Hot and Heavy Matters in the Foundations of Statistical Mechanics

Craig Callender

Abstract Are the generalizations of classical equilibrium thermodynamics true of self-gravitating systems? This question has not been addressed from a foundational perspective, but here I tackle it through a study of the "paradoxes" commonly said to afflict such systems. My goals are twofold: (a) to show that the "paradoxes" raise many questions rarely discussed in the philosophical foundations literature, and (b) to counter the idea that these "paradoxes" spell the end for gravitational equilibrium thermodynamics.

1 Introduction

If one looks into the night sky one struggles to discern a pattern behind all the points of light. The lights look randomly distributed. Looks, of course, deceive. Examined in 3D, instead of in the 2D projection we see, and using probability theory, one finds many empirical statistical patterns lurking behind the panoply of lights. Galaxies, for instance, tend to appear in clusters, and these clusters are themselves scattered roughly as a Poison distribution in space. Within clusters, both the number and spatial distribution of galaxies seem to vary according to particular probabilistic laws. Is there a reason for these empirical distributions? In particular, is there a *thermodynamic* reason that the stars are distributed the way they are?

In asking this question, we engage in a serious idealization, pretending that the stars, etc., can be treated like the particles in an ordinary macroscopic system. We

---

Dedicated to the memory of Urs Hoffmann, from whom I learned much about these topics.

C. Callender (✉)
Department of Philosophy, UCSD, La Jolla, CA 92093, USA
e-mail: ccallender@ucsd.edu

are positioned like causally inert Maxwellian demons, able to see the "particles" in the system, wondering whether the science of thermodynamics results. How do we calculate the expected distribution? Fortunately, we know the mechanics, at least approximately. In astrophysics various stellar systems are well modeled by the classical N-body gravitational equation, e.g., globular star clusters ($N = 10^5$–10^6), galaxies ($N = 10^{12}$–10^{14}), and of course, planetary systems. Because solving the "million-body problem" is analytically intractable, physicists naturally turn to gravitational statistical mechanics.[1] So imagine a system composed of trillions of massive Newtonian particles in empty space. Assuming the particles have no charge, this system is a paradigmatic self-gravitating system, one whose behavior is primarily determined by (classical) gravitational interactions among its constituents. Suppose that there are enough particles to make coarse-graining into macrostates sensible. Impose the appropriate probability distribution over these particles. Now, will the generalizations of classical equilibrium thermodynamics tend to be true of this system? To find out, we first need to know (a) whether statistical mechanical techniques can be applied to such a system, and (b) if so, will the resulting patterns yield thermodynamical behavior?

To motivate these questions, notice that the answers are far from obvious. Pedestrian thermodynamic systems—steam engines and the like—are terrestrial systems, systems embedded in an effectively uniform gravitational field. Here the long-range nature of the gravitational force is not apparent. The system's behavior is primarily determined by interactions via short-range potentials. True, the Coulomb force is infinite in range too, but unlike that force the gravitational force doesn't saturate. In macroscopic systems wherein Coulomb forces dominate, the positive and negative charges effectively cancel out at certain scales and various screening mechanisms exist; for instance, in plasmas charges are screened at distances greater than the Debye length. At such scales the net interaction energy effectively vanishes, and this allows us to treat such systems as if they consist of many statistically independent subsystems—an assumption at the heart of statistical mechanics. Due to this non-saturation and related features, self-gravitating systems face many special difficulties when considered as statistical mechanical or thermodynamic objects.

These difficulties are sometimes dubbed the "paradoxes" of gravitational thermodynamics [21]. To the extent that physicists have dealt with these paradoxes, the community divides over the fate of gravitational statistical mechanics. Some believe that these paradoxes make ordinary equilibrium statistical mechanics inapplicable. In its place various recipes and techniques from statistical mechanics remain, to be sure, but the core foundations from Maxwell, Boltzmann and Gibbs are judged inadequate. By contrast, others try to show that statistical mechanics does apply.

