Statistical Mechanifesto

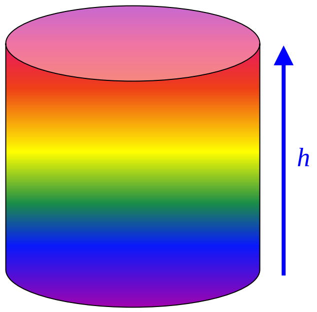
or

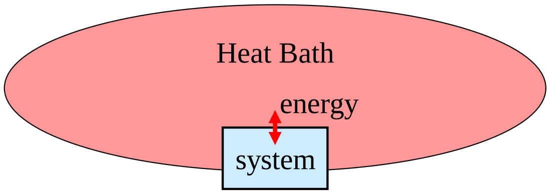
Funky Statistical Mechanics Concepts

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





> “What is most important in studying a physical problem is to grasp it as a problem in *physics*. ... Training in mathematical methods should not be ignored, but it would be a serious mistake if one was to be dazzled by the mathematics and forget the physics.”
> -Ryogo Kubo, *Statistical Mechanics*, p vii.

* 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, 9/2/2024.

2006 values from NIST (now superseded by updates). 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) $\equiv 4.184 \text{ x } 10^{12}$ J = 1 Teracalorie

bar $\equiv$ 100,000 N/m^2

atm $\equiv$ 101,325 N/m^2 = 1.013 25 bar

torr $\equiv$ 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) ρ (= N/V or M/V)
Number of particles	N	ρ (= N/V or M/V)
Number of moles	n	
Helmholtz free energy	A or F	
Gibbs free energy	G	μ

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, $\mathbf{r}$. $\mathbf{E}(\mathbf{r})$ 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 Foundations

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? Is he dense, or an airhead? How energetic is he? How stable is he? How does he respond to pressure? 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 also 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:

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

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

Imagine the archetypal mole of diatomic 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:

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

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?**

$$v \equiv \frac{V}{N} = \frac{.0246\ m^3}{6.02\times 10^{23}} = 4.1\times 10^{-26}\text{ m}^3 \qquad \text{particle spacing} \sim v^{1/3} = 3.4\times 10^{-9}\text{ m or }3.4\text{ nm}.$$

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:

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

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 eV, Mr. Mole is a pretty stable guy.

??Express in kJ/mol.

We can also see this from per-particle values: at 300 K, kT = .026 eV = 4.2×10^{-21} J, which is small compared to the bond energy of ~ 1 eV.

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

$$E=\frac{1}{2}mv_{rms}{}^2 \qquad \Rightarrow \qquad v_{rms}=\sqrt{\frac{2E}{m}}=\sqrt{\frac{2\cdot6.21\times10^{-21}\,J}{28\cdot1.67\times10^{-27}\,kg}}=515\text{ m/s}\,.$$

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:

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

$$mfp\sim\frac{v}{A}=\frac{4.1\times10^{-26}\text{m}^3}{1.6\times10^{-20}\text{ m}^2}=2.6\times10^{-6}\text{m}\sim3\times10^{-6}\text{m}\ \ or\ \ 3\,\mu\text{m}\,.$$

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

$$mttc\sim\frac{mfp}{v_{rms}}\sim\frac{2.6\times10^{-6}\,m}{515\,m/s}\sim5\times10^{-9}\text{s}\ \ or\ \ 5\text{ ns}\,.$$

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. Phase changes are also part of the system thermodynamics.

The Fundamental Postulate of Statistical Mechanics

Thermodynamics began as a study of systems, heat, and energy. Every system has a **macrostate**, which comprises observable properties of the system, that could conceivably be measured in a lab. The macrostate includes 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 macrostate are billions of *microstates* 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 transferring a little energy to it from another molecule, thus slowing down the second molecule . In the sense of microstates, the modified system is different than the original. However, in the sense of macrostates, the difference is imperceptible, and the system is in the same macrostate 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 **microstate** is a single state of a system, including the exact states of all of the system's constituent particles. Classically, a microstate includes each particle's position and momentum. Quantum mechanically, a microstate is an N-particle quantum state that describes the state of every particle in the system. In either case, a microstate cannot be observed or determined in a laboratory. However, a knowledge of microstates 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 has fixed volume, and 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 microstate of a system must satisfy that system's macroscopic constraints. As noted earlier, besides constraints, a macrostate also includes a set of observables. However, knowledge of the system and its macroscopic constraints alone allows us to compute all the microstates (or all those that matter), and then predict the equilibrium macrostate of the system.

The fundamental postulate of Statistical Mechanics is that
"All microstates 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 macrostates (that's really the definition of probability). As we will see, having all

---

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

microstates equally likely does *not* mean that all *macrostates* are equally likely. Far from it. Macrostates are summarizations of microstates, and hence lose a lot of information. This is what makes "most-likely" macrostates from equally-likely microstates.

As an example of a "most-likely" macrostate arising from equally likely microstates, 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 macrostates; for macroscopic systems, "usually" essentially means "always." The probability of observing anything else is so fantastically small that it would be unlikely to happen in a billion billion lifetimes of the universe.

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:

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

The fundamental thermodynamic results are applications of these concepts. We'll talk about all of them. Much of thermodynamics is an accounting of how the (conserved) energy is most likely to be distributed throughout the parts of a system ("energy, energy, who's got the energy?").

A more advanced concept is phase changes, and their critical exponents.

What Is ...?

What is a Thermodynamic System?

What is a thermodynamic system? A system where heat is a "significant" part of its operation. This clearly includes steam engines and internal combustion engines. As counter examples, many systems require simple cooling, but that may be too simple to be "significant". For example, a simple electric motor with a cooling fan: I would not consider that a thermodynamic system.

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 or p), 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 (mass). Now consider a subset of the system (a subsystem). 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 subsystem 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.

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 of systems:

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.)
3. A single-particle quantum state can be a system. Such a state may cover the volume of a larger system, but have a definite momentum.

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 simple with the right choice of "system."

Tiny systems: Note that if you choose a single particle as your "system", then the system's microstates are the single-particle quantum states for that particle: a "single-particle state" is *also* a "microstate," and vice-versa. More later.

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?

There are 3 kinds of "state" in common use in statistical mechanics: a macrostate, a microstate, and a single-particle state.

(1) In general, a **macrostate**, or "system state", is a condition of a system *with definite observables*. What we mean by "condition" depends on the details of the system.

(2) A **microstate** is a multi-particle quantum energy eigenstate of a system (definite energy); it includes position-state (configuration), momentum-state, and any spin state. Almost always, a macrostate has many microstates that are consistent with it. This number of microstates is called the **multiplicity** of the macrostate.

(3) A **single-particle state** is an quantum energy eigenstate for a single particle of a system; it is a solution to the 1-particle Schrodinger equation. For example, the allowed states of the electron in a hydrogen atom are single-particle states. In a large system, such as a big box of gas, a single-particle state is an energy eigenstate for a gas atom. We define the term **wave-function** to be the spatial part of a quantum state; the term **spin-state** specifies, well, the quantum spin-state of the particle, which is a finite dimensional state vector (of dimension $2s + 1$).

A particle's full quantum state is the combination
of the particle's wave-function and its spin-state.

Tiny systems (again): Note that if you choose a single particle as your "system", then its microstates are the single-particle quantum states for that particle. There is then no distinction between a "single-particle state" and a "microstate," and $g(E) = \omega(E)$.

Tiny systems, of just one or a few particles, may have noticeably discrete energies. E.g., an atom has discrete excitations. In this case, we describe the energies with a discrete set of g_j and E_j, which are the microstates.

Later, we'll see that tiny systems do not have a well-defined temperature, but they can be in dynamic equilibrium with a temperature reservoir, which *does* have a well-defined temperature.

Semi-classical systems: Sometimes we use a semi-classical mechanics definition of an energy "(micro)state", but many calculations still must use the energy density of (micro)states (as a function of momenta), derived from quantum mechanics. In the case of a gas, a microstate is an idealized N-particle 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 every particle's (x, y, z, p_x, p_y, p_z). For large energies, we *approximate* the possible states as continuous in position, momentum, and energy.

Be careful to distinguish an energy *state* from an energy *level* (defined below). This applies to both single-particle states and levels, and to microstates and their energy-levels.

What is an Energy Level?

In contrast to a single-particle state, which has a definite energy, there are **energy levels**, which are the set of energies of the single-particle states. "Energy levels" are also called **allowed energies**. Note that there are often many (quantum) states of the same energy (called **multiplicity** [Wan p84b] in statistical mechanics, and "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 its multiplicity g_i. These definitions imply that many sums can be written two ways: in terms of single-particle energy states, or in terms of energy levels:

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

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

Distinguish carefully between energy (eigen)*states*, and energy *levels*:
a single energy level may comprise many (degenerate) energy states.

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.

Even some respectable authorities sometimes say that
"heat is infrared radiation." This is wrong .

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

What is the Thermodynamic Limit?

We frequently speak of the "thermodynamic limit."

The **thermodynamic limit** 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 a density of states, $g(E)$ or $\omega(E)$, taken as a continuous function:

$$Z_1(T) \equiv \int_\infty g(E) e^{-E/kT}\, dE, \qquad Z(T) \equiv \int_\infty \omega(E) e^{-E/kT}\, dE .$$

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:

- Single-particle density of states per unit energy; each particle is in exactly one such state.
- Energy density of states for the system, each single-particle state possibly being occupied by one or more particles (recall that multiple bosons can occupy a single-particle state).
- Energy density of states per unit volume for the system, each single-particle state possibly being occupied by one or more particles.
- Phase-space density of states.
- *k*-space density of states (equivalent to momentum density of states).

We now describe each in turn. We then follow with a derivation of some common densities of state from first principles.

Notation Summary

The most common densities are these:

E_j, g_j	Energy level, and its multiplicity, for a single-particle.
$g(E)$	Energy density of single-particle states, in the continuous approximation. This is sometimes per unit volume of the system.
ω_m	Multiplicity of microstates in system energy level *m*.
$\omega(E)$	Energy density of system microstates, in the exceedingly good continuous approximation.
$\Phi(E)$	Number of (system) microstates with energy < *E*, so $\omega(E) = \partial\Phi/\partial E$.
$\Omega(E)$	Provisional: Multiplicity of system microstates within a small range ΔE of *E*: $\Omega(E) \equiv \omega(E)\,\Delta E \qquad \text{where} \quad \Delta E \text{ is (for now) small but arbitrary} . \qquad (2.1)$ We will refine this definition later in eq. (3.5).

Energy Density of States Per Particle

Energy density of states can be the number of single-particle stationary quantum states (energy eigenstates) per unit energy. As a simple example, consider a hypothetical “harmonic oscillator molecule,” one which oscillates in some way within a harmonic oscillator potential $V(x) = kx^2$. Then from elementary quantum mechanics, the energy states are uniformly spaced at $\hbar\omega$. Therefore, the single-particle density of states is:

$$g_{particle} = \frac{1}{\hbar\omega} \text{ states per unit energy, per particle (for harmonic oscillator)} .$$

The units of $[g_{particle}]$ are $[\mathrm{E}]^{-1}$, or sometimes per unit volume $[\mathrm{E}]^{-1}[\mathrm{L}]^{-3}$. Single-particle energy density of states is usually a function of energy, so we write it $g(E)$, dropping the subscript. We explore all the densities of various states further in a later section.

Energy Density of (Micro)states For a System

The **system energy density of microstates** is generally complicated. It depends on whether the particles are fermions, bosons, or distinguishable particles, and how dense they are. We discuss this more later.

Energy density of microstates is usually a function of the energy of the system, so we write it $\omega(E)$. We define $\Phi(E)$ is the number of states with energy $< E$ (following [Rei]). Then:

$$\omega(E) \equiv \frac{d}{dE}\Phi(E) .$$

Φ is dimensionless, and $[\omega] = [E]^{-1}$. $\omega(E)$ usually increases with E, but there are some important exceptions to this (e.g., negative temperature spin systems). Strictly speaking, $n(E)$ is a sequence of discontinuous steps, at the allowed quantized energies. The energy spacing is incredibly tiny, so we usually use the continuum approximation for $n(E)$, so we can treat it as a smooth, differentiable function.

Many references use Ω to mean different things, so carefully check the definitions in each reference.

For an ideal gas, the density of states per unit volume is *not* a function of temperature. Therefore, for an ideal gas at constant volume, the density of states is independent of temperature. However, in general, the density of states may depend on temperature. For 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. Another example: solid-state systems such as semiconductor crystals have a density of states that depends on the lattice spacing. The lattice spacing is weakly dependent on temperature, and so the density of states is also weakly dependent on temperature. Such a small temperature dependence may often be ignored [Wan p85t].

Energy Density of Microstates Per Unit Volume For a System

The energy density of microstates 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 ("non-interacting" is often a good approximation). Because this density of states is per unit volume, it is independent of the size of the system.

Phase-Space Density of States

In classical SM, you sometimes use phase-space density of states, which is often called (ambiguously) just "density of states." We assume you are familiar with phase-space from classical mechanics. (We'll quantize this shortly.)

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

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

Analogous to the position vector in real space, $\mathbf{r}$, you may think of Γ as a $6N$-dimensional 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:

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

Each point in phase space specifies the positions and momenta (i.e., the classical "state") of *all N* particles of a system, at some instant of time.

But quantum mechanics says that only some regions in phase space are allowed energy eigenstates: a quantum state is actually spread-out over a region of phase space. An energy eigenstate usually has a narrow momentum uncertainty, and a wide position uncertainty. Therefore, each eigenstate can be represented as a small-volume region in phase space, but in practice, we think of the eigenstate as a phase-space point “in the middle” of the allowed region. (We call these allowed states **microstates**.) The set of all allowed quantum states is a collection of such points in phase space. Then the **phase space density of states** is the number of quantum states per unit (hyper-)volume of phase space:

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

It turns out that for large gas systems (i.e., in the thermodynamic limit), the phase-space density of states approaches the constant h^{3N}. [Some references use hand-waving about the uncertainty principle to guess this result, but the logic is not valid. We do a better job later.]

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); 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 $d\Gamma_s$ 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., constant-volume gases, 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 specifies the positions and momenta (i.e., the classical “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.”

k-space Density of States

All of our density of states derive from a single quantum principle: the wave-function has to fit an integer number of half-waves into the system "box". Equivalently, the wave-function must go to zero at the edges of the system. This means that *all* systems have quantized momentum: nonrelativistic, relativistic, or any other dispersion (energy-momentum) relation. Since $\mathbf{p} = \hbar\mathbf{k}$, we treat **p** and **k** as essentially equivalent.

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.) **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 6D space representing its position and momentum. For the gas, we would have N points in the 6D space. This 6D space is 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:"

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

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 distribution of momenta for particles in an ideal gas is a μ-space particle density such that the density is uniform over all x, y, z positions (in the absence of an external potential). 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. In contrast, for a gas in a container with significant gravity, the μ-space particle density would be a function of height z, and of (p_x, p_y, p_z).

Declaration of Independence: What Makes a Macrostate?

We often write state functions as functions of 3 independent variables, e.g. entropy is $S(N, V, T)$. However, we get to choose which 3 variables are our independent variables. Typically, for idealized isolated systems, we would specify (1) one of E, T, or S to indicate how hot something is, and (2) V for its volume or else p for its pressure, and (3) N for number of particles, or n for moles, or M for mass. So we could write:

$$\begin{array}{llll} A(N,V,T), & E(N,V,T), & \text{and } S(N,V,T) & \text{or} \\ A(n,V,S), & E(n,V,S), & \text{and } T(n,V,S) & \text{or} \\ A(N,V,E), & T(N,V,E), & \text{and } S(N,V,E), & \text{etc.} \end{array}$$

[The volume of an isolated system is usually fixed, because expanding or contracting involves work that exchanges energy with the environment.] There may also be other external parameters, such as electric-field, or magnetic H- or B- fields.

To fully specify the state of a system, you must specify at least one extensive variable. To specify only the intensive properties of a simple system, you must specify two independent intensive variables.

Non-equilibrium systems: 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 changes during the approach to equilibrium. Mathematically then, T (of the bath) is fixed, but one of E or S is still a non-equilibrium state parameter of the system, so we might have:

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

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.

Computing Single-Particle Densities of States

Density of states is a crucial concept for statistical mechanics. The term "density of states" is much abused; densities of states come in many forms, but in quantum and statistical mechanics, "density of states" always refers to number density of stationary quantum states per unit of some parameter space. Recall that stationary quantum states are energy eigenstates. We consider these common single-particle densities of states:

- 1D harmonic oscillator
- 1D particle in a box
- 1D free particle
- 2D free particle
- 3D free particle

See *Quirky Quantum Concepts* for a similar, but more detailed, discussion of densities of states. We consider N-particle density of states (i.e., density of microstates) in the next section.

We consider only the density of spatial states (wave-functions). Fermions, e.g. electrons, would have twice as many states, because each spatial state has 2 spin-states. Similarly, photon states are twice as dense, because each spatial state has 2 polarization states.

Our analysis refers to 3 different "spaces": real physical space (1D, 2D, or 3D), k-space, and phase-space. Since $p = \hbar k$, we take momentum-space to be essentially equivalent to k-space. Phase-space is a "tensor product" of real-space and momentum-space:

$$V_{\text{phase-space}} = (Vol)(V_{\text{momentum-space}}) \quad \Rightarrow \qquad dV_{\text{phase-space}} = d^3r \; d^3p \; .$$

Bad Example (because of unbound states, and $\rho(E = 0) \to \infty$: 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:

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

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

Inverting the equation gives $n(E)$, and then differentiating gives $\rho(E)$:

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

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

[This is a somewhat tricky example, because $E < 0$, and E approaches zero from below. That means $\rho(E) \to \infty$ as $E \to 0$, which is true for the bound electron states of hydrogen.]

Another bad example: For a quantum particle in a 1D 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 $E^{-1/2}$, i.e. the energy gaps *increase* with energy. In this case, the continuum approximation fails because the energy states do not approach a continuum with increasing energy. We don't discuss such systems in this work.

Some references define the density of states (either single-particle or microstates) as $\rho(E) \equiv \sum_s^{\text{states}} \delta(E - E_s)$. This is not a useful form, because to perform a calculation, we need to integrate some function times $\rho(E)$: $x = \int f(E)\rho(E)\, dE$. To evaluate the integral, we need $\rho(E)$ as an analytic function (at least piecewise). (Also, the δ-functions are still approximations, because all systems have some uncertainty in their energy levels.)

1D harmonic oscillator: Perhaps the simplest density of states is that for a 1D harmonic oscillator. We write the number of states per unit energy as $\rho(E)$. Since the energy spacing is constant, $\hbar\omega$, the number of states per unit energy is (Figure 2.1a):

$$\rho(E) = \frac{1}{\hbar\omega} \qquad \text{(1D harmonic oscillator)} .$$

The units are $[E]^{-1}$, i.e. inverse energy, equivalent to "quantum states per unit energy". The momentum (and therefore wave vector, k) of the particle in any stationary state is uncertain, and does not figure into this density of states.

Multidimensional and multiparticle harmonic oscillators are essential to the study of solids. Their density of microstates behaves quite differently than this 1D case, but we do not address them in this section.

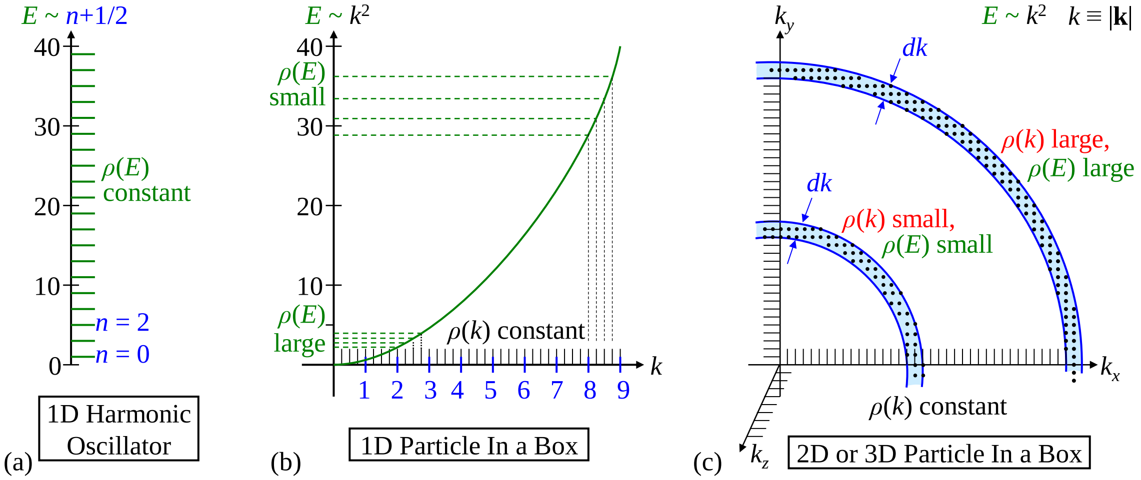


Figure 2.1 Densities of states.

1D particle in a box

Another simple density of states is 1D particle in a box. We will see that the result *per unit energy* is the same as for a free particle, and use this as a warm up for 3D. Each state has a purely real wave-function, which must be 0 at the boundaries. This leads to quantized wave-functions (ignoring normalization):

$$\psi(x)=\sin(kx)=\frac{1}{2i}\left(e^{ikx}-e^{-ikx}\right) \qquad \textit{where} \quad k=n\pi/L,\ \ L=\text{length of box},\ \ n=\text{quantum number}\,.$$

Each energy eigenstate is a superposition of + and – momentum eigenstates. Because k is easily quantized (an integer number of ½ wavelengths: $kL = n\pi$; Figure 2.2a), we first write the density of states parameterized by k. At first, we let k take only positive values, knowing that each positive value of k describes a state which is a standing-wave superposition of both $+k$ and $–k$ momenta. We wrote k as a function of n above, but for a density of states, we must invert this relation to get number density per unit of positive k:

$$n(+k)=\frac{kL}{\pi} \qquad \Rightarrow \qquad \rho_+(k)=\frac{dn}{dk}=\frac{L}{\pi}, \qquad \text{in units of } [\text{rad/m}]^{-1}=[\text{m}]\,. \tag{2.2}$$

But taking into account that each state consumes a $+k$ and $–k$ value, the final (two-sided) density per unit k (of either sign) is only half the above:

$$\rho(k)=\frac{1}{2}\frac{dn}{dk}=\frac{L}{2\pi} \qquad \text{(1D particle in box)}\,. \tag{2.3}$$

Thus, for a given size box, the density of energy eigenstates per unit k is constant. Note that there is no upper bound to k (or E).

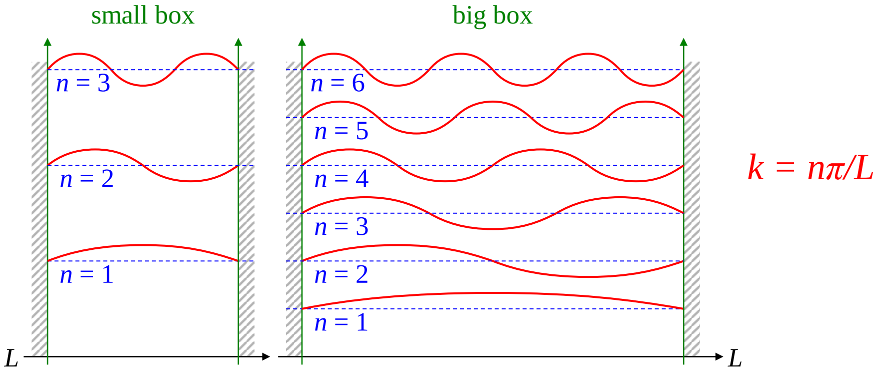


Figure 2.2 Quantum states for a for particle-in-a-box, showing that the density of states, $\rho(k)$, is proportional to the box size. In 3D, $\rho(k)$ is proportional to the volume, where $k \equiv |\mathbf{k}|$.

Rewriting our two-sided formula for $n(k)$ in terms of momentum introduces a general principle:

$$n(k)=\frac{kL}{2\pi}=\frac{pL}{2\pi\hbar}=\frac{\text{phase-space-volume}}{2\pi\hbar}\,.$$

For large systems, the number of quantum states enclosed by boundaries of p and the real-space volume is closely given by the "volume" of the phase-space , divided by $2\pi\hbar$ ($= h$).

It's as if each quantum state takes up h units of phase space volume [Gos p177]. In this case, phase space extends over positions 0 to L, and momenta $–\hbar k$ to $\hbar k$. The number of states in this phase-space volume is:

$$\#\ \textit{states} \equiv n(k)=\frac{V_{\text{phase-space}}}{2\pi\hbar}=(L-0)\big(\hbar k-(-\hbar k)\big)=\frac{2L\hbar k}{2\pi\hbar}=\frac{Lk}{\pi}\,,$$

in agreement with $n(k) = kL/\pi$, equation (2.2) above, since the quantum number n counts the total number of states up to k.

We also notice that a bigger box allows for more states per unit k, since more k values are allowed (Figure 2.2b). Equation (2.3) shows that the density of states per unit k is proportional to the size of the box. So sometimes we say, generally, that the density of states per unit k, per unit size L, is a constant $1/2\pi$ (still in 1D):

$$\frac{\rho(k)}{L} = \frac{1}{2\pi} \qquad \text{in units of } [k]^{-1}[L]^{-1} \;=\; [\text{rad/m}]^{-1}[\text{m}]^{-1} = \text{dimensionless} .$$

It is usually convenient to have the density of states as a function of energy, $\rho(E)$, rather than $\rho(k)$. This allows for easily integrating things like $\int dE\,(...)\,\rho(E)$, which appears in many SM formulas. We can find $\rho(E)$ from $\rho(k)$ by a change of variables:

$$\rho(E)\,dE = \rho(k)\,dk \qquad \Rightarrow \qquad \rho(E) = \rho(k)\frac{dk}{dE} .$$

Then:

$$E = \frac{(\hbar k)^2}{2m} \qquad \Rightarrow \qquad k(E) = \frac{(2mE)^{1/2}}{\hbar}, \qquad \frac{dk}{dE} = \left(\frac{1}{2}\right)\frac{\sqrt{2m}}{\hbar}E^{-1/2} .$$

$$\rho(E) = \underbrace{\frac{L}{2\pi}}_{\rho(k)}\,\underbrace{\frac{1}{2}\frac{\sqrt{2m}}{\hbar}E^{-1/2}}_{dk/dE} = \frac{L}{h}\sqrt{\frac{m}{2E}} \qquad \text{(1D).} \qquad (2.4)$$

We see (Figure 2.1b) that in 1D, while $\rho(k)$ is constant), $\rho(E)$ is large at low E, and decreases with E (opposite of the 3D case, below). As for all energy densities of states, the units are $[\text{E}]^{-1}$, equivalent to “quantum states per unit energy”.

For a massless particle, e.g. a photon, $\rho(k)$ is the same, but the E-k relation (dispersion relation) is now:

$$E(k) = \hbar\omega = \hbar ck \;\; \Rightarrow \qquad k(E) = \frac{1}{\hbar c}E, \qquad \frac{dk}{dE} = \frac{1}{\hbar c} \qquad \Rightarrow \qquad \rho(E) = \rho(k)\frac{dk}{dE} = \frac{L}{\pi\hbar c} .$$

Notice that since E is always positive (kinetic energy), we used the 1-sided $\rho(k)$ to find $\rho(E)$. Equivalently, we could have used the 2-sided $\rho(k)$, and then doubled the result since $E(k) = E(-k)$ are two states of the same energy.

1D Free Particle

The particle is essentially free, but we consider it to be in a large box, and take the limit as $L \to \infty$. Since it has well-defined momentum (k is known), its position is therefore spread widely (uncertainty principle). Since there is no potential, the energy eigenstates are also momentum eigenstates. We simply use the density of states for a particle in a box of length L, but at the end of any calculation, we take the limit as $L \to \infty$. For any realistic calculation, this limit must be well-defined, and be proportional to L. Thus from (2.3) and (2.4):

$$\rho(k) = \lim_{L\to\infty}\frac{L}{2\pi} \qquad \rho(E) = \lim_{L\to\infty}\frac{L}{2\pi\hbar}\sqrt{\frac{m}{2E}} \qquad \text{(1D particle in large box)} . \qquad (2.5)$$

We take the two-sided view of k ($+k$ and $-k$ are separate states), since left-going and right-going free particles are separate states.

