, because there are lots of ways to be disordered, and only a few ways to be ordered.

**Thermodynamics is backwards:** in thermodynamics, temperature is fundamental, and entropy is derived from it. In SM, entropy is fundamental and temperature is derived from it. Before you learn SM, forget all your thermodynamics. The fundamental nature of entropy is why lots of theoretical work is done at constant entropy, instead of constant temperature.

## Refrigerators

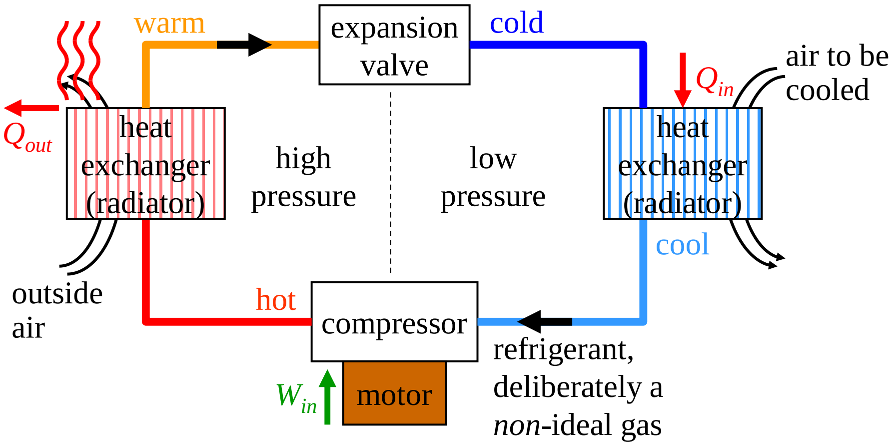


**Figure 6.1** A typical refrigerator.

At a high level, compressor/expander refrigerators work by:

1. "storing" the ability to get cold in the refrigerant, by compressing it. Then ...
2. bringing the refrigerant to ambient temperature with a simple heat exchanger, and ...
3. releasing the ability to get cold from the refrigerant (by expanding it), which brings it below ambient temperature.

One way to implement a refrigerator is to essentially "store" the ability to get cold in the refrigerant. We do this by compressing it, which makes it hotter. If we expanded it now, it would cool back to near where it started. But instead, we first run the refrigerant through a heat exchanger with the ambient, cooling the compressed gas down to ambient. But it's still compressed, and will cool off when we expand it. So we have, in a sense, "stored" the ability to get cooler in the gas. We now expand the gas (in an expansion valve, which is often called the "evaporator"). As the gas expands, it cools. We run the cold gas through a pipe in the box, and voila, the box gets cold. This also warms the gas slightly to the box temperature. We then compress the mostly-cold gas to complete the cycle, and return us to where we began.

More TBS.

### Why does a refrigerant get cold when it expands?

Because it has attractive inter-molecular forces. The energy vs. distance curve looks like Figure 6.2. The internal energy is fixed (if expanded quickly or adiabatically). But the potential energy *in*creases as the molecules separate, so the kinetic energy *de*creases. For unbound particles (a gas), temperature is proportional to kinetic energy. Lower kinetic energy means colder. QED.

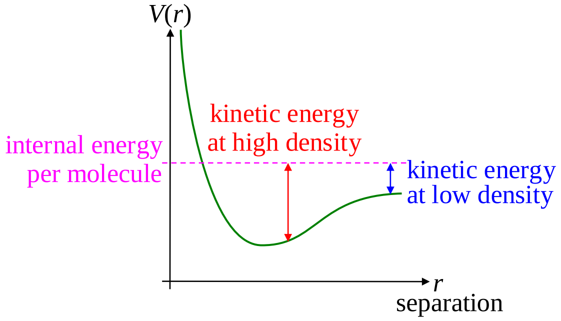


**Figure 6.2** Energy vs. distance for refrigerant molecules.

This means the refrigerant is *not* an ideal gas, because it has intermolecular forces. In fact, we want it *very* non-ideal, because stronger forces make the gas cool more when it expands.