Here is the chemist Rowlinson expressing a common sentiment:

> "[Thermodynamics] is essentially a human science; it started with steam engines and went on to describe many physical and chemical systems whose size is of the order of a metre. Its laws are not truly a theory but a highly condensed and abstract summary of our experience of how such systems behave. We have,

---

[1]For an entry to the relevant literature, see [8–10, 13, 21, 42, 49].

therefore, no right to expect them to apply to other quite different systems, whether extremely large or extremely small. They clearly are inapplicable to the solar system or to galaxies. Here gravity is the dominant force; there is no equilibrium, the energy is no longer proportional to the amount of material, and so there are no extensive functions. Clausius's famous remark that the energy of the universe is constant but its entropy is increasing to a maximum is derived from the behaviour of a closed adiabatic system of constant volume. The universe is neither closed in any classical sense, nor of constant volume. Clearly classical thermodynamics is not a useful branch of science in cosmology; we have extrapolated too far from its human-sized origins." [46, p. 873]

Similarly, Hut [23] writes that there are "fundamental problems that prevent a full thermodynamic treatment of a self-gravitating system of point particles" and that "the traditional road to equilibrium thermodynamics is blocked" (p. 41). Impressed by one of the paradoxes, Binney and Tremaine [3] likewise deny that a "fundamental physical principle" (p. 266), by which they mean Maxwell-Boltzmann statistical mechanics, is responsible for the configuration to which stellar systems settle. The claim that self-gravitating systems have no equilibrium, in particular, is the norm rather than the exception.

Against this opinion, one reads that

"...statistical mechanics of gravitating systems is a controversial subject. However, our modern understanding of statistical mechanics and thermodynamics does handle gravitational interactions rigorously with complete satisfaction." [29, p. 545]

"...on the contrary...[the] usual tools and ideas of statistical mechanics do apply to such systems, both at equilibrium and out of equilibrium." [4, p. 19]

"...not all was well with these three arguments [three of the above problems]." [49, p. 301]

None of these authors feel that the problems mentioned spell the end for gravitational thermodynamics.

No one denies that many statistical mechanical techniques can be applied to stellar systems. But are these merely residual mathematical tools with the core physical principles left behind? Answering this question relies on an antecedent division of statistical mechanical "core" and "peripheral" principles, and this is bound to be controversial. For myself, I believe that if one thinks of the core principles as those provided by Boltzmann, and elucidated by others (e.g. [35, 36]), then one will not see these gravitational paradoxes as a threat to statistical mechanics. They are surprising, but what they show is the delicate balance of factors that must obtain for textbook Gibbsian statistical mechanics to hold. If one does not mistake this textbook theory for statistical mechanics, then one may come away feeling that Boltzmann's original insights hold up even in the gravitational case. I'll leave this argument for another time. Fortunately, it's of interest how to respond to these challenges—regardless of how we categorize these principles.

In what follows I primarily wish to introduce the exciting foundational problems posed by the so-called "paradoxes" of gravitational statistical mechanics. As we'll

see, discussion of these paradoxes will raise many questions rarely discussed in the philosophical foundations literature: the meaning of equilibrium, our understanding of phase transitions, the (non)equivalence of Gibbsian ensembles, and the use of the thermodynamic limit. The paper's primary goal is to advertise the problems for further study by those in the philosophical foundations of physics—and hence it's longer on questions than on answers. However, where possible I do hope to make some progress on carving out space for gravitational thermodynamics; that is, although I'm certain that gravitational self-interaction pushes the envelope of a thermodynamic description, I am skeptical that the paradoxes spell the end for gravitational thermodynamics. But I am confident that these issues promise to teach us much about the foundations of statistical mechanics and the emergence of thermodynamics.

2 The System

Our first system is one composed of N particles, each of equal mass m, evolving in three-dimensional Euclidean space and interacting via a potential V. Such a system's Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + \frac{c}{2} \sum_{i=1, i \neq j}^{N} V(\mathbf{q}_i - \mathbf{q}_j) \qquad (1)$$

where c is a coupling parameter and $(\mathbf{q}_i, \mathbf{p_i})$ are the canonical coordinates of the i-th particle. Long-range forces are typically associated with two-body potentials $V(r) = 1/r^s$, where s is less than or equal to the spatial dimensionality of the system. In the case of self-gravitation, $s = 1$, $V(r) = 1/r$, and $c = -Gm^2$, so (1) becomes

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} - \frac{Gm^2}{2} \sum_{i=1, i \neq j}^{N} \frac{1}{|\mathbf{q}_i - \mathbf{q}_j|} \qquad (2)$$

