

Heat Capacity in Bits

P. Fraundorf*

Physics & Astronomy, U. Missouri-StL (63121), St. Louis, MO, USA

(Dated: July 14, 2002)

Statistical physics since Shannon has shown, following the 19th century work of Gibbs, that physical units for temperature kT defined statistically (via $1/T \equiv dS/dE$) are energy per “nat” of information uncertainty. Consequences of this for heat capacities are explored here for quadratic systems, and systems for which equipartition has little meaning. We show for *any* system that total thermal energy E over kT (an integral or average heat capacity when $T > 0$) is the log-log derivative of multiplicity with respect to energy, as well as (for *all* b) the number of base- b units of information lost about the state of the system per b -fold increase in the amount of thermal energy therein. Similarly the work-free instantaneous heat capacity C_v/k is a “local version” of this log-log derivative equal (for example) to bits of information lost per 2-fold increase in *temperature*. This makes C_v/k independent of both: (i) the energy zero, unlike E/kT , and (ii) one’s choice of the Lagrange multiplier for energy (e.g. kT versus $1/kT$) to within a constant, explaining why it’s usefulness may go well beyond the detection of phase changes and quadratic modes. From UMSStL-CME-94a09pf.

PACS numbers: 03.30.+p,01.40.Gm,01.55.+b

I. INTRODUCTION

The adjective deep has been applied to pedagogically simplifying approaches that offer both (a) wider applicability, and (b) reduced algorithmic complexity. Minkowski’s approach to special relativity via the metric equation is a classic example. Initial impressions of it as “superfluous erudition”¹ were later eclipsed by it’s uses in the development of general relativity², in metric interpretation of other forces³ (e.g. string theory), and in simplifying introductions to space-time^{4,5}.

Similarly, the communication-theory insights of Shannon⁶ led Jaynes^{7,8} and others in the middle of last century to clarify the distinction in thermal physics between “describing the dice” (i.e. physical description) and “taking the best guess” (the “gambling theory” part). Thus for example, the 1st law of thermodynamics “describes the dice” by saying energy is a conserved quantity. By comparison, the zeroeth law, along with the Boltzmann and Gibbs factors for a given system, are “dice-independent” tools of statistical inference⁹ in that they generally describe the role of energy’s Lagrange multiplier (temperature) during thermal equilibration, in the context of various entropy maximizations.

Such statistical insights for example allow students to infer equipartition and the ideal gas law, as well as their range of validity, from simple assumptions about systems to which they apply. The approach is becoming increasingly popular in undergraduate texts^{10–16} as a way to deepen student understanding as well as their physical intuition in matters involving the “codes” of biology¹⁷ and computer science¹⁸. The heart of the approach is Shannon’s observation that entropy S is a measure of statistical uncertainty, most naturally expressed in infor-

mation units (like bits or bytes or nats) rather than in historical units (e.g. Joules per Kelvin). It then follows, of course, that temperature (i.e. the reciprocal of uncertainty slope $1/T \equiv dS/dE$) might be quantified in units of energy per unit information uncertainty (e.g. room temperature $\approx \frac{1}{40}$ eV per nat) and that heat capacities (i.e. dE/dT), normally also in Joules per Kelvin, may be assigned information units as well.

What might be the *physical* meaning of a heat capacity without reference to historical temperature units? Bits of what? An answer to this question (something any student of these modern textbooks might ask) does not appear to be common knowledge among physics teachers, so we outline an answer here. Insight for students into the mechanism that underlies the utility of heat capacity, in calculations involving thermal equilibrium, might better prepare them for applying the same principles of statistical inference to more complex systems i.e. to those not necessarily at equilibrium, or involving conserved quantities other than energy.

II. DESCRIBING THE DICE

The dice of thermal physics are usually physical systems capable of accomodating thermal energy (as well as other quantities that may be conserved, like volume and particles) in a *multiplicity* of ways. This multiplicity is itself the key to understanding, particularly when systems are seen from the (micro-canonical) vantage point of the conserved quantities *per se*. For example, many gases, liquids, and solids behave over some of their temperature range as though multiplicity (Ω) is proportional to $E^{\frac{\nu N}{2}}$, where E is thermal energy, N is the number of molecules,

and ν is the number of ways to store thermal energy per molecule. Such systems are called “quadratic”, since the proportionality results from a “sum of squares” connection between E and the state coordinates involved in storing energy.

Some features of a system which is quadratic at all energies (a Sakur-Tetrode model Argon gas at 1 atmosphere) are illustrated for later reference in Column A (Panels A1 through A4) of Figures 1 and 2. In Figure 1 you’ll find plotted E versus kT , and entropy S versus $\ln[kT]$ as well as its slope. In Figure 2, you’ll find plotted S versus $\ln[E]$ as well as its slope (both rotated by 90°), and heat capacity C_v versus E/kT . For an ideal quadratic system, all of these plots are straight lines except for the last, which plots to a single point. Columns B and C contain similar plots for other systems, to be discussed below.

III. TAKING THE BEST GUESS

Application of gambling theory to physical systems begins with the question: Where might we expect to find a conserved quantity X which has been randomly (as far as we know) shared between systems for so long that prior information about the whereabouts of X is irrelevant? In the jargon of the field, this is the same as asking: Where is X likely to be *after equilibration*?