## The Carnot Conspiracy

The Carnot cycle (or Carnot engine) is an important theoretical construction in thermodynamics, which sets the maximum possible efficiency for a heat engine to convert heat into work. Virtually all fossil fuel and nuclear energy conversions today are forms of heat engine: cars, trains, planes, and oil, coal, and nuclear power plants. The Carnot cycle is used in engineering practice as a standard for comparison: heat engine efficiency is sometimes stated in terms of "percent of Carnot," which is to say, what fraction of the theoretical maximum efficiency is achieved. For example, if the Carnot (theoretical maximum) efficiency for the given hot and cold temperatures is 60%, and the actual engine achieves 30% efficiency, then it is "50% of Carnot." This is a typical efficiency for small-scale conversion such as cars. Large power plants can afford more expensive designs that achieve 70-80% of Carnot, which is only about 34% - 40% conversion efficiency [Wolfson & Pasachoff, "Physics for Scientists and Engineers," 3rd Ed., Addison Wesley Longman, Inc., 1999, p551].

Carnot engine operation also illustrates several important concepts in thermodynamics. The Carnot engine is an idealization which can never be achieved in practice, and if it were, would take an infinite amount of time to run.  We describe the Carnot engine, from concepts, through construction, operation, and implications.  We then briefly describe the Carnot refrigerator, and give conclusions.

This section assumes you understand the concept of reversible (isentropic) processes.

**The Carnot Concept:**  In general, low entropy implies the ability to get useful work done.  Increasing entropy wastes the opportunity for useful work.  For example, if I have a hot gas, and a cold gas, I can make a heat engine which takes heat from the hot gas, exhausts heat to the cold gas, and creates useful work in the process.  However, if I simply mix the two gasses, I get no work, and end up with a warm gas that cannot (by itself) do work.  Such mixing increases entropy, and thus wastes the opportunity for work.  Any time I increase entropy, I am wasting some opportunity for work.  Therefore, if I make an engine in which every step is isentropic (reversible), I waste nothing, and achieve the maximum possible efficiency of converting heat to work.

**The Carnot Contraption:**  The Carnot engine uses a 4-step cycle, which converts heat from a hot reservoir into useful work and exhaust into a cold reservoir.  Each of the 4 steps is isentropic (reversible), and hence the entire cycle is isentropic.  Therefore, the Carnot engine has the maximum possible efficiency of converting heat to work.  As illustrated below, we can imagine the engine as a piston enclosing a gas, which expands and contracts during the cycle.

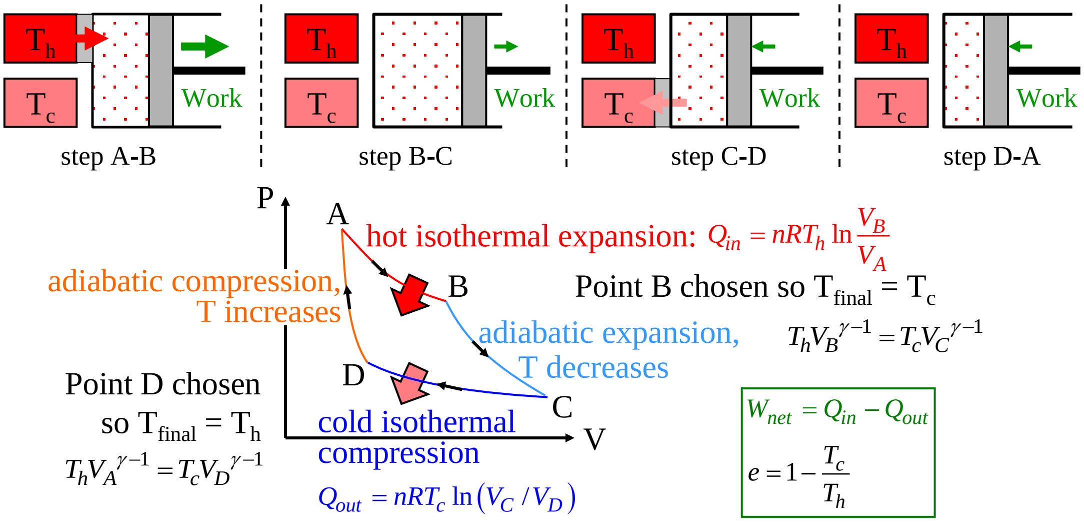


**Figure 6.3**  The Carnot engine (aka Carnot cycle).  Ideal gas formulas are given, but the engine operation is independent of the working fluid.

