

Gambling Theory "Roots" of the 0th Law

Discoveries in the 1940's and 1950's uncovered surprising connections between: (i) the foundations of thermal physics, (ii) the behavior of information stored in codes, and (iii) the mechanisms by which life on earth puts codes to work. These connections, along with discovery of the transistor, set the stage for today's information age, and for the technological and economic possibilities that genetic biotechnology, and molecular nanotechnology, may offer (for better and worse) in the decades ahead. These interdisciplinary developments have at once simplified, and broadened the scope of, thermal physics.

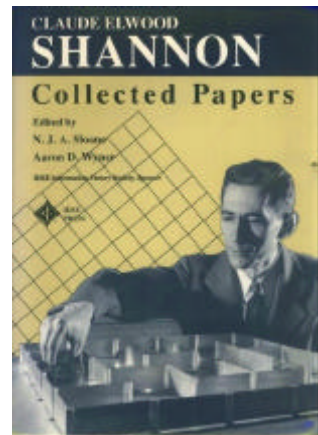
Although "information theory" insights have profoundly changed graduate and senior thermal physics texts, they have only begun to modernize the 19th century perspective of introductory texts. We correct for this with some optional material, designed to strengthen the physical insight of adventurous teachers and students concerning the connection between thermal physics, information, and replicable codes.

The story begins with the way that statistical inference (the science of taking the best guess) connects uncertainty and temperature. Basic thermal physics equations (like *the ideal gas law* and *equipartition*) are "pulled from the blue" in traditional courses. We here show how they follow from Newton's laws plus models of the way our uncertainty about the state of a system (like a bottle of gas) depends on quantities you can measure (like energy and volume). Along the way find a snapshot of the strange way that non-ideal materials (like water) extract thermal kinetic energy from the potential energy wells of bound states, giving up freedom and heat at the same time. A primer on the units by which we quantify information (e.g. bits and bytes) is encountered as well.

From this point, thermodynamic "possibility" calculations from Laws 1 and 2 (a most powerful gift provided us by thermal physics) are extended beyond home heating and automotive design to include energy flow in biology. They also extend to the study of information engines like computers, video cameras, and information-processing organisms. Thus for example, limits follow on how much the data on your computer's hard disk can be compressed. Even the connection between information in codes, and the thermal physics of life, follows naturally, providing insight into the way replicable codes (in the form of both genes and ideas) continue to evolve in their symbiosis with us. Thus the approach of information theory in this century shifts the focus from the thermal physics of the past, to an exploration of its connections to our future.

Note: What follows is relatively compact, so that students and/or teachers can go lightly, or deeply, in the areas they find of particular interest. Since this approach is in early stages of its formulation, suggestions on improving it are quite welcome as well.

[Describing the Dice \(Accessible States\)](#): During work in 1948 on the problem of noise in communication lines, Bell Labs researcher Claude Shannon encountered some equations that looked suspiciously like those used in thermodynamics. This connection between “information theory” and thermodynamics lead to demonstration in the 1950's, by Edwin Jaynes and others, of a deep *distinction* between the physical and statistical parts of thermal physics. It became clear what parts of thermal physics involved “describing the dice”, and what parts involved “taking the best guess”. This insight justified and expanded the applicability of powerful “ensemble” methods developed in the late 1800's by J. Willard Gibbs, and today illuminates the physical relationship between organisms and codes of all sorts (including genetic and digital).



To begin with, Shannon and Jaynes found that the zeroth law follows from a basic assumption in *statistical inference*, the science of decision-making with limited information. This interdisciplinary field, launched by gambler physician Gerolamo Cardano's 1564 book on games of chance, continues to evolve in colorful and spirited fashion. In “gaming” as in other places where statistical inference is applied, one assumes that: *In the absence of information to the contrary, all possible states of a system* (e.g. sides up on a tossed coin, numbers showing on thrown dice, marble positions on a spun roulette wheel, cards in your opponents hand just after the deal, or accessible positions and velocities for a gas atom) *are equally likely*.

Claude E. Shannon of Bell Labs clarified the connection between entropy and uncertainty with “information theory” that he applied to optimizing communication channels. Data compression, cryptography, reliability sciences, & how we understand coding in biological systems, are a few of the areas where his work has had major impact.

cf. <http://it.ucsd.edu>, and
<http://www.research.att.com/~njas/doc/shannonbio>
and <http://www-lmmb.ncifcrf.gov/~toms/>

To see how this assumption and the zeroth law connect, consider a bottle full of gas atoms. Suppose we could count the number of ways Ω that a total energy E might be divided up among those atoms. For example, one way of arranging the kinetic energy in a bottle of gas might be to put all of that energy into the x-direction motion of one single molecule, leaving little or none for the other gas atoms in the bottle. This, of course, is not the most likely way to divide up the energy. One might instead give each molecule equal velocities in x, y and z, or equal velocities in all but one direction, etc. One might further imagine that for most systems, the more energy there is to divide up, the more ways there are to do it. That's about as much as intuition says, at first glance.



Edwin T. Jaynes, at Stanford and Washington U., discovered that thermal physics breaks neatly into two parts, namely “describing the dice” and statistical inference, with key Phys. Rev. papers in the 1950's, and then championed the application of Bayesian methods in fields far and wide.

cf. <http://bayes.wustl.edu/etj/etj.html>
and <http://omega.albany.edu:8008/JaynesBook.html>

For many systems this multiplicity function takes the form $\Omega(E) \propto E^{N/2}$ over much of their energy range, where N is the number of atoms or molecules, and v is the number of “modes” of energy storage (degrees of freedom or “ways to wiggle”) per molecule. Since N is a huge number for macroscopic systems, the addition of a tiny bit of energy *drastically* increases the number of ways that this energy can be distributed. Why the multiplicity function looks like this for many substances will be covered in the problems. But, why do we care about this multiplicity function at all? This is like asking: “Why do we care how many aces there are in the deck?” when playing poker. Such knowledge affects what cards others around us are likely to play next, and hence our actions too.

Measuring Uncertainty (Nats and Bits): Shannon noticed that the log of the multiplicity had certain advantages, since the size of the resulting quantities was smaller and more manageable. This logarithm $S = k \ln(\Omega)$ measures *uncertainty* about the state of a system in “information units” determined by the value of k . If we choose k to be $1/\ln(2)$, S is in *bits* of information. For example, if there are two equally possible answers to a test question ($\Omega = 2$), then one’s uncertainty about the answer is $S = \ln(2)/\ln(2) = 1$ bit. The number of possible “different answer sets” for an eight question true-false test is $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^8 = 256$. The uncertainty, however, is $S = k \ln(2^8) = 8 k \ln(2) = 8$ bits or one *byte*. Another way to say this is that if k is $1/(8 \ln 2)$, then S is in *bytes*. When you buy a computer monitor capable of 24-bit or 3-byte per pixel color, the number of possible colors is $\Omega = 2^{24} = 256^3 = 16777216$ or “17 million colors”. Since 17 million colors sounds much more impressive than 24 bits, 17 million is used in advertisements. On the other hand, if you buy a computer with 32 megabytes of RAM, the number of possible states for the memory in this system is $\Omega = 256^{32000000}$, a number too large for most folks to have any feelings for! If k is 1, then S is in “Shannons” or *nats*, the latter because simple natural logarithms are used.