The science of decision-making in the presence of uncertainty (i.e. statistical inference or “gambling theory”) suggests that the best bet is the distribution of X that can happen in the most ways, provided there is no reason to prefer one way over another. In the jargon of probability theory, this is the recommendation: Assume equal *a priori* probabilities, when evidence to the contrary is not available.

For example, if systems A and B have a total energy E to share between them, then the best bet after equilibration will be that value of $E_A = E - E_B$ that has the largest total multiplicity $\Omega = \Omega_A \Omega_B$. Setting to zero the derivative of Ω with respect to E_A , one finds that this maximum requires that $\frac{1}{\Omega_A} \frac{d\Omega_A}{dE_A} = \frac{1}{\Omega_B} \frac{d\Omega_B}{dE_B}$, and hence that $\frac{d(\ln \Omega)}{dE}$ be the same for both systems! These derivatives (and most others in this paper) are taken under “microcanonical constraints”, i.e. they are partial derivatives with other extensive quantities (like volume or number of particles) held constant.

Information theorists, of course, define the logarithm of multiplicity as uncertainty via the equation $S = k \ln \Omega$, and measure it in {nats, bits, bytes, or J/K} if k is $\{1, \frac{1}{\ln 2}, \frac{1}{\ln 256}, 1.38 \times 10^{-23}\}$, respectively. Thus the best bet for any two systems sharing a conserved quantity X , in the absence of information to the contrary, is that X will rearrange itself between the two systems until each system’s *uncertainty slope* ($\frac{dS}{dX}$) has reached a common value. This is a quantitative version of the zeroth law of thermodynamics, based purely in the science of sta-

tistical inference, which applies to *any systems* sharing conserved quantities.

Of course, when energy E is the quantity shared randomly between systems, the uncertainty slope $\frac{dS}{dE}$ is the reciprocal temperature or *coldness*¹⁴ $\frac{1}{T}$. Hence temperature is a property which signals the propensity of a system for sharing of energy thermally. For example, calculating the uncertainty slope for quadratic systems from the multiplicity given above yields the widely useful equipartition relation: $\frac{E}{N} = \frac{\nu}{2} kT$.

When V is the quantity shared randomly between systems, the uncertainty slope $\frac{dS}{dV}$ is the *free-expansion coefficient*¹⁴ 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 $\frac{dS}{dV}$ yields the ideal gas equation of state $PV = NkT$. When N is the quantity shared randomly, the uncertainty slope $\frac{dS}{dN}$ is the *chemical affinity*, equal to $-\frac{\mu}{T}$ at equilibrium. From this, for example, reaction equilibrium constants may be calculated.

This quantitative version of the zeroth law applies to all thermal systems which equilibrate, including spin systems (like magnets) capable of population inversions and hence negative absolute temperatures. Moreover, as a theorem of statistical inference not involving energy at all, it applies also to thermally unequilibrated systems sharing other conserved quantities (even money, for example), provided the only prior information we have is how the multiplicity of ways that quantity can be distributed depends on the amount of that conserved quantity to begin with! If we have other kinds of information, such as knowledge of a system’s temperature but not its total energy, then the broader class of maximum entropy strategies in statistical inference (e.g. Gibb’s canonical and grand ensembles) predict the distribution of outcomes we can expect there as well.

IV. THERMAL ENERGY OVER KT

A closer look shows that the statistical definition of temperature above can be rewritten as:

$$\frac{1}{kT} \equiv \frac{\partial(\ln \Omega)}{\partial E} = \left\{ \frac{\partial(\ln \Omega)}{\partial(\ln E)} \right\} \frac{1}{E}. \quad (1)$$

The quantity in curly brackets is the log-log derivative of multiplicity with respect to thermal energy. We can also think of this as the “instantaneous exponent” of energy in the expression of multiplicity as energy to some power, or as the slope of the multiplicity versus energy curve on a log-log plot.

Rearranging the equation yields something that looks very much like the familiar equipartition theorem, except that the relation applies to all thermal systems under conditions of maximum ignorance (i.e. at equilibrium):