Many references [Bay p264t; Sak p339m; P&C p4m, p406m; Tow p406m] find the density of states by considering a box of size L, and imposing *periodic* boundary conditions: the complex wave-function at the right and left edges must be the same. We are unaware of any physical justification for this requirement, but it leads to the same result. Periodic boundary conditions means an integer number of cycles fits in the box. This yields a different quantization condition for k:

$$kL = n2\pi \qquad \Rightarrow \qquad n(k) = \frac{kL}{2\pi}, \qquad \rho(k) = \frac{dn}{dk} = \frac{L}{2\pi}, \text{ units} = [\text{m}] \quad \text{(periodic BC).}$$

Notice that now both $+k$ and $-k$ are separate states, so the number of states up to a given *magnitude* of k is the same as a particle in a box. Therefore, the (phase-space volume)/($2\pi\hbar$) is again the number of quantum states enclosed by a given magnitude of k. Then $\rho(E)$ is also the same as for a particle in a box. For massless particles, $\rho(E)$ is also the same: $\rho(E) = L/\pi\hbar c$.

2D free massive particle

Each of k_x, and k_y are quantized separately as in the 1D free particle. The 2D wave-vector **k** then must sit on a lattice point in 2D **k** space (Figure 2.1c, without k_z). The k-space 2D volumetric (really area) density of these points, i.e. points per unit k-volume, is:

$$\rho(\mathbf{k}) = \rho_x(\mathbf{k})\rho_y(\mathbf{k}) = \frac{L_x}{2\pi}\frac{L_y}{2\pi} = \frac{Area}{(2\pi)^2} \qquad \text{in units of } [\text{rad/m}]^{-2} = [\text{m}]^2 .$$

where *Area* is the "volume" of the box in *real* 2D space (not k-space). Much like the 1D case, we usually quote the density of states per unit k-volume per unit (real-space) "volume":

$$\frac{\rho(\mathbf{k})}{Vol} = \frac{1}{(2\pi)^2} \qquad \text{in units of } [\text{rad/m}]^{-2}[\text{m}]^{-2} = \text{dimensionless} .$$

Though derived for a rectangular area,
this result is quite general for any shape large area [P&C p4m].

In 2D space (x, y), phase space is 4-dimensional, (x, p_x, y, p_y). Define $k \equiv |\mathbf{k}|$. The phase space, up to some maximum k_{max}, goes over *Area*, and over the 2D k-space disc up to radius k_{max}.

In 2D, for large k (enclosing many lattice points), we closely approximate the number of states by the k-space "volume" times the k-space density of states:

$$\#\,states \equiv n(k) = \int_0^k d^2k'\,\rho(k') = \int_0^k \underbrace{dk'\,2\pi k'}_{\text{thin annulus area}} \frac{Area}{(2\pi)^2} = \pi k^2 \frac{Area}{(2\pi)^2} .$$

(For small k, this approximation for number of states is less accurate.) To prepare for finding $\rho(E)$, we switch from k to momentum p:

$$\#\,states = \pi \frac{p^2}{(2\pi\hbar)^2} Area = \frac{\left(V_{\text{momentum-space}}\right)\left(Area\right)}{h^2} = \frac{V_{\text{phase-space}}}{h^2} \qquad where \quad p \equiv \hbar|\mathbf{k}| .$$

To find the energy density of states, $\rho(E)$, for free (spinless) particles (or particles in a 2D box), we change variables as before, but now there are many **k** vectors with the same energy (all the **k** on a circle of radius $k \equiv |\mathbf{k}|$):

$$\rho(E) = \rho(k)\,2\pi k \frac{dk}{dE}, \qquad where \quad k = \frac{(2mE)^{1/2}}{\hbar}, \quad \frac{dk}{dE} = \frac{(2m)^{1/2}}{2\hbar} E^{-1/2} .$$

$$\rho(E) = \frac{Area}{(2\pi)^2} 2\pi \frac{(2mE)^{1/2}}{\hbar} \frac{(2m)^{1/2}}{2\hbar} E^{-1/2}$$

$$= Area \frac{2\pi}{(2\pi\hbar)^2} m \qquad \text{in units of } [E]^{-1} .$$

In 2D we see that $\rho(E)$ is constant with E, unlike the 1D case (Figure 2.1b).

3D massive particle, free or in a box

Each of k_x, k_y, and k_z are quantized separately as in the 1D free particle. The 3D wave-vector **k** then must sit on a lattice point in 3D **k** space (Figure 2.1c). The k-space volumetric density of these points, i.e. points per unit k-volume, is:

$$\rho(\mathbf{k}) = \rho_x(\mathbf{k})\rho_y(\mathbf{k})\rho_z(\mathbf{k}) = \frac{L_x}{2\pi}\frac{L_y}{2\pi}\frac{L_z}{2\pi} = \frac{Vol}{(2\pi)^3} \qquad \text{in units of [rad/m]}^{-3} = [\text{m}]^3 .$$

where *Vol* is the volume of the box in *real* 3D space (not *k*-space). Much like the 1D case, we usually quote the density of states per unit *k*-volume per unit (real-space) volume:

$$\frac{\rho(\mathbf{k})}{Vol} = \frac{1}{(2\pi)^3} \qquad \text{in units of [rad/m]}^{-3}[\text{m}]^{-3} = \text{dimensionless} . \tag{2.6}$$

Though derived for a rectangular volume,
this result is quite general for any shape large volume [P&C p4m].

The "volume of phase-space" relation extends to higher dimensions as:

$$N \equiv \#\text{ of generalized coordinates of the system}, \qquad \dim(\text{phase-space}) = 2N$$

$$\#\ states = \frac{V_{\text{phase-space}}}{h^N} .$$

In 3D space (x, y, z), phase space is 6-dimensional, (x, p_x, y, p_y, z, p_z). Define $k \equiv |\mathbf{k}|$. The phase space, up to some maximum k_{max}, goes over *Vol*, and over the 3D *k*-space ball up to radius k_{max}.

For large *k* (enclosing many lattice points), we closely approximate the number of enclosed states by the *k*-space volume times the *k*-space density of states:

$$\#\,states \equiv n(k) = \int_0^k d^3k'\,\rho(k') = \int_0^k \underbrace{dk'\,4\pi k'^2}_{\text{thin shell volume}} \frac{Vol}{(2\pi)^3} = \frac{4}{3}\pi k^3 \frac{Vol}{(2\pi)^3} .$$

(For small *k*, this approximation for number of states is less accurate.) To prepare for finding $\rho(E)$, we switch from *k* to momentum *p*:

$$\#\,states = \frac{4}{3}\pi\frac{p^3}{(2\pi\hbar)^3}Vol = \frac{\left(V_{\text{momentum-space}}\right)\left(Vol\right)}{h^3} = \frac{V_{\text{phase-space}}}{h^3} \qquad where \quad p \equiv \hbar|\mathbf{k}| .$$

To find the energy density of states, $\rho(E)$, for spinless particles (free or in a 3D box), we change variables as before, but now there are many **k** vectors with the same energy (all the **k** on a sphere of radius $k \equiv |\mathbf{k}|$):

$$\rho(E) = \rho(k)4\pi k^2\frac{dk}{dE}, \qquad where \quad k = \frac{(2mE)^{1/2}}{\hbar}, \qquad \frac{dk}{dE} = \frac{(2m)^{1/2}}{2\hbar}E^{-1/2} .$$

$$\rho(E) = \frac{Vol}{(2\pi)^3}4\pi\left(\frac{2mE}{\hbar^2}\right)\frac{(2m)^{1/2}}{2\hbar}E^{-1/2} \tag{2.7}$$

$$= Vol\,\frac{4\pi\sqrt{2}}{h^3}m^{3/2}E^{1/2} \qquad \text{in units of } [E]^{-1} \qquad \left[\text{Goods 1.3.106 p66}\right].$$

The units are equivalent to "quantum states per unit energy". In 3D we see that $\rho(E)$ increases with E, opposite to the 1D case (Figure 2.1b).

3D massless particle, free or in a box: As always, the *k*-space density is the same, but the massless *E*-*k* relation is $E = c\hbar k$:

$$\rho(E) = \rho(k) 4\pi k^2 \frac{dk}{dE}, \qquad where \quad k = \frac{E}{\hbar c}, \qquad \frac{dk}{dE} = \frac{1}{\hbar c}.$$

$$\rho(E) = \frac{Vol}{(2\pi)^3} 4\pi \left(\frac{E^2}{\hbar^2 c^2} \right) \frac{1}{\hbar c} \qquad .$$

$$= Vol \frac{4\pi}{(2\pi\hbar)^3 c^3} E^2 \qquad \text{in units of } [\text{E}]^{-1} .$$

This is very different behavior than the massive density of states, increasing much more rapidly with E.

It can be proved that:

For large volumes or high energies,
the density of states per unit k per unit volume is essentially $\rho(k)/Vol = 1/(2\pi)^3$,
regardless of the shape of the volume or the boundary conditions applied.

This shape-independence follows from short-wavelength considerations similar to the WKB approximation (and historically, led to the "old" (now abandoned) Wilson-Sommerfeld quantization rule).

Note The k-volume density of states per unit *Vol* (real-space volume) is a constant $1/(2\pi)^3$, leading to integrals of the form $d^3k/(2\pi)^3$. Another related situation looks similar: the inverse Fourier transform (from k basis to x basis) also leads to integrals of the form $d^3k/(2\pi)^3$ (depending on the normalization choice). However, the integrands and limits are usually different, and you should be careful to understand the distinction.

Ideal Gas Density of (Micro)states

The concept of entropy requires an understanding of the ideal gas density of microstates, as a function of energy, to validate the approximations needed to define entropy. We develop that background here.

Consider a system of N identical, non-quantum ideal gas particles. In other words, the particle density is small enough that quantum statistical effects are negligible. We can derive its density of microstates $\omega(E)$ generalizing the method for single-particle density of states. Consider a system state of constant energy KE; it satisfies:

$$KE = \frac{1}{2m}\left(p_{1,x}^2 + p_{1,y}^2 + p_{1,z}^2 + p_{2,x}^2 + \ldots + p_{3N,x}^2 + p_{3N,y}^2 + p_{2N,z}^2 \right) = \frac{1}{2m}\sum_{i=1}^{3N} p_i^{\,2} .$$

In $3N$-dimensional momentum space, this describes a hypersphere of constant KE. For example, when N = 1, we have an ordinary sphere (called a 2-sphere), with surface area $4\pi R^2$. For any N, the number of microstates in a small interval dE is the (hyper)surface area of the hypersphere times dE, much like Figure 2.1, but in $3N$-dimensional momentum space. In general, the hypersphere surface area is given by $S_d R^d$ ($d \equiv 3N - 1$), and [from Wikipedia 2023-02-11]:

$$S_d = \frac{2\pi^{(d+1)/2}}{\Gamma\big((d+1)/2\big)} \qquad where \quad d \equiv 3N - 1.$$

$$\text{E.g.,} \qquad S_2 = \frac{2\pi^{3/2}}{\Gamma(3/2)} = \frac{2\pi^{3/2}}{(1/2)\pi^{1/2}} = 4\pi, \qquad \text{using } \Gamma(x) = (x-1)\Gamma(x-1), \text{and } \Gamma(1/2) = \pi^{1/2}.$$

[Since $\Gamma(d) = (d - 1)!$, for large N (and d), we can use Stirling's factorial approximation also for Γ(large x), but we won't need that here.]

We find the N-particle $\omega(E)$ by continuing the generalization to higher dimensions that led from the 1D single-particle density (2.5) to the 3D single-particle density (2.7), and at first ignoring quantum indistinguishability. Here, **k** is a vector in $3N$-dimensional k-space, and the k-space density of states is (cf. (2.6)):

$$\omega(\mathbf{k}) = \frac{(Vol)^N}{(2\pi)^{3N}} \quad \textit{where} \quad Vol \equiv \text{real-space 3D volume} . \tag{2.8}$$

For $k \equiv |\mathbf{k}|$, the number of states between k and $k + dk$ is:

$$\omega(k)\, dk = S_d k^d \left(\frac{Vol}{(2\pi)^3} \right)^N dk . \tag{2.9}$$

Then in terms of energy $E = (\hbar k)^2/2m$:

$$\begin{aligned} \omega(E) &\equiv \frac{dn}{dE} = \frac{dn}{\underbrace{dk}_{\rho(k)}} \frac{dk}{dE} = \underbrace{S_d k^d \left(\frac{Vol}{(2\pi)^3} \right)^N}_{\rho(k)} \frac{(2m)^{1/2}}{2\hbar} E^{-1/2} \\ &= S_d \left(\frac{Vol}{(2\pi)^3} \right)^N \frac{(2m)^{1/2}}{2\hbar} \left(\frac{\sqrt{2mE}}{\hbar} \right)^d E^{-1/2} = S_d (Vol)^N \frac{(2m)^{3N/2}}{2h^{3N}} E^{(3N-2)/2} . \end{aligned} \tag{2.10}$$

For $N = 1$, we get $d = 2$, $S_d = 4\pi$, and this reduces to our single-particle result (2.7).

To account for the indistinguishability of the particles, we must reduce this result by a factor of $N!$.

Ideal Gas Multiplicity

In our final definition of macroscopic entropy (3.6), we will approximate the multiplicity $\Omega(E)$ as the total number of microstates of all energies up to E, called $\Phi(E)$. This multiplicity is in-turn well-approximated from the volume inside a k-space hypersphere of dimension $3N$ [Wikipedia 2023-06-05]:

$$V_{3N}(R) = \frac{\pi^{3N/2}}{\Gamma\left(3N/2+1\right)} R^{3N} .$$

The k-space density of states is still (2.8), so the number of enclosed microstates in a ball of radius k is:

$$\Phi(k) = \frac{V^N}{(2\pi)^{3N}} \frac{\pi^{3N/2}}{\Gamma\left(3N/2+1\right)} k^{3N} \quad \textit{where} \quad V \equiv \text{real-space 3D volume} .$$

Then in terms of total system energy $E = (\hbar k)^2/2m \sim NkT$:

$$k = \frac{\left(2mE\right)^{1/2}}{\hbar} \qquad \Rightarrow \qquad \Phi(E) = V^N \frac{\pi^{3N/2}}{\Gamma\left(3N/2+1\right)} \frac{\left(2mE\right)^{3N/2}}{h^{3N}} . \tag{2.11}$$

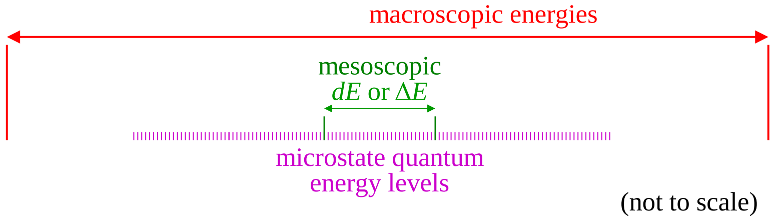
To account for the indistinguishability of the particles, we must reduce this result by a factor of $N!$.

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 microstates. It follows from the calculations above, and it's very important. The calculations above counted discrete energy states. Therefore, it seems like we must avoid constraining our system energy too tightly, or we run into the "granularity" of allowed quantum energies. The standard "trick" is to (at first) allow each system some small *constant* latitude in energy, ΔE [Hua p176b, Kub p7b, Rei p??]. ΔE must be large compared to quantum energy spacings (Figure 2.3), i.e. must include a large number of quantum microstates so they look like a continuum. It may be comparable to our differential energy changes:

$$\text{microstate quantum spacing} << dE \sim \Delta E << \text{observable energy changes.} \tag{2.12}$$



(not to scale)

Figure 2.3 Qualitative view of the three scales of energy: quantum, mesoscopic, macroscopic.

This allows us to use the energy continuum approximation ($\Delta E >>$ quantum spacing), while still constraining the system to almost fixed energy, at least far below measurement accuracy. In practice, this is a very realistic requirement. However, for a proper definition of entropy in a system at a fixed temperature (canonical ensemble), we will later remove any need for (or reference to) ΔE.

As numerical examples of energy densities, we preview some results derived later, starting with (2.13). For a single electron in a 1 cm^3 box, at room temperature (300 K):

$g(E) \sim 10^{21}$ states per eV. We could take $dE \sim 10^{-10}$ eV. Macroscopic energy $\sim kT \sim 10^{-3}$ eV.

For a tiny system of only $N = 10$ electrons in a 1 cm^3 box, at 300 K, we use (2.10). The total system energy is $(3/2)NkT$. The numbers are so large, they exceed the capability of most handheld calculators:

$\omega(E) \sim 10^{198}$ states per eV. We could take $dE \sim 10^{-10}$ eV. Macroscopic energy $\sim NkT \sim 10^{-3}$ eV.

With 10^{198} states per eV, the continuum "approximation" is more than justified. Most systems have $N >> 10$, often $N > \sim 10^{20}$. Then the density of microstates $\omega(E)$ is incomprehensibly large.

Despite these fantastic numbers, statistical mechanical theory
agrees extremely well with experiment.

Impartial Derivatives and the Constants of Nature

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:

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

However:

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

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:

Let $w(x, y) = xy$ with an equation of state $x = yz$.

Then $w(x, z) = x^2 / z$

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

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

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

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

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

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:

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

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:

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

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 as $M \to \infty$ of the average of M trials. 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: What Do You Seek?

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 among the new set of allowed microstates of the system.

The equilibrium macroscopic state is the most likely macrostate for the system.
It has the largest probability-weighted sum of microstates.

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.

Other Numerical Guideposts

It is important to have a numerical sense for common thermodynamic quantities. Following the idea of Mr. Mole, we compute here some other numerical reference points. We routinely use the conversions:

$$\text{eV} = \text{J} / 1.602\times10^{-19}, \qquad \text{K} = 11{,}606 \cdot \text{eV} .$$

Thermal energy

1. What is kT at room temperature (300 K):

$$kT = 1.38\times10^{-23}\ \text{J/K}\left(300\ \text{K}\right) = 4.14\times10^{-21}\ \text{J} = 0.0258\ \text{eV} = 300\ \text{K}$$

Single-Particle

2. What is the ground-state energy of an electron in a 1 cm^3 box, and a 1 m^3 box? The minimum 1D wave-number k fits π radians in the box, so in 3D:

$$\begin{aligned} &k_x = \pi / .01 = 314\ \text{rad/m}, \quad p_x = \hbar k = 2.08\times10^{-31}\ \text{kg-m/s} \\ &E = 3\frac{p_x^{\,2}}{2m} = 7.1\times10^{-32}\ \text{J} = 4.5\times10^{-13}\ \text{eV} = 5.2\times10^{-9}\ \text{K} \qquad \text{(electron)}. \end{aligned} \tag{2.13}$$

Note that in a 1 m^3 box, the momenta are 100× smaller, so the energies are $10^4\times$ smaller. The energy scales inversely with the mass, so for a proton ($m = 1.67\times10^{-27}$ kg):

$$E = 3.9\times10^{-35}\ \text{J} = 2.4\times10^{-16}\ \text{eV} = 2.8\times10^{-12}\ \text{K} \qquad \text{(hydrogen)} .$$

These are extremely small energies.

For a photon, $E = c|\mathbf{p}|$. If the box is conducting, the photon goes to zero at the faces, just like the other particles, so:

$$E = \sqrt{3}\, c p_x = 1.1\times10^{-22}\ \text{J} = 6.7\times10^{-4}\ \text{eV} = 7.8\ \text{K} \qquad \text{(photon)} .$$

3. What is the single-particle energy density of states for an electron at room temperature (300 K) in a 1 cm^3 box? We'll use (2.7) in SI units to start:

$$E = \frac{3}{2}kT = \frac{3}{2}\left(1.38\times10^{-23}\text{ J/K}\right)(300\text{ K}) = 6.21\times10^{-21}\text{ J}.$$

$$\rho = (.01)^3 \frac{4\pi\sqrt{2}}{\left(6.63\times10^{-34}\text{ J-s}\right)^3}\left(9.11\times10^{-31}\text{ kg}\right)^{3/2} E^{1/2} \qquad (2.14)$$

$$= 4.2\times10^{39}\text{ J}^{-1} = 6.7\times10^{20}\text{ eV}^{-1} = 5.8\times10^{16}\text{ K}^{-1} \qquad (\text{electron}).$$

This is already a very high density of states. Furthermore:

- Since $\rho \propto Vol$, a 1-liter container has 1,000x higher density of states.
- Since $\rho \propto m^{3/2}$, a hydrogen atom has 7.8×10^4 times higher density of states.
- Since $\rho \propto T^{1/2}$, at 1 K, the density of states is only $\sqrt{300} = 17$x smaller.

4. What is a typical k value for an electron in a 1 cm^3 box, at room temperature? We compute the 1D value for k_x (or k_y or k_z), which is quantized as $n\pi/L$:

$$k_x = \frac{\sqrt{2mE}}{\hbar} = \frac{\sqrt{2m(1/2)kT}}{\hbar} = \frac{\left(9.11\times10^{-31}\cdot kT\right)^{1/2}}{\hbar} = 9.26\times10^{7}\text{ rad/m}$$

$k \equiv |\mathbf{k}|$ is $\sqrt{3}$ times bigger, and is not be quantized as $n\pi/L$.

Multiparticle System

5. What is the density of microstates for 10-electrons in a 1 cm^3 box, at room temperature? The numbers are so large, they exceed the capability of most handheld calculators:

$$N = 10,\ d = 29$$

$$\omega(E) = \frac{2\pi^{15}}{14!}\frac{\left(10^{-6}\text{ m}^3\right)^{10}}{2}\left(\frac{2\cdot 9.11\times10^{-31}\text{ kg}}{h^2}\right)^{15}\left(10\cdot 6.21\times10^{-21}\right)^{14}$$

$$= \left(6.6\times10^{-4}\right)\left(5\times10^{-61}\right)\left(4.1\times10^{36}\right)^{15}\left(6.21\times10^{-20}\right)^{14} = \left(6.5\times10^{21}\right)\times10^{195} \qquad (2.15)$$

$$= 6.5\times10^{216}\text{ J}^{-1} = 1.04\times10^{198}\text{ eV}^{-1} = 9.0\times10^{193}\text{ K}^{-1}$$

6. What is a typical k value for this 10-electron system, at room temperature? Note that k here is the magnitude of the 30-dimensional system $\mathbf{k}$ vector.

$$k = \frac{\sqrt{2mE}}{\hbar} = \frac{\sqrt{3mNkT}}{\hbar} = \frac{\left(3\cdot 9.11\times10^{-31}\cdot 10kT\right)^{1/2}}{\hbar} = 1.6\times10^{8}\text{ rad/m}.$$

While $\mathbf{k}$ is quantized to k-space lattice points, $k \equiv |\mathbf{k}|$ is not quantized to $n\pi/L$, and for large values such as that above, is nearly a continuum.

3 Equilibrium, the One True Entropy, and Temperature

Entropy is almost surely the most misunderstood concept in statistical mechanics and thermodynamics. We introduce it here in stages: starting with the big picture, and continuing to more detailed definitions, and consequences. Recall eq. (2.1): **multiplicity** Ω is the number of ways (microstates) that a given system can meet a given macrostate (i.e., observable conditions, or "observables"), such as energy, volume, temperature, magnetization, pressure, etc. In other words, we speak of the multiplicity of a *set of constraints* (a macrostate, aka set of conditions), which is the number of microstates that satisfy those constraints. Different constraints have different multiplicities (and therefore different entropies).

Not So Equal: A Preview of Equilibrium

We often hear, "Systems evolve to highest entropy." In general, they do not.
Or, "Systems evolve to lowest energy." In general, they do not.

Consider a warm glass of water on a table. It cools down. That's *lower* entropy (not higher). Consider a cold glass of water on a table. It warms up. That's *higher* energy (not lower).

If systems don't, in general, evolve to highest entropy, or lowest energy, what *do* they evolve to?

Systems spontaneously evolve to lowest *free energy*.

A full discussion must wait a few chapters. But for now, **free energy** is a combination of real energy, entropy, and temperature. The simplest form of free energy is the Helmholtz free energy: $F \equiv E - TS$. Since E, T, and S are (macro)state functions, free energy is also a state function (independent of the path a system took to get in the macrostate). When a system is in thermal equilibrium its free energy is minimized compared to other nearby macrostates. If the system is in thermal equilibrium with an environment, then additionally, its temperature is the same as the environment's.

In the special case of an isolated system, it *does* evolve to the state of highest entropy. Since an isolated system has fixed energy E, highest entropy is also lowest free energy.

In the special case of a system where thermal energy is negligible compared to system energy (such as a die on a table: it does not visibly fluctuate), the system evolves to lowest energy. In such a system thermal energy is negligible, $TS << E$, and lowest energy is also lowest free energy.

Board Game Entropy

Perhaps the simplest familiar example of entropy is in throwing a pair of dice. What is the *most likely* outcome (total of dice)? Board game players know it is 7. Why? Do dice have a preference for the number 7? Is this magic? No: 7 is most likely simply because there are more ways to get the sum of two dice to be 7 than any other number. In fact, there are 6 ways to get two dice to sum to 7. We call the sum of the dice our "macrostate." We say the **multiplicity** of the macrostate 7 is 6: there are 6 ways to meet the condition that the sum is 7.

Another example: there's only 1 way to get two dice to sum to 2 (snake eyes). The multiplicity of the macrostate 2 is 1. We call the multiplicity Ω, and it is a function of the macrostate. Thus:

$$\Omega(\text{sum=7}) = 6, \qquad \Omega(\text{sum=2}) = 1, \qquad \Omega(\text{sum=3}) = 2 \ldots .$$

With dice, there is still a large chance that you'll see an outcome which is not *most likely*. You don't always roll a 7; you see other numbers quite a bit. But for large numbers of particles (like trillions of molecules), the law of large numbers turns *probabilities* into *certainties*: the states you observe are the most-likely states, and you will observe no others.

Dealing directly with multiplicity often proves inconvenient, so we usually work with its logarithm instead. We define **entropy** as simply a measure of multiplicity on a log scale:

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

Thus the entropy of a set of constraints (a macrostate) is well-defined. It is up to the analyst to choose appropriate constraints for the task at hand.

We usually 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 may mislead a scientist.

A better macroscopic visualization of entropy is *uniformity of distribution* throughout a system.

More on uniformity later.

We will see that entropy is a major driver of energy, thermodynamics, and the universe. Furthermore, we will show that:

Systems do *not* tend toward maximum entropy;
they do *not* tend toward minimum energy.

They *do* tend toward a minimum of a combination of energy and entropy (called free-energy), which we develop fully later.

Cutting Through the Static

We frequently hear discussions of entropy that include dice (above), or coins, or cards. But they are often (usually?) misleading. A typical example is: here is a deck of cards that I just shuffled, and here's one sorted in order. Which has higher entropy? This is not meaningful, because both decks are in a single, specific order. Each has only 1 way to be in that order, so $\Omega = 1$ for both. Therefore, $S = k \ln 1 = 0$ for both.

A meaningful question is: what is the entropy of a shuffled deck of cards? There are 52! ways for a deck to meet the condition of being shuffled. Therefore, a shuffled deck of cards has $S = k \ln 52! = 2.2\times10^{-21}$ J/K. (This is tiny.)

We call a specific configuration of a system (such as a specific ordering of cards) a **microstate**. By definition, the multiplicity of every microstate is 1. It is not meaningful to ask for the entropy of a microstate.

Putting the Dynamics in Thermo

There is a deep flaw in looking at a specific, discrete macroscopic system (cards, coins, dice, ...) and talking about its "entropy." Thermodynamic systems are *dynamic*: they constantly swirl amongst trillions of configurations, driven by their thermal energy. In contrast, discrete macroscopic systems don't swirl: cards don't shuffle themselves, dice don't roll themselves, coins don't flip themselves .... That's important.