(2) is of course Newton's famous N-body problem in gravitation theory. Although our emphasis is on gravity, gravity is not necessary for many of the puzzles that follow. There are non-gravitational systems that display some of the same peculiar behaviors, e.g., so-called geophysical fluid dynamical models ($s = 0$), some non-neutral plasmas ($s = 1$), dipolar ferroelectrics and ferromagnetics ($s = 3$), as well as some spin systems and toy models. (For information on these, and references, see e.g., [13].) Interestingly, many of the peculiar physical properties found in large stellar systems are also found in very tiny systems like nanoclusters. This is no accident. If we consider a system "small" if its spatial extension is smaller than the range of its dominant interaction, then a self-gravitating system is small too. In any case, the astrophysical applications of (1) make the gravitational case the most interesting.

3 Divergences

Self-gravitating systems have what might be called infrared and ultraviolent divergences. Let's begin by arguing that these problems aren't terminal.

The infrared divergence arises from the fact that the gravitational potential r^{-1} has infinite range. Since there is no shielding, strictly speaking the microcanonical density of states diverges. The density of states $g(E)$ is given by

$$g(E) = \frac{1}{N!} \int \delta(E - H(q, p)dqdp \qquad (3)$$

where E is the fixed energy, H the Hamiltonian, and N the number of systems. The above integral diverges if the range of spatial integrations is infinite. Since the Gibbs entropy $S(E)$ is essentially the logarithm of $g(E)$, this problem implies that the entropy diverges too, thereby preventing a sensible thermodynamics from emerging.

The ultraviolet divergence, by contrast, arises over short distance interactions. Here the problem is the local singularity of the Newtonian pair interaction potential. Two classical point particles can move arbitrarily close to one another. As they do so, they release infinite negative gravitational potential energy. Partition functions, which need to sum over all these states, thereby diverge. The entropy and other quantities are again threatened (see [28, 48]).

Consider the ultraviolet divergence. Strictly speaking, from the foundational perspective, this is a big problem, but it also affects Coulomb systems equally from a foundational perspective (that is, we can imagine systems composed of like-charged constituents without screening mechanisms). From that perspective, there is nothing new here. If we consider the less abstract question, however, and acknowledge that real-world Coulomb systems do have screening mechanisms, then there is a difference between the two. However, if we let a little bit of the real world into our question, that can also help. After all, we know that ultimately classical mechanics gives way to quantum mechanics. We don't want thermodynamics to fail in gravitational contexts just because classical theory ultimately fails. Or if we think of the practical application in astrophysics rather than the classical N-body problem, then the stellar radius provides a natural cutoff. Hence, it seems permissible to replace the gravitational potential with a softening or small cut-off potential. For instance, one can regularize the potential with the following replacement

$$\frac{1}{|\mathbf{q}_i - \mathbf{q}_j|} \rightarrow \frac{1}{\sqrt{(\mathbf{q}_i - \mathbf{q}_j)^2 + \eta^2}} \qquad (4)$$

where η is a softening parameter that bounds from below the interaction potential. This softening will prevent arbitrarily large momenta from developing.

The infrared divergence of $g(E)$ occurs because the integral is taken over an infinite range of position space. However, as has been pointed out by others, there is nothing special here. The same divergence occurs for the non-gravitating ideal gas in a spatially infinite universe. (That issue was independently a subject of controversy among philosophers of science; see, e.g., [43].) Since we don't take this divergence to prohibit a proper thermodynamics obtaining, we shouldn't in the gravitational case either. The solution in the non-gravitational case is to put the gas "in a box" and the same can be done in the gravitational case. Sometimes it's said that this is unwarranted in the gravitational case; after all, tidal effects and evaporation will always occur. However, I'm skeptical this is much different than in the non-gravitational case.

What we need to do is focus on the time and space scales. If a system is relatively isolated—if, say, its rate of evaporation is proportionally small—we can conceptually put a box around it. We can do so when the interactions with the "box" are much less dominant within the time scales of interest than the interactions among constituents. Since no screening is perfect even in the non-gravitational case, I don't see this justification as different in kind here.

The divergence problems, therefore, can be solved with plausible idealizations and approximations. The only stumbling block lay in mistakenly thinking such idealizations weren't needed all along in non-gravitational systems.