So far, we’ve discussed the physical basis for units used to measure information or its absence (uncertainty) in mathematics and computer science. It even applies (Fig. 2) to your knowledge on tests! In thermal physics, k is usually set to Boltzmann’s constant $k_B \equiv 1.38 \times 10^{-23}$ J/K. In this case, S takes on the historical units for *entropy* in thermodynamics, i.e. joules per degree Kelvin (J/K). Since J/K is a measure of uncertainty, like bits and bytes, this suggests that temperature has physical units of energy per unit information. That is because reciprocal temperature ($1/T$) is precisely *the rate of uncertainty change per unit of added thermal energy*. To see why, we first ask why such rates are important.

Example Problem 1: In throwing a pair of ordinary 6-sided dice, what are the odds (i.e. the multiplicity ratio) against throwing a seven?

Answer: As you can see from Fig. 1, 6 of the 36 possible tosses show seven spots. Hence the “odds against” are (36-6) to 6, or 5 to 1. By comparison, the odds against snake eyes (two ones showing) are 35 to 1.

Example Problem 2: According to the label on its box, a backup tape can store 4 megabytes of data uncompressed and “8 megabytes of data compressed”. Which of these two numbers is a property solely of the tape, and makes no assumptions about your data? Hence what is the multiplicity of possible “records” that the tape is really able to hold?

Answer: While also providing insights underlying all modern compression methods, Shannon showed that compressibility depends on *properties of your data*, like the number of repeated characters! Claims that your data can be compressed by a factor of two are speculative at best, and false advertising at worst. Hence your tape is really capable of assuming only $\Omega = e^{S/k} = e^{4000000/0.18034} = \text{Exp}[2.218 \times 10^7]$ states. Still, this number is huge! *Bonus:* What is x , if you write this huge number in the form $\Omega = 256^x$?

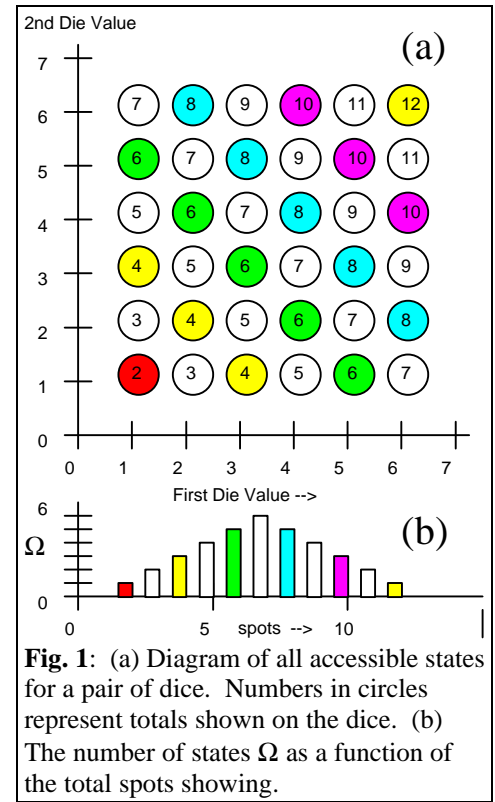


Fig. 1: (a) Diagram of all accessible states for a pair of dice. Numbers in circles represent totals shown on the dice. (b) The number of states Ω as a function of the total spots showing.

Note: Conversions between units: The use of various values for the constant k in choosing information units has the consequence that 1 [Kelvin] = 1.3807×10^{-23} [joule/nat]; and 1 [nat] = 1.4427 [bit] = 0.18034 [byte] = 1.3807×10^{-23} [joule/K]. Thus Boltzmann’s constant converts between historical and natural units for both absolute temperature, and uncertainty!

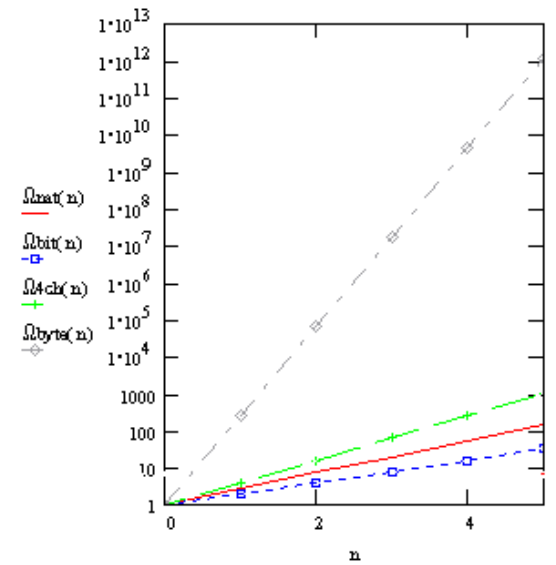


Fig. 2 - If you have n questions each with C choices per question, the total number of ways to complete the test is $\Omega = C^n$. We graph this as function of n , to illustrate multiplicity versus test size for 1-nat, 1-bit (true-false), 2-bit (4 choice), and 1-byte question tests.

Finding the Best Guess (Equilibration and Law Zero): I imagine two systems A and B. Each has an accessible states function and hence an uncertainty S with known dependence on a shared but otherwise conserved quantity X . How much of the total amount of X available for sharing, call it X_{tot} , might we expect to end up in system A if we allow the two systems to exchange X between them randomly, long enough for the initial distribution of X to become irrelevant (i.e. until they “equilibrate”)? Gambling theory tells us (since all possible ways are equally likely) that the best bet is that state of sharing which can happen in the most ways, i.e. for which $S_{tot} = S_A + S_B$ is a maximum. Since at the maximum the first derivative is zero, this means that $0 = \frac{dS_{tot}}{dX_A} = \frac{dS_A}{dX_A} + \frac{dS_B}{dX_A} = \frac{dS_A}{dX_A} - \frac{dS_B}{dX_B}$, or $\frac{dS_A}{dX_A} = \frac{dS_B}{dX_B}$. This follows when S_{tot} is at its peak, since transferring some X from A to B increases the uncertainty of B as much as it decreases the uncertainty of A.

For example, the systems described in Fig. 3 reach a maximum in total uncertainty when individual system uncertainty slopes are equal. Therefore, in the absence of information to the contrary, systems allowed to share X randomly will likely redistribute X until their “uncertainty slopes with respect to change in X ”, i.e. dS/dX , are the same. This is the statistical form of the zeroth law of thermodynamics. Note that it works for any conserved variable, not just for energy!

Energy, Temperature, & Equipartition: When energy is the quantity shared randomly between systems, dS/dE is the system’s coldness or reciprocal temperature $1/T$. Thus temperature, the reciprocal of an uncertainty slope with respect to energy, is useful because systems in thermal contact randomly transfer energy until a common T is reached (cf. Fig. 4). Similarly, if two objects have much different temperatures, we expect lots of energy to change places on thermal contact. That’s why most children early on determine, by trial and error, what “hot” means!