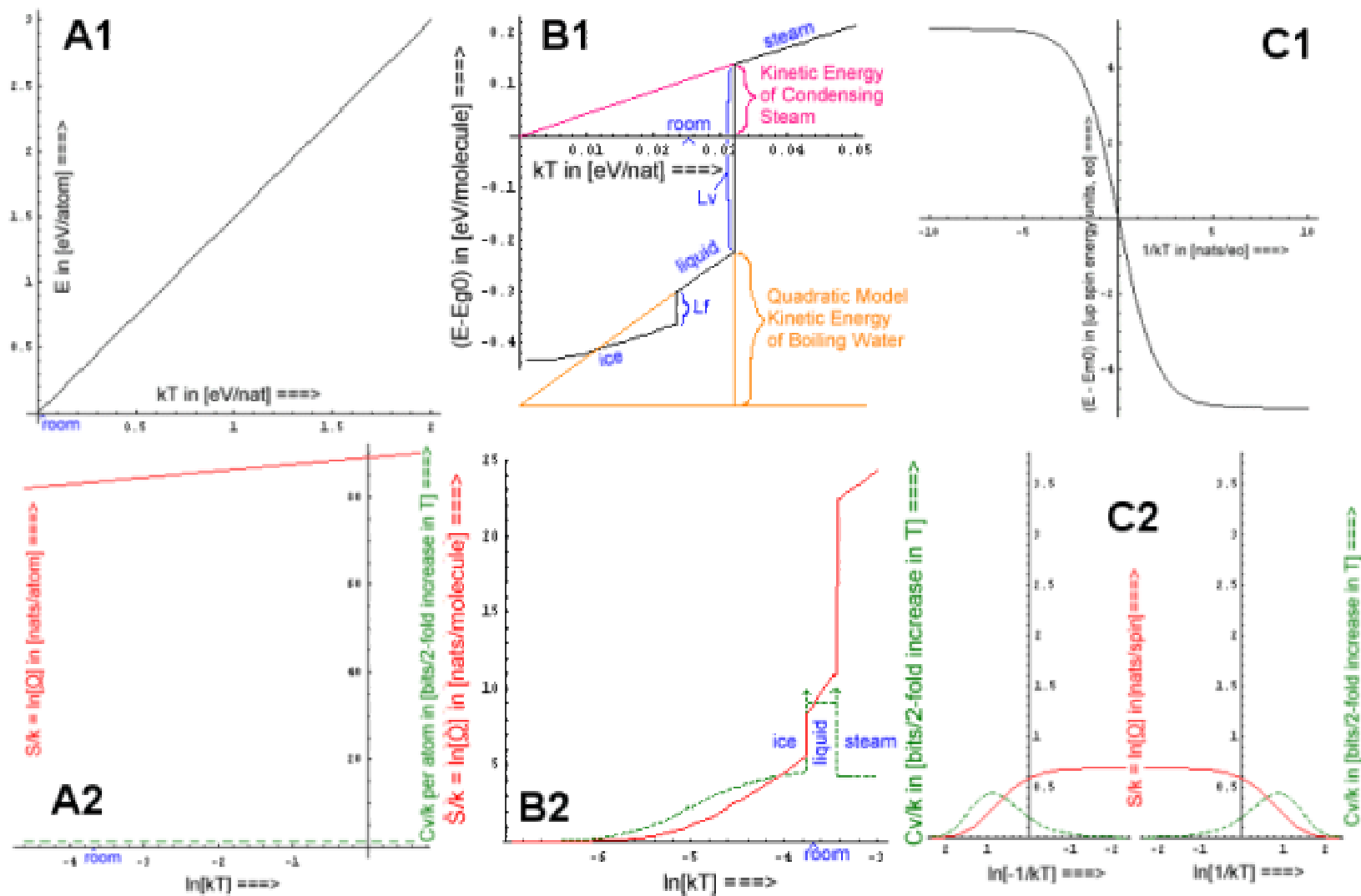


FIG. 1: Plot of E versus kT , and S versus $\ln(T)$, in natural units for (i) an ideal (Sakur-Tetrode) monatomic Argon gas at 1 atmosphere, (ii) a steam, water, and Debye-model ice system ($T_{Debye} = 333\text{K}$ assuming sound speed near 3500m/s) at 1 atmosphere, and (iii) a two-state paramagnet consisting of 10 non-interacting spins. All quantities are “per atom”, “per molecule”, and “per spin” respectively.