Note that the P-V diagram tells us nothing about temperature, or time.  Recall that all reversible operations must be slow, since they are quasi-equilibrium, and take place with only infinitesimal temperature differences and/or pressure differences.

Our cycle starts at point A, where the piston begins fully compressed, and the gas temperature is a smidge below $T_h$; the outside pressure is a smidge below the gas pressure.  The engine reaches out and touches the hot reservoir, so heat slowly flows in, and from the tiny pressure difference, the gas slowly expands and the piston pushes out, slowly doing work.  This continues to point B.  Since heat flows in at constant temperature, the entropy of the engine increases, and the entropy of the hot reservoir *decreases* by the same amount:

$$\Delta S_{A\text{-}B} = \int \frac{\text{đ}Q}{T} = +\frac{Q_{in}}{T_h} \qquad\qquad \text{engine entropy increases}\,.$$

Thus the entropy of the universe remains the same.  Note that this process is independent of the working fluid of the engine; in particular, it need not be an ideal gas.

At B, the engine releases its touch of the hot reservoir, and continues to slowly expand with no heat flow in or out (adiabatically).  The outside pressure (magically) follows the P-V diagram, always staying an infinitesimal amount below the gas pressure. This expansion does more work until point C, which is chosen to be the point at which the adiabatic expansion leaves the gas a smidge above $T_c$.

At point C, the engine reaches out and touches the cold reservoir. Heat flows slowly into the reservoir. The outside pressure is changed to slightly above the gas pressure (an infinitesimal change), so the piston slowly compresses, taking in work (but less than the outward work done already). Since heat flows out of the engine, the entropy of the engine decreases, and that of the cold reservoir increases, by:

$$\Delta S_{C\text{-}D} = \int \frac{đQ}{T} = -\frac{Q_{out}}{T_c} \qquad \text{engine entropy decreases .}$$

The entropy of the universe is still unchanged, and now the entropy of the engine has been restored to that of point A. Since $T_c < T_h$, and $|\Delta S_{C\text{–}D}| = |\Delta S_{A\text{–}B}|$, it must be that $Q_{out} < Q_{in}$, and net work is done on the piston and outside world.

At point D, the engine releases its touch on the cold reservoir. The outside pressure continues to slowly increase just enough to compress the piston, adiabatically, to point A. Point D is carefully chosen so that the temperature of the gas at point A is infinitesimally below $T_h$. This D-A step takes in more work. The cycle has now returned to its initial state.

The net effect of one cycle is that heat is transferred from the hot to cold reservoirs, and more work was done on the outside world than was taken in. The engine converted heat to (potentially useful) work. The entire process is reversible.

**The Carnot Contrivance:** Most discussions of the Carnot cycle omit the difficult requirements that the temperatures and pressures must be within a smidge of the outside conditions. While the engine itself can control its own gas temperature (by choosing the points B and D), the engine requires some kind of magic outside pressure regulator, which follows the gas P-V curve, but slightly lower for the expansion steps, and slightly higher for the compression steps.

**The Carnot Cooler:** The Carnot cycle can be run in reverse, such that it takes work in, and moves heat from the cold to hot reservoir, like a refrigerator. All that is necessary is that the infinitesimal temperature and pressure differences be reversed, so that the piston expands from A-D, heat comes in from the cold reservoir from D-C, the piston compresses from C-B, and heat goes into the hot reservoir from B-A. Each step in the cycle is still reversible.

**The Carnot Conclusion:** Because any increase in entropy represents a lost opportunity for doing useful work, a fully reversible (isentropic) process is necessarily maximum efficiency, because there is no entropy increase. The Carnot cycle is such an idealized process which, in principle, comes arbitrarily close to zero entropy increase, and therefore is the maximum possible efficiency in converting heat to work.

# 7 Etcetera

## Temperature and Internal Energy

The question of internal energy vs. temperature is a very important one. It's only for systems of non-interacting particles that an isothermal process causes no change in internal energy. Better stated, "For a system of negligible particle interaction around the mean particle separation, the internal energy is proportional to temperature." Note that this is true without reference to any "process." Hence, a process need not be isothermal, so long as the final temperature equals the initial temperature. By "negligible" we mean that the interaction potential difference over, say, 1/10 to 10 times the mean particle spacing is small compared to the particle kinetic energy. Inside 1/10 the mean spacing, we say the particles are colliding elastically.