Let’s see quantitatively what this means for a system with multiplicity $\Omega(E) \approx (E/E_0)^{vN/2}$, where E_0 is some reference energy independent of E . Uncertainty is $S(E) = vN/2 k [\ln(E) - \ln(E_0)]$, so that $dS/dE = vNk/(2E)$. Solving this for E/N , and substituting $1/T$ for dS/dE , gives $E/N = (v/2) kT$. This “principle of equipartition” says that the average energy for each “degree of freedom” (i.e. way to store thermal energy e.g. in x-direction kinetic energy or y-directional potential energy) is $1/2 kT$. The quantity kT (temperature measured in the physical units energy per nat of uncertainty increase) hence tells the “size” of typical random chunks of thermal energy (and hence the “color of heat”) at a specified temperature. For example, $kT = k_B \times 293K \approx 1/40$ eV/nat at room temperature (so that radiant heat photons have energies in the infrared range), while at the surface temperature of the sun, kT is around 2 eV/nat (putting most solar photons in the visible energy range).

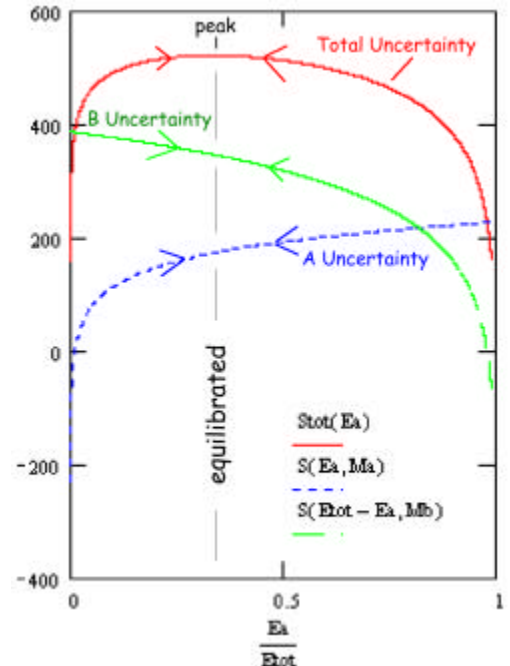


Fig. 3 – Uncertainties for two systems A: blue, $\Omega_a=(E/\epsilon_0)^{99/2}$, B: green, $\Omega_b=(E/\epsilon_0)^{99}$, and the sum of the two (red), as a function of energy in system A as they share a total energy of $E_{tot}=10000\epsilon_0$. Note that the sum uncertainty peaks for E_a around $E_{tot}/3$, making this the most likely “equilibrium” value of E_a . This is also where the slopes of the blue and green plots (dS/dE_a) become equal and opposite.

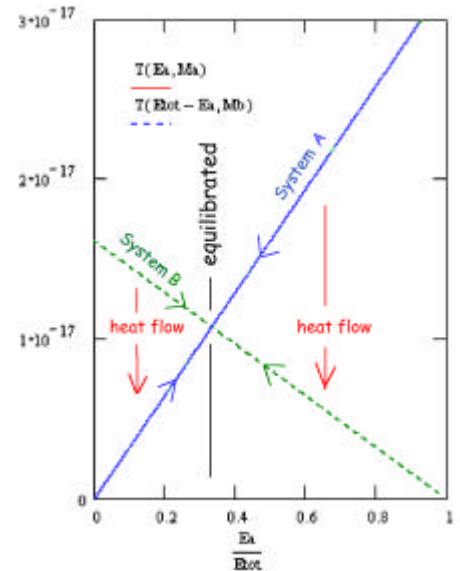


Fig. 4 –Temperatures for the systems of Fig. 3, obtained from the “equipartition relation” given in the text. Note that the temperatures of the two systems approach a common value (as do the uncertainty slopes) when $E_a = E_{tot}/3$.

You can see from its derivation here that this principle only works if a system's number of accessible states obeys $\Omega(E) \propto E^{vN/2}$. By the way, the $\frac{1}{2}$ in the exponent of many systems arises (as shown in the problems) because, for example, kinetic energy is proportional to velocity *squared*. Thus systems to which this equipartition rule applies are called "quadratic". **Note:** We measure energy in units of *electron volts* (1 [eV] $\equiv 1.6 \times 10^{-19}$ [joules]) when units suited to single atoms are appropriate.

Volume, Pressure, & The Ideal Gas Equation of State: For ideal gases, it is also easy to determine the dependence of multiplicity on volume. Since molecules in an ideal gas are so widely separated that they hardly interact, the number of accessible positions for each molecule is simply proportional to volume V . For an N -molecule gas, the number of position states is then proportional to V^N . Thus for an ideal gas, we can write $\Omega(E,V) \propto V^N E^{vN/2}$, or $S(E,V) = Nk \ln(V) + vN/2 k \ln(E) + k \ln(C)$ where C is independent of V and E . The volume uncertainty-slope dS/dV is called the *free expansion coefficient*, and is at equilibrium in general equal to $dS/dV = dS/dE dE/dV = (1/T) (FdX)/(AdX) = (1/T)(F/A) = P/T$. For an ideal gas from the equation for S above, this means that $P/T = dS/dV = Nk d(\ln V)/dV = Nk/V$. Hence the ideal gas law, $PV = NkT$, arises from knowledge of a system's multiplicity function as well.

Quick Check:

Q. Wait a minute! How can we treat particles as if they were dice with only a *finite* number of allowed states?

A. We actually had to wait until the discovery of quantum mechanics and "Planck's constant" before we could count the allowed states of, for example, gas particles in a bottle. To apply gambling theory to the gas, however, and to do most of what we do in this book, we (like J. Willard Gibbs well before the discovery of quantum mechanics) only need to know what the number of allowed values is proportional to, and not its quantitative value. That's because when we take derivatives of the uncertainty $S = k \ln W$, constant factors in the multiplicity (that is, proportionality constants) drop out.

Example Problem 3: How many states are accessible when "tossing" three coins, and how much uncertainty does one have about the outcome before they are tossed?

Answer: The possibilities are TTT, TTH, THT, THH, HTT, HTH, HHT, and HHH, where H means "heads" and T means "tails". Hence $\Omega = 2 \times 2 \times 2 = 2^3 = 8$. In other words, the number of accessible states for individual coins are *multiplied* to get the total multiplicity. The uncertainty is $S = k \ln \Omega = 2.079$ [nats] if we use $k=1$. If we use $k=1/\ln(2)$, then $S = k \ln(2^3) = 3 k \ln(2) = 3 \ln(2)/\ln(2) = 3$ [bits]. Since the uncertainty for one coin is $S_1 = k \ln(2) = 1$ [bit], for N coins the uncertainty is $S = k \ln(2^N) = N k \ln(2) = N S_1 = N$ [bits]. Hence the uncertainties for individual coins *add* to get the total uncertainty.

Example Problem 4: I imagine a gas of $N_g = 33$ atoms in thermal contact with a nano-crystal of $N_p = 33$ atoms, both contained in a two-liter thermos whose walls transfer no heat in or out. Suppose the multiplicity for both systems is of the form $\Omega = (E/\epsilon)^{vN/2}$. For the gas, $\epsilon = \epsilon_g = 10^{-40}$ [joules], and each gas atom has $v = v_g = 3$ "ways to store energy" or degrees freedom. For the crystal, $\epsilon = \epsilon_c = 10^{-21}$ [joules], and each atom has $v = v_c = 6$ degrees freedom. If $E_{\text{tot}} = 10^{-18}$ [joules] or 6.25 [eV] of energy is available for sharing thermally between gas and crystal, what will the temperature, energy, and state uncertainty of each be after equilibration?