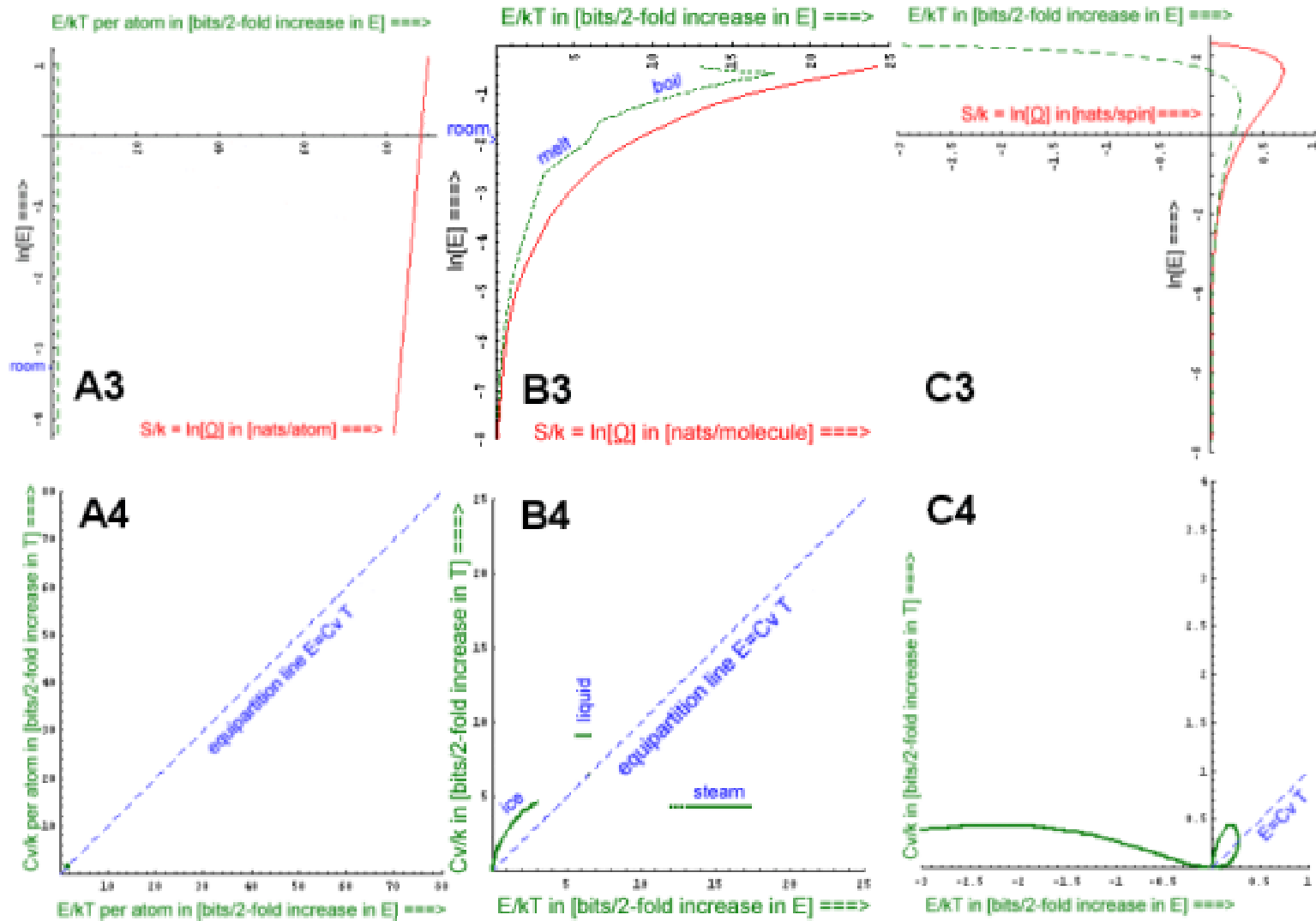


FIG. 2: Plot of S versus $\ln(E)$, and C_v/k versus E/kT , in natural units for (i) an ideal monatomic Argon gas at 1 atmosphere, (ii) a steam, water, and Debye-model ice system ($T_{Debye} = 333K$ assuming sound speed near $3500[m/s]$) at 1 atmosphere, and (iii) a two-state paramagnet consisting of 10 non-interacting spins. All quantities are “per atom”, “per molecule”, and “per spin” respectively.

$$\frac{E}{kT} = E \frac{\partial(S/k)}{\partial E} = E \frac{\partial \ln \Omega}{\partial E} \equiv \xi \stackrel{E \geq 0}{=} \left\{ \frac{\partial \ln \Omega}{\partial \ln E} \right\} = \left\{ \frac{\partial(\log_b \Omega)}{\partial(\log_b E)} \right\} \forall b \in \{R+\}. \quad (2)$$

For quadratic systems, it is easy to see that our log-log derivative is nothing more than half the number of degrees of freedom ($\frac{\nu N}{2}$). For any system, however, ξ measures the instantaneous energy exponent, as well as the number of *nats of information lost about the state of the system per e-fold increase in thermal energy* of the system. The last term in the equality string simply notes that the value is independent of the base b of the logarithms used, provided the same base is used in numerator and denominator. Thus we can also think of ξ as the number of bits of information lost per 2-fold increase in thermal energy, or more generally the number of base- b units of information lost per b -fold increase in thermal energy. In our search for the meaning of heat capacity in natural units, this is our first big clue.

Before we move on, we should also point out something that applies if our energy origin has been chosen so that $E \rightarrow 0$ as $T \rightarrow 0$, something we *might* expect for a measure of thermal energy. In terms of the no-work (e.g. constant volume) heat-capacity $C_v \equiv \frac{\partial E}{\partial T}$, we can write:

$$\xi = \frac{E}{kT} = \frac{\int_0^T C_v dT}{kT} \stackrel{T \geq 0}{=} \frac{\int_0^T C_v dT}{k\Delta T} \equiv \left\langle \frac{C_v}{k} \right\rangle, \quad (3)$$

with the middle equality applying only for systems NOT in a population inversion, so that absolute temperature $T > 0$ and $\Delta T = T - 0 = T$. Thus when absolute temperature is positive, ξ is a *heat capacity average* over temperatures ranging between T and absolute zero.

We've shown here that the log-log derivative of multiplicity, with respect to energy, has a simple information theoretic interpretation, and is elegantly given in natural units by $\frac{E}{kT}$ as well. From the perspective of an experimentalist, however, it has one glaring disadvantage: *It's numeric value depends on our choice for the zero of thermal energy.*

To illustrate the problem, consider the cooling of water until it becomes ice. As water cools initially, the temperature drop per unit energy removed is roughly constant. One might easily say: "This looks like a quadratic system with about 18 degrees of freedom per molecule", so ξ must be about 9 bits per 2-fold increase in thermal energy. Then, at the freezing point, the temperature stops dropping as energy continues to be removed, suggesting a quadratic system with nearly infinite degrees of freedom! Once all is frozen, of course, temperature continues its drop, this time suggesting a quadratic system with about 8 degrees of freedom per molecule, or ξ closer to 4 bits per 2-fold increase in thermal energy! Since kT may change little during this experiment, how can E/kT be jumping around so much?