We now derive that the internal energy of a non-interacting gas is proportional to temperature, following these steps: $E \alpha p^2$; $T$ is a derivative of entropy; entropy, $S$ is a known function of density of states; find $S$; find $T$; find $E$. $T$ and $E$ are proportional.

We choose our zero of potential at the mean spacing (same as at infinity, because the interaction potential is negligible beyond that distance). So the total particle energy is all kinetic:

$$E = \frac{p^2}{2m} \quad \Rightarrow \qquad dE \propto p\, dp$$

Now we recall that the *definition* of temperature is [see Introduction to Entropy, earlier]:

$$T \equiv \left( \frac{dS}{dE} \right)^{-1} .$$

(The derivation is nontrivial, but straightforward, and we take it as a given for now.) We also recall that we can *define* S to be (within an additive constant)

$$S \equiv k \ln \Omega \approx k \ln \left( \rho V \delta E \right)$$

where $\rho$ is the density of states in states/volume/energy,
$\delta E$ is a finite, but small, energy range (how to justify this??),
$V$ is an arbitrary system volume.

So now we must derive the density of states as a function of energy, and then we can simply plug into our formulas for entropy and temperature. Recall that the number of states in a small region of momentum $\delta p$ is the number of quantized lattice points (allowed momenta) in the spherical shell in momentum space of thickness $\delta p$. That is

$$\rho(E)\, \delta E \propto p^2 \; \delta p$$

where $\delta p$ is the variation in momentum giving rise to $\delta E$ variation in energy.

But from the first equations above, $p\, \delta p \propto \delta E, \;\; and \;\; p \propto E^{1/2}$, so

$$\rho(E)\, \delta E \propto p (p\, \delta p) \propto E^{1/2} \delta E \, ,$$

and at any arbitrary volume, $V$:

$$S \propto \ln\left(\rho(E)\,V\delta E\right) = \ln\left(E^{1/2}\,V\delta E\right) = \frac{1}{2}\ln E + \ln\left(V\,\delta E\right) = \frac{1}{2}\ln E + const.$$

$$\frac{dS}{dE} \propto \frac{1}{2E} \propto \frac{1}{E}$$

$$T \equiv \left(\frac{dS}{dE}\right)^{-1} \propto E$$

independent of volume. QED.

When there is particle interaction, $E$ is not proportional to $p^2$, and the above derivation fails; however since $KE$ (kinetic energy) *is* proportional to $p^2$, the above derivation is valid if we replace $E$ with $KE$. Hence, for “interacting” particles (with significant potential energy changes on the scale of the mean particle spacing), $T$ is not proportional to $E$, but $T$ α $KE$. This gives rise to the high-school definition of temperature as “a measure of the kinetic energy of particles in a gas.”

For example, for repulsive particles, if the gas expanded, the average potential energy between particles goes down, even if heat flow keeps their kinetic energies the same. So the total internal energy goes down, while the temperature remains the same. In this case, temperature is *not* proportional to internal energy, but *is* proportional to KE.

However, even the kinetic energy definition breaks down for bound particles (liquids and solids), hence the more general definition $T \equiv \left(\frac{dS}{dE}\right)^{-1}$.

## Negative Temperature (Negative β)

Negative temperature is hotter than positive temperature. Increasing $T$ just beyond $+\infty$ takes you to $-\infty$. Sounds ridiculous? That’s because $T$ isn’t always the best parameter of thermal equilibrium; sometimes $\beta$ is. $\beta$ can be a more natural measure of temperature than $T$, because negative temperatures go smoothly through $\beta = 0$, not $T = 0$ (in fact, $T$ jumps discontinuously from $+\infty$ to $-\infty$) [Wan p71m] [Pat fig 3.15 p81] [Kub p148m].