Answer: Uncertainty slopes and temperatures, as shown in the text plus Figs. E & F, will reach a common value after equilibration. If $\Omega = (E/\epsilon)^{vN/2}$, then $S \equiv k \ln \Omega = vNk/2 \ln(E/\epsilon)$, and therefore $1/T \equiv dS/dE = vNk/(2E)$. If the slopes reach a common value, then $v_g N_g k / (2E_g) = v_c N_c k / (2E_c)$ *after equilibration*. We also know that $E_{\text{tot}} = E_g + E_c$. Replacing E_c with $E_{\text{tot}} - E_g$ in the equilibration equation above allows us to solve for E_g , getting $E_g = E_{\text{tot}} v_g N_g / (v_g N_g + v_c N_p) = 10^{-18} \times 3 \times 33 / (3 \times 33 + 6 \times 33) = 3.333 \times 10^{-19}$ [joules] or $3.333 \times 10^{-19} / 1.6 \times 10^{-19} = 2.083$ [eV]. Then E_c is simply $E_{\text{tot}} - E_g = 4.167$ [eV]. The final temperature (from the expression for $1/T$) is $kT = 2E_g / (v_g N_g) = 2 \times 2.083 / (3 \times 33) = 0.042$ [eV/nat]. You can check this final temperature by calculating kT for the crystal, to see if it is the same. For historical temperature units, divide this by Boltzmann's constant (8.63×10^{-5} [eV/nat K]) to get $T = 487$ K (or 214 C). *That's hot!* Finally, how do final uncertainties compare? For the crystal, $S_c = v_c N_c k / 2 \ln(E_c/\epsilon_c) = 644$ nats. For the gas, $S_g = v_g N_g k / 2 \ln(E_g/\epsilon_g) = 2453$ nats. Note that the state uncertainty (or entropy) of the gas is much larger, although the temperature of gas and crystal are the same. Perhaps this is because gas atoms are free to roam about!

Non-ideal substances, and Changes of Phase: Many systems, over part of their temperature range, obey the equipartition relation between energy and temperature derived above for “quadratic” accessible states. Changes in energy and temperature are described by a constant “no-work” heat capacity $C_v = dE/dT = \nu Nk/2$. This behavior occurs when adding heat increases the size of a fixed number of atomic velocities or displacements. However, energy versus temperature for real substances seldom obeys equipartition over a wide range of temperatures. This is illustrated with the graph of energy versus temperature for water, along with related plots of uncertainty and its derivatives in Fig. 5. Although many details of this plot may be worth a look, we simply note here that at two specific temperatures (the boiling and freezing points of water), removal of energy from H_2O results in no temperature change at all! Moreover, the energy extracted from steam on condensation exceeds the “kinetic energy” of the molecules in the gas just before!

Condensation is not magic. In fact, water molecules simply fall into each other’s potential wells (much like planetoids sweeping each other up during the early days of formation of our solar system). Thermal energy is given up by this “accretion process” (that’s one reason why the surfaces of early solar system planets were hot), as water molecules allow themselves to become trapped with a set of molecules in the liquid phase. To liberate them once more (i.e. to evaporate the water), of course, we must add heat. That, for example, is how we cool off by sweating: heat we want to get rid of is used to buy back the freedom of bound water molecules that we place on the surface of our skin.

Possibilities for Laws 1 and 2: The relationships above arise when we apply gambling theory to the fact that quantities like energy and volume are sometimes “randomly shared” but otherwise conserved. Hence temperature and uncertainty prove to be useful concepts whenever we apply Newton’s laws to systems about which the details are unknown. Random sharing of energy, in particular, is called “heat flow”. By comparison

“work” is done when energy is transferred in a controlled manner (e.g. by a force F acting through distance dX , or by a pressure P causing a volume change dV). If we define δQ as the heat flow *into a system* and δW as the work that system does *on its environment*, then conservation of energy gives us the standard form of the first law of thermodynamics: Total energy change $dE = \delta Q - \delta W$.

The second law of thermodynamics is also grounded in common sense: I imagine a system that you put together in a prescribed manner, so that you know many facts about its state. If you proceed to isolate that system from the rest of the world (i.e. make it totally inaccessible), what is likely to happen to your uncertainty about the details of that system from year to year, or century to century? Common sense (mine at least) suggests that knowledge outside about what’s inside that system will either not change, or will decrease because of either internal change or forgetfulness outside. It is unlikely, however, that outside knowledge about the details inside of an isolated system will increase. Mathematically, we might say that for an isolated system, uncertainty can only increase irreversibly with time. We write this as $\delta S_{irr}/\delta t \geq 0$.

This law provides insight into the nature of heat. Consider a system that is not isolated, but is allowed to share energy and volume with the world around. We can write $dS = \partial S/\partial E|_{V,N} dE + \partial S/\partial V|_{E,N} dV + \delta S_{irr}$. Since $\partial S/\partial E|_{V,N} = 1/T$ and $\partial S/\partial V|_{E,N} = P/T$ from before, we can rearrange this (with help from the first law) to get: $dS = \delta Q/T + \delta S_{irr}$. Thus heat flows are energy transfers which, unlike work, act to increase our total uncertainty about the state a system is in.

These two laws together offer a powerful strategy for assessing what is possible, and what is not. To apply it, imagine a “steady-state” engine (something that may go through cycles but to first order keeps returning to exactly the state in which it began). Since the engines themselves do not change, our equations can treat the world outside the engines as a self-contained universe to which the first and second laws apply, i.e. for which $dE = 0$ and $\delta S_{irr} > 0$.