4 No Equilibrium?

Having tidied up our system, the next question to ask is whether the concept of equilibrium makes sense.[2] In thermodynamics, equilibrium is that state of a macroscopically isolated system wherein the thermodynamic observables remain constant. Crucially, when left to themselves, systems tend to head to equilibrium, and also, equilibrium is transitive, i.e., if system A is in equilibrium with B, and B with C, then A is also in equilibrium with C. Statistical mechanically, equilibrium is more complicated. According to a Boltzmannian view, it is usually defined as the macrostate with the largest volume in $6N$-dimensional phase space. In the Gibbsian picture, by contrast, equilibrium is instead a feature of a probability distribution, and it is defined as a distribution that is stationary in time, thus mimicking the thermodynamic definition.

Many physicists voice skepticism about there being an equilibrium state for self-gravitating systems. The reasons for this judgement hang on a motley assortment of facts, and its sometimes hard to tease out a precise argument.

To begin, suppose the N-body system is *not* put in a box. Then if the system is isolated—and hence one uses the microcanonical ensemble—one can prove that there is no entropy maximum. Since equilibrium is sometimes defined as the state of maximum entropy, this result implies that there is no equilibrium.[3] If the system is instead connected to a heat bath—and hence one uses the canonical ensemble—one can similarly prove that the Gibbs free energy has no maximum. Since equilibrium in the canonical case is associated with the maximum of the Gibbs free energy, one again has an argument that there is no equilibrium. These arguments, of course, simply depend on the above infrared divergence argument. If we don't put the system in a box, at least conceptually, then even short-range "normal" systems may have no entropy maximum in the microcanonical ensemble and no free energy maximum in the canonical ensemble. We should put systems in a box, for thermodynamics is really about systems that are approximately stationary with respect to some thermodynamic observables for certain spatiotemporal scales.

Matters become more interesting when we do put the system in a box. If we consider the system isolated, then there are various arguments, stemming from both fundamental theory and approximations—and verified in computer simulations—that if

---

[2]See Callender [7] for a non-equilibrium counterpart of this paper.

[3]With self-gravitating systems in mind, some have challenged this identification [34].

we begin with a system in a stably bound "core-halo" state, as it develops with time the central core will become increasingly concentrated without end. The core loses energy to the surrounding "halo" of particles, but because of its negative heat capacity, it will get hotter as it contracts. If the surrounding halo can't warm quickly enough, then we get runaway instability. The system gets more and more inhomogeneous, both in configuration and velocity space, without limit. Hence its entropy increases without limit. For any state—since we're dealing with point particles—there is always a higher entropy state found by making the core denser and the halo more diffuse.

This claim is also bolstered, and indeed, initially inspired, by the notorious *gravothermal catastrophe*. Try to find the density distribution $f(x, v)$, defined on 6-dimensional μ-space, that maximizes the Boltzmann entropy subject to constraints on the total mass, energy, linear and angular momentum. It turns out that this variational problem has no solution in the gravitational case for some constraints. In fact, it has no solution even when we artificially confine the system to an arbitrarily shaped "ball" of physical space and allow radial asymmetries [45]. Hence the catastrophe. Discovered by Antonov [1] and dramatically described by Lynden-Bell and Wood [39], the catastrophe has been confirmed in many computer simulations. It shows that for some values of the constraints there is no distribution that maximizes Boltzmann entropy.

In the canonical case, there is a similar catastrophe, sometimes called the isothermal catastrophe. This is a state wherein, in configuration space, all of the particles sit at the same material point. It is represented by a Dirac peak in the canonical ensemble. Equilibrium then corresponds to this energy non-conserving state that maximizes the Gibbs free energy.

True equilibrium, then, is said to correspond to one or the other of these unobtainable states: a Dirac peak where the boxed system is attached to a heat bath, or an endlessly evaporating-halo, shrinking-core system when the system is isolated. Workers in the field are then quick to point out that nonetheless, if the energy is high, there are many meta-stable states (local entropy maxima) that function effectively like equilibrium in the sense of having stationary thermodynamic observables for long time periods. Although in principle the gravothermal catastrophe can occur in a finite amount of time, this amount of time is calculated to be longer than the age of the universe. The meta-stable states are then regarded as effectively the equilibrium states.