Entropy is defined *only* for *dynamic* systems with a given set of constraints.

Other Kinds of Constraints

There are many other properties of a system that one might constrain. For example, a box of gas of spin-1/2 particles with a magnetic dipole moment. We impose an external magnetic field, and the system responds with a net magnetization M. We can now ask, "What is the entropy of a system with N particles, energy E, volume V, and net magnetization M?" Then:

$$\Omega = \Omega(N,V,E,M) \qquad \Rightarrow \qquad S = S(N,V,E,M)\ .$$

This entropy is smaller than the entropy of the same system when M is unconstrained. (This is the basis of some real-world magnetic cooling processes [ref??].)

Equilibrium, Entropy, and Temperature In a Nutshell

We briefly introduce equilibrium, entropy, and temperature here. A later section gives more insight. Figure 3.1 shows an isolated system, comprising two subsystems, A and B. The subsystems may be dissimilar, and their details are not important. They are in thermal contact, and therefore can exchange energy. By definition of "isolated", the energy E of the total system is fixed. The question is: in equilibrium, how will the energy be divided between subsystems A and B? Note that thermal equilibrium is a dynamic equilibrium, with small random fluctuations of energy wandering throughout a system. Over time, though, there is no net change in macroscopic properties.

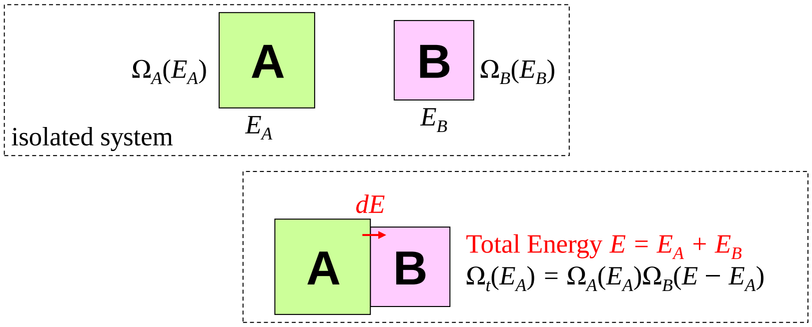


Figure 3.1 Two subsystems (A and B) brought into thermal contact.

Recall the equal likelihood postulate (p12): all microstates consistent with a given macrostate are equally likely. This implies the system will evolve to the equilibrium macrostate with the most number of microstates, i.e. with the highest multiplicity $\Omega_t = \Omega_A\Omega_B$. From the definition of entropy:

$$S \equiv k \ln \Omega : \qquad S_t = k \ln \Omega_t = \underbrace{k \ln \Omega_A}_{S_A} + \underbrace{k \ln \Omega_B}_{S_B} \qquad \textit{where} \quad k \equiv \text{Boltzmann's constant} . \tag{3.1}$$

The macrostate of highest multiplicity is also the macrostate of highest entropy. This is why isolated systems evolve toward the macrostate of maximum entropy.

Note that the multiplicities of A and B depend on how much energy each has:

$$S_A = S_A(E_A), \qquad S_B = S_B(E_B) = S_B\left(E - E_A\right).$$

Pursuing our question: in equilibrium, how much energy is in A? In equilibrium, the total entropy S_t is maximum, so it is stationary with respect to small exchanges of E_A. As shown in Figure 3.1, imagine that at equilibrium, a small amount of energy randomly moves from A to B. There will be no first-order change in entropy S_t (multiplicity Ω_t), because it is at a maximum. Therefore, S_A will decrease a small amount, and S_B will increase by the *same* amount, so the sum is constant. For A the energy change is $-dE$, and:

$$\Delta S_A = -\frac{\partial S_A}{\partial E} dE, \qquad \Delta S_B = \frac{\partial S_B}{\partial E} dE \qquad \Rightarrow \qquad \frac{\partial S_A}{\partial E} = \frac{\partial S_B}{\partial E} .$$

This defines a temperature parameter: $\beta \propto \partial S/\partial E$. In thermal equilibrium, the temperature parameter of A must equal that of B. For consistency with our common notion of temperature (e.g., from ideal gas laws), we define the temperature parameter T inversely to β:

$$T \equiv \left(\frac{\partial E}{\partial S}\right)_V . \tag{3.2}$$

Then the total system is in equilibrium when the component subsystems have the same temperature:

$$\text{In equilibrium:} \qquad T_A = T_B \qquad \text{or equivalently} \qquad \beta_A = \beta_B .$$

Summary: We define entropy as simply multiplicity measured on a log scale. Equilibrium occurs at maximum total multiplicity (maximum total entropy), which implies that at equilibrium, $\partial S_A/\partial E = \partial S_B/\partial E$. This defines temperature parameters $\beta \equiv (\partial S/\partial E)/k$, and $T \equiv \partial E/\partial S$. Therefore, at equilibrium, $T_A = T_B$.

Equilibrium, Entropy, and Temperature: More Insight

[?? This section needs big overhaul.]

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.

For counting states, we provisionally defined (eq. (2.1)):

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

In this simple definition of $\Omega(E)$, our choice of ΔE introduces an arbitrary additive constant to the value of entropy; such a constant is unimportant in most analyses. The exact value of ΔE has no bearing on the following derivations; therefore we don't need to specify ΔE exactly. The only requirement is that it satisfy the energy continuum approximation (2.12):

quantum spacing << ΔE << observable energy changes.

(Nonetheless, we later define "absolute entropy" in the discussion around eq. (3.5), which eliminates ΔE for good.)

Consider again the system in Figure 3.1. 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 (sub)systems which can exchange energy is called **thermal equilibrium**. Along the way to equilibrium, we again define temperature as in (3.2).

In the nutshell section above, it may seem contrived that we used logarithms without motivation. So let us now back up, and derive the same result directly from multiplicities, in two slightly different ways; we will see that in both cases, logarithms arise quite naturally. (We give further motivation for logarithms shortly.) We proceed by providing detail to these steps:

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

Each subsystem has its own energy density of microstates as a function of energy, here called $\Omega_A(E)$ and $\Omega_B(E)$. Figure 3.2 illustrate typical functions for $\Omega_A(E)$ and $\Omega_B(E)$. Most density of microstates functions are strictly increasing with E [but see *Negative Temperature* on p110]. Both subsystems A and B must have energy between 0 and the total energy E_t. Figure 3.2c shows the numbers of microstates where subsystem A has energy E_A, which we call $\Omega_t(E_A)$ (an abuse of notation).

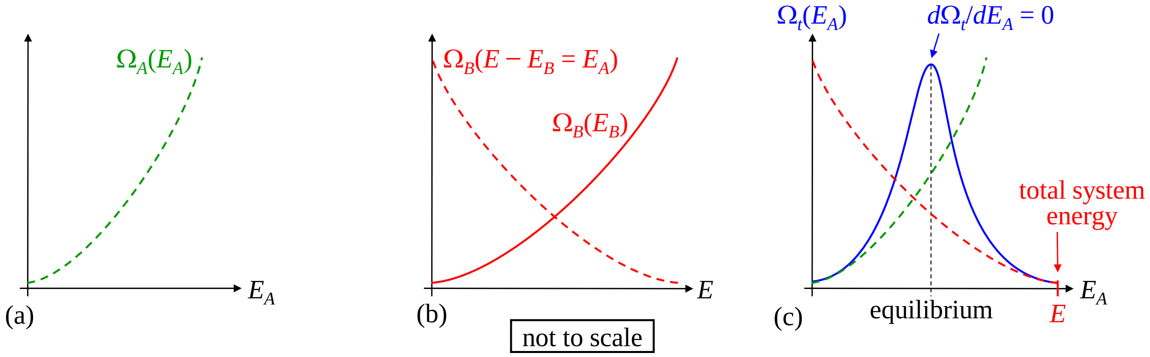


Figure 3.2 Qualitative view of density of microstates vs. energy for (a) subsystem A, (b) subsystem B, and (c) the combined system, as a function of A's energy, E_A.

The fundamental postulate of SM says that the *total system* equilibrium macrostate is the one with the most microstates having total energy E_t. From the discussion of combining density of states, we have:

$$\Omega_t(E_A) = \Omega_A(E_A)\Omega_B(E_t - E_A) .$$

At equilibrium, this is maximum (over all values of E_A). Therefore, if a small amount of energy, dE, moves from A to B, then Ω_A decreases a little, Ω_B increases a little, and Ω_t is unchanged (to first order). For this to happen, the fractional *decrease* in subsystem A's density of states must equal to the fractional *increase* in subsystem B's density of states. E.g., if Ω_A decreases by 1%, then Ω_B must *increase* by 1% for the product to remain constant (to first order):

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

The logarithm arises naturally, and as in (3.2), we find an equilibrium temperature parameter defined by $\partial S/\partial E$ (or equivalently, by $\partial E/\partial S$).

We can repeat this derivation even more mathematically, by formally setting the derivative of $\Omega_t(E_A)$ to zero (Figure 3.2c). For A, the energy change is $-dE$; for B, the energy change is $+dE$. At equilibrium:

$$\frac{\partial\Omega_t}{\partial E_A} = \frac{\partial}{\partial E_A}\left[\Omega_A(E_A)\Omega_B(E - E_A)\right] = 0$$

$$\frac{\partial}{\partial E_A}\Omega_A(E_A)\Omega_B(E - E_A) + \Omega_A(E_A)\frac{\partial}{\partial E_A}\Omega_B(E - E_A) = 0 .$$

Since $E_A + E_B = E$ = const, $dE_B/dE_A = -1$. Then:

$$\frac{\partial\Omega_A(E_A)}{\partial E_A} + \Omega_A(E_A)\left[-\frac{\partial\Omega_B(E_B)}{\partial E_B}\right] = 0$$

$$\frac{\dfrac{\partial\Omega_A(E_A)}{\partial E_A}}{\Omega_A(E_A)} = \frac{\dfrac{\partial\Omega_B(E_B)}{\partial E_B}}{\Omega_B(E_B)} \qquad \Rightarrow \qquad \frac{\partial}{\partial E}\ln\Omega_A(E) = \frac{\partial}{\partial E}\ln\Omega_B(E)$$

as before. And still, 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. This universal definition of **temperature** applies to *any* system, so we drop the subscripts A and B:

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

For any system, temperature is a function of its internal energy, $T(E)$, though the specifics of the function vary radically for different kinds of systems.

What Entropy Is

The quantity ln Ω appears frequently in statistical mechanics. We therefore define it, when scaled by the Boltzmann constant, k, as a system state function, **entropy**:

$$S(E) \equiv k \ln \Omega(E) \qquad \Rightarrow \qquad T(E) \equiv \left[\frac{dS}{dE} \right]^{-1} = \frac{dE}{dS} . \tag{3.3}$$

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

Entropy is a state function that is a measure of the multiplicity of a macrostate, on a logarithmic scale. Entropy is a measure of multiplicity, and nothing else.

We show later how this state-counting definition of entropy, $S = k \ln \Omega$, is related to the thermodynamic one: for a reversible process, $dS = đQ / T$.

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

Why Entropy *S* Is a Logarithm

There are several 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) It makes entropy *extensive*: 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:

$$S_t = S_A + S_B .$$

(3) It allows for a simple definition of temperature that agrees with that from ideal gases: $\dfrac{1}{T} \equiv \dfrac{\partial S}{\partial U}$.

(4) 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:

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

What Entropy Is and Is Not

Entropy Is Not Disorder, but Uniformity of Distribution

The idea 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," or "disordered," 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 precisely define a "well-mixed" system by requiring all local groups of particles to be smaller than some size. Thus most people would say the blobby, partially mixed system, being "disordered," is high entropy, and the well-mixed system, being "ordered," is low entropy, exactly the opposite of the physics. And opposite to our perception of the papers, where "neatness" is the state of *lower* entropy. The physics hasn't changed, only our aesthetics have.

The not-well-mixed system has lower entropy than the well-mixed system: the well-mixed system is more uniform, hence higher entropy. Fortunately, to make a nice cake we *want* uniform distribution of particles; we want higher entropy. We take advantage of the system's tendency toward higher entropy by mixing it essentially randomly until it becomes "uniform."

Remember that "uniformity of distribution" applies only within the constraints/conditions of the system: imagine a row of coins in small cylindrical holes, so each lies flat in a hole. What is the entropy of a system where the first half of them are all heads. This "uniformity" might misleadingly suggests high entropy. In fact, the first half being heads is a constraint, and the "system" must now be taken as including *only* the second half: only those coins that are unconstrained. Adding constraints to a system always lowers the entropy, because it essentially defines a new system with fewer ways (microstates) that satisfy the additional constraints (otherwise, they wouldn't be constraints).

Non-Uniformity Means Available Energy: Faster Than The Wind

In 2021, a video [https://www.youtube.com/watch?v=jyQwgBAaBag] claimed to show that a wind-powered car can be made to travel directly downwind *faster* than the wind itself (abbreviated DWFTTW). The car was made of wheels, gears, and a large propellor in the back. As thermodynamicists, do we think the claim is plausible? [This video, by the car's designer, Rick Cavallaro, provides several good insights: https://www.youtube.com/watch?v=X6oJpnSJyV8]

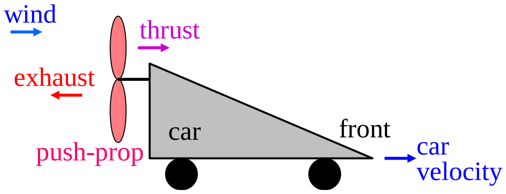


Figure 3.3 (a) The car is propelled by a push-prop, and can travel directly downwind faster than the wind.

I was skeptical, but open minded (as any good scientist should be). In fact, a UC Physics professor bet $10,000 that such a feat was impossible. However, it occurred to me (much later) that the DWFTTW car is a fun example of a simple, very general physical principle. From this principle, one can *predict* (with no math) that traveling FTTW is clearly possible, and consistent with the laws of physics, and also that in the absence of losses, there is no fundamental limit to how much faster than the wind the car can go. These are exactly the conclusions that a more-detailed aeronautical analysis derived [according to the follow-up YouTube video].

[We consider this same claim in *Funky Classical Mechanics Concepts*. I like that we can look at this from multiple perspectives: a simple mechanical perspective, and a statistical-mechanical/thermodynamic perspective.]

I'll bet with some thought, this physical principle can be used in other fun and surprising ways. The principle is:

> Any discrepancy in the physical conditions between two regions of space can be used to do useful work. The process of extracting such work will reduce the discrepancy.

In the case of the wind car, we have that the velocity of the air is different than the velocity of the ground. This is true in all reference frames: ground, wind, car, etc., and true no matter how fast the car is moving. By establishing a "link" or "device" between the wind and ground, the car extracts work from the difference in velocity. In the process, it reduces the discrepancy in velocity: the velocity of wind exiting the propeller is closer to the velocity of the ground (in the ground frame, the wind is slowed).

So long as the ambient discrepancy is available to the car, it can extract continuous power, which means over time an arbitrarily large amount of energy. Therefore, (without losses) it can accelerate to any speed. We reach this conclusion simply from this general principle of physics, and with no math at all.

Perhaps the best known applications of this principle are in thermodynamics. Differences in temperature can create electric potential differences in a thermocouple, or drive a heat engine. Differences in particles' chemical potential run an electric battery. And so on. Differences in conditions necessarily have lower entropy than uniform conditions; one extracts work at the expense of increased entropy. Thermodynamics includes the study of using such differences to do useful work.

Another example of extracting useful work from a difference of conditions is evaporative cooling. We discuss this more later (TBS).

Entropy Is Not About Knowledge

Here's an example which at first seems to pose a paradox, but which is readily understood by the "constraint" definition of entropy. Recall that a "macrostate" is a set of constraints, maybe pressure, volume, mass, etc., and entropy is a (macro)state function: a given system in a given macrostate has some definite entropy.

Before we cautiously give the following example, note that in thermodynamics, entropy is a property of a system in *dynamic* equilibrium. Such a system is constantly swirling through countless microstates, all of which are compatible with the macrostate (i.e., the constraints of macroscopic p, T, N, etc.). The number of such microstates gives us the macrostate entropy. Analogies involving decks of cards, or mixtures of flour and sugar, are not in any dynamic equilibrium: the cards remain in a fixed order, the flour and sugar sit statically in a container. They are only dynamic if we add some artificial dynamic process (e.g., shuffling the cards, or shaking the flour/sugar container).

Bear in mind the limitations of statistical mechanical analogies involving macroscopic "particles" such as playing cards, or grains of sugar.

So forewarned: Suppose I hand you a shuffled deck of ordinary playing cards[2]. The "system" is defined to be the ordinary deck of cards, and the "state" is "shuffled." It could be in any one of 52! configurations, so its entropy is ln(52!).

Someone once said, "Now you look at the cards. Because you looked at them, the entropy has changed. Now that you know what the order is, the entropy is 0." This seems odd, perhaps paradoxical, because how can looking at a system change its entropy? It can't.

The "constraints" definition easily explains these events, with no weirdness. When I hand you a shuffled deck of cards, and ask, "What is the entropy?", the system is the ordinary deck of cards, along with the state that it is shuffled. The constraint is the macrostate "shuffled". Its entropy is ln(52!). If you look at the cards, that changes nothing. The entropy of "a deck of cards, shuffled," is not affected by what you look at.

The confusion is that after looking at the cards, many people unconsciously change the macrostate, by changing the constraints: they re-define the system/state to be "an ordinary deck of cards, constrained to be in the order I've seen." There is only one way to configure a deck in a given order, so its entropy is ln(1) = 0. The constraints are part of the macrostate: different constraints are different macrostates, with different entropies.

Entropy is very different than uncertainty.

---

[2] Thanks to Wendell Misch for this example.

Order From Entropy, and Zero Gravity on Earth

As another, but somewhat rare, example of how "order" can be a misleading proxy for entropy, we consider a system of long, thin particles, such as straight-chain molecules, e.g. a liquid crystal [Frenk]. [Thanks to James J. Orgill (YouTube: The Action Lab) for this idea and reference.] We show below that at high density, there are more ways to arrange the particles when they are nearly parallel than when they are uniformly distributed in angular orientation.

This is similar to packing your suitcase: you get more leftover space if you fold and pack your clothes neatly, than if you through them in as a crumpled mess. If you take your neatly packed suitcase to zero-gravity, the clothes will have room to float around (many possible configurations of clothing). Your crumple-packed suitcase may well have no room at all, and when taken to zero-gravity, the clothes won't move (few possible configurations of clothing). Thus the more "ordered" clothes result in a higher-entropy system, and the "disordered" clothes result in a lower-entropy system.

We performed a somewhat crude experiment to demonstrate this effect. To avoid corruption of the results by gravity, we need to perform the experiment in zero-gravity. We cannot feasibly get 3-dimensional zero-gravity on earth, but we *can* get approximate 2D zero-gravity: any flat, horizontal surface has zero-gravity laterally. If we limit ourselves to a monolayer in height, we effectively have a 2D system with zero-gravity.

We performed the following experiment with finishing nails as particles. Be advised that the experiment is (1) not exact, (2) must be done carefully, and (3) has several sources of error, which we discuss below. However, we think it illustrates the effect of a high entropy system appearing to have more "order" than a low entropy one.

We placed 15 finishing nails in a round pan (Figure 3.4a). We toss the nails vertically, with as little lateral shaking as possible. We keep tossing until we get a monolayer of nails (no nails stacked on each other). Once stopped, we have a snapshot of the nails in a "random" configuration. Figure 3.4a shows a typical configuration, and as expected, both the location and orientation of the nails appears fairly uniformly distributed. The locations and the orientations are the two "sources" of multiplicity/entropy in this system.

Figure 3.4b shows a similar result for roughly double the number of particles (~27), and therefore double the density of particles. We still see a wide distribution of location and orientation, though it looks to me as though there is a little bit of "bunching" of nails with similar orientations.

Figure 3.4c roughly doubles the density again (~56 nails). There are now clearly local groups of nails with similar orientations.

Figure 3.4d roughly doubles again (~93 nails). The bunching of nails into groups of similar orientation is indisputable. Looking at the pattern, we can see why: nails that are perpendicular cannot be stacked into as a high a density as nails that are nearly parallel, or equivalently, nearly parallel nails leave more room (locations) for the nails to be located. Having *fewer* orientations (i.e., being nearly parallel) creates *more* available locations, and in this case, the increased number of locations outweighs the decreased number of orientations. There are simply more ways to place the nails when many of them are parallel, than when the orientations are uniformly distributed. More ways = higher entropy, and in the absence of energy differences, higher entropy wins.

Difficulties with this experiment: we require a monolayer to avoid any gravitational potential energy differences, which would counteract some of the effects of entropy. For Figure 3.4a and b, we achieved a perfect monolayer. However, Figure 3.4c and d show a small number of stacked nails. We simply could not find a shake trial with zero stacking, so we settled for *nearly* zero.

Figure 3.4 Sample configurations after careful “random” shaking: (a) 15 particles. (b) ~27 particles. (c) ~56 particles. (d) ~93 particles.

Another problem is the edges of the container. We first tried this with a square cardboard box, but found the nails bounce off the sides inelastically, and preferentially parallel to the sides. This completely invalidates the results. Using a round container at least eliminates any pre-existing orientation preference. Also, using a metal pan causes “harder” bouncing, which is more elastic, and leaves the nails less likely to align with the edge at the point of impact. Even so, you cannot completely eliminate the tendency for nails to somewhat align with the edges after impact. Another source of error is the slight, but inevitable, sideways shaking that results from shaking by hand. Any sideways motion causes the nails to roll, and favors nails pointing perpendicular to the rolling direction.

Despite these difficulties, we think the results here are essentially valid, and accurately represent the concept of entropy tradeoff causing the appearance of “order” in a high-entropy system: at high densities, the increase in location-entropy more than offsets the decrease in orientation-entropy. In other words, higher total entropy favors aligned orientations.

References

[Frenk] Daan Frenkel, *Order through entropy*, Nature Materials, Vol 14, January 2015, p9.

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.

Entropy, in contrast, is defined even for "lumpy" systems. For example, a hydrogen *ground-state* electron has two spin states, and one spatial state; its entropy is therefore $k \ln 2$. An excited p-orbital electron has six states (2 spin for each of 3 spatial), and entropy $k \ln 6$. Smoothness is irrelevant.

Absolute Entropy and the (Ir)relevance of Δ*E*

For a microcanonical system (fixed energy), there is a deeper story about the meaning of the "multiplicity" $\Omega(E)$, and the use of ΔE in our provisional definition (2.1), repeated here:

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

Strictly Absolute Entropy

Some systems are simple enough that they have resolvably discrete energies, and multiplicities that can be exactly counted directly. Then there is no need to even consider a density of states $\omega(E)$; we simply define:

$$\Omega(E) \equiv \text{exact \# of microstates with energy exactly } E. \tag{3.4}$$

Examples are: (1) an atom in a magnetic field whose orbital magnetic dipole moment interacts with an external B-field; (2) an electron whose intrinsic magnetic dipole moment interacts with an external B-field; (3) a set of N such electrons, even with N macroscopic. The combinatorics of N electrons are simple enough to count $\Omega(E)$ exactly. (4) A 1D particle in a box; (5) a harmonic oscillator; (6) a small set of N identical harmonic oscillators. In this last case (to our knowledge), the counting becomes infeasible quickly as N gets larger.

For these kinds of systems, entropy is exactly extensive. For a total system comprising two subsystems:

$$\Omega_{tot}\left(E_A + E_B\right) = \Omega_A(E_A)\Omega_B(E_B) \quad \Rightarrow \quad S_{tot}\left(E_A + E_B\right) = k \ln \Omega_{tot} = k \ln \Omega_A + k \ln \Omega_B = S_A + S_B .$$

Note that such systems cannot have E different than those on the enumerated list of energies, so the multiplicities of non-enumerated energies are irrelevant (you can call them 0, if you like). In this super-idealized fixed-energy model, we need not consider quantum superpositions of energy states (in fact, we claim that we never need to consider them).

Virtually Absolute Entropy for Some Macroscopic Systems

For systems of nearly-free particles (e.g., a gas in a box), it is infeasible to directly count multiplicities [Pathr p18-20]. This motivated our provisional definition of multiplicity $\Omega(E) \equiv \omega(E)\ \Delta E$. As a simple example, consider a gas with N particles and fixed volume V. Then $\Omega(E) \equiv \Omega(E, V, N)$. We focus here on how Ω varies with E, because when we eventually upgrade to a system at fixed temperature, energy is the main variable of interest: energy is the only parameter in the Boltzmann factor. For now, we define the other parameters, such as N and V as fixed, for simplicity. For a simple box of gas, that is all the relevant parameters. (Later, we consider how S varies with other parameters, such as V.)

Since entropy $S(E) \equiv k \ln \Omega(E)$, the choice of ΔE defining Ω introduces an additive constant in entropy. Most applications depend only on entropy *changes*, so an additive constant has no effect. For example, standard chemical entropies are defined relative to a zero point chosen for laboratory convenience. In these applications, the exact value of ΔE has no bearing on the physics, or on subsequent derivations; therefore we don't need to specify ΔE exactly. The only requirement is that it satisfy the energy continuum approximation (2.12):

microstate quantum spacing << ΔE << observable energy changes.

Since even a trivial system has a microstate spacing ~ 10^{-20} eV (see (2.14)), this constraint is easily met.

Sadly, the advent of *absolute* entropy breaks this simplicity. The Third law of Thermodynamics says that all systems approach zero absolute entropy (or nearly 0) as $T \to 0$ K (and this has important consequences). The additive constant in entropy introduced by ΔE can no longer be tolerated. We need a new definition for multiplicity $\Omega(E)$. Conceptually, however, we must maintain the fact that entropy is a measure of *absolute* multiplicity, and this will preserve the absolute character of entropy S. (Later, it also allows a consistent definition of entropy $S(T)$ for a canonical system (fixed temperature, rather than fixed energy), see (4.2)).

A first step in "fixing" Ω is to define it more carefully as a count of states, rather than a rough estimate. We do this by considering the behavior of $\omega(E) \sim E^{3N/2}$ (see (2.10)). Given ΔE, the true multiplicity in the energy interval $(E - \Delta E, E)$ is (Figure 3.5a):

$$\Omega(E) = \int_{E-\Delta E}^{E} \omega(E')\, dE' .$$

(For this integral, we must take $dE \ll \Delta E$, but again, this is readily accomplished.)

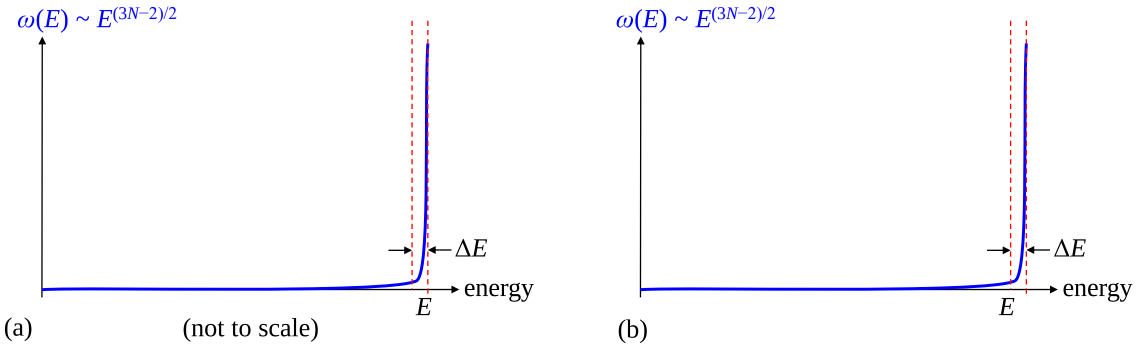


Figure 3.5 (a) Qualitative view of ideal gas density of microstates vs. system energy. For large N, essentially all the microstates below energy E have energies immeasurably different than E. (b) Mathematically more convenient comparison puts E on the left edge of the sliver.