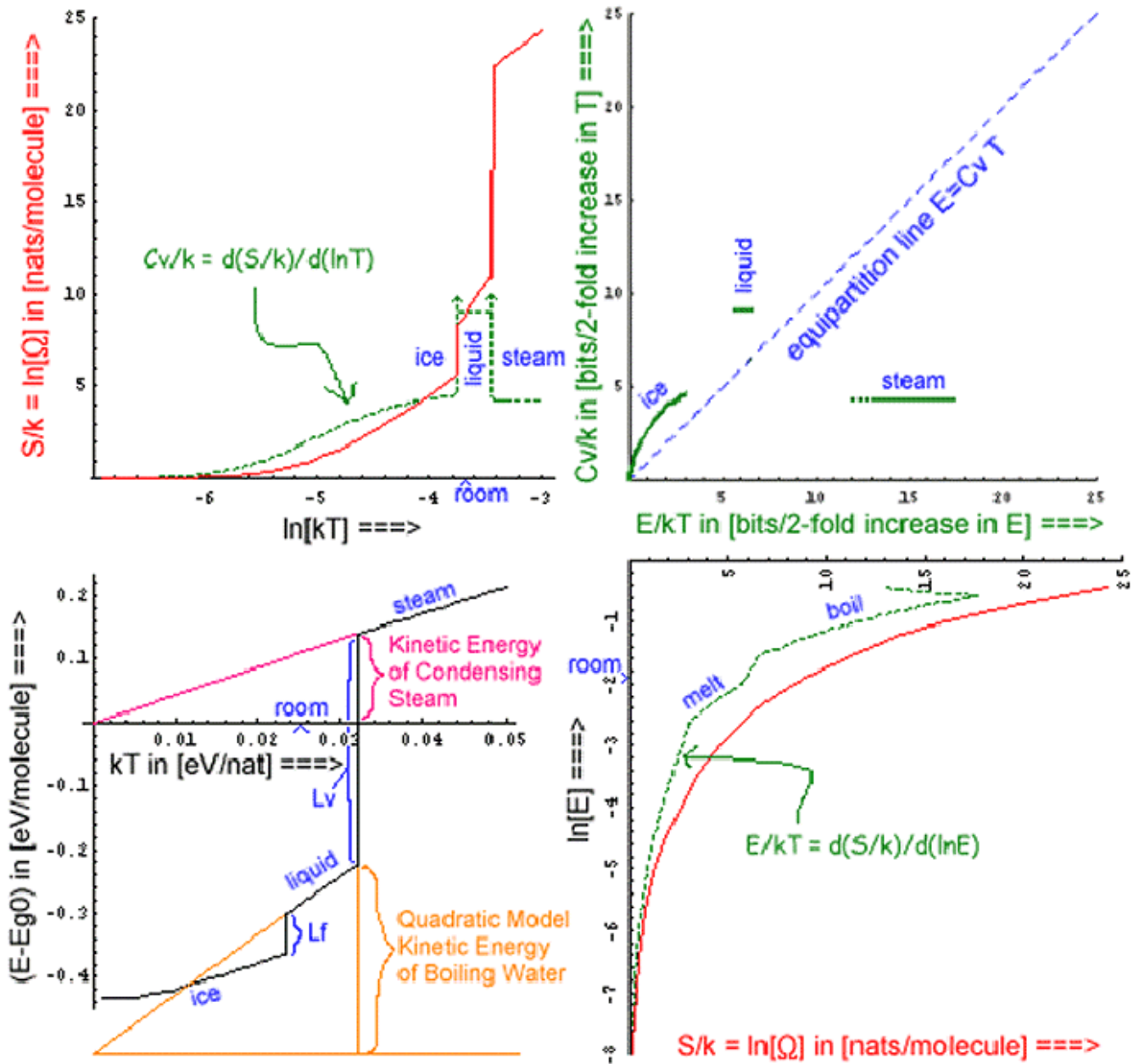


Fig. 5: Plot of E vs. kT (lower left), C_v/k vs. E/kT (upper right), and entropy plots associated therewith for H₂O steam, water, and ice. All quantities are "per molecule" values. Note how the energy of thermal motion for water increases as water molecules condense and become bound. Also note how equipartition applies only far from phase changes and then only when an imagined zero for the thermal energy is picked.

Requirements for an engine to be possible then become: **Heat OUT + Work OUT = Heat IN + Work IN & Uncertainty OUT - Uncertainty IN = Irreversibility Loss ≥ 0** . Written in this form, the equations put physical limits on: (i) the work available from a "heat engine" operating on fossil fuel or via photosynthesis; (ii) work needed to keep the frost on a 6-pack or to pump winter heat from the outside in; (iii) thermal energy needed to warm your home with a flame in winter, or to cook a pizza when it's really cold outside; and (iv) irreversible entropy increases that result from an oven leaking heat, ice melting in your tea, or your coffee getting cold. The answers to some of these calculations may knock your socks off!

As we show below, possibility calculations also put limits on the operation of molecular and macroscopic information engines, like copy machines, DNA polymerase, and even researchers. Such "engines" create *mutual information* resources whose state mirrors the state of nature somewhere else in space and time.

Since the possibility limits do not depend on how in detail you perform the task at hand, they are excellent tools for the detection of "snake oil peddlers", i.e. folks purporting to offer solutions (like perpetual motion machines) that will never materialize, regardless of their originality. The downside, of course, is that when these calculations tell you something might work, they provide few clues concerning how in practice to pull it off.

Quick Quiz: Home heating is oft done by irreversibly thermalizing heat from a flame. If each joule of heat from a 1000K flame were used to pump heat in from 273K air outside, how much room temperature (295K) heat might instead be gained for home heating? For more on this, cf. <http://bayes.wustl.edu/etj/articles/thermal.ency.pdf>

[Heat Engines and Energy Flow in Biology:](#)

Although seldom mentioned in physics books, plants act as heat engines by converting radiant energy from the sun to potential energy in the form of combustible biomass. That energy powers our metabolism, our automobiles, plus our oil-burning power plants. In the process, this energy (following

the 2nd law) is irreversibly thermalized (converted to ambient heat and eventually re-radiated into space).

Life's power stream is ignored in many discussions of global energy production and consumption, as in the 1993 World Almanac (Scripps Howard Publishing). But it is much larger than the flows they do mention, it is dearer to our hearts (giving them the energy to beat), and it is more at risk of being ignored by those occupied with short-term economic goals! Work in the 1950's by ecologists Eugene and Howard Odum, and more recently by Vitousek et al. (*Biosciences* **36-6** (1986) 368-373), indicate that of the 7×10^{17} watts of solar radiation incident on the earth, somewhere between 10^{14} and 10^{15} watts goes into primary production. This primary production is the lion's share of the power stream accessible to life on our planet. For example, nuclear power in 1993 provided only about 3×10^{11} watts by comparison.

Terrestrial life's power stream is something worth keeping track of. An entertaining exercise is to ask how many nanoseconds of this power stream we "thermalize" as consumers, with each daily activity. Assuming a 10^{15} watt power stream for reference, a 100 watt bulb burning for an hour thermalizes about 0.36 nanoseconds of that power stream, an apple thermalizes about 0.5 nanosecond, a vegetarian meal around 4 nanoseconds, and a gallon of gasoline about 100 nanoseconds. About 4 microseconds of life's power stream, by comparison, is needed to lift you alone (forget the spaceship) out of earth's gravitational potential well. In this sense, "life power stream time" is a way to measure energy usage from a global, renewable-resource, point-of-view.

Quick Quiz: How many nanoseconds of life's power stream do you consume per day in (a) food, (b) electricity, (c) gasoline, and (d) natural gas? What is the cost of each of these resource types in dollars per power stream nanosecond?

[Mutual Information, and Engines that make it:](#)

Nineteenth century physicists considered entropy S to be, like energy E , a purely "extensive" quantity. If you had twice as much of something (twice in all respects), then you had twice as much energy AND twice as much entropy. The 20th century connection between uncertainty and entropy makes S something

quite different than E. In fact, entropy is only extensive when we don't consider (or don't have) information about correlations between sub-systems. Uncertainty about the whole is generally less than the sum of uncertainties about the parts. For the case of two sub-systems A and B, you can write total uncertainty as $S_{AB} = S_A + S_B - I_{AB}$. Here the non-negative quantity I_{AB} is *mutual information*, information that A in some way mirrors B even if the details of both are unknown.

For example, if answers to 3 true-false questions are written on sheet A, then one may have $S_A = 3$ bits of uncertainty about what's on sheet A. If sheet B also contains answers to 3 true-false questions, then you might have $S_B = 3$ bits of uncertainty about what's on sheet B. Your total uncertainty S_{AB} is then $S_{AB} = S_A + S_B = 6$ bits. If you further know that sheet B was copied from sheet A (this is the mutual information), then only three yes-or-no answers can determine what's on both sheets. Hence $S_{AB} = 3$ bits, which means that the mutual information about the relationship between A & B, given above, is $I_{AB} = S_A + S_B - S_{AB} = 3$ bits as well. This makes sense, because I_{AB} basically says yes, for the 1st, 2nd, and 3rd questions on each sheet, to the question: "Do sheets A and B say the same?"