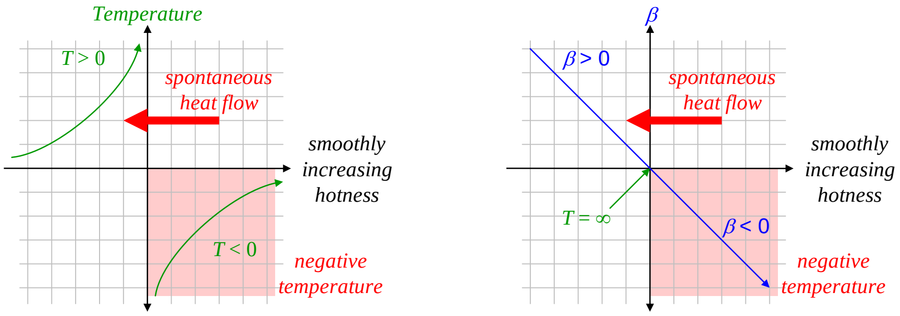


**Figure 7.1** Why $\beta$ is a better parameter for temperature than $T$, including negative temperature.

Negative temperature, or better, negative $\beta$, means that the probability of occupancy of an energy level *increases* with energy, instead of decreasing as with positive $\beta$ [Wan p68t]. For most systems, this is impossible because there are an infinite number of increasing energy levels: kinetic energy alone has an ever increasing density of states. If the probability of occupancy increased with energy, the system would contain an infinite amount of energy, which is unphysical. That is:

$$\Pr\left(E_r\right) \propto e^{-E_r/kT} \qquad \Rightarrow \qquad T > 0, \text{ or else probability blows up}.$$

However, some systems have a *finite* number of energy states, or a decreasing density of states with energy that integrates to a finite number. A spin system is an example of a system with a bounded energy density of states, and a finite number of energy states. Such systems can easily have internal energies that make higher energy states *more* likely to be occupied than lower energy states. Such a system has a negative $\beta$.

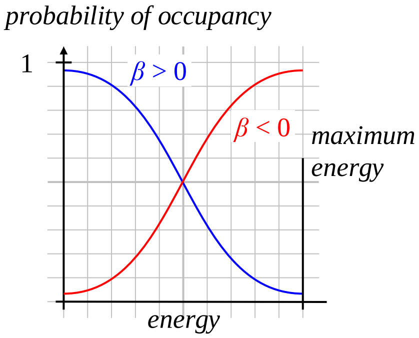


**Figure 7.2** Probability of occupancy vs. system energy for a finite system.

$$\beta \propto \frac{1}{T} = \frac{\partial S}{\partial E} .$$

Now, even for $\beta < 0$, larger $\beta$ is always colder, smaller $\beta$ is always hotter. Heat flows spontaneously from low $\beta$ to high $\beta$ [Wan p69m][Pat p83]. Ricipricaly, applying heat to the system lowers $\beta$. Heating beyond $\beta = 0+$ smoothly changes the system temperature to $\beta = 0-$.

To create such a system in a solid, there must be good isolation of the spins from the lattice. Otherwise, the lattice steals the energy, and the spin-subsystem reverts to the Boltzmann distribution at the lattice temperature.

# 8 Appendices

## Future Statistical Mechanifesto Topics

1. Why change in free energy is the maximal useful work to be gotten from change in system state (and that's why it's called "free" (available) energy).
2. Show how for bosons, multiple quantum states of the same energy affect the microscopic statistics. Since all energy levels can have any number of particles, show how having two levels at the same energy matters. It is due to distinguishability.
3. How can we treat Coulomb repulsion as just collisions, but not interactions? Because, over 99.9% of the volume, $\delta V << kT$.
4. Derive the Fermi-Dirac and Bose-Einstein energy distributions.
5. Grand potential (Schroeder)
6. Endothermic reactions: $N_2O \rightarrow N_2 + ½O_2 \;\Rightarrow\; S \rightarrow 1.5S$
7. Legendre transformations
8. Greg's example of a system of 3 doubly degenerate energy levels.
9. Note: avg over time = avg over ensemble. Ergodicity.
10. Photon gas: $E \sim VT^4$ (extensive). Related to $P = \sigma T^4$ ? For any ultra-relativistic particle, $PV = E/3$. For classical gas, $PV = 2E/3$. Photon eq. of state: photon reflects off wall, solid angle of impact is $2\pi$ steradians.

$$Force = \frac{\Delta p}{\Delta t} = \frac{2p\cos\theta}{(2L/c\cos\theta)} = \frac{pc}{L}\cos^2\theta$$

$$\left\langle \cos^2\theta \right\rangle = 1/3 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\pi/2}\cos^2\theta\sin\theta\,d\theta\,d\phi \qquad \Rightarrow \qquad P = \frac{pc}{3V} \;\Rightarrow\; \qquad PV = \frac{1}{3}E$$