Not only is $\omega(E) \sim E^{3N/2}$ huge, but it changes rapidly with energy (Figure 3.5a). This means any ΔE bigger than a tiny threshold is immeasurably different from any other. In fact, we could just as well choose $\Delta E = E$, and count *all* the states of energy from 0 up to E [Pathr p18]:

$$\Omega(E) = \int_{0}^{E} \omega(E')\, dE' . \tag{3.5}$$

(This is the same as $\Phi(E)$). To verify this claim, let us quantify how rapidly $\omega(E)$ grows, to a good estimate. Purely for mathematical convenience, we redefine our thin sliver to be between E and $E + \Delta E$ (Figure 3.5b); we ask, "For what width ΔE does the thin sliver contain many more states than all the states up to E?" $\omega(E)$ goes as a power p of E, so ΔE must satisfy:

$$\int_{E}^{E+\Delta E} E'^{p}\, dE' \gg \int_{0}^{E} E'^{p}\, dE' \qquad \Rightarrow \qquad \frac{(E+\Delta E)^{p+1} - E^{p+1}}{p+1} \gg \frac{E^{p+1}}{p+1} .$$

Canceling denominators, and using the binomial theorem on the left side:

$$(p+1)E^{p}\Delta E + \frac{(p+1)p}{2} E^{p-1}\Delta E^{2} + \ldots \gg E^{p+1} .$$

All the terms on the left are positive, so if the inequality is satisfied for the first term, it is also satisfied exactly. Then we can solve for a sufficient ΔE:

$$\Delta E >> \frac{E}{p+1} .$$

Now $p \sim 3N > \sim 10^{20}$, so we can choose:

$$\Delta E >\sim \frac{E}{\sqrt{N}} \sim \frac{E}{10^{10}} .$$

Or we can just as well choose (3.5); it makes no measurable difference [Pathr p20, Huang p134]. Thus:

Even for isolated systems (i.e., the microcanonical ensemble),
we define an absolute entropy that is independent of any choice of ΔE.

[One must use caution in interpreting [Pathr p20] that says for system energy E, we could just as well use an energy range of $E - \Delta E/2$ to $E + \Delta E/2$, with ΔE small. Energy levels in the upper half (above E) have enormously large multiplicities, which overwhelm the total, supposed to be taken right at E. Nonetheless, after taking the logarithm (to find S), the discrepancy becomes negligible (for macroscopic systems).]

Notice that the above argument requires only a macroscopic system, but puts no restrictions on the nature of its constituent particles (or subsystems). The result applies even to macroscopic systems with finite-dimensional single-particle states, such as spin systems, orbital angular momentum, or finite multi-level subsystems of any kind. In the special case of a macroscopic system with directly countable multiplicity, (3.5) agrees with the exact result (3.4), to high accuracy.

The final definition of entropy is then:

$$S(E) \equiv k \ln \Omega(E) = k \ln \Phi(E) \qquad (3.6)$$

This result shows that the extensivity of entropy, and the absolute nature of our definition of entropy, are true (to a high degree of approximation) in the thermodynamic limit, but not always for small systems [Pathr p20, Huang p134].

We use the definition (3.5) later to define entropy of a realistic system in thermal equilibrium (eq. (4.2)).

As an aside, note that a macroscopic system has unimaginably tight energy spacing (even a simple 10 electron system in 1 cc has spacing ~ 10^-200 eV at room temperature, see (2.15)). Such a tight spacing means the idealization of a perfect container is completely unphysical. Fortunately, quantum mechanics is such that perturbations of the state energies do not affect the *number* of quantum states (only their energies), so counting the states in our super-idealized system still gives essentially the correct number.

What About the Uncounted States?

In the discussion around (3.5), we assumed that all the microstates were those of indivisible point particles, given by (2.10). But an ideal gas is made of atoms or molecules, which certainly *are* divisible. Some uncounted microstates include ionized atoms, dissociated molecules, and even dissociated nuclei, but we ignored them. Have we erred?

Of course, our approach is an approximation which is quite valid for realistic gases at moderate temperatures (ionization energy >> kT). Consider ionized atoms. Those microstates with ionized atoms have a bunch of energy tied up in the ionization energy, leaving significantly less energy to be shared by other atoms. Since $\omega(E)$ is a *very* steep function of E, the remaining atoms have drastically fewer microstates available to them. Therefore, microstates including ionized atoms are negligibly small in number.

However, at high temperatures (ionization/dissociation energy ~ kT), these dissociated microstates (which have higher particle numbers due to ionization/dissociation) become significant. This is the basis of metal smelting: at high enough temperatures, the increased entropy of dissociated oxygen favors the reaction from metal-oxide → pure-metal + O_2. Similar arguments underly the Saha ionization equation.

We discuss other uncounted degrees of freedom after deriving the Sackur-Tetrode entropy equation.

Sackur-Tetrode: The Ideal Entropy?

Since we have the multiplicity for an ideal gas (2.11), it is fairly straightforward to find its entropy, $S(E) = k \ln \Omega(E)$. We start with Stirling's approximation for the Γ function in the multiplicity:

$$\ln \Gamma(x+1) \approx x\left(\ln x - 1\right), \quad x >> 1 \qquad \Rightarrow \qquad \ln \Gamma\left(\frac{3N}{2}+1\right) \approx \frac{3N}{2}\left[\ln\left(\frac{3N}{2}\right)-1\right].$$

We divide by $N!$ to account for the indistinguishability of the particles, which assumes a low-enough particle density that quantum statistics are insignificant. Then take the natural log, using Stirling's approximation for $\ln N!$:

$$\ln \underbrace{\frac{\Phi(E)}{N!}}_{\Omega(E)} \approx N \ln\left(\frac{V\left(2\pi m E\right)^{3/2}}{h^3}\right) - N\frac{3}{2}\left[\ln\frac{3N}{2}-1\right] - N\left[\ln N - 1\right].$$

Note that $\ln 10^{20} = 46$, so we must keep all terms above. We expect the entropy to be extensive, but this formula is not obviously so, because terms like $N \ln N$ are not extensive. An extensive quantity must equal N times a function of intensive quantities (including constants). In the above equation, we have $E^{3/2}$ in the first term, and essentially $(3N/2)^{3/2}$ in the second, so we can combine them into an extensive quantity:

$$N \ln E^{3/2} - N \ln\left[\left(\frac{3N}{2}\right)^{3/2}\right] = N \ln\left[\left(\frac{2}{3}\frac{E}{N}\right)^{3/2}\right];$$

this is extensive because E/N is intensive. Similarly, we have $\ln V$ in the first term, and $\ln N$ in the third:

$$N \ln V - N \ln N = N \ln \frac{V}{N},$$

which is extensive, because V/N is intensive. Then:

$$\ln \frac{\Phi(E)}{N!} \approx N \ln\left[\frac{V}{N}\frac{1}{h^3}\left(\frac{4}{3}\pi m \frac{E}{N}\right)^{3/2}\right] + N\frac{3}{2} + N.$$

Finally, the absolute entropy is the Sackur-Tetrode equation for a monatomic ideal gas:

$$S(E,N,V) = Nk\left\{\ln\left[\frac{V}{N}\frac{1}{h^3}\left(\frac{4}{3}\pi m\frac{E}{N}\right)^{3/2}\right]+\frac{5}{2}\right\} \qquad \left[\text{Sch 2.49 p77}\right]. \qquad (3.7)$$

This is valid for large N (because of Stirling's approximation), and when the single-particle state occupation numbers << 1 (because we ignored quantum statistics). It is extensive, as expected. It includes only the translational degrees of freedom. (It is often written in other, equivalent forms.)

Sample calculation: absolute entropy of 1 mole of H_2 at 25 C, 1 bar = 100 kPa): we need the specific volume v (volume per particle), and energy per particle u. From basic thermodynamics:

$$m_{H_2} = 3.32\times 10^{-27} \text{ kg}, \qquad u = \frac{3}{2}kT = 5.65\times 10^{-21} \text{ J},$$

$$\text{Use: } pv = kT: \qquad v = \frac{kT}{p} = 3.77\times 10^{-26} \text{ m}^3 \quad \Rightarrow \qquad S_{H_2} = 117 \text{ J/K}.$$

This compares well with the tabulated value of 131 J/K. [NB: despite the fact that IUPAC Standard Temperature and Pressure (since 1982) is 273.15 K = 0 C and 100 kPa = 1 bar, the standard entropy of H_2 is usually given at 25 C. Shockingly, most tables do not list the temperature, and so foment much confusion.]

A similar calculation for He ($m = 6.646\times 10^{-27}$ kg) gives:

$$S_{\text{He}} = 126\ \text{J/K} \qquad \text{vs.} \qquad 126.0\ \text{(tabulated)}.$$

Therefore, we attribute the H_2 discrepancy to its somewhat non-ideal behavior, since He is much more ideal than H_2.

What About Other Degrees of Freedom?

Many gas particles have internal degrees of freedom, due to electron spin, nuclear spin, molecular orbital angular momentum, etc., which add to the entropy. Defining g as the degeneracy of such degrees of freedom, the added entropy is:

$$S_{\text{int}} = k \ln\left(g^N\right) = Nk \ln g, \qquad S_{total} = S_{trans} + S_{\text{int}}.$$

How significant is S_{int}, compared to the entropy from translational motion? This depends on temperature, because $\Phi(E)$ decreases as E (and therefore T) decrease. We can find a temperature scale by approximating at what T does $S_{\text{trans}} = S_{\text{int}}$? $\Phi(E)$ also depends on volume, N, and the particle mass m. For illustration, we choose $V = 1\ \text{cm}^3 = 10^{-6}\ \text{m}^3$, $N = 10^{20}$, $E \approx (3/2)NkT$ (where k here is Boltzmann's constant, not wave-number), and molecular H_2 gas ($m = 3.3\times10^{-27}$ kg).

Substituting into the Sackur-Tetrode entropy (3.7), and then solving for T:

$$Nk \ln g = Nk\left\{\ln\left[\frac{V}{N}\frac{1}{h^3}\left(\frac{2\pi mNkT}{N}\right)^{3/2}\right] + \frac{5}{2}\right\}$$

$$g = \frac{V}{Nh^3}\left(2\pi mkT\right)^{3/2} + \frac{5}{2}$$

$$T = \left(\frac{\left(g - 5/2\right)N}{V}\right)^{2/3}\frac{h^2}{2\pi mk}.$$

For molecular H_2, if it has an angular momentum of $L = 1$, then $g = 3$, yielding $T = 0.21$ K. The Sackur-Tetrode equation may not be valid at such low temperatures, because the effects of quantum statistics are significant there. Nonetheless, this sets a rough scale for T to make internal multiplicity g a significant contributor to entropy S. Such a temperature is easily achieved in modern laboratories, showing that other degrees of freedom *can* make a significant increase in the system entropy.

From the form of T here, we see that the entropy of internal degrees of freedom increases with smaller volume (T gets bigger), increases with N, and decreases with increasing particle mass m. Note that g itself may depend on temperature, such as when the rotational states of diatomic gas molecules become excited around 100 K. The effect on entropy increases with g.

We can turn the question around, and ask: for our system at room temperature, how big does g have to be for S_{int} to equal S_{trans}? Plugging into the equation for g above, we get $g = 28{,}000$. Therefore, at ordinary temperatures, the entropy from internal degrees of freedom is insignificant.

General Equilibrium

The arguments above defining entropy and temperature generalize to any two systems exchanging any conserved thing which affects the multiplicity of a macrostate, 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:

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

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

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

Roughly speaking:

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 μ
any conserved thing, X	$\sim \partial S/\partial X$	defines some equilibrium parameter

We expand on these relations later.

The chemical potential μ 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 μ.

Entropy For Systems Out Of Equilibrium

One might be tempted to say that entropy is defined only for systems in thermal equilibrium. But there is a fundamental contradiction between declaring that restriction, and then comparing the entropies of systems that are clearly not in thermal equilibrium (as in (3.1)). This is the most egregious failure of many references on entropy. Equilibrium is defined as "stationary macroscopic properties" (properties don't change with time). Therefore, by definition, any system which spontaneously time evolves was not initially in thermal equilibrium. By restricting entropy to equilibrium, it is then meaningless to ask if a system's entropy has changed.

We have resolved this problem by defining entropy for a system with a given set of *constraints*. Perhaps the simplest constraint is that the system be in thermal equilibrium, but other constraints describe out-of-equilibrium systems. One often defines non-equilibrium constraints in terms of "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. This is simplest if the "local" entities are discrete, bounded subsystems, as in our prior discussions of Figure 3.1. We supposed the system had two well-defined subsystems, each of which has its own entropy. Then we discussed how the total system entropy varied as energy flowed between the subsystems. Often, though, we must consider a system as a continuum of mesoscopic volumes, each in local thermal equilibrium.

Seeing Entropy: How Long Must I Wait?

When Are Entropy Differences Significant?

We can define that two systems have "significantly" different entropies if a simple tabletop, earthbound experiment at realistic temperatures can demonstrate the difference. We describe some examples shortly. It is not meaningful to compare entropy differences to absolute entropy. A huge baseline entropy has no effect when systems with small entropy *differences* interact. As often stated in thermodynamics, the zero of entropy is usually irrelevant: it is (usually) entropy *differences* that matter. Furthermore, an appropriate quantitative test of entropy differences compares them in light of the time it takes for the system to evolve at the temperatures of interest. (This evolution time is related to the height of the free energy barriers between quasi-equilibrium states).

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.

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 our definition of “significant”, 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 dry 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 realistic image 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.

The Shake Test: Which Has Higher Entropy?

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 significant barriers. This focuses in on the entropy alone. When we stop shaking, we remove that temporary energy, recalling that the 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 has 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: Driving Change

A key point is that entropy is not the whole 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 equilibrium thermodynamics inapplicable.

Furthermore, even without large energy barriers, the importance of entropy differences must take into account the temperature. A tiny entropy difference of 1×10^{-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 we must say that those entropy differences are significant. 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. The shake test for "significant" entropy differences intentionally avoids free-energy barriers: "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).

"Small" Entropy Changes: What Is the Significance of This?

Many references state that even big factors in the multiplicity $\Omega(E)$ are negligible when we compute entropy with the logarithm:

$$S \sim \ln B\Omega' = \underbrace{\ln B}_{\text{negligible?}} + \ln \Omega' \qquad \textit{where} \quad \Omega \equiv B\Omega' .$$

But we must be careful with such an approximation, because sometimes we exponentiate the entropy, such as in computing the macrostate probability (see (5.2)):

$$\text{RelPr}(\text{macrostate } r) = (...)\exp\left(TS / kT\right) = (...)\exp\left(\ln \Omega\right) = (...)\Omega .$$

Therefore, any *factor* that is significant in Ω can lead to a *term* that is significant in *S*, no matter how "small" that term might seem.

Of course, if we're only seeking the location of the *peak* in macrostate probability, then *any* factor (big or small), can be neglected:

$$\arg\max_x \Omega(x) = \arg\max_x \Omega'(x) .$$

Nonetheless, we recommend caution before throwing out seemingly-small terms in a logarithm.

4 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 Boltzmann distribution is universal: it applies to photon quantum states, atoms, galaxies, etc. The only condition for the Boltzmann distribution to apply is that 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.

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. Therefore the energy of a given atom follows the Boltzmann distribution. Since all the atoms are in the same environment, all atoms' energies follow the Boltzmann distribution.

As a counter example, we might ask, how likely is a given quantum *state* to be occupied by a massive particle? Will it follow the Boltzmann distribution? No, because to be occupied, it must take a particle from the reservoir, which means it violates the condition needed to use Boltzmann (it exchanges particles *and* energy). (We consider the microscopic distribution of such systems later.) In contrast, photons are freely created and destroyed, and their numbers are *not* conserved. In fact, photons are a *mechanism* of energy exchange. Thus, we can fill a photon state without depleting any photons from the reservoir. Therefore, photon quantum states *do* follow the Boltzmann distribution.

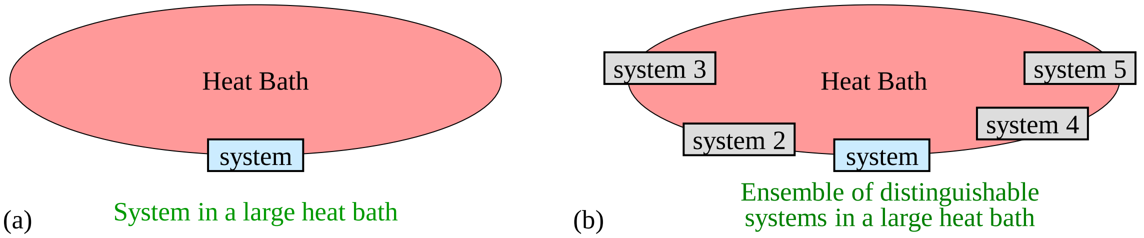


Figure 4.1 (a) A system in thermal contact with a heat reservoir (heat bath). (b) A real or hypothetical ensemble of identical systems.

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 s_1 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 , $\Omega_B(E_t)$:

$$\mathrm{RelPr}(s_1) = \Omega_B(E_t) .$$

Consider another single microstate s_2 of the system, of higher energy E_{s2}. Its relative probability is:

$$\text{RelPr}(s_2) = \Omega_B(E_t - E_{s2}) \ .$$

Recall that the definition of temperature tells us something about how Ω_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 $d \ln \Omega_B$:

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

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

$$\frac{\Omega_B(E_t - E_{s2})}{\Omega_B(E_t)} = \frac{\text{RelPr}(s_2)}{\text{RelPr}(s_1)} = e^{-E_{s2}/kT}$$

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:

$$\text{RelPr}(s_2) = e^{-E_{s2}/kT} \ .$$

Or, for any single microstate:

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

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

$$\text{Pr}(\text{microstate } s) = \frac{e^{-E_s/kT}}{\sum_{r=1}^{\text{microstates}} e^{-E_r/kT}} \qquad \textit{where} \quad E_s, E_r \equiv \text{energies of microstates } s \ \& \ r \ .$$

Be careful to distinguish the Boltzmann distribution from the Maxwell-Boltzmann distribution. More on this later.

Finally, if the Boltzmann distribution seems to suggest that a system "seeks lower energy," remember that (1) the Boltzmann distribution refers to a single system *micro*state, not system macrostate; 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. But we've already seen that the equilibrium state is, in fact, not the state of lowest system energy.

Meet the Partition Function

The denominator above is much more important than just a normalization factor, so we give it a special name:

$$Z(T) \equiv \sum_{r=1}^{\text{microstates}} e^{-E_r/kT} \qquad \textit{where} \quad r, E_r \text{ are microstates and their energies} \ .$$

Z(*T*) is the famous **partition function** of a system. It applies to any sized system from a single particle, to a large system; the "states" you sum over are always system microstates, but recall that a single-particle system has microstates that are also single-particle states. Notice, though, that the normalization factor, *Z*(*T*), is a function of temperature. Therefore:

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

The partition function is dimensionless. We study the partition function in more detail later, and learn that it defines *all* the thermodynamic properties of a system.

For discrete states, we can also write the partition function as a sum over system *energy levels*:

$$Z(T) \equiv \sum_{r=1}^{\text{microstates}} e^{-E_r/kT} = \sum_{m=1}^{\substack{\text{system}\\\text{energy}\\\text{levels}}} \omega_m e^{-E_m/kT} \qquad \textit{where} \qquad \begin{cases} r,\, E_r \text{ are microstates and their energies, } \textit{and} \\ E_m,\, \omega_m \text{ are energy levels and their multiplicites} . \end{cases}$$

The distribution ω_m of allowed system energies E_m affects the absolute probabilities of microstates only through the normalization factor. However, the *relative* probabilities of individual microstates or system energies are *in*dependent of the allowed energies.

The partition function readily generalizes to a system with continuous energies:

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

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

Meet the Macrostates

For macroscopic systems, we are interested in the probability of the system being in some observable *macrostate*, i.e., having a specified *energy* *E*, rather than being in a specified microstate. Then for discrete energies, we can simply sum the probabilities of all system microstates of energy *E*:

$$\Pr(\text{system energy} = E) = \frac{\omega_E\, e^{-E/kT}}{Z(T)} \qquad \textit{where} \qquad \omega_E \text{ is the multiplicity of system energy } E .$$

However, macroscopic systems have essentially continuous energies with density $\omega(E)$ (a classical system), so we use the continuum equivalent of the above, with a narrow range ΔE near *E*:

$$\Pr(E < \text{system energy} < E + \Delta E) = \frac{\omega(E)\Delta E\, e^{-E/kT}}{Z} ,$$
$$\textit{where} \qquad \omega(E) \equiv \text{energy density of microstates} ,$$

and

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

Macroscopic Properties as Averages

Thermal equilibrium between two systems is a dynamic equilibrium: there are small fluctuations in the properties of the systems, but over time, no macroscopic properties change significantly. (The study of fluctuations quantifies these statements, but we do not pursue that here.) But fluctuations require that we define our macroscopic properties as averages over these fluctuations. Usually, we can readily define these averages by using the Boltzmann distribution to provide the probabilities. For example, suppose we know the pressure as a function of the energy $p(E)$ of an isolated system. Then the average pressure of a system in thermal equilibrium at temperature *T* is:

$$\left\langle p(T)\right\rangle = \sum_{r}^{\text{microstates}} \underbrace{\frac{e^{-E_r/kT}}{Z(T)}}_{\Pr(E)} \underbrace{p(E_r)}_{\text{pressure}} .$$

Usually, this average value is simply called *the* pressure of the macrostate, and its average nature is understood. This works for any property that is well-defined as a function of energy of the isolated-system (microcanonical ensemble):

$$\left\langle X(T)\right\rangle = \sum_{r}^{\text{microstates}} \underbrace{\frac{e^{-E_r/kT}}{Z(T)}}_{\Pr(E)} \underbrace{X(E_r)}_{\text{property } X} \qquad \textit{where} \quad Z(T) = \text{constant at fixed temperature} . \qquad (4.1)$$

We now choose energy E as the independent variable (instead of temperature). In the thermodynamic limit, the fluctuations of any property X are tiny (Figure 4.2), so the average value of any property X of a system with average energy $<E>$ is just the value of X for an isolated system with energy exactly equal to $<E>$:

$$\underbrace{\left\langle X(E)\right\rangle}_{\substack{\text{canonical}\\ \text{distribution}}} = \underbrace{X\left(\left\langle E\right\rangle\right)}_{\substack{\text{microcanonical}\\ \text{system}}} \qquad \textit{where} \quad X \text{ is any property of the system} .$$

This concept is straightforward for most properties, but entropy raises a problem.

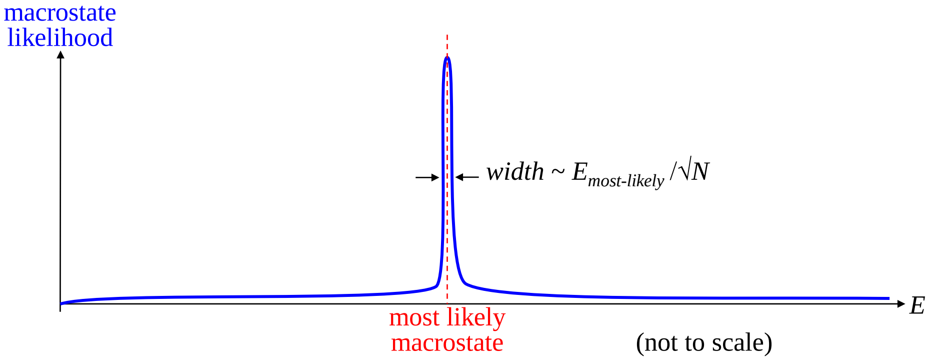


Figure 4.2 Likelihood system has energy E, for a system in equilibrium with a heat bath, i.e. a system at fixed temperature. Each value of E defines a macrostate.

Multiplicity of What, Exactly? Canonical Equilibrium

For a system in equilibrium with a heat bath, average entropy poses a unique problem: just what are we averaging? We said earlier, “Multiplicity is what matters, and nothing else” (p77). To find the average entropy, do we average the multiplicity and then take the log? Or do we take logs of multiplicity, and average them? The difference between the two is vitally important.

Physically, our original claims still stand: multiplicity is what matters, not its logarithm; for an isolated system, maximum multiplicity is the equilibrium point. These remain true for a system in equilibrium with a heat bath (Figure 4.1a): maximum multiplicity is the equilibrium point. But multiplicity of what, exactly?

Entropy as an Average

Before defining multiplicity more carefully, let’s quickly dispense with entropy. The entropy of a canonical system (a system in equilibrium with a heat bath) is defined the same as all other properties, with (4.1), and the definition of absolute multiplicity (3.5):

$$\left\langle S(T)\right\rangle=\sum_{r}^{\text{microstates}}\underbrace{\frac{e^{-E_r/kT}}{Z(T)}}_{\Pr(E)}S(E_r)=\frac{k}{Z(T)}\sum_{r}^{\text{microstates}}e^{-E_r/kT}\ln\Omega(E_r)\,. \qquad (4.2)$$

Canonical entropy is the average of the logarithms, so it no longer directly relates to the average multiplicity. Nonetheless, entropy remains an important and widely-used property of a canonical system.

Equilibrium of a Canonical System

As noted above, for a canonical system in equilibrium, entropy is *not* maximized. But the *multiplicity* is still maximized, and for a canonical system, the *average* multiplicity is maximized:

$$\left\langle \Omega(T)\right\rangle=\frac{1}{Z(T)}\sum_{r}^{\text{microstates}}e^{-E_r/kT}\underbrace{\Omega(E_r)}_{=\Phi(E_r)}\,.$$

For a canonical system (i.e., in equilibrium with a heat bath at fixed temperature),
the average of the log of Ω gives the entropy, $S \propto$ <ln Ω>;
but the log of the average Ω is maximized in equilibrium.

We show in the Free Energy chapter that the log of the average multiplicity is the (negative) free energy: $A \equiv -kT$ ln<Ω>. Equilibrium occurs at maximum average multiplicity, which is *minimum* free-energy.

5 Free Energy and Chemical Potential

Introduction to Free Energy

Free energy is of fundamental importance to all of statistical mechanics, thermodynamics, and chemistry. Thermodynamic systems can often exchange heat with the environment or other systems. A system can import heat from the environment and do useful work with it. Free energy is an energy quantity that accounts for the availability of (and demand for) energy from outside the system. As a result, free energy tells you how much useful work a system can do during various processes.

Free energy 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).

Beware that systems with different constraints require different definitions of free-energy. Each system has one and only one meaningful definition of free energy. Some system constraints are fairly common, so we give them standard names, such as Helmholtz free-energy, or Gibbs free-energy. Other systems have free-energies with no standard names.

We start with an example, and an overview of the properties of free energy, then we derive it.

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 (Figure 5.1). We find that the equilibrium position has the water soaked part way up the towel (Figure 5.1c). The system on the left is minimum energy, since all the water is as 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 microstates). 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 macrostate is the most likely.

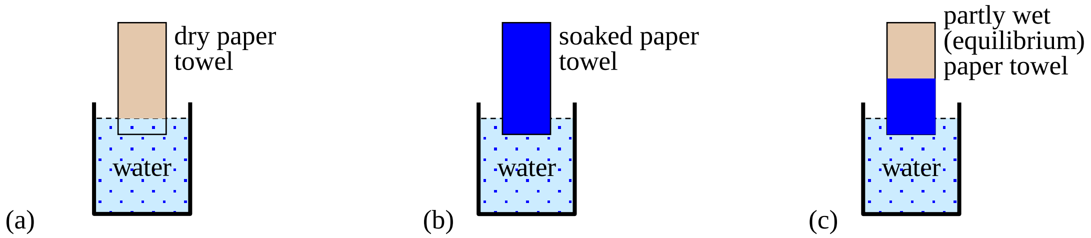


Figure 5.1 (a) System with minimum energy. (b) System with maximum entropy. (c) System with minimum free-energy (in equilibrium).

The equilibrium position is neither minimum energy nor maximum entropy;
it is minimum free-energy.