Quick Quiz: What is the mutual information in: (a) two sheets of paper containing the same 9 decimal-digit social security number, and (b) two strands of DNA with the same sequence of 96 bases, each of which takes one of 4 possible values?

Creation of mutual information can be figured into "possibility problems" by simply subtracting I_{AB} before and after from the **Uncertainty In** and **Uncertainty Out** terms, respectively where appropriate. This predicts an energy cost for creating mutual information, equal to $W_{\text{required}} = I_{AB}/kT$ where T is the ambient temperature at which one's "information engine" dumps the heat created by thermalizing that work. This means that during the transcription of DNA, as well as observing the stars or photocopying, for each bit of mutual information recorded, energy available for doing work must be converted to heat. It also provides clues to why video camera CCD's can record more bits of information per pixel, if cooled to low T.

Storing Mutual Information in Evolving Codes:

Thus in the mid-20th century, communication scientists discovered that messages written in replicable code are a great way to store mutual information that is in turn subject to the laws of thermodynamics. At the same time, Watson and Crick were finding that living cells have been tracking adaptations to their environment, with messages written in nucleic acid code, since the origin of life on earth.

These were the clues needed for later realization that codes and organisms share a symbiotic existence, and that as microbiologist Lynn Margulis and animal behaviorist Richard Dawkins point out, in multi-celled animals it is the genetic codes (as distinct from the organisms) that evolve. These codes evolve through successful replication, thus in effect recording mutual information about ongoing discoveries of what works, and what doesn't, in the environment of the day.

The plot thickened further when McLuhan in the 1960's, among others, pointed out that major changes for humans in the past two millenia accompanied changes in the way codes in the form of ideas, not genes, are able to replicate. These changes included the discovery of phonetic alphabets, the invention of moveable type, the development of broadcast media, and of course most recently the establishment of a world wide web.

The resulting distinction between "the evolution of self-consistent idea sets", and "the modeling of real-world organisms and objects with ideas", is the distinction with which we began this note (between Jayne's "logic of gambling theory" and "the physics of describing the dice"). Medical clinicians have long recognized a third challenge: respecting a subject's individuality along with their relation to statistics. Thus the storage of mutual information in replicable codes, the thermodynamic laws which limit those processes, and the ways in which such codes have evolved and continue to evolve, are very exciting subjects. They will continue to bring together many physicists, chemists, biologists, as well as computer, medical, and social scientists, in the century to come.

ReCap: In this essay, you've seen how information units in computer science, and thermodynamic units for entropy, share a common origin in the science of gambling theory or statistical inference. This provides a way to quantify information (and the lack thereof) about answers to tests, the state of a computer's memory, or even the state of gas atoms in this room, using *either* uncertainty S or accessible states Ω . You have also seen how to figure the dependence of energy on temperature (like equipartition) and of volume on pressure (like the ideal gas law) given the accessible states function for a given system, much as your next play in a card game may be determined by knowledge of the deck being used in the game. Thermodynamics is in this fundamental sense gambling. Sometimes it's predictions, however, have overwhelmingly good odds, since the number of particles in most macroscopic systems is on the order of Avogadro's number. We've further illustrated, in the problems or figures or text, how these insights connect the laws of thermal physics to: (i) inverted population sub-systems at negative absolute temperature, like lasers and spin systems and dominos standing on end, (ii) the storage and communication of information in files and transmittable messages, (iii) the evolution of genetic codes, and replicable ideas, in their symbiotic relationship to organisms and other "steady-state" excitations.

SECTION SUMMARY

Question: Where might we expect to find a *conserved quantity* X which has been randomly shared between systems for so long that prior information about the whereabouts of X is irrelevant? In other words, where is X likely to be *after equilibration*?

Answer: The most likely X distribution is that which can happen in maximal ways, and hence is one whose sub-system *uncertainty slopes* $\frac{dS}{dX}$ are equal. Here entropy or *uncertainty* $S \equiv k \ln \Omega$, where accessible states or *multiplicity* Ω is the number of ways a (sub-)system might accomodate quantity X , with information units defined by $k = \{1, \frac{1}{\ln 2}, 1.38 \times 10^{-23}\}$ for S in {nats, bits, or J/K}, respectively.

Example 1: When energy E is the quantity shared randomly between systems, the uncertainty slope $\frac{dS}{dE}$ is the reciprocal temperature or *coldness* $\frac{1}{T}$. For ideal gases and many solids, $\Omega \propto E^{vN/2}$, where N is the number of molecules and v is the number of degrees freedom or ways to store thermal energy per molecule. Solving this for $\frac{dS}{dE}$ gives the *equipartition theorem*: $\frac{E}{N} = \frac{v}{2} kT$.

Example 2: When volume V is the quantity shared randomly between systems, the uncertainty slope $\frac{dS}{dV}$ is the *free-expansion coefficient* equal to $\frac{dS}{dE} \frac{dE}{dV} = \frac{1}{T} \frac{F dx}{A dx} = \frac{P}{T}$ at equilibrium. For an ideal gas, $\Omega \propto V^N$. Solving this for $\frac{dS}{dV}$ gives the *equation of state*: $PV = NkT$.

Question: Is it possible to create a steady-state "engine" capable of a specified set of energy in/out and uncertainty in/out requirements?

Answer: Perhaps, if and only if **Heat OUT + Work OUT = Heat IN + Work IN**,
and **Uncertainty OUT - Uncertainty IN = Irreversibility Loss ≥ 0** .

Example 1: Work available from a heat engine (e.g. operating on fossil fuels or via photosynthesis)...

$Q_{\text{exhaust}} + W_{\text{out}} = Q_{\text{hot}}$, and $Q_{\text{exhaust}}/T_{\text{exhaust}} \geq Q_{\text{hot}}/T_{\text{hot}}$,
implies that *Carnot Efficiency* $W_{\text{out}}/Q_{\text{hot}} \leq 1 - T_{\text{exhaust}}/T_{\text{hot}}$.
Case Study: An automobile engine, with $T_{\text{hot}} = 429\text{K}$
and $T_{\text{exhaust}} = 300\text{K}$, has an efficiency limit of 30%.

Example 2: Maximum astrophysical (or other) observation rates, per observer per meal...

$Q_{\text{ambient}} = W_{\text{food}}$, and $Q_{\text{ambient}}/T_{\text{ambient}} - I_{\text{recorded}} \geq 0$,
limits mutual information to $I_{\text{recorded}} \leq W_{\text{food}}/kT_{\text{ambient}}$
Case Study: Human observers with typical caloric intake themselves record less than 10^{24} Mb/day.
(Aside: Some of us, yours truly included, produce MUCH LESS!)

GOAL Strategy for Solving Equilibration Problems

Gather...

- Take a guess at the form of multiplicity for the systems involved, e.g. $\Omega \propto (E-E_0)^{3N/2}$.
- What conserved quantities (like E for thermal interaction, V for mechanical interaction) are shared to equilibration?

Organize...

- Guess the behavior of accessible states, uncertainty, and energy versus temperature, for the systems involved. For example, $\Omega \propto (E-E_0)^{3N/2} \Rightarrow \frac{S}{k} = \frac{3}{2}N \ln(E-E_0) + \text{const} \Rightarrow \frac{1}{T} \equiv \frac{dS}{dE} = \frac{3Nk}{2(E-E_0)}$ or $E = E_0 + \frac{3}{2} N kT$.