The answer of course is that these inferences follow not by measuring total energy E , but only changes in energy. Moreover, in the process our preferred zero of thermal energy has been shifting about. We can see the effects of this more explicitly if we plot energy versus temperature for water, as shown in the top center (Panel B1) of Figure 1.

Column B of Figures 1 and 2 illustrates some uses of physical (as distinct from historical) temperature units for water. Note in Panel B1 how the quadratic model "energy of thermal motion" for water *increases* as water condenses from a gas to a liquid. Also note how equipartition applies only far from phase changes and "freeze out" zones, and then only if a fictitious zero for the thermal energy is chosen. Also note that the first row in Fig 2 plots S versus $\ln[E]$ and it's slope (dashed), which from equation 2 is nothing other than E/kT . These graphs were rotated so as to share a common horizontal axis with plots in the second row of Figure 2.

The question then is, can we modify our estimate for the log-log derivative of thermal energy so as to reflect only data on temperature changes over a limited energy range? Such a quantity might allow us to probe the ways that thermal energy is being accommodated specifically, one energy range at a time.

V. INSTANTANEOUS HEAT CAPACITIES

The no-work (e.g. constant volume) instantaneous heat capacity, in natural units, can be written in terms above as:

$$\frac{C_v}{k} \equiv \frac{\partial E}{k\partial T} = \frac{\partial}{\partial T} [\xi T] = \left[1 + T \frac{\partial}{\partial T} \right] \xi = T \frac{\partial(S/k)}{\partial T}. \quad (4)$$

Here T is absolute temperature in any units you like, and the partials are taken with work parameters (like volume) held constant.

This quantity has an interesting property. If we define thermal energy E as a difference between total energy U and a specified "zero thermal energy" origin U_o , i.e. as $E \equiv U - U_o$, then it is easy to see that $C_v \equiv \frac{\partial E}{\partial T}$ is independent of our choice for U_o . Thus, although ξ obviously depends on one's choice of U_o , C_v does not.

To see what C_v actually measures, let's suppose that we have a quadratic system whose multiplicity obeys $\Omega = \left(\frac{U-U_1}{\varepsilon_o}\right)^{\frac{\nu N}{2}}$. It then follows simply that $\frac{S}{k} = \frac{\nu N}{2} \ln\left(\frac{U-U_1}{\varepsilon_o}\right)$, $\frac{\partial S}{\partial E} = \frac{\nu N}{2(U-U_1)}$, $\frac{U-U_1}{kT} = \frac{\nu N}{2} = \frac{C_v}{k}$. Thus $\frac{C_v}{k}$ estimates not $\frac{U}{kT}$ but $\frac{U-U_1}{kT}$, where U_1 is the "true origin" of thermal energy for this quadratic system. This is illustrated

in Column A of Figures 1 and 2, which plots for an ideal monatomic gas (the classical quadratic system) the same quantities plotted for water in Column B. Note also that S versus $\ln[kT]$ is plotted in the second row of Figure 1, along with its slope (dashed) which from equation 4 is nothing other than the no-work heat capacity C_v . Thus the second rows of Figures 1 and 2 share a common vertical parameter.

If the system is not simply quadratic (e.g. if it has phase changes, modes of energy storage which freeze out, etc.), then $\frac{C_v}{k}$ is simply a local estimate of thermal energy over kT , under the quadratic assumption. Thus $\frac{C_v}{k}$ modifies the log-log derivative of multiplicity with respect to energy, by combining it with its rate of increase per e-fold change in temperature, to yield an estimate of $\frac{U-U_1}{kT}$, where U_1 is a zero of thermal energy determined by assuming that the locally-measured log-log derivative of multiplicity is constant down to $T = 0$.

This observation also provides a different perspective on the mechanism by which heat capacity blows up during a phase change. Thermal energy over kT (or the log-log derivative of multiplicity) of course should have no singularities in it, since both energy and $\frac{1}{kT}$ are expected to be finite for finite systems. It is thus the second term in the two-term expression for heat capacity above, namely the temperature derivative, that provides the instability. Discontinuous shifts in the locally-inferred energy zero (U_1) on which heat capacity is based during a phase change thus, via this second term, also cause heat capacity to become singular.

Returning to Panel B1 of Figure 1, this raises the interesting question: *Does the thermal energy of steam increase, decrease, or go negative, when it condenses to water?* From the above, we can see that it of course decreases (perhaps even goes negative) if one's "zero of thermal energy" is held constant, since steam loses the latent heat of vaporization when it condenses, bringing its total energy down. However, because the specific heat of water at boiling is higher than that for steam, energy of random motion measured with respect to our locally-inferred zero of thermal energy (e.g. at 100C) actually goes up! In other words, a small part of the binding energy, liberated when water molecules fall into the potential well of their neighbors, goes to *increase* the energy of random motion in the condensed phase relative to that available to particles in uncondensed gas!