Should Dirac peaks and runaway core-halo systems count as equilibrium? One reason against doing so no doubt stems from a combined intuitive repugnance coupled with skepticism about the coherency of the presumed equilibrium states. Equilibrium, it is thought, shouldn't be represented by monstrosities like those states. Kiessling [28], however, counsels that our ordinary notion of thermal equilibrium originates from the non-gravitational cases; therefore, it shouldn't bias us against unorthodox states appearing in more unfamiliar systems (205). With the intuitive concerns put aside—as I think they should be—the question becomes whether such "singular" states are well-defined and possibly part of a functioning statistical mechanics. As it turns out—and this is a major result in the field—the Dirac peak in

the canonical case is well-defined and acts like an equilibrium state.[4] One can object that, for the astrophysical problem, the microcanonical ensemble is more appropriate (galaxies and clusters aren't attached to heat baths). Yet if we focus on the foundational question, whether equilibrium exists for our modified N-body problem, then a powerful argument can be made that there is.

That said, I want to suggest another tack. I don't want to minimize the differences between the self-gravitating case and the normal case. There are real differences, for instance, in the maximization problem that inspires the gravothermal catastrophe. However, I think they can be overblown. I think there is 'no equilibrium' for these systems only if one employs an overly strict notion of equilibrium.

Equilibrium, in thermodynamics and statistical mechanics, is an idealization. Though defined as a stationary state, if the system is mechanically Hamiltonian, then there can't really be any *truly* stationary states. Poincaré recurrence implies that states will eventually evolve out of equilibrium if ever they weren't in equilibrium. Equilibrium holds only with respect to certain observables, spatial scales and temporal scales [6]. As Ma [40] stresses, in the span of a few seconds, the measured volume and temperature of boiled water poured into a tea cup will not change significantly, and hence we can regard the system as being in equilibrium. Wait a while and eventually the water's temperature will decrease until it is roughly equal to room temperature; once again we have equilibrium. But in a few days the volume will vary as evaporation sets in. Then we must wait until the cup is empty for another equilibrium state. Wait even longer and the molecules in the cup will evaporate, and on the scale of many years the system will again fall out of equilibrium. Statistical mechanics' identification of equilibrium with a stationary probability distribution pretends that there is what Ma called "absolute equilibrium." But absolute equilibrium is a fiction and not at all connected to the states we normally call equilibrium states in thermodynamics. From this perspective, the meta-stable states are equilibrium enough.

Besides, the Dirac peak and core-halo states are, like the first worry, the result of the above divergence worries. In this case, the runaway behavior stems from the absence of a short-distance cut-off. And as mentioned, in the practical issue, there is abundant reason to introduce a cut-off, e.g., at stellar radii. In the purely theoretical case, it's of course interesting that these states result when no short-distance cut-off is imposed. But if one does not impose a priori demands on what an equilibrium state—or perhaps better—absolute equilibrium state should look like, then when they are well-defined there is nothing to object to.

Does the lack of an entropy maximum state hurt the Boltzmannian picture? No. Two points are important to make here.

---

[4]Since the momentum and configurational equilibrium densities factorize in the canonical ensemble, yielding a trivial Gaussian measure over the momentum sector (see [28, p. 213]), Kiessling (and others) focus on the configurational sector. The question posed is whether the canonical equilibrium measure can describe this collapsed state in $3N$-dimensional configuration space. Kiessling's proof that it can for three-dimensional systems examines modified gravitational potentials with smoothed out singularities. He shows that in the limit, as these potentials approach the exact Newtonian potential, the canonical ensemble converges on a superposition of Dirac distributions each describing a collapsed state. Which exact spot will suffer the collapse is equal to a probability density that depends on the external potential and a Boltzmann-like factor. Kiessling is then able to show that this work connects up with thermodynamics in the mean field limit.

First, the demonstration of the catastrophe is within astrophysics and awash in idealizations. These idealizations are perfectly natural ways of modeling some astrophysical phenomena (see [24]), but they do limit the applicability of this result for our foundational question. The Boltzmann entropy maximized in the gravothermal catastrophe is the one over the one-particle distribution function f in 6-dimensional μ-space, i.e., $S(f) = -k_B \int \int f \log f \, d^3\mathbf{x} d^3\mathbf{v}$, not the volume corresponding to the macrostate M in $6N$-dimensional Γ-space, i.e., $S = -k_B \log |\Gamma_M|$. The former is fine as an approximation of many stellar systems, for they don't have many close encounters or collisions. But at many stages of the above process additional physical mechanisms will become important. If we think of the logarithm of the $6N$-dimensional volume as the "real" entropy, then we've only shown that an approximation for some spatial and temporal scales doesn't have an equilibrium point. Of course, how one should define macrostates if the usual prescription yields trouble is a problem, but not one unique to the gravitational case.