11. Fermi level of non-interacting ultra-relativistic electrons:

$$N = \underset{\substack{spin\\factor}}{2}\int_0^{k_f}\frac{V}{(2\pi)^3}\,d^3k = \frac{Vk_f{}^3}{3\pi^2} \qquad \Rightarrow \qquad k_f = \left(3\pi^2\frac{N}{V}\right)^{1/3}$$

$$E_0 = 2V\int\frac{\hbar ck}{(2\pi)^3}\,d^3k = \underbrace{\frac{3}{4}\hbar ck_f}_{\langle E\rangle}\,N$$

12. $d^3k \rightarrow 4\pi k^2 dk$
13. Gaussian integral approximation (saddle point method)
14. S = -k sum($p_i$ ln $p_i$) [Wan p87]
15. Isothermal work done on a system decreases its entropy [Wan p103]

## References

[Bai] Baierlein, Ralph, Am. J. Phys., Vol. 69, No. 4, April 2001.

[Car] Carter, Ashley, *Classical and Statistical Thermodynamics*, Prentice-Hall, Upper Saddle River, NJ, 2001, ISBN 0-13-779208-5.

| | |
|---|---|
| [CRC] | *Handbook of Chemistry and Physics*, 60th ed., CRC Press, Boca Raton, FL, 1980, ISBN 0-8493-0460-8. |
| [Hil] | Hill, Terrell L., *An Introduction to Statistical Thermodynamics*, Dover Publications, Inc., New York, 1986, ISBN 0-486-65242-4. |
| [Hua] | Huang, Kerson, *Statistical Mechanics*, 2nd ed., John Wiley & Sons, Inc., 1987, ISBN 0-471-81518-7. |
| [Kub] | Kubo, Ryogo, *Statistical Mechanics*, North-Holland Personal Library, Amsterdam, May 3, 1990, ISBN-13: 978-0444871039. |
| [Pat] | Pathria, R. K., *Statistical Mechanics*, 2nd ed., Butterworth-Heinemann, Burlington, MA, 1996, ISBN 0 7506 2469 8. |
| [Rei] | Reif, Federick, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, 1965 (International edition 1985), ISBN 0-07-085615-X. |
| [Sch] | Schroeder, Daniel V., *An Introduction to Thermal Physics*, Addison-Wesley, 2000, ISBN 0-201-38027-7. |
| [Wan] | Wannier, Gregory H., *Statistical Physics*, Dover Publications, 1987, ISBN 048665401X. |
| [W&P] | Wolfson, Richard & Jay M. Pasachoff, *Physics for Scientists and Engineers*, 3rd Ed., Addison Wesley Longman, Inc., 1999. ISBN 0536-17026-6. |

---

# Glossary

Definitions of common Statistical Mechanics terms:

| | |
|---|---|
| $<x>$ | the average (sometimes called "expectation") value of '$x$' |
| absolute probability | a number in [0,1] (inclusive) that quantifies how likely an event is to occur. Given a large sample of trials, the fraction of occurrences of the event approaches its absolute probability. Zero means the event will never happen; one means the event happens with certainty. See also relative probability. |
| adiabatic | without heat transfer |
| Boltzmann distribution | Given a system in thermal equilibrium with a heat bath (i.e. at some fixed temperature), the Boltzmann distribution gives the *relative* probability of the system being in a *single* microstate of known energy. |
| Bose-Einstein distribution | The average occupation number of a single quantum state in a system of many bosons, which may be > 1. Cf "Bose-Einstein distribution" and "Maxwell-Boltzmann occupation number." |
| classical system | a hypothetical system obeying classical mechanics. This is useful as an approximation to real (i.e. quantum) systems. In the high energy limit, classical mechanics gives real results. |
| combinatorics | the mathematical study of permutations and combinations of things (elements of sets). |
| density of states | In SM, usually means "phase-space density of states." This is the number of states per unit "volume" (really hyper-volume) of phase space. Can also mean "energy density of states": # states per unit energy. |
| dynamic variable | generalized coordinate or momentum |
| energy state | micro-state |
| energy eigenstate | micro-state |
| ensemble | a real or imaginary set of "systems." Typically, the set is very large. |