VI. BEYOND EQUIPARTITION

The observations above suggest that introductory texts might consider highlighting $\frac{C_v}{k}$ per molecule in natural units for common substances, with no apology for the fact that it is near but not exactly half-integral in many cases. After all, thermal energy may not have the same access to all molecules, all of the time. This quantity nonetheless provides deep insight into the relationship between uncertainty and thermal energy.

In shifting the focus from historical temperature units to the multiplicities which underlie our inferences, we can say that both ξ and $\frac{C_v}{k}$ measure bits of uncertainty per 2-fold increase in energy for *all* physical systems, with respect to their respective choices for energy origin. In this sense, they represent physical quantities like degrees freedom, but with wider applicability. After all, *degrees freedom* presumes not only multiplicities that are linear with energy on a log-log plot (like a high-temperature Einstein solid), but it also presumes quadratic energies (i.e. energies proportional to a sum of squares of some "randomly-occupied" coordinates of state). The idea that "every active coordinate gets $\frac{kT}{2}$ ", as we show below, even more strongly resists extension to systems with one or more entropy maxima. We begin, however, with a non-quadratic example of less drastic proportion.

A. Debye Solids

The Debye heat capacity of a solid is one case where $\frac{C_v}{k}$ depends strongly on temperature¹⁰. In the Debye low temperature limit, one has $E = \frac{6N}{2} \frac{\pi^4 T^3}{5\theta^3} kT$, so that $\xi = \frac{6N}{2} \frac{\pi^4 T^3}{5\theta^3}$, while $\frac{C_v}{k} = \frac{6N}{2} \left(\frac{\pi^4 T^3}{5\theta^3} + T \frac{3\pi^4 T^2}{5\theta^3} \right) = 4 \frac{6N}{2} \frac{\pi^4 T^3}{5\theta^3}$. Here of course, θ is the Debye temperature related to the density and speed of sound in the solid. Note that in this limit, only a quarter of $\frac{C_v}{k}$ comes from equipartition ($\frac{E}{kT}$), the remaining three quarters from the time derivative of $\frac{E}{kT}$ (in effect, from the unfreezing of new modes of energy accommodation). As you can see from Panel B1 of Figure 1, such unfreezing is associated with a lowering of the thermal energy zero locally referenced by the heat capacity. Thus attempts to infer ξ from the heat capacity by assuming that $\xi \simeq \frac{C_v}{k}$ yield a 4-fold overestimate of the number of degrees of freedom! This overestimate decreases as temperatures work themselves up to and beyond the Debye temperature θ of the solid, as illustrated in Panel B1 of Figure 1 for temperatures well above θ . There in the high temperature limit, $\xi \simeq \frac{C_v}{k} \simeq 3N$, as one expects from a classical lattice model above with $\Omega \propto E^{3N}$.

B. Two-State Paramagnets

A system more challenging to the traditional interpretation of "degrees freedom" is that for a system of N half-integral spins (i.e. a two-state paramagnet) of orientation energy ε . I like it because, as Dan Schroeder says¹⁶, "it forces us to think primarily in terms of entropy rather than temperature". Begin with any system whose energy-storing coordinates (e.g. displacements in a potential field) in practice have an upper limit on the amount of energy they'll accommodate. As long as energy is low enough that no single coordinate approaches the maximum value, then we may well find behavior very much like that of the systems discussed above.

However, when individual coordinate energies begin to approach their maximum (this happens quickly for two-state paramagnets whose coordinates accommodate but one unit of energy), things change fundamentally. In particular, there will be but one way (neglecting degeneracies) for the system to store the maximum amount of energy (namely when each of the coordinates is fully energized). With but one way to accommodate either minimum energy or maximum energy, multiplicities will approach 1 at both endpoints of the continuum. Since large systems may have many ways to store intermediate amounts of energy, multiplicity (and entropy) as a function of system energy will have a maximum (or *maxima*) somewhere between. At such maxima, $\frac{dS}{dE} \equiv \frac{1}{kT}$ will be zero, while on their high-energy side $\frac{dS}{dE} \equiv \frac{1}{kT}$ will be negative (signaling “population-inverted” states not accessible by thermal contact with reservoirs at positive absolute temperature).

Taylor-expanding to second order about the energy ($E_{S \max}$) of such multiplicity maxima gives $S(E)$ in the neighborhood as $S(E_{S \max}) + \frac{1}{2}A(E - E_{S \max})^2$, where $A \equiv \left(\frac{d^2S}{dE^2}\right)_{E=E_{S \max}} < 0$. Hence for energies near $E_{S \max}$, $\frac{dS}{dE} \equiv \frac{1}{kT} \simeq A(E - E_{S \max})$. Thus deviations from $E_{S \max}$ are negative at positive absolute temperature, and (for small deviations at least) are proportional to reciprocal temperature $\beta \equiv \frac{1}{kT}$, as illustrated in Panel C2 of Figure 1. Of course, this law was discovered by Pierre Curie in experimental study of magnetization, and bears his name.