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

Putting Entropy to Work: A Brain Teaser

As another example, consider a container of ideal gas (say, air), at ambient temperature (Figure 5.2).[3] Now we compress it (say, adiabatically in an insulated container), thus putting energy into the system of air (Figure 5.2a). Doing so results in a smaller volume of air at higher pressure, and higher temperature (Figure 5.2b). Now remove the insulation, and let that air cool to ambient temperature (Figure 5.2c). Thus the system

---

[3] Thanks to Rick Cavallaro for this interesting puzzler.

lost, as heat outflow, some of the energy that we put in. But we still have compressed air, and we can do work with it. Question: what fraction of the energy that we put in gets lost as heat out?

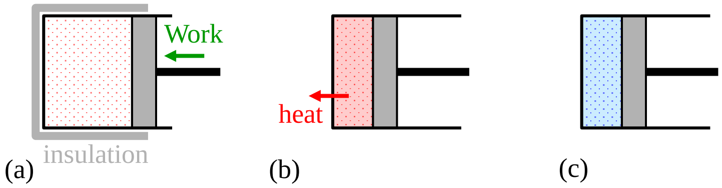


Figure 5.2 (a) Compressing a gas adiabatically in a cylinder. (b) Cooling the gas back to ambient. (c) The resulting system has high pressure, and so can do work.

The answer may surprise you: we lose 100% of the energy that we put in, but still have a container of compressed air that can do work. How so?

The internal energy of an *ideal* gas depends only on the temperature, and nothing else (not pressure, volume, etc.). The final gas is at the same temperature as the initial gas, so its internal energy is the same as when we started: *all* the work-energy put into compressing the gas has gone out again. However, at fixed energy, entropy increases (or decreases) with volume, so the entropy of the system after compression and cooling is *lower* than before. Also, we have a discrepancy of conditions: inside the container is high pressure, outside it's lower. So we can extract useful work by letting the volume expand.

Two simple ways we can expand are (1) adiabatically, and (2) isothermally. For adiabatic expansion, we insulate the container and let it expand. The gas cools down as it does work: the internal energy of the gas is converted to outgoing work.

For isothermal expansion, we let it expand slowly, with heat flowing in from the ambient as it expands, keeping the system temperature constant. That heat is the source of energy that gets turned into expansion work. Note that the system entropy increases as it expands, and that entropy increase draws in heat from the ambient.

Free Energy Preview:

In thermodynamics, for non-cyclic processes, we quantify the amount of energy that is available (free) to do work: it is the "free-energy" state function. For this system, we can compare the initial free energy F_i to that after compression and cooling, F_f. We use the Helmholtz free energy: $F = U - TS$. As noted, $U_i = U_f$, so the lower entropy after compression and cooling results in a higher free energy: thus the system has access to some energy that is "free" (available) to do work. In doing such work, the system free energy goes back down to its initial value F_i, but the environment's entropy has irreversibly increased during the cooldown.

> [Aside: See later sections on free energy, but for now: the ambient pressure is constant, so we *could* use the Gibbs free energy, but we use the simpler Helmholtz free energy by including the atmospheric pV work in the total work that we calculate. Also, in the full cycle (after expansion), the atmospheric pV work has canceled. Finally, since the temperatures are the same in the initial and final states, we can use an arbitrary entropy reference point.]

Work and Free Energy for Gas Compression

We now quantify this process with a numeric example, which is more than just an exercise in calculator proficiency. It emphasizes some important concepts that are omitted from our qualitative discussion so far. For each step of the calculation, we describe the physical meaning of the result. Figure 5.3 is the p-V diagram. We take the initial conditions to be: $p_i = 100$ kPa (about 1 atm); $T_i = T_f = 300$ K (about room temp); $V_i = 1$ m^3, and $V_f = 0.5$ m^3 (a hefty air tank). We want to compute the work needed to compress the gas, and compare it to the change in free-energy for the compression. We will compute several other relevant quantities, as well.

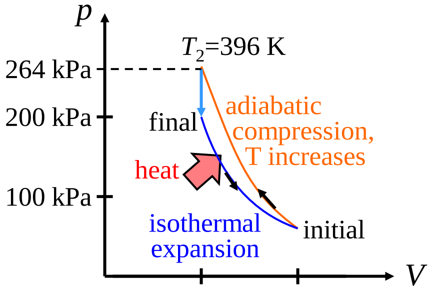


Figure 5.3 p-V diagram for the compressed gas brain teaser. Heat flows in during the isothermal expansion. Work done is the area under the curves.

The work to compress the gas is (where positive means *leaving* the system):

$$W = \int_{V_i}^{V_f} p\, dV .$$

We eliminate p in favor of V with the equation of state for an adiabatic compression of an ideal gas:

$$pV^{\gamma} = const = p_i V_i^{\gamma} \qquad where \quad \gamma \equiv C_p / C_V \qquad \Rightarrow \qquad p(V) = p_i V_i^{\gamma} / V^{\gamma} .$$

Then:

$$W = p_i V_i^{\gamma} \int_{V_i}^{V_f} V^{-\gamma}\, dV = \frac{p_i V_i^{\gamma}}{-\gamma+1} V^{-\gamma+1} \Big|_{V_i}^{V_f} = \frac{p_i V_i^{\gamma}}{-\gamma+1} \left(V_f^{\,-\gamma+1} - V_i^{\,-\gamma+1} \right) .$$

(Check the units on that.) Using γ (air) = 1.4, the work is:

$$W = \frac{(100\text{ kPa})\left(1\text{ m}^3\right)^{1.4}}{-0.4} \left(0.5^{-0.4} - 1^{-0.4} \right) = -79.9\text{ kJ} .$$

Note that we're integrating "backwards" because $V_f < V_i$, and the work is negative, as expected, because that is our convention for work *entering* the system.

We are now at point 2 in the diagram, and the gas is hot. But its entropy is unchanged from the initial state, because the adiabatic compression exchanged no heat: $dS = đQ / T = 0$. We do not yet know the temperature (T_2) or pressure. However, if we wanted, we could get all our work back by simply reversing the compression process, turning it into an adiabatic expansion. But that is not our goal.

We will let the gas cool back to 300 K, and its entropy will decrease during that process. To compute that ΔS, we need the gas temperature just after compression, T_2. We get that from the alternate form of the equation of state for adiabatic compression of an ideal gas:

$$TV^{\gamma-1} = const = T_i V_i^{\gamma-1} \qquad \Rightarrow \qquad T(V) = T_i V_i^{\gamma-1} / V^{\gamma-1} .$$

$$T_2 = (300\text{ K})\left(1\text{ m}^3\right)^{0.4} \left(0.5\text{ m}^3\right)^{-0.4} = 396\text{ K}$$

We find $\Delta S = S_f - S_i$ from:

$$\Delta S = nC_V \ln \frac{T_f}{T_2} \qquad where \quad n = \frac{p_i V_i}{RT_i}; \quad C_V = \frac{5}{2} R$$

$$\Delta S = \frac{p_i V_i}{T_i} \frac{5}{2} \ln \frac{T_f}{T_2} = \frac{(100\text{ kPa})\left(1\text{ m}^3\right)}{300\text{ K}} \frac{5}{2} \ln \frac{300\text{ K}}{396\ K} = -231\text{ J/K}$$

The entropy has gone down, as expected for cooling.

Finally, since the internal energy is the same in the initial and final states, the change in free energy is:

$$\Delta F_{fi} = \cancel{\left(U_f - U_i\right)} - T\left(S_f - S_i\right) = -\left(300\text{ K}\right)\left(-231\text{ J/K}\right) = +69.3\text{ kJ} .$$

We now have the ability to do 69.3 kJ of work, which is less than the 79.9 kJ of work we put into the process. We see this lost work in the p-V diagram (Figure 5.3) as the area between the compression and expansion curves. What happened? Why can't we get *all* the work back?

We lost some ability to do work when we let the gas cool: this process lowered the entropy of the gas, but it also *increased* the entropy in the environment by even more. This exemplifies how even a slow process can change net entropy (see discussion around eq. (9.1)). How much entropy did the environment gain? The environment is large, so its temperature is essentially constant at 300 K. The heat lost by the gas is the heat gained by the environment. Then the environment's entropy gain is:

$$\Delta S = \frac{Q_{cooldown}}{T_{env}} = \frac{nC_V \Delta T}{T_{env}} = \frac{5}{2}\frac{p_i V_i}{T_i}\frac{\Delta T}{T_{env}} = \frac{5}{2}\frac{\left(100\text{ kPa}\right)\left(1\text{ m}^3\right)\left(96\text{ K}\right)}{300\text{ K}}\frac{1}{300\text{ K}} = +267\text{ J/K} .$$

In total, the *net* entropy of the universe increased by 267 – 231 = 36 J/K. This increase is to be expected, because the cooldown transferred heat across a finite temperature difference (starting at 396 K, and dumping into 300 K).

Whenever total entropy increases, we lose some ability to do work.

Note that the *temperature* increase of the environment is negligible, but its *entropy* increase is significant.

As an aside, note that free energy is a *state function*: it is independent of how you got to that state. Therefore, any compression/heat-flow process that leads to the same final volume and temperature will end up with the same free energy.

As a check, let's compute the work we get by now expanding the gas isothermally (at 300 K):

$$W = \int_{V_i}^{V_f} p\, dV = \int_{V_i}^{V_f} \frac{nRT}{V}\, dV = p_i V_i \int_{V_i}^{V_f} \frac{dV}{V} = p_i V_i \ln\frac{V_f}{V_i} = \left(100\text{ kPa}\right)\left(1\text{ m}^3\right)\ln 2 = 69.3\text{ kJ} .$$

Exactly what we predicted from the free-energy. The energy to do this work comes into the system as thermal energy from the environment, as the system's entropy increase draws it in. Note that such an isothermal process has no *net* entropy change: the environment loses as much entropy as the system gains: $\text{đ}Q_{env,out}/T = \text{đ}Q_{gas,in}/T$. With no net entropy change, the process wastes no opportunity to do work

What about the free-energy at point 2? It is exactly 79.9 kJ above the initial free-energy:

$$\Delta F_{2i} = \Delta U_{2i} - T \underbrace{\Delta S_{2i}}_{0} = 79.9\text{ kJ} - T \bullet 0 = 79.9\text{ kJ} .$$

This confirms the earlier note that, at point 2, we can get *all* our work back: we've wasted no opportunity because we have not increased the entropy of anything. It is the net entropy increase during cooldown that costs us some opportunity to do work.

In summary: The adiabatic compression cost 79.9 kJ of work, and increased the energy and free-energy of the gas by exactly the same 79.9 kJ. There is no entropy change of anything during this compression. The cooldown cost a *net* (system + environment) entropy increase of 36 J/K, which lowered the free-energy of the gas to 69.3 kJ above F_i. In effect, the net entropy increase cost us the opportunity to do about 10.6 kJ of work. If we extract the available work isothermally, there is no further *net* entropy change, and we get all 69.3 kJ of work.

Choosing A Different Path

Instead of the adiabatic compression, followed by cooling, we could compress the system isothermally: the system goes isothermally from initial to final state. Then there is no net entropy change (the process is reversible: entropy lost by the system is gained by the environment). The work done in compression is less, and it is exactly 69.3 kJ. The compression is the time-reverse of the isothermal expansion we described above. The final state is the same with either path, so the final energy, entropy, and free energy are the same. The difference is that there was no net entropy gain, no wasted energy during compressionl if we extract work from the compressed system, the work taken out is exactly the same as the work put in (69.3 kJ).

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 the various forms 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 the quantity that is minimized
> when a constrained system reaches equilibrium.

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. Free energy is a state function, so how we get to the final state is not important. 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. The connection between the two is subtle, and one might imagine that in some alternate universe, these two phenomena might have been given by two different state-functions.

The Free-Energy Principle: Multiplicity Defies Energy

The free-energy principle 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, nor the maximum entropy; it is the *most likely* state, which takes both energy *and* multiplicity into account.

Consider a macroscopic system which has discrete (i.e. quantized) microstates, each with a definite energy. The energies are microscopically close together, and indistinguishable at the scale of laboratory measurements, but QM says the microstates are discrete. (It is usually easier to think of them as discrete states, so this is a rare case where QM simplifies our lives.) We can label each system microstate with an index, such as s. Note that many states may have the same energy. The number of states with a given energy is the **multiplicity** of that energy level (sometimes called the "degeneracy" of the energy).

When any (macroscopic or distinguishable microscopic) system is in thermal equilibrium with a reservoir, the reservoir temperature characterizes the system (or equivalently, the parameter $\beta \equiv 1/kT$, with units of inverse energy). The Boltzmann distribution gives the relative probability of the system being in a particular quantum-state, or microstate, as a function of the state's energy, and β, according to the **Boltzmann factor**:

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

This formula states that for systems which can exchange energy with a large reservoir, we are more likely to find the system in a microstate with a lower energy than another microstate with higher energy. (We'll assume $\beta > 0$ for now, which is by far the most common case. We'll address systems of indistinguishable particles later, in the Fermi-Dirac and Bose-Einstein distributions.)

Example: At room temperature, $\beta \approx 1/(26\text{ meV})$. Consider two microstates with energies $E_1 = 0$, and $E_2 = 18$ meV. Then:

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

Microstate 1 is twice as likely to appear as microstate 2 (Figure 5.4a). In the figure, each blue box is a single microstate. The height of the box is the relative probability (Boltzmann factor) for its energy; higher energies have smaller relative probabilities. If a system has only these two microstates, then we can normalize the probabilities to find Pr(microstate 1) = 2/3, and Pr(microstate 2) = 1/3. Therefore:

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

Then the average system energy of this 2-state system is:

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

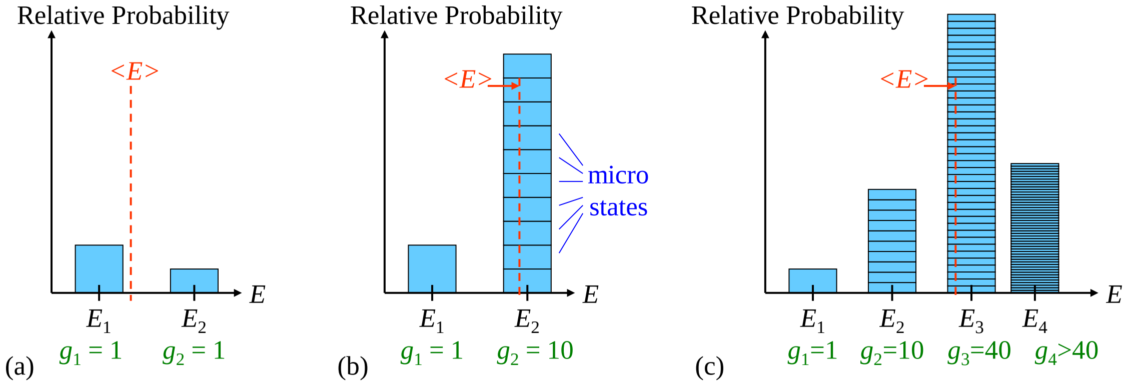


Figure 5.4 (a) A system of two microstates. (b) A system of 11 microstates. (c) A system with more energy levels, and still more microstates.

If a similar system had the same two energies, but there were 10 microstates with energy E_2 (Figure 5.4b), the average system energy would be much closer to E_2. All microstates of the same energy have the same relative probability, and the relative probability for a system energy level E_m is the sum of the relative probabilities of all its microstates; this is just its multiplicity, ω_m, times the Boltzmann factor:

$$\text{RelPr}(E_m) = \omega_m e^{-E_m/kT} \qquad \Rightarrow \qquad \text{RelPr}(E_2) = 10(0.5) = 5 .$$

Then the probability normalization factor is $Z = 6$, and the average energy is much higher than before:

$$\langle E \rangle = \frac{1}{6}E_1 + \frac{5}{6}E_2 = \frac{1}{6}(0) + \frac{5}{6}(18 \text{ meV}) = 15 \text{ meV} .$$

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

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

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

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

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

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

Even though the probability of being in any *given* high energy microstate is only half that of being in a *given* low energy microstate, the multiplicity of the high energy microstate vastly outweighs the small energy "advantage" of the lower energy microstate. Multiplicity defies energy, and raises the average system 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).

With more energies and microstates (Figure 5.4c), the average system energy is close to the dominant energy level, where dominance is given by the *product* of the Boltzmann factor times the multiplicity.

We will see shortly how much energy a given multiplicity is worth. This quantification of multiplicity extends the concept of entropy:

> Entropy is a measure of multiplicity, on a logarithmic scale, which in turn,
> is a measure of a macrostate's ability to acquire more energy from the reservoir.

In practice, for large systems, we usually treat the micro- and macrostates as continuous; the continuum approximations to the above analysis are straightforward.

Example: The Statistical Chain

Consider a massless chain of N links, hanging vertically with a constant weight F pulling on the bottom (Figure 5.5). 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 ?

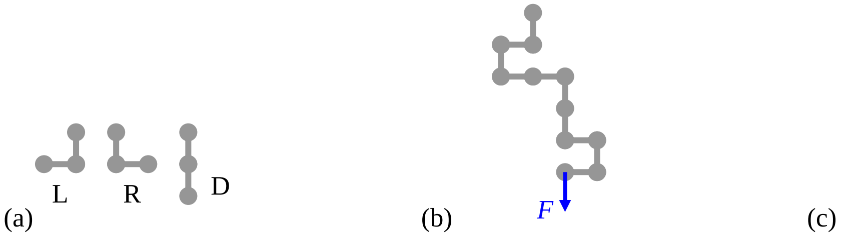


Figure 5.5 Statistical chain. (a) Three possible states of a link. (b) Example chain configuration.

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 macrostate of lowest energy, and it has only one microstate.

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, so the average length is simply *Nl* / 3. Notice that this is *not* the macrostate of lowest energy, since some links are Left or Right, and have more energy than Down links. This macrostate has many microstates (microstates of the same length).

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 way, we are temporarily considering each link to be a "system." In this (or the larger full-chain) 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:

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

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

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

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:

$$E = -NFl \qquad (\text{low } T).$$

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

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

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 use Ω(*T*) ≡ the number of microstates available at temperature *T* (rather than using any *density* of microstates 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 for small displacements, 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 macrostate of lowest energy, and the number-of-microstates (multiplicity) function Ω(*T*) combined with the Boltzmann distribution makes one particular macrostate (length) more likely than any other. At high temperature (*β* → 0), the spring constant goes to infinity: with infinite thermal energy, the chain can lift any weight.

Entropic elasticity is a significant part of the elasticity of a rubber band, though it also has components of elasticity from inter-molecular forces [ref??]. Note that:

Rubber bands shrink when heated (and expand when cooled).
This is characteristic of entropic elasticity.

It's an experimental result that shows that a significant component of rubber-band elasticity is entropic.

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

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

We first note that the order of magnitude for any fractional fluctuation is almost always ~1/√N. This gives us ~10% length (or energy) fluctuation for 100 links, and ~10^{-10} for 10^{20} links. For the exact answer, we use 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:

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

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

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

At high T, $p = 1/3$, and the correction factor to 1/√N is only √2. As T goes down, the fractional fluctuation increases to a point. From there, as T gets even colder, the chain hangs lower with little energy to make it fluctuate. At $T = 0$, $p = 1$, and the chain length is Nl, with zero fluctuation.

8. The entropy is computed from the multiplicity of the macrostate, i.e. the average number-of-microstates of the system that match the macrostate (4.2):

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

For large N, the 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:

$$n_D{}^* \approx \bar{n}_D = N\frac{e^{\beta Fl}}{2+e^{\beta Fl}} \qquad \text{(most likely number of down links)}. \tag{5.1}$$

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

$$\Omega(T) = \binom{N}{n_D{}^*} \equiv N \text{ choose } n_D{}^* = \frac{N!}{n_D{}^*!\left(N-n_D{}^*\right)!}$$

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

Eq. (5.1) shows that $n_D{}^* > N/3$, so if N is large, so is $n_D{}^*$. Then we can use Stirling's approximation for the factorials.

Return to statistical chain after free-energy and partition function, and compute those??

How Much Energy Is Multiplicity Worth? Definition of Free Energy

We've seen that multiplicity allows a system to offset higher energy. So how much energy is multiplicity worth? In other words, how much energy surplus can multiplicity overcome? We find the answer directly from the Boltzmann factor. Our basis for comparison is a macrostate with multiplicity 1; it has a relative probability of exp(–E/kT). Without loss of generality, we can take $E = 0$, so RelPr = 1. Consider a different macrostate with energy δE and multiplicity Ω. To have the same probability as our reference macrostate, we must have:

$$\text{RelPr}(E_0 \equiv 0) = e^{-E_0/kT} \cdot 1 = 1, \qquad \text{RelPr}(\delta E) = \Omega e^{-\Delta E/kT} = N_1 = 1 \qquad \Rightarrow \qquad \delta E = kT \ln \Omega .$$

In other words, our multiplicity Ω is “worth” $\delta E = kT \ln \Omega$ units of energy.

Note that how much energy Ω is worth (δE) depends on the temperature. At higher temperature, multiplicity is worth more energy, because more thermal energy makes it easier to reach higher energy macrostates. Also, for any T, the energy-equivalent of multiplicity is proportional to ln Ω., i.e. it’s proportional to entropy. Specifically, from $S = k \ln \Omega$:

$$\delta E = TS .$$

We can see the effect of multiplicity on macrostate probability directly. The relative probability of a macrostate with energy E and multiplicity Ω is (using the Boltzmann factor):

$$\begin{aligned} \text{RelPr}(E, \Omega) = \Omega e^{-E/kT} &= \exp(-E/kT + \ln \Omega) = \exp\left[-\left(E - kT \ln \Omega\right)/kT\right] \\ &= \exp\left[-\underbrace{\left(E - TS\right)}_{\text{free energy}}/kT\right]. \end{aligned} \tag{5.2}$$

In other words:

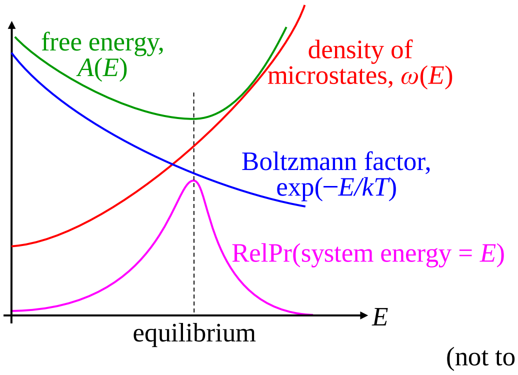
The quantity $A \equiv E - TS$ gives us the *macro*state relative probability,
much like E gives us the *micro*state relative probability, using the Boltzmann formula.
A is called the **free energy**.

Helmholtz Free Energy

The free energy of a system starts with the internal energy U, and de-rates it by the energy equivalent of multiplicity. In other words, it gives you an energy equivalent to tell you the macrostate relative probability from the Boltzmann formula. We illustrate with the simplest free energy: Helmholtz free energy.

The simplest constraint on a system 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:

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



(not to scale)

Figure 5.6 Qualitative view of the product of density of microstates and the Boltzmann factor to produce the relative probability of the system having a given energy, and the free energy associated with that probability. Equilibrium is at the probability maximum, which is the free-energy minimum.

This is just an example of the “multiplicity defies energy” concept shown above. First, recall that the relative probability of the system having discrete energy E_m is:

$$\text{RelPr(system energy is } E_m) = \omega_m e^{-E_m/kT} \qquad \textit{where} \qquad \begin{aligned} E_m &\equiv \text{discrete energy of the system,} \\ \omega_m &\equiv \text{multiplicity of energy } E_m . \end{aligned}$$

The energy (macrostate) with the maximum relative probability is the most likely energy. Finding this state involves dealing with two parameters: ω_m and E_m. It'd be nice if we could combine these two into a single function of macroscopic state variables (which we call **free energy**) which combines the energy and the multiplicity. To do this, we first bring the factor of ω_m up into the exponent:

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

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

$$\text{RelPr(system energy} = E_m) = \exp\left(\frac{-\left(E_m - kT\ln\omega_m\right)}{kT}\right) = \exp\left(\frac{-\left(E_m - TS_m\right)}{kT}\right) \qquad \textit{where} \qquad S_m \equiv k\ln\omega_m$$

$$= \exp\left(-A_m / kT\right) \qquad \textit{where} \qquad A_m \equiv E_m - TS_m$$

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

$$\delta E = kT \ln \omega_m = TS .$$

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

$$\text{RelPr}(E) = \Omega(E) exp(-E/kT) .$$

The higher system energy comes from the heat bath (Figure 4.1a), 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 *TS* term in the Helmholtz free energy represents real energy that goes into the system from the heat bath.

The most likely macrostate is that with the highest relative probability, $e^{-A_m/kT}$, which is that when the free energy A_m is minimized.

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

An explicit statement of absolute macrostate probability of a small energy range *dE* is:

$$\Pr\left[\text{energy} \in \left(E, E + dE\right)\right] = \frac{\omega(E)\, dE\, e^{-E/kT}}{Z(T)}$$

$$\textit{where} \qquad \omega(E)dE \text{ is the multiplicity of states in the range} .$$

Entropy as a Heat Sponge: Energy Absorbed From the Heat Bath

TBS.

We show that the entropy of a macrostate of a system at temperature *T* defines the heat energy the system has absorbed from the heat-bath, through the equation (derived later):

$$Q = TS .$$

Therefore, if a system undergoes a process taking it from macrostate 1 (state 1) to another (state 2) at the same temperature, the heat entering the system is:

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

TBS.

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 *internal* energy lost by the system , 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, because the *TS* term in the free energy represents
real energy that goes into the system from the heat bath..

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 go out as useful work. This is exactly the case with a car battery (Figure 5.7): the useful (electrical) work out is *more* than the internal energy loss, because the car battery entropy increases while discharging, so it absorbs some heat from the environment.

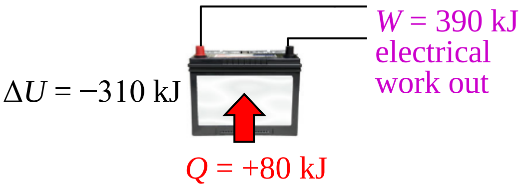


Figure 5.7 A car batter converts chemical energy *and* ambient heat to useful work, thereby producing *more* work than the decrease in internal energy.

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.

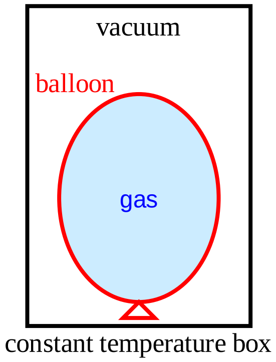
TBS: example of burning hydrogen from [Sch].

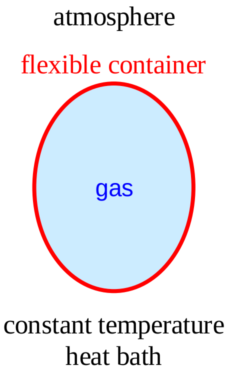
Gibbs Free Energy and Enthalpy

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]:

$$H \equiv U + pV .$$

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":

$$\text{free energy} = U + pV - TS\ .$$

We can think of $(U + pV)$ 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:

$$G \equiv U + pV - TS\ .$$

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]:

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

$$\Rightarrow \qquad d(\text{free-energy}) = dU + \sum_i F_i \; dX_i - T \; dS$$

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

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:

$$A \equiv U + pV + \left(-\mathbf{H} \bullet \mathbf{M}\right) - TS \, .$$

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 1).
Fixed temperature	Helmholtz: $F = U - TS$	Heat is exchanged with heat bath.
Fixed temperature and pressure	Gibbs: $G = U - TS + pV$	Heat is exchanged with heat bath, work can be done and stored in the ambient pressure.
Other (general) energy stores	free-energy = $U + \sum_i F_i X_i - TS$	Heat is exchanged with heat bath, work can be done and stored in various state-dependent ways (Note 2).

Notes

1. $-S$ isn't really an "energy", but temperature is not relevant for a system that cannot exchange heat, so we don't define the free energy as $-TS$. We could instead consider $A = U - TS$ as the free energy of an isolated system, since U is fixed, and T doesn't matter.
2. The F_i are generalized forces, and so unrelated to Helmholtz free-energy F.