Analyze...

- Sum ΔX values to zero. For 2 systems here, equate energy lost by hotter system A to heat gained by cooler B, in terms of initial T for each and a final equilibrium T_{eq} , i.e. $\Delta E_A + \Delta E_B = 0 \Rightarrow \frac{3}{2}kN_A(T_A - T_{\text{eq}}) = \frac{3}{2}kN_B(T_{\text{eq}} - T_B)$.
- Then solve for equilibrium temperatures, pressures, etc. that result. Here we get $T_{\text{eq}} = \frac{N_A T_A + N_B T_B}{N_A + N_B}$.

Learn...

- Check by substitution into the equations for X (in this case E as a function of T) to see that all sub-systems sharing X end up at the same dS/dX value (e.g. have the same T_{eq} when X is energy).
- Examine limiting behaviors, e.g. what happens when $T \rightarrow \infty$.

Some ShortCuts

For many systems between phase changes, $\Omega \propto (E-E_0)^{vN/2} \Rightarrow \frac{E-E_0}{N} = \frac{v}{2} kT$.

For ideal gases, $\Omega \propto V^N \Rightarrow PV = NkT$.

Read More About It

One of the first books to describe the importance of *reciprocal temperature* for ordinary folks was **Science by degrees - Temperature from 0 to 0**, by Castle, Emmenish, Henkes, Miller & Rayne (Walker & Co., Westinghouse Search Book Series, 1965). One of the first senior undergraduate texts, organized from the ground up with these principles in mind, was **Thermal Physics**, by Charles Kittel and Herbert Kroemer (W. H. Freeman, NY, 1980). The latest in the series of senior undergraduate thermal physics texts built on this foundation may be **An Introduction to Thermal Physics** by Daniel V. Schroeder, to be published in 2000 by Addison-Wesley.

On the world wide web, you might enjoy the tutorials for molecular biologists (plus other *Biological Information Theory and Chowder Society* folk) at the National Cancer Institute links of Tom Schneider (<http://www-lmmb.ncifcrf.gov/~toms/>), the course notes for digital logicians at <http://scitec.uwichill.edu/bmp/online/p31q/p31q.htm>, our *Information Physics* page (<http://newton.umsl.edu/infophys>), links to recent Bayesian studies (<http://bayes.wustl.edu>), and Ed Jaynes' book on probability as logic (<http://omega.math.albany.edu:8008/JaynesBook.html>). For a taste of continuing debates in the statistical inference community, check out Steve Goodman's editorial (and later reactions thereto) in the 1999 *Annals of Internal Medicine* 130, 7 (cf. <http://www.annals.org/issues/v130n7/full/199904060.00011.html>).

End-of-Section Questions

1. If you know the answers (for sure) to a quiz of ten true-false questions, do you have 10 bits of information, or 10 bits of uncertainty? Why?

End-of-Section Problems

Units

1. You are given a multiple-choice question with 5 possible answers, three of which you can rule out immediately. How much uncertainty about the answer to this question remains, in [nats] and in [bits]?
2. The interior temperature of a certain oven is 1500C. What is the temperature in natural units (say in electron volts per nat), and hence what is the typical size of thermal energy "chunks" that it will be prone to share? Are our eyes able to detect chunks of light (photons) with this energy?
3. A old-school physicist, with no interest in "modern" units, visits a computer store and asks for 2.45 [femto-joules per Kelvin] of dynamic RAM for her PC. How much random access memory does she really want?
4. How much is uncertainty increased, by adding 1 eV of thermal energy to *any system* at room temperature?

Accessible States

5. When throwing a pair of icosahedral (20-sided) dice with sides having between 1 and 20 spots, which spot-total facing up on the dice can occur in the largest number of ways? How many ways can this spot-total occur?
6. You have a parking lot with 12 spaces, and 5 cars to place in those 12 spaces. How many accessible states does this system have?
7. * I imagine a gas of atoms allowed to move in only x and y directions, with velocity components which are the integer multiples of some fundamental speed v_0 , as shown in Fig. 6. The concentric circles in Fig. 6 correspond to states of energy E equal to some integer multiple of the fundamental energy $\epsilon_0 = \frac{1}{2}mv_0^2$. If multiplicity is proportional to the perimeter of these circles for large E , what is M in the equation $\Omega \propto E^M$?

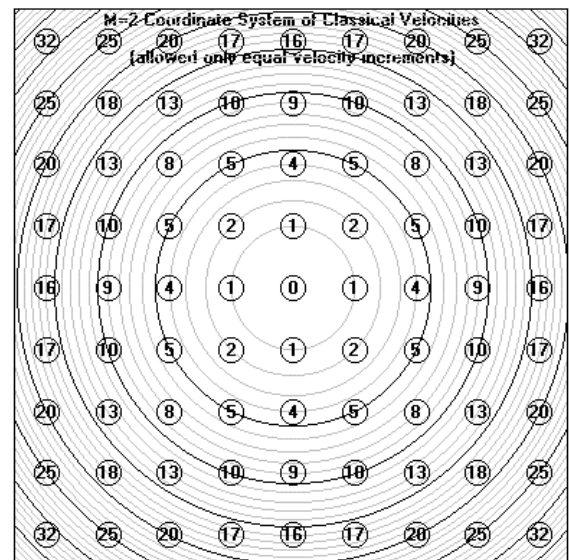
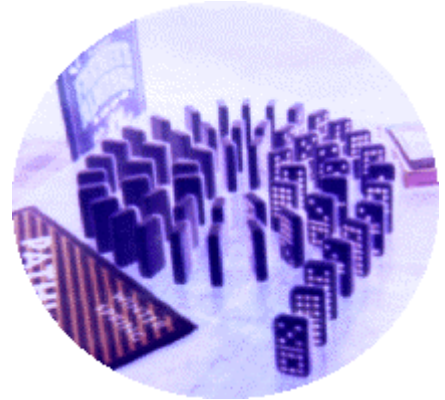


Fig 6: Accessible states, labeled with energy in units of ϵ_0 , for a single atom "gas" allowed random velocities only in x and y directions ($v=2$). Note how the annular energy-level rings get thinner and thinner, since speed (distance from 0) is $\propto \sqrt{\frac{E}{\epsilon_0}}$.