Column C of Figures 1 and 2, for a two-state paramagnet of 10 non-interacting spins, differs from the plots in Columns A and B in that the two left side plots use $1/kT$ in place of kT because of its more natural mapping of inverted and uninverted population states. All quantities are “per spin” values, and of course the energy units e_o depend on both the spin magnetic moment and the

magnetic field. Note in particular how the “equipartition line” seems even less relevant here, even though heat capacity remains positive as reciprocal temperature dips below zero. For virial systems (like a gravity bound gas), negative heat capacities make the concepts of equipartition and degrees-freedom even less apropos, even though the statistical relationships shown here remain intact.

For two-state paramagnets, $A = -\frac{1}{N\varepsilon_B^2}$ and $E_{S \max} = \frac{N\varepsilon_B}{2}$, where ε_B is the energy of alignment per spin (magnetic moment times magnetic field strength). Thus for E near $E_{S \max}$, $\frac{C_v}{k} \simeq N\varepsilon_B^2\beta^2 > 0$, while $\frac{E - E_{S \max}}{kT} \simeq -N\varepsilon_B^2\beta < 0$. Since β for these systems may be positive or negative, ξ will be negative for some β values regardless of our choice of the thermal energy zero! Although negative “degrees freedom” may cause discomfort for some, a negative value for ξ should disturb no one since, to paraphrase a related comment by Schroeder¹⁶, there’s no law of physics guaranteeing that there will not be fewer ways to distribute energy, as more energy is added.

In fact, as we now know, temperature and reciprocal temperature are simply different forms for the Lagrange multiplier that characterizes a system’s willingness to share thermal energy^{7,9}. Systems, like these spin systems, capable of taking on (and sharing energy from) negative absolute temperature states show clearly that for them reciprocal temperature has more fundamental significance, and that the “absolute zeros” of temperature (approached from negative or positive directions) are indeed at opposite ends of a continuum¹⁹. But if reciprocal temperature is more fundamental, our instantaneous no-work heat capacity should be no less simply connected to the log-log derivative via reciprocal temperature. Rearrangement of the equation above shows that indeed this is the case:

$$\frac{C_v}{k} = -\beta^2 \frac{\partial E}{\partial \beta} = -\beta^2 \frac{\partial}{\partial \beta} \left[\left\{ \frac{\partial \ln \Omega}{\partial \ln E} \right\} \frac{1}{\beta} \right] = \left[1 - \beta \frac{\partial}{\partial \beta} \right] \left\{ \frac{\partial(\ln \Omega)}{\partial(\ln E)} \right\} = -\beta \frac{\partial \ln \Omega}{\partial \beta} = -\beta \frac{\partial(S/k)}{\partial \beta}. \quad (5)$$

Had we historically adopted as our measure of “willingness to share energy” some other power of the uncertainty slope, say $\gamma \equiv \beta^\alpha$, the instantaneous heat capacity is more generally the log-log derivative of multiplicity with respect to that measure, times $-\alpha$. Hence log-log derivatives of multiplicity with respect to the Lagrange multiplier of a conserved quantity (i.e. instantaneous heat capacities when the conserved quantity is energy) are, within a constant, independent of the specific multiplier used (e.g. temperature or its reciprocal).

This $\alpha = 1$ version of the relation now lets us simplify our perspective on entropy maxima. In the continuum

(Stirling) approximation for spin system accessible states Ω , and measuring energy from the “low-energy side”, equation (1) yields $E \simeq N\varepsilon/(1 + e^x)$ where $x \equiv \varepsilon/kT$, so that $\frac{E}{kT} = Nx/(1 + e^x)$ and $\frac{C_v}{k} = [x/(1 + e^x)]^2 e^x N$. Here of course, ε/kT takes on positive and negative values, ranging from around $+\ln(N)$ to $-\ln(N)$ respectively for orientation energies E with allowed values from 0 to $N\varepsilon$. As you can see from the plot in Fig. 3, again $\frac{C_v}{k}$ overestimates $\frac{E}{kT}$ at low temperatures (high values of ε/kT), although the estimate becomes exact when ε/kT decreases to around $x \equiv \frac{\varepsilon}{kT} \cong 1.279$ (the solution of $e^{-x} = x - 1$). After this $\frac{C_v}{k}$ underestimates $\frac{E}{kT}$, which

begins to decrease as x decreases and T increases from this point.

All of this switches again when $\frac{E}{kT}$ passes through zero, since $\frac{C_v}{k}$ remains positive while the change in uncertainty per e -fold increase in energy ($\frac{E}{kT} \equiv \xi$) becomes negative since uncertainty about the system state *decreases* with added energy past this point. Also, of course, average heat capacity goes to zero and no longer equals $\frac{E}{kT}$, since the average must be obtained piecewise when temperature (unlike reciprocal temperature) breaches the discontinuity from plus infinity to minus infinity. Thus in addition to information units for heat capacity, we gain from this approach a way to minimize consternation over negative degrees of freedom (e.g. for spin and virial systems), and to visualize the limits of equipartition as well. Although the concept of equipartition offers a way to think about the size of energy fluctuations as well as the effects of heat flow on temperature, as you might infer from the plots in Row D of Figure 2 its usefulness taken literally is limited only to the simplest of models.