Second, what ultimately matters to the Boltzmannian is that the transition probabilities for going from one macrostate to the other are calculated via the relative size of the corresponding volumes in $6N$-dimensional phase space. This can be done even if there is no unique maximum entropy macrostate. Recall that the Boltzmannian picture works, in principle, for equilibrium *and* nonequilibrium systems. As Lavis [35] stresses, it's best to regard equilibrium not as an all-or-nothing state, but instead as a spectrum of states corresponding to a spectrum of sizes of volumes in phase space. How "big" should the volume corresponding to a macrostate be before we dub it 'equilibrium' in this picture? The answer is to some extent arbitrary. From this perspective, I don't see why meta-stable states can't be regarded as equilibrium states for the same reasons cups of coffee at room temperature can be.

Finally, let me end this section by acknowledging that there may be threats to equilibrium from other directions. It may be that some features we desire of equilibrium cannot obtain for some self-gravitational systems. I'll describe such a possibility later. But as that threat hangs on many open questions and the usual arguments against equilibrium are wanting, we'll tentatively conclude that equilibrium does make sense in self-gravitating systems.

5 Non-extensivity

The problem we want to consider arises from the fact that entropy and energy seem to not even be approximately extensive or additive for self-gravitating systems. Yet without extensivity or additivity holding, it's not clear how much of equilibrium thermodynamics or statistical mechanics survives.

Before seeing why, we must first clarify terms, as many inequivalent definitions of extensivity exist in the literature and this proliferation leads to a number of (apparently) conflicting claims about the relationship between extensivity and additivity. Often used interchangeably, the two concepts can be different (suitably defined) and this difference can be important for the systems of interest.

Thermodynamics and statistical mechanics assume that, when a system is large enough, it can be divided into composite subsystems; alternatively, that one macro-

scopic system can be combined with another into a third macroscopic system. Extensivity and additivity are features that apply to the interrelationships among the subsystems of a compositive system. Additivity is normally straightforward. Let x be the system and $x_1, \ldots, x_n$ be its subsystems. A function f over these is additive just in case $f(x) = f(x_1) + f(x_2) + \cdots + f(x_n)$. But extensivity is defined many inequivalent ways in thermodynamics [16]. Intuitively put, extensive quantities are those that depend upon the amount of material or size of the system (such as the mass, internal energy, entropy, volume and various thermodynamic potentials), whereas intensive quantities are those that do not (such as density, temperature, and pressure). Commonly it's claimed that extensive quantities are those that are halved when a system in equilibrium is divided into two equal-sized subsystems. Alternatively, we might say a function f of thermodynamic variables is extensive if it is homogeneous of degree one. If we consider a function of the internal energy U, volume V, and particle number N, homogeneity of degree one means that
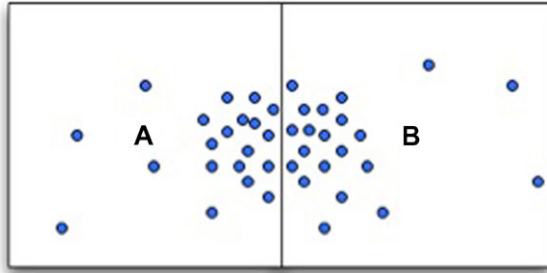
$$f(aU, aV, aN) = af(U, V, N) \tag{5}$$

for all positive numbers a. Either claim neatly encodes the experimental facts. Consider a box of gas in equilibrium with a partition in the middle and consider the entropy, so that $a = 2$ and $f = S$. Then (5) describes the experimental fact that if we halve the system we also halve the entropy. The same goes for the energy. Thermodynamically, we can take energy or entropy to be the dependent variable, $U = U(S, V, N)$ or $S = S(U, V, N)$; if one is not extensive then the other is not. In neither sense of extensivity, however, does extensivity strictly imply additivity [16]. Additivity does imply extensivity, by contrast.