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??

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 one or more final (product) particles. For example:

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

These reactions need not be chemical in nature. When they are, 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. "Chemical potential" is a misnomer, since the reaction need not be chemical. Some sources use the term "electrochemical potential" to try to broaden the term's scope, but it is still inadequate. As we will show, a better term is "free-energy per particle", which one might shorten to "particle potential."

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 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:

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

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* ($F \equiv U - TS$); and (2) a system at constant temperature and constant pressure (which therefore includes the reversible pV work) is characterized by *Gibbs free energy* ($F \equiv U + pV - TS$). 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:

$$\mu \equiv \frac{\partial A}{\partial N} \quad \textit{where} \quad A \equiv \text{appropriate free energy for the system} . \tag{5.3}$$

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

Thus, we can say μ measures the relative probability of adding a particle to that system.

Example: for a system at fixed T and V:

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

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

Example: For a system at fixed T and p:

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

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

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.

Many sources define $\mu \equiv (\partial U/\partial N)_{S,V}$. This is equivalent to our definition through some identities [Reif p313], but it is much less useful. Definition (5.3) is directly applicable to most situations. Note that it need not specify which variables to hold constant, since that is implicit in the one-and-only applicable free-energy for the system, A.

> [The absolute or relative nature of chemical potential is subtle. Note that free energy A may include an arbitrary additive constant (per particle) for its energy term, and often *another* arbitrary additive constant (per particle) for its entropy. These constants therefore affect μ. However, in some cases, μ is an absolute quantity, without an additive constant.]

For some applications, only differences in chemical potential are important, and such applications tolerate an additive constant in μ. (See below.)

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, and it's the cost of *adding* such a particle to a system (rather than creating it).

As previously noted, "probability rules," not energy; equivalently, "free energy" rules (not energy). Therefore, we can think of the chemical potential of a particle as the "probability cost" of adding a particle, measured by free energy. When $\mu > 0$, the system evolves toward fewer particles (if it can), to lower its free energy. When $\mu < 0$, the system evolves toward more particles, again to lower its 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:

$$đQ = T\,dS .$$

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:

$$\text{probability cost} \leftrightarrow dE - T\,dS = dA \qquad \textit{where} \quad dA \equiv \text{the change in free energy}.$$

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.

Systems with different kinds of particles: Many systems are a mixture of different kinds of particles, which we label here with the subscript *i*. Then instead of *N* to specify the number of particles, we have a set of $\{N\} \equiv \{N_1, N_2, ...\}$ to specify the number of each kind of particle. Then each kind of particle has its own chemical potential:

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

Gibbs Free Energy and Chemical Potential: 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:

$$\text{probability cost of one particle} = dE + pdV - TdS \equiv dG \qquad \rightarrow$$

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

For reasons I don't understand, many references now consider a special case: 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:

$$\mu = G / N \qquad \text{(system at constant } T, p)\,. \tag{5.4}$$

This situation of each particle "feeling" the same environment comes about for *any* system when 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. (Such additive constants are extensive: they scale with the system size; they are constant only for a given system size.) Therefore, the μ in (5.4) 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.

Secondly, it is easy to misinterpret the statement to mean that chemical potential is independent of system state. In fact, chemical potential *is* a function of the system state. For example, for a fixed-volume system, μ depends on pressure, *N*, etc [Sch p165].

And finally, for systems of multiple particle types, each type has its own chemical potential. The description 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. Once again, we see that the definition (5.3), $\mu \equiv \partial A/\partial N$, avoids such confusion.

Maxwell relations for chemical potential: TBS.

Chemical Potential of Photons

Photons are different than most other particles because their number is not conserved. Systems freely create (radiate) or absorb photons in any number. Therefore, in equilibrium, the chemical potential of a photon must be zero: if it were positive, the system would absorb some photons and decrease the free-energy. If it were negative, the system would radiate more such photons, and again decrease the free energy. The photon number is stable (i.e., in equilibrium) only when there is no free-energy benefit to either absorb or create photons.

Each photon frequency is considered a separate particle "type". Therefore, each frequency has a chemical potential of zero.

Comments on the Chemical Potential

At least one reference says μ is due to inter-molecular forces. The potential energy of such forces are sometimes part of μ, 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 (KE) per particle, potential energy (PE) per particle, possibly other energies, and entropy per particle:

$$\mu = (KE) + (PE) - TS \qquad \text{for one particle}.$$

[?? However, recall that the potential energy and the entropy S both include arbitrary zero points. If we change the zero of either, don't we also change the zero of μ, 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 μ changed by that amount. However, the μ for photons and neutral particles such as atoms, molecules, and neutrons, are not affected. Consider gas particles when gravity is significant, or charged particles in an electric field.]

Free Energy for Unusual Systems

To illustrate the generality of free energy, we consider here an unusual kind of system that expands against a spring. TBS.

6 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*:

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

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

The continuous energy approximation: As before, since the detailed structure of the quantum energy states is not important, we aggregate discrete quantum states into a continuous energy density of states. For single particles:

$$g_j \to g(E)dE \qquad \textit{where} \quad g_j \equiv \text{quantum multiplicity of state } j$$

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

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

$$Z_1(\beta) = \sum_{j=1}^{\infty} g_j e^{-\beta E_j} \qquad \to \qquad Z_1(\beta) = \int_0^{\infty} dE\ g(E) e^{-\beta E} .$$

For a full system, we generally know the hamiltonian, which gives *E* as a function of *p* and *x*:

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

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

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

$$Z(\beta) \equiv \sum_{s=1}^{\text{microstates}} e^{-\beta E_s} = \int_{\substack{\text{phase}\\ \text{space}}} \frac{dx_1 dp_1 \ldots dx_f dp_f}{h^f} H(p_i, x_i) .$$

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 the 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, according to $\delta W = F\, dX$. 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).

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.

Example: 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:

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

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:

$$H_i = H_f \qquad \Rightarrow \qquad U_i + p_i V_i = U_f + p_f V_f \qquad \Rightarrow \qquad U_f - U_i = p_i V_i - p_f V_f .$$

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

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

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

$$= 239\text{ C}$$

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. The "heat of vaporization" of water at 1 atm is 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.

The Partition Function

The partition function was briefly introduced earlier:

$$Z(T, ...) \equiv \sum_{r=1}^{\text{microstates}} e^{-E_r/kT} \quad \textit{where} \quad r, E_r \text{ are microstates and their energies}.$$

It is dimensionless, and may be a function of any system parameters, such as *N*, *V*, *P*, etc. For brevity, we generally omit the extra parameters, and write it only as a function of *T* (or *β*). Taken as a function of temperature *T*, the partition function *Z*(*T*) defines *all* thermodynamic properties of a system. In practice, we often use it 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. The use as normalization constant is trivial, though, and simply restates a basic probability definition that doesn't require the bother of defining any partition function. Using the partition function to compute average energy is more interesting, because it starts us down the path of using it to compute all other thermodynamic properties, as well.

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

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, in principle, all the way down to zero. Differing much from the average is extremely unlikely for large systems, but for mathematical expedience, it is still counted as a possibility in *Z*(*T*).

When our "system" is a single particle, we often note that by calling the partition function $Z_1(T)$, where *T* is the temperature of the heat bath. For any system, we can call its partition function *Z*(*T*).

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.

This following discussion 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):

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

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 system energy value. For a system with discrete energies:

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

This readily generalizes to a system with continuous energies:

$$Z(\beta) \equiv \int_{E_{\min}}^{\infty} dE \, \underbrace{\omega(E)\exp(-\beta E)}_{\text{RelPr(energy } E)} \qquad \textit{where} \quad \omega(E) \equiv \text{energy density of microstates} . \tag{6.1}$$

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

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

$$\text{Pr}\left(\text{energy} = E_m\right) = \frac{\text{RelPr}(E_m)}{\displaystyle\sum_{j=1}^{\text{microstates}} \text{RelPr}\left(E_j\right)} = \frac{\omega_m \exp(-\beta E_m)}{Z(\beta)}$$

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 an important first step in understanding its broader use. Recall from basic statistics that for any discrete random variable, say E, its average is:

$$\left\langle E \right\rangle = \sum_i \text{Pr}(E_i)\, E_i \; . \tag{6.2}$$

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

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

This is exact for all systems, tiny or large. But purely as a mathematical trick, we notice that the factor:

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

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

So far, it gives us nothing new, simply using Z for something already well-known. As another step, notice that:

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

This is still exact for all systems.

Partition function tells all: In the end, given the partition function $Z(\beta)$ for all values of β, 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 microstates $\omega(E)$. Therefore, knowing the partition function and knowing the density of microstates are equivalent. How are the two related? Recall the partition function for continuous systems, (6.1):

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

This is the definition of the Laplace transform of $\omega(E)$, with transform variable β. 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 $\omega(E)$,
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 $\omega(E)$ to find the partition function, so this Laplace transform relationship isn't usually very useful. [Pathria's example where it is useful??] We show shortly that the partition function is a convenient intermediate step to go from the density of microstates $\omega(E)$ to the free-energy $A(T, ...)$. Then the free-energy gives us other properties from simple derivatives.

Example: Thermodynamic zipper

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

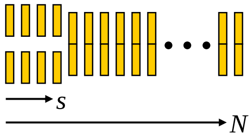


Figure 6.1 Zipper requires energy ε to open each link.

(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:

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

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

This is exact, for all *N*.

(b) The average number of open links:

$$\begin{aligned}\langle s\rangle &= \frac{\langle E\rangle}{\varepsilon} = \frac{-1}{\varepsilon}\frac{\partial}{\partial\beta}\ln\left(Z(\beta)\right) = \frac{-1}{\varepsilon}\frac{\partial}{\partial\beta}\left[\ln\left(1-\exp\left(-(N+1)\beta\varepsilon\right)\right) - \ln\left(1-\exp\left(-\beta\varepsilon\right)\right)\right] \\ &= -\frac{1}{\varepsilon}\left[\frac{\left(N+1\right)\varepsilon\exp\left(-(N+1)\beta\varepsilon\right)}{1-\exp\left(-(N+1)\beta\varepsilon\right)} - \frac{\varepsilon\exp\left(-\beta\varepsilon\right)}{1-\exp\left(-\beta\varepsilon\right)}\right] \\ &= -\frac{\left(N+1\right)\exp\left(-(N+1)\beta\varepsilon\right)}{1-\exp\left(-(N+1)\beta\varepsilon\right)} + \frac{\exp\left(-\beta\varepsilon\right)}{1-\exp\left(-\beta\varepsilon\right)}\end{aligned} \qquad (6.3)$$

This is exact, for all *N*.

(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 $<s>$, but a 1st order expansion of $(1 - \exp(-x)) = x + O(x^2)$ resolves the problem:

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

Whoa! Can this be right? Yes. We'll analyze the high-temperature limit in two different ways. Consider the temperature of our zipper, as a function of energy (from (3.2)):

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

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 ε. 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 $<s> = 0$. This is an unusual system (constant entropy), and such weird systems fairly often have bizarre behavior like this.

A slightly 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. Let's compute the average energy. From the above probabilities, and also from (6.3):

$$\langle E \rangle = \frac{8\cdot 0 + 4\cdot 1 + 2\cdot 2 + 1\cdot 3}{15}\varepsilon = \frac{11}{15}\varepsilon, \qquad \langle E \rangle = -\frac{4\cdot(1/16)}{15/16} + \frac{1/2}{1/2} = \frac{11}{15}\varepsilon .$$

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. TBS.

Free Energy and the Partition Function

When considering the possible macrostates of a system, what really matters is the *most-probable* macrostate, 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)-microstate-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):

$$E_{\text{most-likely}} = \underset{\substack{E_m = \text{energy} \\ \text{levels}}}{\arg\max} \left(\omega_m e^{-\beta E_m} \right) \quad \text{or} \qquad E_{\text{most-likely}} = \underset{E}{\arg\max} \left(\omega(E) e^{-\beta E} \right),$$

where "arg max" means pick the argument E_m that maximizes the following expression. 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 likelihood = (Boltzmann-factor) × (#-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. Therefore, we use an approximation to the most-likely macrostate (Figure 4.2): 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. For large systems, it is indistinguishable from the "exact" answer.

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:

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

Here's why: for large N, there will be only a few terms in the partition function which dominate 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:

$$Z(\beta) = \sum_{m=1}^{\text{microstates}} \omega_m \exp\left(-E_m / kT\right) \approx \sum_{m}^{\substack{\text{microstates near} \\ \text{maximum}}} \omega_m \exp\left(-E_{max} / kT\right)$$

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

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

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

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 ω_m near E_{max} with a single ω_{eff}, since they all have essentially the same energy, E_{max}:

$$Z(T) \approx \omega_{eff} \exp\left(-E_{max} / kT\right) \qquad \text{where} \quad E_{max} \text{ is the dominant energy level in the sum}.$$

Now we write ω_{eff} in terms of energy by putting it into the exponential:

$$Z(T) \approx \exp(\ln \omega_{eff}) \exp\left(-E_{max} / kT\right) \qquad \text{writing } \omega_{eff} \text{ as an exponential}$$

$$= \exp\left(\frac{-E_{max} + kT \ln \omega_{eff}}{kT}\right) \qquad \textit{where} \quad kT \ln \omega_{eff} \text{ writes } \omega_{eff} \text{ in terms of energy}$$

$$= \exp\left(-\frac{E_{max} - TS(E_{max})}{kT}\right) \qquad \textit{where} \quad S(E_{max}) \equiv k \ln \omega_{eff}$$

$$= \exp\left(-\frac{A(T)}{kT}\right) \qquad \textit{where} \quad A(T) \equiv \text{free energy of the macrostate at temp } T$$

Once again, we're simply quantifying how multiplicity (ω_{eff}) defies energy, and measuring multiplicity in units of energy. Solving for $A(T)$ yields the free energy from the partition function:

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

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 tiny 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. Computing $Z(T)$ within the system constraints automatically give the one-and-only meaningful free-energy for the system.

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.

Distinguishable Subsystems

We start with a simple case: we define a new system as the combination of two smaller subsystems, A and B. For now, we take them to be distinguishable, so no exclusion principle or quantum statistics applies. 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:

$$\text{RelPr}(A \text{ in } S_A \text{ with energy } E_A, \text{and } B \text{ in } S_B \text{ with energy } E_B) = \text{RelPr}(A \text{ in } S_A)\text{RelPr}(B \text{ in } S_B)$$

$$= e^{-E_A/kT} e^{-E_B/kT} = e^{-(E_A+E_B)/kT}$$

This is nothing more than the statement that the combined system, with energy $E_A + E_B$, follows the Boltzmann distribution. A few short steps of algebra then shows [Wan p86t]:

$$\begin{aligned}
Z_{AB}(T) &\equiv \sum_{r}^{\substack{\text{microstates}\\ \text{of } A}} \sum_{s}^{\substack{\text{microstates}\\ \text{of } B}} \exp\left(-\left(E_r + E_s\right)/kT\right) \\
&= \sum_{r}^{\substack{\text{microstates}\\ \text{of } A}} \sum_{s}^{\substack{\text{microstates}\\ \text{of } B}} \exp\left(-E_r/kT\right)\exp\left(-E_s/kT\right) \\
&= \left(\sum_{r}^{\substack{\text{microstates}\\ \text{of } A}} \exp\left(-E_r/kT\right)\right)\left(\sum_{s}^{\substack{\text{microstates}\\ \text{of } B}} \exp\left(-E_s/kT\right)\right) = Z_A(\beta)Z_B(\beta)
\end{aligned}$$

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

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

$$Z_N(T) = \left(Z_1(T)\right)^N \qquad \text{(identical, distinguishable, noninteracting subsystems)}. \qquad (6.4)$$

As an example, consider an Einstein solid (a crystal of harmonic oscillators) made of spin-1/2 fermions with a magnetic dipole moment. Besides the quantum spatial state (oscillator) energy, there can be spin energy, either from an external magnetic field or from adjacent-spin interactions. The spin energy is independent of the spatial state (in fact, we can consider the spatial and spin components as separate subsystems). Then the total energy separates into a sum of spatial + spin, and therefore the partition function factors into a product of spatial × spin:

$$E(T) = E_{spatial}(T) + E_{spin}(T) \qquad \Rightarrow \qquad Z(T) = Z_{spatial}(T)Z_{spin}(T).$$

Identical Subsystems

If the subsystems are indistinguishable, such as atoms, then 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):

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

This assumes that the particles are so sparse that multiple-occupation effects are negligible. 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.

However, correct Boltzmann counting will be critically important later, for the grand partition function, where N is a thermodynamic variable.

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:

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

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:

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

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].

7 The Ensembles

Whole chapter under construction.

There are 3 main classes of systems we consider, distinguished by their constraints. An **ensemble** is a (usually 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

These 3 classes of systems lead to 3 kinds of ensembles:

Microcanonical ensemble (MCE)	N fixed, E "essentially" 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 (except for a few unphysical concoctions). This is equivalent to saying that given enough time, the system will visit a representative set of microstates. (However, in practice, it cannot visit every microstate, because that would take inconceivably many multiples of the age of the universe to happen.)

Microcanonical Ensemble

A **microcanonical ensemble** is a set of systems, where each is (almost) completely isolated from its surroundings: no energy or particle exchange, no work done. The microcanonical ensemble is an unphysical idealization used to develop methods of calculating thermodynamic quantities from microscopic physics. There are some problems defining entropy [e.g., Pathr sec. 1.4], so theorists usually allow some tiny variation in system energy (here called ΔE). In the discussion surrounding eq. (3.5), we show in detail that absolute entropy can be defined (independent of ΔE). [I'm not sure the "ensemble" idea is a good one.]

TBS.

System of (nearly) fixed energy [Hua p176b, Kub p7b], volume, particle number. Don't appeal to "classical" thermodynamics; pretend we don't know thermodynamics. T is defined and it is *intensive* (see definitions of equilibrium, temperature, etc., 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; it must be allowed to fluctuate.

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

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

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 (reservoir) at a fixed temperature *T*, and which can also exchange particles with the reservoir. Thus heat and particles can both be exchanged with the reservoir, 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 reservoir is irrelevant, and in particular, it need *not* be other identical systems. Only its temperature and chemical potential matter.

TBS.

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

$$\Phi = -kT \ln \mathcal{Z} = -pV$$

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 $z = \exp(\beta\mu)$:

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

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

Classically:

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

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

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

Quantum mechanically:

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

$$= \sum_{N=0}^{\infty} \sum_{\{n_j\}} \exp\left[-\beta n_j\left(\varepsilon_j - \mu\right)\right]$$

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

$$\Phi = -kT \ln \mathcal{Z} = -pV \qquad \text{(grand canonical ensemble)}$$

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:

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

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

$$\text{pdf}(x) = \frac{e^{-\beta E(x)}}{Z} \qquad \textit{where} \quad Z(\beta) = \int_{all\ x} e^{-\beta E(x)}\, dx .$$

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

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

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 $\langle E \rangle = kT/2$.

Some references overcomplicate the derivation, and miss 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 any 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):

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

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

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

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 PDF for the dynamic variable. This process also allows you to answer questions about the 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-spring oscillator, there are two dynamic variables: position of the ball, and momentum of the ball. In a set of *N* such balls-and-springs in thermal contact, energy can be exchanged between any of the *2N* dynamic variables. **Thermal equilibrium** is by definition the stable macrostate 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, the derivation of [Rei] 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: [Rei] 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 β, but is a function of β. Rewrite the numerator in terms of the denominator (we do this all the time in Stat Mech):

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

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

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

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

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

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

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

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

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

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

$$\langle E \rangle = -\frac{d}{d\beta}\left(-\frac{1}{2}\ln\beta + const\right) = \frac{1}{2}\left(\frac{1}{\beta}\right) = \frac{1}{2}kT$$

Wasn't that fun?]

8 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.]

Summary of the 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
Maxwell	TBS	The speed (or momentum) distribution of particles in an ideal gas.	An ideal gas with no potential (e.g., no gravity).	
Boltzmann	$e^{-E/kT}$	*relative* probability of being in a single microstate, of energy *E*.	any system of any size or complexity, quantum or classical. Most important!	an atom moving in a gas, electronic excitations of an atom in a gas, a magnetic dipole in a lattice
Maxwell-Boltzmann	$\dfrac{1}{e^{(E-\mu)/kT}}$ $= e^{-(E-\mu)/kT}$	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 boson 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
Black body (Planck spectrum)	$u(E) =$ $\dfrac{8\pi}{(hc)^3}\dfrac{E^3}{e^{E/kT}-1}$	the energy density per unit energy of thermal photons.	Black body radiation, i.e. photons in equilibrium with a black body.	

Be careful to distinguish the "Maxwell distribution", the "Boltzmann distribution", and the "Maxwell-Boltzmann distribution". They are all different, though some sources (confusingly) use the term "Boltzmann distribution" to also mean "Maxwell-Boltzmann distribution".

The black body radiation case is somewhat unusual, because photons have no rest mass, and their numbers are not conserved. They are freely created (radiated) and absorbed by charges, with no threshold energy. This gives them a different distribution than massive bosons.

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, μ, serves as a natural energy reference level. The chemical potential may include an arbitrary additive constant of energy per particle. This allows the Fermi-Dirac distribution to be written with *absolute* probabilities:

$$\langle \text{occupation number} \rangle = \Pr(\text{single-particle state s with energy } E_s \text{ is occupied})$$

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

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, but no such state need exist.

μ can be determined by the density of single-particle states $g(E)$, the number of particles in the system N, and the temperature, T. Given fixed constraints (e.g., volume), there is a unique μ 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 μ.

Physicists usually 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 μ the Fermi energy at all temperatures: $\varepsilon_F(T) \equiv \mu(T)$, i.e. the Fermi level is just another name for chemical potential.

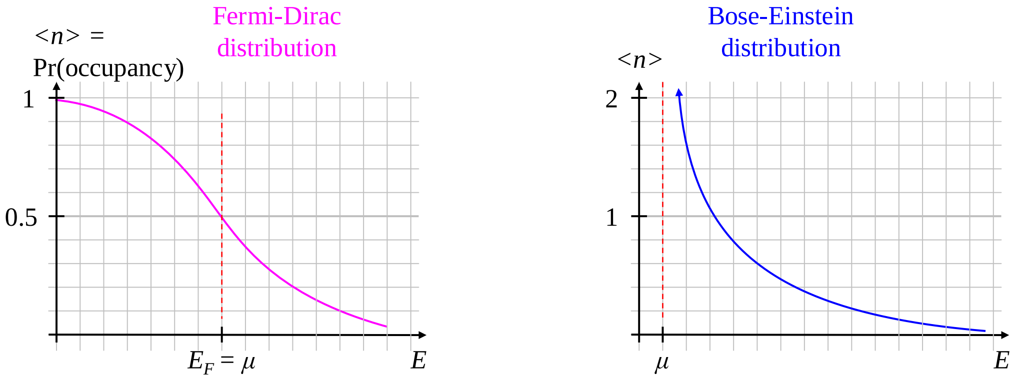


Figure 8.1 (Left) Fermi-Dirac distribution: <occupation number> = Pr(occupancy). (Right) Bose-Einstein distribution: <occupation number> is unbounded.

We derive the Fermi-Dirac distribution from the definition $\mu = \partial A/\partial N$ [TBS].

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, μ, 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:

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

The behavior at low energy E_s is important: 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 $E > \mu$. It blows up when $E = \mu$, and would be negative for $E < \mu$, which is physically meaningless. Therefore:

For bosons, μ 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. Each frequency of photon is a different particle species.

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

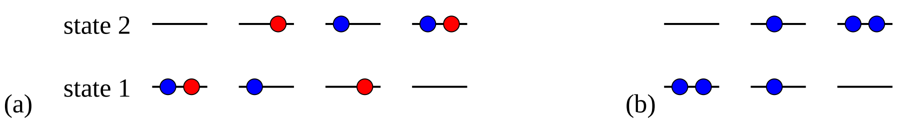


Figure 8.2 (a) Classical states of two particles. (b) Quantum states of two identical particles.

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

We now derive the black body radiation spectrum (and thus the occupation number formula) from the Boltzmann distribution applied to photons, which are bosons. However, photons do not have a fixed particle number, since they are easily created (radiated) and destroyed (absorbed); this is distinct from say, ^{4}He where the number of atoms is fixed.

Occupation numbers of boson states: If we treat each EM radiation mode ($\mathbf{k}$, $\boldsymbol{\varepsilon}$) 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 distinguishable system held at fixed temperature (say, by a heat bath), and it applies to boson excitation levels, because it relies only on the properties of the heat bath, and not at all on properties of the system.

Note that in this context, we consider a photon state, say |2>, as a single "system" which is either occupied or not; we do *not* consider it as 2 separate photons. Then $N \equiv N_{\mathbf{k},\boldsymbol{\varepsilon}}$ here is excitation level, not number of system particles. Then for a given mode ($\mathbf{k}$, $\boldsymbol{\varepsilon}$) of frequency $\omega = c|\mathbf{k}|$, the total number of excitations over all N is:

$$\text{RelPr(photon mode is}\left|N\right\rangle) = e^{-E_N/kT} \qquad \Rightarrow \qquad \text{RelPr}\left(\left|N\right\rangle\right) = e^{-N\hbar\omega/kT}$$

$$\underbrace{\left\langle N_{\mathbf{k},\boldsymbol{\varepsilon},tot}\right\rangle}_{\text{total \# of excitatons of mode }\mathbf{k},\boldsymbol{\varepsilon}} = \sum_{N=0}^{\infty} N\,\Pr(N) = \frac{\displaystyle\sum_{N=0}^{\infty} Ne^{-N\hbar\omega/kT}}{\displaystyle\sum_{N=0}^{\infty} e^{-N\hbar\omega/kT}}, \qquad N \equiv N_{\mathbf{k},\boldsymbol{\varepsilon}}\,.$$

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:

$$1 + x + x^2 + x^3 + \ldots = \frac{1}{1-x} = \frac{1}{1-e^{-\hbar\omega/kT}} \qquad \left(\text{denominator}\right).$$

Now consider the numerator:

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

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

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

The final average number of excitations is numerator/denominator:

$$\left\langle N_{\mathbf{k},\boldsymbol{\varepsilon},tot}\right\rangle = \frac{e^{-\hbar\omega/kT}}{\left(1-e^{-\hbar\omega/kT}\right)^{2}} \bullet \left(1-e^{-\hbar\omega/kT}\right) = \frac{e^{-\hbar\omega/kT}}{1-e^{-\hbar\omega/kT}} \qquad \text{multiply top \& bottom by } e^{+\hbar\omega/kT}$$

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

Equivalent to the above, it can be shown that the chemical potential of photons in thermal equilibrium is $\mu = 0$ [Schr p??]. Then we can think of each mode as a single-photon quantum state. Photons are bosons, so each state has an occupation number, N_{occ}. The average N_{occ} using the B-E distribution with $\mu = 0$ is that just derived.

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

Note that $<N_{\mathbf{k},\boldsymbol{\varepsilon},tot}>$ may be fractional, and greater or less than 1.

The Amazing Maxwell-Boltzmann Occupation Number Distribution

The "Maxwell-Boltzmann distribution" is 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:

$$(E-\mu) > (few)kT \qquad \Rightarrow \qquad e^{(E-\mu)/kT} >> 1 \quad \Rightarrow$$

$$\underbrace{\frac{1}{e^{(E-\mu)/kT}+1}}_{\text{Fermi-Dirac}} \approx \underbrace{\frac{1}{e^{(E-\mu)/kT}-1}}_{\text{Bose-Einstein}} \approx \underbrace{\frac{1}{e^{(E-\mu)/kT}}}_{\text{Maxwell-Boltzmann}} \quad << 1.$$

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 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 likely 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:

$$\Pr(\text{single-particle energy } E_s) \propto e^{-\beta E_s} \qquad \text{or} \qquad \text{RelPr}(E_s) = e^{-\beta E_s}.$$

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 $g(E)$ of single-particle 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, $g(E)$! However, the *absolute* probabilities *do* depend on $g(E)$ 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 exceedingly 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 likely 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 speed distribution: The Maxwell speed distribution describes the distribution of speeds (or momenta) of nonrelativistic molecules in an ideal gas. You can derive it from the M-B energy distribution and $g(E)$ by noting that *speed* $= (2E/m)^{1/2}$.