| | |
|---|---|
| entropy | the multiplicity of a macro-state measured on a logarithmic scale. Entropy is a function of the macro-state of a system which helps compute thermal equilibrium parameters. |
| extensive | (briefly) a property that is proportional to system size (e.g., mass). |
| Fermi-Dirac distribution | The average occupation number of a single quantum state in a system of many fermions, which must be < 1. Cf "Bose-Einstein distribution" and "Maxwell-Boltzmann occupation number." |
| free energy | Where entropy is a measure of a system's ability to defy energy, free energy incorporates that defiance in units of energy, and is the actual quantity minimized by a system in equilibrium. There are several kinds of free energy, because each set of constraints on a system requires its own definition of free energy. |
| Gibbs distribution | = Boltzmann distribution = Maxwell-Boltzmann distribution [Wan ch. 4.??] |
| heat | random energy which can be transferred between different elements of the system. Commonly, heat is random kinetic energy of particles transferred by collisions. |
| ideal | particles have insignificant interactions over most inter-particle distances, and significant interactions (i.e., collisions) only at tiny fractions of the average inter-particle distance. |
| intensive | (briefly) a property that does not scale with the system size (e.g., $T$). |
| isobaric | constant pressure |
| isochoric | constant volume |
| isothermal | constant temperature |
| macro-state | An observable state of a system, which could conceivably be measured in a lab. The macro-state includes observables such as temperature, pressure, volume, magnetization, etc. See also micro-state. |
| Maxwell-Boltzmann speed distribution | The distribution of particle speeds in a gas in thermal equilibrium. |
| Maxwell-Boltzmann energy distribution | The distribution of particle energies in a gas in thermal equilibrium. |
| Maxwell-Boltzmann occupation number | The approximate average occupation number of a given quantum state in a sparse system (average occupation number << 1). |
| micro-state | A single state of a system, including the states of all of the system's constituent particles. Classically, a micro-state includes each particle's position and momentum. Quantum mechanically, a micro-state is an $N$-particle energy eigenstate. It is a wave function that describes the state of every particle in the system. In either case, a micro-state usually cannot be observed or determined in a laboratory. However, a knowledge of micro-states from physical theories is essential to applying Statistical Mechanics to predict observable phenomena. |
| multiplicity | the number of micro-states with a given macro-state |
| normalization factor | a number by which you can divide (or perhaps multiply) relative probability to get absolute probability. |
| occupation number | The average number of particles in a state, for some fixed $N$ total particles. The occupation number is a relative probability of occupancy for the energy. |
| PDF | Probability distribution function: $\Pr\left(x < X < x + dx\right) = pdf_X(x)\, dx$ . |
| phase-space | a $6N$ dimensional space of coordinates and their conjugate momenta. Each particle has 3 position coordinates (e.g., $x$, $y$, $z$), and 3 momenta (e.g., $p_x$, $p_y$, $p_z$). |
| probability | absolute probability. |

| | |
|---|---|
| quantum system | a system obeying quantum mechanics, as all real systems do. |
| relative probability | a number that quantifies how likely an event is to occur, out of a pool of *N* samples. Given *N*, absolute probability = (relative probability) / *N*. Even without *N*, you can compare the relative probabilities of two events, to determine the ratio of their absolute probabilities. See also absolute probability. |
| smidge | an infinitesimal amount |
| usually | always. In SM, saying a system is "usually" in the most-likely macro-state pretty much means it is "always" in that state. |

---

## Formulas

Generalized factorial: $a! \equiv \int_0^\infty x^a e^{-x}\,dx, \qquad a \text{ real}$

Gamma function: $\Gamma(a) \equiv (a-1)! = \int_0^\infty x^{a-1} e^{-x}\,dx, \qquad a \text{ real}$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots, \quad |x| < 1 \quad or \qquad \ln x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \cdots, \quad 0 < x \le 2$$

Stirling's approximation for factorials (or Γ( ) functions):

For large *a*: $\ln a! \approx a(\ln a - 1) \qquad or \qquad a! \approx (a/e)^a$

$S = k \ln \Gamma(E, \Delta E) \qquad \Gamma(E, \Delta E) \equiv \text{phase space volume for } \ E < E_{system} < E + \Delta E$ (Hua 6.27 p134)