VII. SUMMARY

In short, we've looked here at natural (as distinct from historical) units for the common thermodynamic quantities, so that we might explore the possibility that common uses of T , as a measure of "willingness to share thermal energy", have inherited their present emphasis partly because they predate our present understanding of multiplicity (the W in $S = k \ln W$ on Ludwig Boltzmann's tombstone). Heat capacity is a particularly knotty concept in this regard, since for most of us it has always been a change in energy "per degree Kelvin". In fundamental units, if heat capacity has any dimensions at all they are information units, since for example "change in energy per unit change in energy per bit" leaves us with nothing but bits in the bargain.

We point out a simple interpretation here. To begin with, thermal energy E , divided by temperature $T \equiv \frac{\partial E}{\partial S} \equiv \beta^{-1}$, is fundamentally $\frac{E}{kT} \equiv E \frac{\partial S}{\partial E}$, i.e. the log-log derivative of multiplicity with respect to energy (e.g. a measure of the bits of uncertainty increase per two-fold increase in energy). The instantaneous no-work heat capacity then becomes $\frac{C_v}{k} \equiv T \frac{\partial S}{\partial T} = -\beta \frac{\partial S}{\partial \beta}$, an estimate of $\frac{E}{kT}$ with an energy-zero inferred from the "local slope" of the log-log plot. We show that these two quantities bear a simple relationship to each other, regardless of the variable (e.g. T or $\frac{1}{kT}$) chosen to keep track of a system's willingness to share energy thermally.

The former of the quantities, namely $\frac{E}{kT}$, plays the role of "degrees freedom over two" in quadratic systems, but is dependent on the energy zero, and regardless can take on negative values in systems with entropy maxima. Its limitations are those of the concept of equipartition itself. The latter quantity, namely $\frac{C_v}{k}$, provides deep insight into the ways a system accommodates *new* thermal energy. Because these quantities are: (i) here defined only in

terms of state multiplicity and the conserved variable being shared (e.g. energy), and (ii) relatively independent of the form chosen for the Lagrange multiplier in the problem (e.g. temperature), their *analogs* in problems that involve the sharing of other conserved quantities (e.g. volume, particles, or even dollars) may be easier for students to recognize and put to use.

ACKNOWLEDGMENTS

This work has benefited indirectly from support by the U.S. Department of Energy and the Missouri Research Board, as well as by Monsanto and MEMC Electronic Materials Companies. It has benefited more directly from my acquaintance with students and colleagues (among them David Hestenes, Art Hobson, Ed Jaynes, John Mallinckrodt, Tom Moore, Dan Schroeder, Bruce Sherwood, and Edwin Taylor) committed to never-ending work on the details of content modernization.

REFERENCES

- * pfraundorf@ums1.edu
- [1] A. Einstein, *Einstein's 1912 manuscript on the special theory of relativity - a facsimile* (George Braziller, Inc., New York, 1996), p. 34.
 - [2] A. Einstein, *Annalen der Physik* **49**, 50 (1916).
 - [3] L. Randall, *Science* **296**, 1422 (2002).
 - [4] E. Taylor and J. A. Wheeler, *Spacetime physics* (W. H. Freeman, San Francisco, 1963), 1st ed., contains some material not found in the 2nd edition.
 - [5] E. F. Taylor, *Am. J. Phys.* **66**, 369 (1998), the Oersted Medal lecture.
 - [6] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL, 1949).
 - [7] E. T. Jaynes, *Phys. Rev.* **106**, 620 (1957).
 - [8] E. T. Jaynes, *Phys. Rev.* **108**, 171 (1957).
 - [9] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (Prentice Hall, New York, 1989), chap. 2.
 - [10] L. A. Girifalco, *Statistical Physics of Materials* (Wiley & Sons, New York, 1973), p. 72.
 - [11] C. Kittel and H. Kroemer, *Thermal Physics* (W. H. Freeman, New York, 1980), p. 39.
 - [12] K. Stowe, *Introduction to Statistical Mechanics and Thermodynamics* (Wiley & Sons, New York, 1984), p. 116.
 - [13] D. S. Betts and R. E. Turner, *Introductory Statistical Mechanics* (Addison-Wesley Publishers Ltd., Wokingham, England, 1992).
 - [14] C. Garrod, *Statistical Mechanics and Thermodynamics* (Oxford University Press, New York, 1995).
 - [15] T. A. Moore, *Six Ideas that Shaped Physics: Unit T* (McGraw-Hill, New York, 1998).
 - [16] D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, San Francisco, 2000).

- [17] T. D. Schneider, *Nucleic Acids Res* **28**(14), 2794 (2000).
- [18] S. Lloyd and J. Paul Penfield, *Hot bits: Seeing the world in terms of information and entropy* (1999), national Electrical Engineering Department Heads Association Annual Meeting, Destin FL.
- [19] J. Castle, W. Emmenish, R. Henkes, R. Miller, and J. Rayne, *Science by degrees: Temperature from Zero to Zero*, Westinghouse Search Book Series (Walker and Company, New York, 1965).