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):

$$w_{abs} = B_{abs} N M P_0 \qquad \text{(photons/s)} . \tag{8.1}$$

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:

$$w_{em,stim} = B_{em} N M P_e \qquad \textit{where} \quad N \equiv \#\text{ ambient photons}. \tag{8.2}$$

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:

$$w_{em,spon} = A M P_e \qquad \textit{where} \quad A \equiv \text{decay constant of proportionality}. \tag{8.3}$$

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

$$w_{abs} = w_{em,stim} + w_{em,spon} \quad \Rightarrow \quad B_{abs} N P_0 = B_{em} N P_e + A P_e = \left(N B_{em} + A\right) P_e . \tag{8.4}$$

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 $T \to \infty$:

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

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

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 from the equilibrium of rates, (8.4), 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:

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

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

In 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 (8.1) - (8.3), 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):

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

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

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

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 single-particle density of states by the average energy of each state, at a temperature *T*: (this needs cleaning up the notation for densities??)

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

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 ω 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:

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

As we've seen before, this implies that each state occupies a volume in phase space of $(2\pi\hbar)^3 = h^3$.

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 |**k**|, we change variables to *E* in two steps: (1) we convert $\rho_{\mathbf{k}}$ to the density per unit *magnitude* of **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 **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 ω:

$$u(\omega) = \hbar\omega N(\omega)\rho_E(E)\frac{dE}{d\omega} \qquad \textit{where} \quad E = \hbar\omega, \quad \frac{dE}{d\omega} = \hbar$$

$$= \frac{\hbar\omega}{e^{\hbar\omega/kT}-1}\cdot\frac{(\hbar\omega)^2 V}{\pi^2(\hbar c)^3}\hbar = \frac{\hbar}{\pi^2}\cdot\frac{\omega^3}{c^3}\cdot\frac{1}{\left(e^{\hbar\omega/kT}-1\right)}V.$$

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 8.3). The ground state kinetic energy is *not* zero; for the idealized particle in a cube, it is:

$$KE_0 = n_x^2 + n_y^2 + n_z^2 = 3 \text{ (in some units)} \qquad \textit{where} \quad KE \equiv \text{kinetic energy}.$$

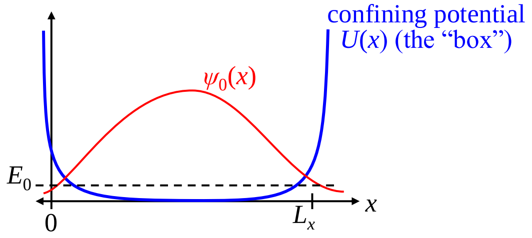


Figure 8.3 The ground state wave-function is always about a half-wavelength, regardless of the size of the box, or the details of the confining potential.

In a large system, this is tiny, but conceptually, it is finite. The 1st excited state has kinetic energy:

$KE_1 = 1^2 + 1^2 + 2^2 = 6$ (twice E_0).

We choose $U(\mathbf{r}) = 0$ everywhere, so:

$E_n = KE_n$, and in particular: $E_0 = KE_0$.

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 of a box. 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:

$$\psi_0(x,y,y) = N\left(\sin\left(k_x x\right)\sin\left(k_y y\right)\sin\left(k_z z\right)\right) \quad where \quad k_x = \pi / L_x, \text{ etc.}$$

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:

$$N = \sum_{s}^{\text{microstates}} \frac{1}{e^{(E_s - \mu)/kT} - 1} \qquad \text{(exact)} .$$

At warm temperatures, we can approximate the sum of average occupation numbers with an integral (Figure 8.4, left):

$$N \approx \int_0^\infty g(E) \frac{1}{e^{(E-\mu)/kT} - 1} dE \qquad where \quad g(E) \equiv \text{single-particle density of states} .$$

At every temperature, the ground state occupation number N_0 is finite, but above some critical temperature T_c, it is nearly zero (negligibly small).

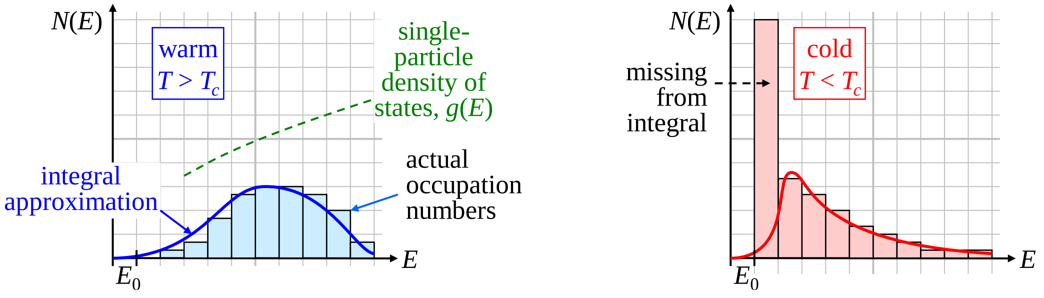


Figure 8.4 (Left) The integral approximation to the occupation number sum works well at high *T*. (Right) The integral approximation to the occupation number sum fails at low temperatures.

However, at very low temperatures, the integral approximation for N fails because the ground state occupation number is mostly excluded from the integral (Figure 8.4, 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}):

$$N = N_0 + N_{ex} \approx N_0 + \int_0^\infty g(E) \frac{1}{e^{(E-\mu)/kT} - 1} dE . \qquad (8.5)$$

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 (8.5), 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:

$$\rho_k(k) = 4\pi k^2 \frac{V}{(2\pi)^3}$$

$$g(E) = \rho_k(k)\frac{dk}{dE}, \qquad E = \frac{\hbar^2 k^2}{2m}, \qquad k = \frac{\sqrt{2mE}}{\hbar} \qquad \Rightarrow \qquad \frac{dk}{dE} = \frac{\sqrt{2m}}{\hbar}\left(\frac{1}{2}E^{-1/2}\right)$$

$$g(E) = \underbrace{4\pi\left(\frac{2mE}{\hbar^2}\right)\frac{V}{(2\pi)^3}}_{\rho_k}\underbrace{\frac{\sqrt{2m}}{\hbar}\left(\frac{1}{2}E^{-1/2}\right)}_{dk\,/\,dE} = \underbrace{4\sqrt{2}\pi\frac{m^{3/2}}{(2\pi\hbar)^3}}_{\text{constant}} V\,E^{1/2}\,. \qquad \text{[Schroeder 7.123 p316]}$$

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.

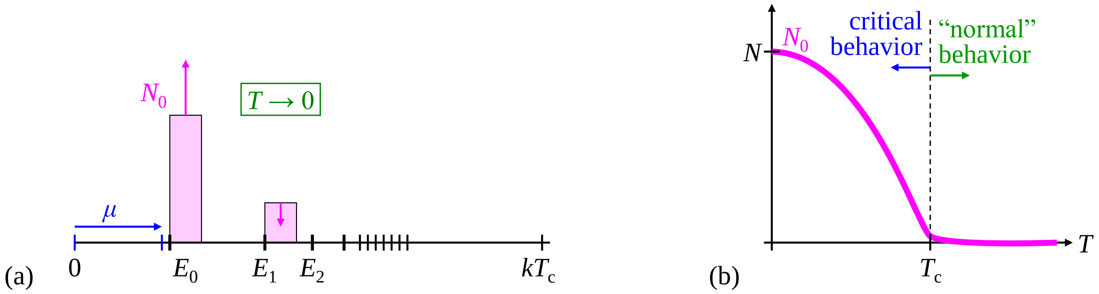


Figure 8.5 (a) As T decreases below T_c, $\mu \to E_0$, and $N_0 \to N$. (b) Ground state occupation number N_0 vs. temperature, showing the critical region.

At low temperatures, as more particles go into the ground state, μ must approach E_0. N_{ex} can then be upper-bounded (and approximated) by letting $\mu(T) \to E_0$:

$$N_{ex} < N_{upper} \approx \int_0^\infty g(E)\frac{1}{e^{(E-E_0)/kT}-1}\,dE = C\int_0^\infty \frac{E^{1/2}}{e^{(E-E_0)/kT}-1}\,dE\,.$$

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, μ is very close to E_0 (Figure 8.5a), 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 $T < T_c$:

$$N_{ex} \approx N_{upper} = N\left(\frac{T}{T_c}\right)^{3/2} \qquad T < T_c,$$

$$N_0 = N - N_{ex} = N\left(1 - \left(T/T_c\right)^{3/2}\right), \ T < T_c .$$

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" enough to call this a higher-order phase transition.

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 8.5b). 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 [ref??].

9 Thermodynamics

Thermodynamics isn't really Statistical Mechanics, because it's empirically based on observations of macroscopic systems. But SM is the microscopic model that should be able to derive all of thermodynamics, so they are strongly 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:

$$\frac{1}{T}=\left(\frac{\partial S}{\partial E}\right)_{N,V} \qquad \text{is empirical in thermodynamics .}$$

But in Statistical Mechanics, the relation:

$$\frac{1}{T}\equiv\frac{dS}{dE} \quad \text{is a precise } \textit{definition} \text{ in Statistical Mechanics .}$$

Reversible Vs. Irreversible, and Entropy

Reversibility is a huge stumbling block for many people, and many 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 essentially 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.

Note 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. Here again, even if friction conversion to heat is very slow, it is still irreversible.

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

$$dS = \frac{đQ_{slow}}{T} \qquad \text{(changes } to \text{ a system)} . \tag{9.1}$$

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 finite heat input:

$$\Delta U = Q \qquad \text{(no losses)} ,$$

no matter how fast or slowly the heat is transferred. The final system state is the same, whether the heat is added quickly or slowly. Thus, the final entropy is the same, whether heat is added quickly or slowly. However, when heat is added quickly, the system temperature becomes undefined for a while: the system is knocked out of equilibrium, and it has no definite temperature. Therefore, we cannot use (9.1) *on this fast process* to compute the final entropy.

Equation (9.1) requires temperature T be well-defined at all times. Therefore, to compute the final entropy of a system that had heat added quickly, we can simply compute the final entropy of the same system as if the heat were added slowly. This requires us to know it's heat capacity, and be able to find its temperature at every point in the process; if we can, then:

$$\Delta S = \underbrace{\int_i^f \frac{đQ}{T}}_{\substack{\text{reversibly} \\ =\text{slowly}}} \qquad \text{(heat flow } \textit{to} \text{ a system that has no energy losses)} . \qquad (9.2)$$

Furthermore, the state of the system after *any* 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. We can compute the entropy change as if it *were* heat added slowly, using the same equation.

Finally, note that any infinitesimal heat addition to a system, by itself, is *always* reversible. Reversibility or not enters only when we consider a *finite* amount of heat added over some finite time. Therefore, we don't write $dS = đQ_{rev}/T$, because $đQ$ (by itself) is always reversible. A proper equation referring to reversibility must refer to finite differences, and a reversible path of integration, as in (9.2).

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 microstates of a given macrostate (set of observables), 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

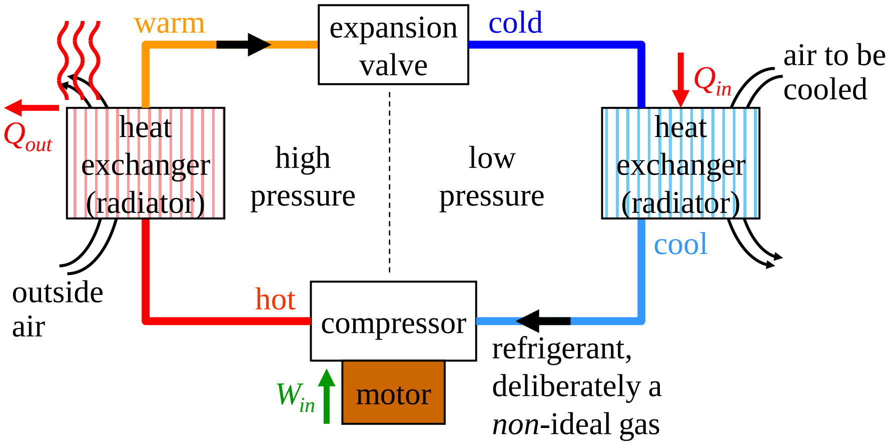


Figure 9.1 A typical refrigerator.

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 9.2. 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.

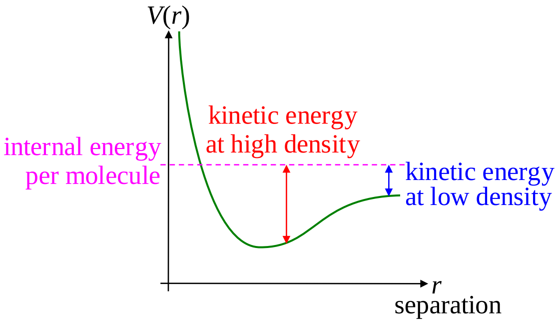


Figure 9.2 Energy vs. distance for refrigerant molecules.

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 the entropy of a closed system (even the universe)
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 gases, 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. We elaborate below, but in short, the engine's entropy increase of heat absorption must be canceled by its decrease on exhausting heat, which fixes the efficiency:

$$\left|\Delta S\right| = \frac{Q_{in}}{T_{hot}} = \frac{Q_{out}}{T_{cold}} \qquad \Rightarrow \qquad \eta = \frac{Q_{in} - Q_{out}}{Q_{in}} = 1 - \frac{T_{hot}}{T_{cold}} .$$

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.

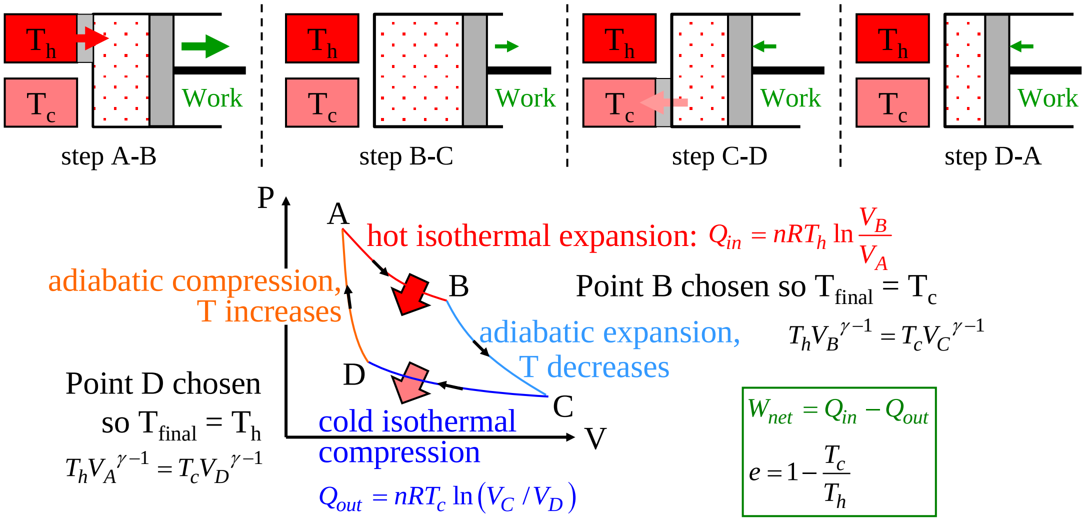


Figure 9.3 The Carnot engine (aka Carnot cycle). Ideal gas formulas are given, but the engine operation is independent of the working fluid.

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:

$$\Delta S_{A\text{-}B} = \int \frac{đQ}{T} = +\frac{Q_{in}}{T_h} \qquad \text{engine entropy increases .}$$

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.

10 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:

$$E = \frac{p^2}{2m} \quad \Rightarrow \qquad dE \propto p\, dp \tag{10.1}$$

Now we recall that the *definition* of temperature is:

$$T \equiv \left(\frac{dS}{dE}\right)^{-1}.$$

We also recall that we can *define* S to be (within an additive constant):

$$S(E) \equiv k \ln \Omega(E) \approx k \ln\left(\omega(E)\, V \delta E\right)$$

where ω is the density of microstates in states/volume/energy,
δE is a finite, but small, energy range (how to justify this??),
V is an arbitrary system volume.

We now consider a single particle, and since all particles are non-interacting, all particles behave the same. So now we must derive the single-particle density of states as a function of energy (we did this earlier??), 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 δp is the number of quantized lattice points (allowed momenta) in the spherical shell in momentum space of thickness δp. That is:

$$g(E)\,\delta E \propto p^2\ \delta p$$

where δp is the variation in momentum giving rise to δE variation in energy.

But from (10.1), $p \propto E^{1/2}$, and $p\,\delta p \propto \delta E$, so:

$$g(E)\,\delta E \propto p(p\,\delta p) \propto E^{1/2}\delta E\ .$$

Then at any arbitrary volume, V:

$$S \propto \ln\left(g(E)\, V \delta E\right) = \ln\left(E^{1/2}\, V \delta E\right) = \frac{1}{2}\ln E + \ln\left(V\ \delta E\right) = \frac{1}{2}\ln E + const.$$

$$\frac{dS}{dE} \propto \frac{1}{2E} \propto \frac{1}{E}$$

$$T \equiv \left(\frac{dS}{dE}\right)^{-1} \propto E$$

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 $T \propto KE$. 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 β is. β 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].

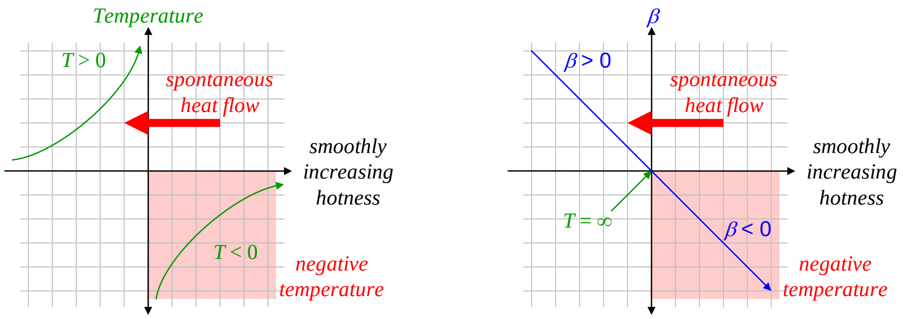


Figure 10.1 Why β is a better parameter for temperature than T, including negative temperature.

Negative temperature, or better, negative β, means that the probability of occupancy of an energy level *increases* with energy, instead of decreasing as with positive β [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, for systems with unbounded energy:

$$\Pr\left(E_r\right) \propto e^{-E_r/kT} \qquad \Rightarrow \qquad T > 0, \text{ or else probability blows up}.$$

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 β.

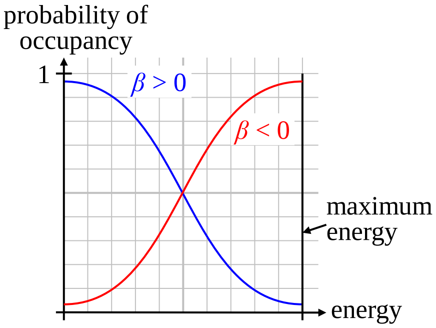


Figure 10.2 Probability of occupancy vs. system energy for a finite system.

$$\beta \propto \frac{1}{T} = \frac{\partial S}{\partial E} .$$

Now, even for $\beta < 0$, larger β is always colder, smaller β is always hotter. Heat flows spontaneously from low β to high β [Wan p69m][Pat p83]. Reciprocally, applying heat to the system lowers β. 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.

Spin Systems (and Other Finite Dimensional Quantum State Spaces)

Spin-dependent systems are very common in the real world, e.g. magnetic systems. The spin, and therefore magnetic moment, of a particle is described by a finite dimensional quantum state space, e.g. spin-1/2 particles have a 2D state space, and atomic orbitals have a $2l + 1$ dimensional state space. Also, in some cases, bound particles have only a small number of significant quantum states, and can be approximated as if they have only that small number of allowed states (often just 2 or 3). Systems of single-particles in a low-dimensional state space create distinct phenomena compared to free-particle systems, and therefore require analysis methods specific to them. However:

The underlying statistical mechanical concepts of Boltzmann factors, partition functions, and thermodynamic limits apply to spin systems just the same as other systems.

We now consider a few simple systems to illustrate some basic concepts. We start with a non-interacting system, which also defines our notation, and then proceed to the simplest 1D interactions, the Ising model.

This section assumes a basic familiarity with quantum matrix mechanics, and relies on understanding partition functions.

Noninteracting Spins (Subsystems)

Consider a particle with two magnetic moment states, $\sigma_{up} = +1$, and $\sigma_{down} = -1$. It is immersed in a magnetic field of magnitude B. The spins are eigenstates of energy. The single-particle hamiltonian can be written, in the |+> and |–> basis:

$$\hat{H}_1 = \begin{bmatrix} -\mu B & 0 \\ 0 & +\mu B \end{bmatrix} \qquad \text{where} \quad |\sigma\rangle = \begin{bmatrix} \sigma_+ \\ \sigma_- \end{bmatrix}, \ \mu \equiv \text{magnetic dipole moment} . \tag{10.2}$$

The signs denote that the low-energy state has the magnetic moment aligned with the magnetic field, and the high-energy state is anti-aligned (it takes work to twist the particle away from the magnetic field).

Now put this "system" in contact with a heat bath. Recall that such a tiny system has no temperature, so $\beta \equiv$ temperature of the heat bath. We can write its partition function by inspection:

$$Z_1(\beta) = e^{+\beta\mu B} + e^{-\beta\mu B} = \sum_{\sigma=\pm 1} \exp(-\beta\mu B\sigma) .$$

Since the hamiltonian is diagonal, its elements are the energies of the single-particle states, and the partition function can be written in the useful and compact form of a trace:

$$Z_1(\beta) = \operatorname{Tr} e^{-\beta \hat{H}_1} = \operatorname{Tr}\exp\left(\begin{bmatrix} +\beta\mu B & 0 \\ 0 & -\beta\mu B \end{bmatrix}\right) = \operatorname{Tr}\begin{bmatrix} \exp\left(+\beta\mu B\right) & 0 \\ 0 & \exp\left(-\beta\mu B\right) \end{bmatrix}$$

$$= \sum_{\sigma=\pm 1} \exp(-\beta\mu B\sigma)$$

as before. We have used these facts: the exponential of a matrix is a matrix of the same size; the exponential of a diagonal matrix is the matrix of exponentials of the diagonals; the trace of a matrix is the sum of the diagonals.

Now consider a system of N distinguishable such particles, that are noninteracting. Its partition function gives all its thermodynamic properties. The system partition function is the simple composition of subsystem partition functions, (6.4):

$$Z_N(\beta) = \left(Z_1(\beta)\right)^N .$$

In preparation for interacting particles, it is instructive to rederive this noninteracting Z_N directly from the N-particle hamiltonian. The basis states are all the 2^N combinations of the N magnetic moments; e.g. for $N = 3$: |+ + +>, |+ + –>, |+ – +>, etc. The full hamiltonian is the tensor product (aka outer product) of the individual hamiltonians:

$$\hat{H}_N = \hat{H}_1 \otimes \hat{H}_2 \otimes ... \hat{H}_N, \qquad \text{a } 2^N \times 2^N \text{ matrix} , \tag{10.3}$$

where all the $\hat{H}_j$ have the same form, (10.2). So far, our $\hat{H}_j$ are diagonal, and the outer product of diagonal matrices is also diagonal. For example:

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \otimes \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a\begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & b\begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} \end{bmatrix} = \begin{bmatrix} ac & . & . & . \\ . & ad & . & . \\ . & . & bc & . \\ . & . & . & bd \end{bmatrix} . \tag{10.4}$$

Therefore, $\hat{H}_N$ is diagonal, and:

$$Z_N(\beta) = \operatorname{Tr} e^{-\beta \hat{H}_N} .$$

Furthermore, we see from the first equality of (10.4) that the trace of an outer product of diagonal matrices is the product of their traces: $\operatorname{Tr}\left(\mathbf{A} \otimes \mathbf{B}\right) = \left(\operatorname{Tr}\mathbf{A}\right)\left(\operatorname{Tr}\mathbf{B}\right)$. Now we invoke a crucial theorem:

The trace of a matrix is independent of the basis in which it is expressed.
Corollary: the trace of any matrix is the sum of its eigenvalues.

Thus, diagonal or not:

$$\operatorname{Tr}\left(\mathbf{A} \otimes \mathbf{B}\right) = \left(\operatorname{Tr}\mathbf{A}\right)\left(\operatorname{Tr}\mathbf{B}\right) \qquad \text{(for } \textit{any} \text{ matrices } \mathbf{A} \text{ and } \mathbf{B}\text{)}.$$

Finally, then, we can say that for noninteracting subsystems, in *any* basis:

$$Z_N(\beta) = \mathrm{Tr}\, e^{-\beta \hat{H}_N} = \left(\mathrm{Tr}\, \hat{H}_1\right)\left(\mathrm{Tr}\, \hat{H}_2\right)\ldots\left(\mathrm{Tr}\, \hat{H}_N\right) = \left(\mathrm{Tr}\, \hat{H}_1\right)^N = \left(Z_1(\beta)\right)^N , \qquad (10.5)$$

as before. This means that if we can write the hamiltonian of N identical, distinguishable subsystems as a tensor product of N identical single-subsystem hamiltonians, then we can easily find the system partition function from the above equation. This is true, even if $\hat{H}_1$ is not diagonal.

The interacting system below will not be able to use this method exactly, but will be inspired by it.

Interacting Particles (Subsystems): the Ising Model

With all this as background, we can consider the Ising model [Pathr ??]. The Ising model describes particles on some kind of lattice, and each particle interacts only with its nearest neighbors. For simplicity, we consider a ring of N identical interacting 2-state magnetic moment particles (or subsystems). This is a 1-dimensional system. Some long-chain molecules are well-described by this simple model.

The system is immersed in a magnetic field B, and additionally, each particle interacts with its two nearest neighbors. The system hamiltonian is:

$$\hat{H} = -\mu B \sum_{i=1}^{N} \sigma_i - J \underbrace{\sum_{i=1}^{N}}_{\substack{\text{unique} \\ \text{n.n. pairs}}} \sigma_i \sigma_{i+1} = -\sum_{i=1}^{N} \left[\mu B \sigma_i + J \sigma_i \sigma_{i+1} \right]$$

$$\textit{where} \quad \sigma_i = \pm 1, \quad \sigma_{N+1} \equiv \sigma_1 \text{ because it's a ring}.$$

Note that the second summation considers each nearest-neighbor pair exactly once. If $J > 0$, then the spins are lower energy when aligned, called **ferromagnetism**. If $J < 0$, the spins are lower energy when anti-aligned, and the spins tend to alternate + – + ... , called **antiferromagnetism**.

We seek the system partition function, since that gives us all the thermodynamic properties. This is a complicated hamiltonian that is hard to analyze, as-is. If we could write it as in (10.3), we could use (10.5) to simplify it, but the interaction terms do not allow it to be factored as a string of tensor products. So let us see if we can do something similar. Consider the i-th particle in the system, and its single-subsystem hamiltonian:

$$\hat{H}_i = -\left[\mu B \sigma_i + J \sigma_i \sigma_{i+1} \right] \qquad \text{and} \qquad \hat{H} = \hat{H}_1 + \hat{H}_2 + \ldots + \hat{H}_N = \sum_{i=1}^{N} \hat{H}_i$$

$$\textit{where} \quad \sigma_{N+1} \equiv \sigma_1 \text{ because it's a ring}.$$

In the i-th particle hamiltonian $\hat{H}_i$, we have included its interaction with the next particle. This subsystem has 4 basis states, since each of σ_i and σ_{i+1} has two possible values, but as noted, we cannot use this hamiltonian directly. We are forced to invent a "trick" to find the partition function.

> [At this point, some references "symmetrize" the $\hat{H}_i$ [Pat4 (2) p476], but this step is unnecessary. We arrive at the same result without it.]