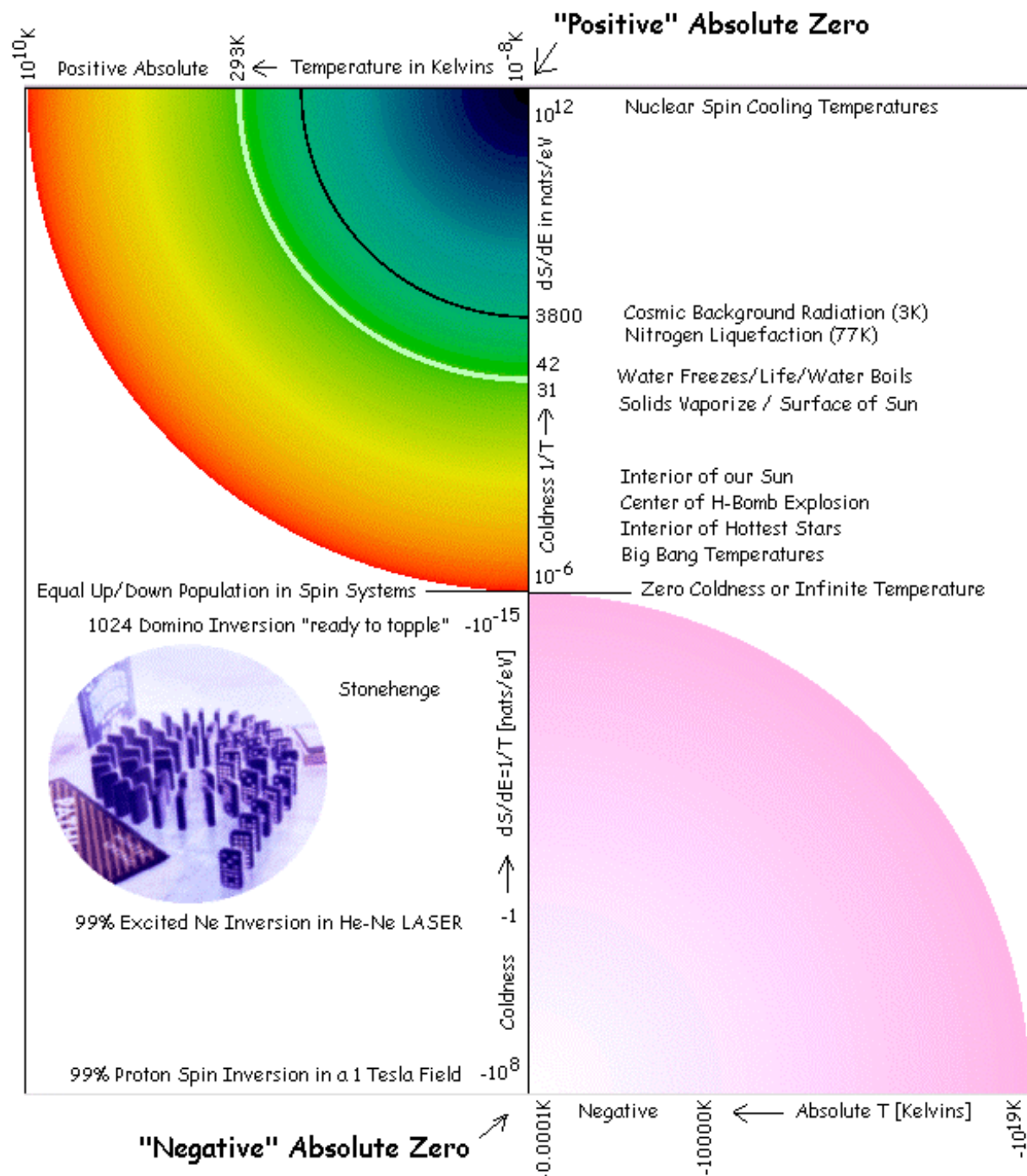
Equilibration

8. Two otherwise-insulated containers of mono-atomic argon gas, one with a half mole of atoms and the other with 3/2 moles of atoms, and each in fixed volumes of 22.4 liters, are allowed to randomly share a total of $E_{\text{tot}} = 7300$ joules of thermal energy between them. Assuming that the number of accessible states for each gas is proportional to $E^{3N/2}$, what equilibrium temperature will they eventually reach?
9. * Two otherwise-insulated containers of mono-atomic argon gas, one with a half mole of atoms and the other with 3/2 moles of atoms, randomly share a total of $E_{\text{tot}} = 7300$ joules of thermal energy and a total of $V_{\text{tot}} = 44.8$ liters through a flexible membrane. Assuming that the number of accessible states for each gas is proportional to $V^N E^{3N/2}$, what equilibrium temperature and volumes will they eventually reach?
10. Two otherwise-insulated blocks of iron, each containing 1 mole of iron atoms, are allowed to randomly share a total of $E_{\text{tot}} = 7300$ joules of thermal energy. Assuming that the number of accessible states for each block is proportional to $E^{6N/2}$, what equilibrium temperature will they eventually reach?
11. Two sets of $N = 100$ dominos (or nuclear spins in a magnetic field) capable of only two positions (up or down) are allowed to randomly share $\frac{1}{4}$ of the energy it would take to stand them all up as follows: Every time a domino (spin) in the first set flips up, a domino (spin) in the second set flips down, and vice versa. Assume for each set that multiplicity $\Omega(E) \propto \text{Exp}\left[-\frac{2(E-E_0)^2}{N\epsilon^2}\right]$, where $\epsilon = 0.001$ joule is the energy to stand one domino (or spin) up and E_0 is the energy if half are standing. What "orientation" temperature will they equilibrate at? This calculation gives rise to a relationship between energy and temperature for magnets, quite different from equipartition, now known as *Curie's law*.
12. * The two sets of dominos in Problem 10 are allowed to randomly share $\frac{3}{4}$ of the energy it would take to stand them all up. What "orientation" temperature will they equilibrate at? Is this temperature higher or lower than that of the previous problem? (*Hint*: The energy uncertainty-slope or reciprocal-temperature for this system behaves much more nicely than does temperature as it was historically defined.)
13. * Suppose the multiplicity of a star follows $\Omega = \{E_0/(E_0 - E)\}^{N/2}$, where E_0 is the depth of the star's "energy well" when cold, and E is thermal energy above that. What happens to the temperature of a 6000K star with binding energy $E_0 = 10^{42}$ joules when its thermal energy is increased from 10^{41} joules to 5.5×10^{41} joules, as its nuclear engines switch to the fusing of elements heavier than hydrogen. (*Hint*: What color are giant stars?)
14. * State uncertainties (entropies) are often huge numbers, and within a few orders of magnitude of the number of particles contained therein! If the entropy of a black hole with energy $E = mc^2$ is something like the number of photons of energy hc/λ it could contain, where wavelength λ is the black hole radius $\approx Gm/c^2$, then $S_{\text{blackhole}}$ is $\approx kGE^2/hc^5$. Here m is the black hole mass, G is the gravitational constant, h is Planck's constant, and c is the speed of light. If S is proportional to E^2 as predicted here, a calculation of dS/dE suggests that black holes might radiate with a temperature that decreases with mass! Although this would not be expected in Newtonian physics since light should not be able to escape, this conclusion is supported by the quantum-tunnelling calculations of Stephen Hawking in the late 1960's. What mass and radius must a black hole have, to radiate near the temperature of boiling water? Would it be a good idea to drop such a black hole into a glass of cold coffee, to warm it up? (Note: Problems 13 and 14 were inspired by problems in the senior undergraduate text on "Thermal Physics" by Daniel V. Schroeder, Addison-Wesley, 2000).



* Problems that are more difficult, but which may lead to a broader understanding.

The "Gambling Theory" Roots of the Zeroth Law



UW-STL Physics & Astronomy. Copyright (1998) P. Fraundorf

Most agree that the absolute zero of temperature is impossible to reach. But few know that there are two absolute zeros of temperature. The second may be approached from the direction of negative absolute temperatures, with a set of dominoes in the living room of your very own home! How so? (*Hint: As suggested by the drawing above, the connection between positive and negative absolute temperatures is best visualized with a look at reciprocal temperature or coldness, and its origins in the science of "making the best guess".*)