$S = k \ln \Omega(E) \qquad \Omega(E) \equiv \text{phase-space density of states}$ (Hua 6.28 p134)

$S = k \ln \Sigma(E) \qquad \Sigma(E) \equiv \text{phase space volume of } \ E_{system} < E$ (Hua 6.28 p134)

$$S = -\frac{\partial A}{\partial T} \qquad where \quad A \text{ is any free energy}$$

$$S = k\left(\ln Z + \beta \bar{E}\right) = k\left(\ln \sum_j^{\substack{energy\\levels}} \Omega(E_j) e^{-\beta E_j} + \beta \bar{E}\right) \approx k\left(\ln\left(\Omega(\bar{E}) e^{-\beta \bar{E}}\right) + \beta \bar{E}\right) = k\left(\ln \Omega(\bar{E}) - \beta \bar{E} + \beta \bar{E}\right)$$

$$= k \ln \Omega(\bar{E})$$

Entropy of ideal gas (Sackur-Tetrode equation):

$$S = Nk\left\{\ln\left[\frac{V}{N}\frac{1}{\hbar^3}\left(\frac{U}{N}\frac{4\pi m}{3}\right)^{3/2}\right] + \frac{5}{2}\right\} = Nk\left\{\ln\left[v\frac{1}{\hbar^3}\left(u\frac{4\pi m}{3}\right)^{3/2}\right] + \frac{5}{2}\right\}, \qquad v \equiv \frac{V}{N}, \quad u \equiv \frac{U}{N}$$

Helmholtz free energy: $A(N,V,T) \equiv U - TS = -kT \ln Z \qquad P = -\dfrac{\partial A}{\partial V}$

Gibbs free energy: $G \equiv U - TS + PV = A + PV = \mu N$ [Car 9.11 p154, & inside cover]

Enthalpy: $H = U + PV$

Chemical potential: $\mu \equiv \dfrac{\partial A}{\partial N} \quad where \quad A \equiv \text{appropriate free energy for the given system}$

|   |   |   |
|---|---|---|
| V | A | T |
| U | ↗ ↖ | G |
| S | H | P |

$$Z(\beta,V,N) \equiv \underbrace{\frac{1}{N!}}_{\substack{\text{indistin-}\\ \text{guishable}}} \frac{1}{h^{3N}} \int_\infty d^{3N}p \, d^{3N}q \, e^{-\beta \mathcal{H}(p,q)}$$ (classical canonical ensemble)

$$= \frac{1}{N!} \int_{E_0}^{\infty} dE \, \Omega(E) e^{-\beta E}$$

$$Z(\beta,V,N) \equiv \sum_{\{n_j\}} \exp\left[ -\beta \sum_{j=1}^{\substack{energy\\ levels}} \varepsilon_j n_j \right], \qquad where \quad \sum_j^{\substack{energy\\ levels}} n_j = N$$ (quantum canonical ensemble)

Ideal gas: $Z(\beta) = \left(2\pi m / \beta\right)^{3N/2} V^N$

$$\mathcal{Z}(\beta,V,\mu) = \sum_{N=0}^{\infty} e^{\beta\mu N} Z(\beta,V,N) = \sum_{N=0}^{\infty} \sum_{\{n_j\}} \exp\left[ -\beta n_j \left( \varepsilon_j - \mu \right) \right]$$ (grand canonical ensemble)

*but* $\mu = \mu(T,V,N)$ may vary within the summation

$\ln \mathcal{Z} = \beta PV$

$$\lambda_T = \sqrt{\frac{h^2}{2\pi mVT}} \qquad \textit{Ideal gas density of states, 3D}: \ \rho(E) = \frac{V}{h^3} 2m^{3/2} E^{1/2}$$

$$E = -\frac{\partial}{\partial \beta} \ln Z \qquad \mathrm{var}(E) = \frac{\partial^2}{\partial \beta^2} \ln Z = kT^2 C_V$$

Fermi pressure (low $T$): $PV = \frac{2}{5} \varepsilon_F N$

Boson critical ρ: $\rho_C = \dfrac{2.612}{{\lambda_T}^3}$

---

# Index

The index is not yet developed, so go to the web page on the front cover, and text-search in this document.