By definition, the system partition function can be written directly as the sum over all microstates of the Boltzmann factors. There are 2^N microstates, one for each combination of $\sigma_1 \ldots \sigma_N$:

$$Z_N(\beta) = \sum_{\sigma_1 = \pm 1} \ldots \sum_{\sigma_N = \pm 1} \exp\left\{ \sum_{i=1}^{N} \left[\beta \mu B \sigma_1 + \beta J \sigma_i \sigma_{i+1} \right] \right\} .$$

Now is when things get tricky. In (10.3), we found the string of products useful, so here, we try something similar. First, we convert the innermost sum to products of exponentials:

$$Z_N(\beta)=\sum_{\sigma_1=\pm 1}\dots\sum_{\sigma_N=\pm 1}\prod_{i=1}^{N}\exp\left[\beta\mu B\sigma_1+\beta J\sigma_i\sigma_{i+1}\right].$$

It's hard to see in the above form, but the exponential can be written as an inner product of vectors around a matrix operator. We define our kets, bras, and state vectors:

$$\sigma_i=+1\to\left|\sigma_i\right\rangle=\begin{pmatrix}1\\0\end{pmatrix}\quad\text{and}\quad\left\langle\sigma_i\right|=\begin{pmatrix}1&0\end{pmatrix},\qquad\sigma_i=-1\to\left|\sigma_i\right\rangle=\begin{pmatrix}0\\1\end{pmatrix}\quad\text{and}\quad\left\langle\sigma_i\right|=\begin{pmatrix}0&1\end{pmatrix},$$

and similar for σ_{i+1}. Then we can write the exponential in linear algebra form:

$$\exp\left[\beta\mu B\sigma_1+\beta J\sigma_i\sigma_{i+1}\right]\equiv\left\langle\sigma_i\right|\mathbf{P}\left|\sigma_{i+1}\right\rangle\qquad\Rightarrow\qquad\mathbf{P}=\begin{bmatrix}e^{\beta\mu B+\beta J}&e^{\beta\mu B-\beta J}\\e^{-\beta\mu B-\beta J}&e^{-\beta\mu B+\beta J}\end{bmatrix}.\tag{10.6}$$

P is called the **transfer matrix**, and is independent of *i*. Note that we're using a notation reminiscent of quantum mechanics, but **P** is *not* a quantum operator. It's just the matrix we need to get the exponential of the *i*-th particle energy, exp($-\beta E_i$). The system partition function now takes the form:

$$Z_N(\beta)=\sum_{\sigma_1=\pm 1}\dots\sum_{\sigma_N=\pm 1}\left\langle\sigma_1\right|\mathbf{P}\left|\sigma_2\right\rangle\left\langle\sigma_2\right|\mathbf{P}\left|\sigma_3\right\rangle\dots\left\langle\sigma_{N-1}\right|\mathbf{P}\left|\sigma_N\right\rangle\left\langle\sigma_N\right|\mathbf{P}\left|\sigma_1\right\rangle.$$

Each inner product is a number. Notice that the sum over the factors $|\sigma_2><\sigma_2|$ is a sum over a complete set of states, which is just the identity operator:

$$\sum_{\sigma_2=\pm 1}\left|\sigma_2\right\rangle\left\langle\sigma_2\right|=\mathbf{1}_{op}.$$

This is also true for σ_3 through σ_N. Therefore, all the summations disappear, except σ_1:

$$Z_N(\beta)=\sum_{\sigma_1=\pm 1}\underbrace{\left\langle\sigma_1\right|\mathbf{P}^N\left|\sigma_1\right\rangle}_{\text{diagonal entries of }\mathbf{P}^N}=\operatorname{Tr}\mathbf{P}^N.$$

Now, the trace is the sum of the eigenvalues of $\mathbf{P}^N$. If the eigenvalues of **P** are λ_1 and λ_2, then the eigenvalues of $\mathbf{P}^N$ are λ_1^N and λ_2^N. Let us find the eigenvalues of **P**. Using the standard eigenvalue approach, for an arbitrary 2×2 matrix, we have the characteristic equation:

$$0=\det\begin{bmatrix}a-\lambda&b\\c&d-\lambda\end{bmatrix}=\left(a-\lambda\right)\left(d-\lambda\right)-bc=\lambda^2-\left(a+d\right)\lambda+ad-bc\ \Rightarrow$$

$$\lambda_{1,2}=\frac{\left(a+d\right)\pm\sqrt{\left(a+d\right)^2-4\left(ad-bc\right)}}{2}.$$

Using the definition of **P,** (10.6):

$$\lambda^2 - 2e^{\beta J}\cosh\left(\beta\mu B\right)\lambda + 2\sinh 2\beta J = 0$$

$$\lambda_{1,2} = \frac{2e^{\beta J}\cosh\left(\beta\mu B\right) \pm \sqrt{4e^{2\beta J}\cosh^2 \beta\mu B - 8\sinh 2\beta J}}{2}$$

$$= e^{\beta J}\cosh\left(\beta\mu B\right) \pm \sqrt{e^{2\beta J}\cosh^2 \beta\mu B - 2\sinh 2\beta J} \qquad \text{use: } \cosh^2 = \sinh^2 + 1$$

$$= e^{\beta J}\cosh\left(\beta\mu B\right) \pm \sqrt{e^{2\beta J}\sinh^2 \beta\mu B + \cancel{e^{2\beta J} - e^{2\beta J}} + e^{-2\beta J}}$$

$$= e^{\beta J}\cosh\left(\beta\mu B\right) \pm \sqrt{e^{2\beta J}\sinh^2 \beta\mu B + e^{-2\beta J}} \qquad \left[\text{Pat4 (8) p477}\right].$$

Then the eigenvalues of $\mathbf{P}^N$ are $\lambda_1{}^N$ and $\lambda_2{}^N$. Hence:

$$Z_N(\beta) = \lambda_1{}^N + \lambda_2{}^N .$$

λ_1 must be greater than λ_2, so in the thermodynamic limit (large N), λ_1 overwhelmingly dominates the partition function. As an example, even if λ_2 is 99% of λ_1, it only takes $N > 459$, for λ_1 to contribute < 1% to the partition function. In other words, $0.99^{459} < 0.01$. For bulk materials, with $N \sim 10^{20}$, λ_1 is utterly negligible.

Now that we have Z_N, we find the free energy $F = -kT \ln Z_N$, and any other properties from standard thermodynamic relations.

Weakly Interacting Gases

Many (if not most) real gas systems involve significantly nonideal gases. The molecules interact, often becoming more attractive as the particles get closer, up to a point, and then strongly repulsive as they get even closer (Figure 10.3). This is why most gases can liquify. The thermodynamic properties of such a system depend on the details of the interaction. Computing the exact properties is almost always impossible, so we use approximations. Weakly interacting gases are defined as those that deviate a small amount from ideality.

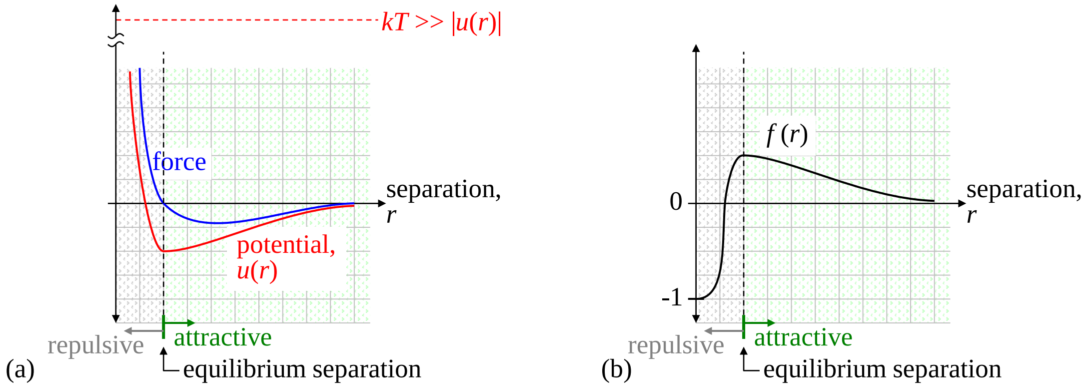


Figure 10.3 (a) Qualitative view of a typical weak interaction for gas particles. (b) Qualitative view of the Mayer f-function.

Cluster Expansion Approximation

A common approach for math problems such as the equation of state for weakly interacting gases is to make an expansion in powers of a dimensionless small parameter; such an expansion has a good chance of converging rapidly, so the approximation simply takes the first few terms of the expansion. A tractable (but nontrivial) such approximation is the **cluster expansion**. The cluster expansion can then be used to compute the virial expansion, which is an expansion of the gas equation of state in powers of its (small) particle density.

Recall that all the thermodynamic properties of a system can be found from its partition function, which gives the relevant free-energy, and derivatives of the free-energy give all the other properties. E.g.:

$$A(T,V,N,...)=-kT\ln Z(T,V,N,...), \qquad \text{and}$$

$$\text{partial derivatives of } \mathrm{A}(\,) \to \text{all thermodynamic properties}.$$

Therefore, we start by finding the partition function. Before we include interactions, recall that a partition function for a gas starts with the translational kinetic energy of the particles, and could also include the internal energies (excited molecular states, rotations, vibrations), or even interactions with external fields. These "extra" energies are not relevant to a cluster expansion, so we illustrate with the simplest system we can (just translational kinetic energy, e.g., monatomic gas).

The cluster approximation uses just 2-particle interactions, ignoring effects such as distortion when 3 or more particles get close. It further approximates that the interactions depend only on the 2-particle *positions*, ignoring effects such as dipole orientation dependence from polar molecules like steam. Finally, it approximates that the interaction potential depends only on the particle separation, i.e. the interactions are "central" (see Figure 10.3):

$$u_{ij}(\mathbf{r}_i,\mathbf{r}_j)\equiv u(r_{ij}), \qquad r_{ij}\equiv\left|\mathbf{r}_i-\mathbf{r}_j\right|. \tag{10.7}$$

The system energy is then:

$$U\approx\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{2m}+\sum_{i<j=2}^{N}u_{ij}(\mathbf{r}_i,\mathbf{r}_j) \qquad \textit{where} \qquad \sum_{i<j=2}^{N}\equiv\sum_{j=2}^{N}\sum_{i=1}^{j-1}.$$

The interaction summation is over all *unique* pairs, *i*, *j*. For weakly interacting gases (Figure 10.3), for all $r >$ equilibrium distance, the $|u(r)| << kT$ (which is also roughly the typical particle kinetic energy).

Using the classical continuum approximation, the system partition function separates into a product of integrals:

$$Z_N(T)=\sum_{\substack{s=\text{micro-}\\\text{states}}}e^{-U_s/kT} \quad\to$$

$$\begin{aligned}Z_N(T)&=\frac{1}{N!}\frac{1}{h^{3N}}\int_\infty d^3p_1\ldots d^3p_N\;d^3r_1\ldots d^3r_N\,\exp\left(\sum_{i=1}^{N}-\frac{\mathbf{p}_i^2}{2mkT}+\sum_{i<j=2}^{N}-\frac{u(r_{ij})}{kT}\right)\\&=\underbrace{\frac{1}{N!}\frac{V^N}{h^{3N}}\int_\infty d^3p_1\ldots d^3p_N\,\exp\left(\sum_{i=1}^{N}-\frac{\mathbf{p}_i^2}{2mkT}\right)}_{\text{ideal gas}}\;\underbrace{\frac{1}{V^N}\int_\infty d^3r_1\ldots d^3r_N\,\exp\left(\sum_{i<j=2}^{N}-\frac{u(r_{ij})}{kT}\right)}_{\text{interactions}\equiv Z_c}\\&\equiv Z_{ideal}\cdot Z_c\,.\end{aligned} \tag{10.8}$$

Z_c is called the **configuration integral**, because it depends only on the system configuration (positions of the particles). We normalize it with V^N, so it is dimensionless. With no interactions, $Z_c = 1$, and the partition function is simply that of an ideal gas.

The momentum integral is the same as for an ideal gas, and is simple to evaluate: we split it into 3*N* independent 1-dimensional integrals, over *x*, *y*, *z*, for all the *N* particles:

$$Z_{ideal} = \frac{1}{N!}\frac{V^N}{h^{3N}}\int_{\infty} d^3p_1 \ldots d^3p_N \, \exp\left(\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{2mkT}\right)$$

$$= \frac{1}{N!}\frac{V^N}{h^{3N}}\int_{-\infty}^{\infty} dp_{1x}dp_{1y}dp_{1z} \ldots dp_{Nx}dp_{Ny}dp_{Nz} \, \exp\left(\sum_{d=1}^{3N}\frac{p_d^2}{2mkT}\right) \qquad \left(d \equiv \text{degree of freedom}\right)$$

$$= \frac{1}{N!}\frac{V^N}{h^{3N}}\left[\int_{-\infty}^{\infty} dp \, \exp\left(\frac{p^2}{2mkT}\right)\right]^{3N} = \frac{1}{N!}\frac{V^N}{h^{3N}}\left[\sqrt{2\pi mkT}\right]^{3N} \qquad \left(\text{use} \int_{-\infty}^{\infty} e^{-ax^2} = \sqrt{\frac{\pi}{a}}\right).$$

As always, the partition function is dimensionless.

So far, we've made the three approximation described above eq. (10.7). We now consider weak interactions, defined as $|u(r)| << kT$. In (10.8), we defined the configuration integral Z_c:

$$Z_c \equiv \frac{1}{V^N}\int_{Vol} d^3r_1 \ldots d^3r_N \, \exp\left(\sum_{i<j=2}^{N} u(r_{ij})/kT\right) = \frac{1}{V^N}\int_{Vol} d^3r_1 \ldots d^3r_N \prod_{i<j=2}^{N} \exp\left(-u(r_{ij})/kT\right).$$

For small $|u(r)| << kT$, the exponentials are *near* 1, so it is convenient to expand them in terms of small deviations from 1. So we define these deviations as the functions $f_{ij}(\mathbf{r}_i, \mathbf{r}_j)$:

$$f_{ij} \equiv f_{ij}(\mathbf{r}_i, \mathbf{r}_j) \equiv \exp\left(-u(r_{ij})/kT\right) - 1 \qquad \text{or} \qquad \exp(-u(r_{ij})/kT) \equiv 1 + f_{ij}.$$

Formally, each function f_{ij} is a function of $\mathbf{r}_i$ and $\mathbf{r}_j$. Figure 10.3b qualitatively illustrates a typical Mayer f-function. Then:

$$Z_c = \frac{1}{V^N}\int_{Vol} d^3r_1 \ldots d^3r_N \prod_{i<j=2}^{N}\left(1 + f_{ij}\right) = \left(1 + f_{12}\right)\left(1 + f_{13}\right) \ldots \left(1 + f_{N-1,N}\right). \qquad (10.9)$$

This form of Z_c is called the **cluster expansion**. Evaluating this product is difficult. If $f_{ij} << 1$, then we can approximate it, but even then, the approximation is complicated. To help compute this approximation, a graphical tool is used (similar to Feynman diagrams).

To recap all the approximations we've made in Z_c:

- The interactions are pair-wise (involve only two particles at a time).
- The interactions depend only on particle *positions* (not orientations, or speeds, etc.).
- The interactions are central (depend only on particle separation).
- The interactions are weak: $|u(r)| << kT$.

The configuration integral (10.9) has $N(N-1)/2$ binomial factors, because there are $N(N-1)/2$ unique pairs of particles. If we multiply it out into a sum of terms, we can sort the terms by orders of the f_{ij}:

$$Z_c = \frac{1}{V^N}\int_{Vol} d^3r_1 \ldots d^3r_N \left[1 + \sum_{i<j=2}^{N} f_{ij} + \sum_{i<j=2}^{N}\sum_{\substack{k<l=2 \\ k,l\neq i,j}}^{N} f_{ij}f_{kl} + \sum_{i<j=2}^{N}\sum_{\substack{k<l=2 \\ k,l\neq i,j}}^{N}\sum_{\substack{m<n=2 \\ m,n\neq i,j \\ \text{or } k,l}}^{N} f_{ij}f_{kl}f_{mn} + \ldots\right].$$

There are now $N(N-1)/2$ terms in the square brackets. Importantly, each term is *unique* because each of the f_{ij} in (10.9) appears only once. Since the f_{ij} are << 1, we might hope that the first-order terms would be smaller than the second-order, and so on. It turns out to be not that simple, as we now explore. We will find that it is more convenient to group the terms differently than as above.

Consider the first term in Z_c above. For each particle position $\mathbf{r}_i$, the integral over the container volume is just V:

$$Z_c^{(1)} = \frac{1}{V^N} \int_{Vol} d^3 r_1 \ldots d^3 r_N = \frac{1}{V^N} V^N = 1 .$$

The second term in Z_c above is:

$$Z_c^{(2)} = \frac{1}{V^N} \sum_{i<j=2}^{N} \int_{Vol} d^3 r_1 \ldots d^3 r_N \; f_{ij} .$$

For a given (i,j) pair, we have an integral that includes the N–2 positions other than $\mathbf{r}_i$ and $\mathbf{r}_j$, each of which integrates to just the container volume V. This simplifies to:

$$Z_c^{(2)} = \frac{V^{N-2}}{V^N} \sum_{i<j=2}^{N} \int_{Vol} d^3 r_i \int_{r_j \to \infty} d^3 r_j \; f_{ij} = \frac{1}{V^2} \sum_{i<j=2}^{N} \int_{Vol} d^3 r_i \int_{r_j \to \infty} d^3 r_j \; f_{ij} .$$

Since the f_{ij} fall off rapidly with distance, then in the surviving 2-particle integral above, we can replace the limits on $\mathbf{r}_j$ with infinity, which makes the integral easier to evaluate. This is valid in the thermodynamic limit where surface effects of the container are negligible. Because the limits are now at infinity, we can change variables on the inner integral to $\mathbf{r}'_j = \mathbf{r}_j - \mathbf{r}_i$, to eliminate dependence on $\mathbf{r}_i$ (this is equivalent to taking the current value of $\mathbf{r}_i$ as the origin of coordinates for our new integration variable $\mathbf{r}'_j$). This means the integral over $\mathbf{r}_i$ produces a simple factor of V:

$$Z_c^{(2)} = \frac{1}{V} \sum_{i<j=2}^{N} \int_{r'_j \to \infty} d^3 r'_j \; f_{ij}(\mathbf{r}', \mathbf{0}_\mathbf{v}) .$$

In our cluster diagrams, we will use this device to eliminate one position variable from each term in Z_c.

More TBS.

The Virial Expansion

When written per particle, the ideal gas equation of state is:

$$pv = kT \qquad \text{where} \quad v \equiv V/N \text{ is the specific volume} .$$

For nonideal gases, i.e. interacting particles, we can include some corrections to the equation of state as an expansion in a small parameter:

$$\frac{pv}{kT} = \underset{a_1}{1} + \text{corrections} \equiv \sum_{l=1}^{\infty} a_l(T) \left(\lambda^3 / v \right)^{l-1} \qquad \left[\text{Pat310.2.1 p307b}\right]$$

$$\text{where} \quad a_1(T) \equiv 1; \quad \lambda \equiv \frac{h}{\sqrt{2\pi m k T}} \text{ is the so-called "thermal wavelength".}$$

This is an expansion in the dimensionless small parameter (λ^3/v) = ($\lambda^3\rho$). At low density, $\rho = 1/v$ is small, and the equation of state reduces to the ideal one. As the density increases, so do the corrections to ideality. Note that $a_1 \equiv 1$, and the *second* virial coefficient goes with the *first* power of the expansion parameter (λ^3/v). Also, the virial expansion explicitly includes a temperature dependence for the corrections.

More TBS.

11 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), including chemical examples where particles (molecules) are not conserved.
6. Endothermic reactions: $N_2O \rightarrow N_2 + \frac{1}{2}O_2 \;\Rightarrow\; S \rightarrow 1.5S$
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 2π steradians.

$$Force = \frac{\Delta p}{\Delta t} = \frac{2p\cos\theta}{(2L/c\cos\theta} = \frac{pc}{L}\cos^2\theta$$

$$\left\langle \cos^2\theta \right\rangle = 1/3 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\pi/2}\cos^2\theta\sin\theta\, d\theta\, d\phi \qquad \Rightarrow \qquad p = \frac{pc}{3V} \quad \Rightarrow \qquad pV = \frac{1}{3}E$$

11. Fermi level of non-interacting ultra-relativistic electrons:

$$N = \underbrace{2}_{\substack{\text{spin}\\\text{factor}}}\int_0^{k_f}\frac{V}{(2\pi)^3}d^3k = \frac{Vk_f{}^3}{3\pi^2} \qquad \Rightarrow \qquad k_f = \left(3\pi^2\frac{N}{V}\right)^{1/3}$$

$$E_0 = 2V\int\frac{\hbar ck}{(2\pi)^3}d^3k = \underbrace{\frac{3}{4}\hbar ck_f}_{\langle E\rangle}N$$

12. $d^3k \rightarrow 4\pi k^2 dk$
13. Gaussian integral approximation (saddle point method)
14. $S = -k\,\mathrm{sum}(p_i \ln p_i)$ [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.

[Goods] Goodstein, David L., *States of Matter*, Dover Publications, Inc., New York, 1985 (republished 2002).

[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
arg max	$\underset{x}{\arg\max}\left[f(x)\right]$ is the value of x that maximizes $f(x)$.
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."
chemical potential	an unfortunate name for the free-energy of a particle. It need not have anything to do with chemistry.
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).
configuration	a specification of the *positions* of all the degrees of motion, but *not* the momenta. The combination of configuration + momentum-state = phase-space-state. phase-space-state + spin-state (if any) = microstate.

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	ambiguous: may be a single-particle state, microstate, or macrostate.
energy eigenstate	ambiguous: may be a single-particle state, or a microstate.
ensemble	a real or imaginary set of “systems.” Typically, the set is very large.
entropy	the multiplicity of a macrostate measured on a logarithmic scale. Entropy is a function of the macrostate of a system which helps compute thermal equilibrium parameters. Intuitively, entropy measures *uniformity* throughout a system (not disorder or chaos).
extensive	(briefly) a property that is proportional to system size (e.g., mass).
factor	an element of an expression that is multiplied or divided: in “a/b”, a and b are factors.
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.
gases	plural of the noun “gas”.
gasses	third person singular form of the verb “to gas”.
gazillion	technical term for “a very large number”.
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
macrostate	An observable state of a system, which could conceivably be measured in a lab. The macrostate includes observables such as temperature, pressure, volume, magnetization, etc. See also microstate.
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).
microstate	A single state of a system, including the states of all of the system’s constituent particles. Classically, a microstate includes each particle’s position and momentum. Quantum mechanically, a microstate is an N-particle energy eigenstate, and includes any spin-states.

	It is a quantum-state that describes the state (position-momentum-spin) of every particle in the system. In either case, a microstate usually cannot be observed or determined in a laboratory. However, a knowledge of microstates from physical theories is essential to applying Statistical Mechanics to predict observable phenomena. phase-space-state + spin-state (if any) = microstate.
momentum-state	a specification of the *momenta* of all the degrees of motion, but *not* the positions. The combination of configuration + momentum-state = phase-space-state. phase-space-state + spin-state (if any) = microstate.
multiplicity	the number of microstates consistent with a given macrostate
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: $\Pr\left(x < X < x + dx\right) = pdf_X\left(x\right) dx$.
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).
phase-space state	a specification of the *positions* of all the degrees of motion, *and* the momenta. The combination of configuration + momentum-state = phase-space-state. phase-space-state + spin-state (if any) = microstate.
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
spin-state	the spin part of a quantum state, a vector of dimension 2s + 1.
term	an element of an expression that is added or subtracted: in “$a + b$”, a and b are terms.
usually	for macroscopic systems: always. In SM, saying a macroscopic system is “usually” in the most-likely macrostate pretty much means it is “always” in that state.
wave-function	the *spatial* part of a quantum state.

Formulas

Generalized factorial: $a! \equiv \int_0^\infty x^a e^{-x}\, dx, \qquad a \text{ real}$

Gamma function: $\Gamma(a) \equiv \left(a-1\right)! = \int_0^\infty x^{a-1} e^{-x}\, dx, \qquad a \text{ real}$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots, \quad |x| < 1 \quad or \qquad \ln x = \left(x-1\right) - \frac{\left(x-1\right)^2}{2} + \frac{\left(x-1\right)^3}{3} - \cdots, \quad 0 < x \le 2$$

Stirling’s approximation for factorials (or Γ( ) functions):

For large a: $\ln a! \approx a\left(\ln a - 1\right)$ or $a! \approx \left(a/e\right)^a$

NB: Huang's notation is different than ours:

$$S = k \ln \Gamma(E, \Delta E) \qquad \Gamma(E, \Delta E) \equiv \text{phase space volume for } \; E < E_{system} < E + \Delta E \qquad \text{(Hua 6.27 p134)}$$

$$S = k \ln \Omega(E) \qquad \Omega(E) \equiv \text{phase-space density of states} \qquad \text{(Hua 6.28 p134)}$$

$$S = k \ln \Sigma(E) \qquad \Sigma(E) \equiv \text{phase space volume of } \; E_{system} < E \qquad \text{(Hua 6.28 p134)}$$

$$S = -\frac{\partial A}{\partial T} \qquad \textit{where} \quad A \text{ is any free energy}$$

$$S = k\left(\ln Z + \beta \bar{E}\right) = k\left(\ln \sum_{j}^{\substack{\text{energy}\\ \text{levels}}} \Omega(E_j) e^{-\beta E_j} + \beta \bar{E}\right) \approx k\left(\ln\left(\Omega(\bar{E}) e^{-\beta \bar{E}}\right) + \beta \bar{E}\right) = k\left(\ln \Omega(\bar{E}) - \beta \bar{E} + \beta \bar{E}\right)$$

$$= k \ln \Omega(\bar{E}), \qquad \bar{E} \equiv \langle E \rangle$$

Entropy of monatomic ideal gas (Sackur-Tetrode equation (3.7)):

$$S(u, N, v) = Nk\left\{\ln\left[v \frac{1}{h^3}\left(u \frac{4\pi m}{3}\right)^{3/2}\right] + \frac{5}{2}\right\}, \qquad v \equiv \frac{V}{N}, \quad u \equiv \frac{E}{N}.$$

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: $\mu \equiv \dfrac{\partial A}{\partial N}$ *where* $A \equiv$ appropriate free energy for the given system

V	A	T
U	↗ ↖	G
S	H	P

Fix notation here:??

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

$$= \frac{1}{N!} \int_{E_0}^{\infty} dE \, \Omega(E) e^{-\beta E}$$

$$Z(\beta, V, N) \equiv \sum_{\{n_j\}} \exp\left[-\beta \sum_{j=1}^{\substack{\text{energy}\\ \text{levels}}} \varepsilon_j n_j\right], \qquad \textit{where} \quad \sum_{j}^{\substack{\text{energy}\\ \text{levels}}} n_j = N \qquad \text{(quantum canonical ensemble)}$$

Ideal gas: $Z(\beta) = \left(2\pi m / \beta\right)^{3N/2} V^N$

$$\mathcal{Z}(\beta,V,\mu)=\sum_{N=0}^{\infty}e^{\beta\mu N}Z(\beta,V,N)=\sum_{N=0}^{\infty}\sum_{\{n_j\}}\exp\left[-\beta n_j\left(\varepsilon_j-\mu\right)\right] \qquad \text{(grand canonical ensemble)}$$

but $\mu=\mu(T,V,N)$ may vary within the summation

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

$\lambda_T=\sqrt{\dfrac{h^2}{2\pi mVT}}$ Ideal gas density of states, 3D: $\rho(E)=\dfrac{V}{h^3}2m^{3/2}E^{1/2}$

$E=-\dfrac{\partial}{\partial\beta}\ln Z$ $\mathrm{var}(E)=\dfrac{\partial^2}{\partial\beta^2}\ln Z=kT^2C_V$ Fermi pressure (low T): $PV=\dfrac{2}{5}\varepsilon_F N$

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.