Statistical mechanically, matters are subtler. One reads that additivity implies extensivity but not the reverse in one place [9] but that they are logically independent in another [53]. In fact there is little confusion—just different definitions. A common definition of extensivity in statistical mechanics is that a thermodynamic function f is extensive iff its f-density, defined by $f(x_n)/n$, attains a constant in the infinite n limit. This definition neither implies nor is implied by additivity. If, by contrast, all one means by extensivity is that thermodynamic variables are proportional to system size, i.e., then additivity implies extensivity again, but extensivity doesn't imply additivity. There are systems whose energies scale with N yet which do not have additive energies, e.g., the Curie-Weiss model.

However we arrange the logical geography, it's clear that the assumption that energy is extensive is plausible partly because in terrestrial cases we are usually dealing with short-range potentials. At a certain scale matter is electrically neutral and gravity is so weak as to be insignificant. If the potential is short-range and our subsystems aren't too small, then the subsystems will interact with one another only at or in the neighborhood of their boundaries. When we add up the energies of the subsystems, we ignore these interaction energies. The justification for this is that the interaction energies are proportional to the surfaces of the subsystems, whereas the subsystem energies are proportional to the volumes of the subsystems. So long as the subsystems are big enough, the subsystem energies will vastly trump the interaction energies as the number of subsystems increases because the former scale as $(\text{length})^3$ and the latter as $(\text{length})^2$. Of course, we are also assuming that the subsystems sizes aren't

chosen in too contrived a way. Bulk effects will not outpace surface effects if one chooses, say, very thin subsystems with large surface areas and small volumes. Normal systems that are macroscopic and screened after a certain distance at the macroscopic level, however, warrant the claim that energy (etc.) are extensive and additive.

However, if gas molecules are replaced by stars—that is, short-range potentials replaced by long-range potentials—this reasoning doesn't work. The interaction energies may not be proportional to the subsystem surfaces. For short-range potentials, the dominant contribution to the energy comes from nearby particles; but for long-range potentials, the dominant contribution can originate from far away particles. To drive home the point, consider a sphere filled with a uniform distribution of particles. Now add a particle to the origin and consider its internal energy U:

$$U \propto \int_0^R 4\pi r^2 dr \rho r^{-3-\epsilon} \propto \int_0^R dr r^{-1-\epsilon} \qquad (6)$$

One can verify that with $\epsilon > 0$ ("short-range potentials"), the significant contribution to the integral comes from near the particle's origin, whereas with $\epsilon < 0$ ("long-range potentials"), the contribution comes from far from the origin [42]. The interaction energies do not become negligible as the system grows.

As a result, think of what happens in the standard case of a chamber of gas divided into two equal boxes, A and B (see Fig. 1). If the particles are interacting via long-range forces, the particles in box A will feel the particles in box B as much or even more than the particles nearby. Let E_A represent the energy of box A and E_B represent the energy of box B. Because $E = E_A + E_B + f(E_A, E_B)$ where $f(E_A, E_B)$ is the non-negligible interaction energy, it is easy to devise scenarios whereby $E_A = E_B = -a$, where $a > 0$, yet where the energy of the combined system E vanishes. The energy might not be even approximately additive, in the usual sense of additive. Same goes with the entropy.[5] Depending on the Hamiltonian, however, it may be that the energy nonetheless scales with N and hence one sometimes sees the claim that systems are extensive but not additive.

Of course, the problem is really one of degree. Real finite systems aren't perfectly extensive. Nonetheless, there is a genuine physical difference between the gravitating and non-gravitating cases, for one has screening mechanisms the other doesn't.

---

[5]There are ways of re-defining the entropy or energy so that one but not both are non-additive; see [25].

It's important not to let the distinction collapse merely because it's ultimately one of degree. One way of collapsing the distinction appears in "rigorous" treatments of statistical mechanics. Here one tries to regain the exact formal properties of thermodynamics despite the fact that we know it holds only approximately. Elsewhere I have dubbed this practice "taking thermodynamics too seriously" (see [6]) and pointed out that it can often lead us astray. In this case it is simply a matter of semantics. In approaches to statistical mechanics like Ruelle [47], it's commonly said that no finite system is ever extensive. What is meant is that when one defines the entropy and energy via the partition function, one is unable to reproduce the distinction between extensive and intensive for any finite system. No matter how large the system, if it's finite, surface effects contribute to the partition function. Entropy and energy are claimed to be truly extensive only in the thermodynamic limit, where N, V go to infinity while the ratio N/V is held constant. In this field a variable f is extensive if the thermodynamic limit of f is infinity while the thermodynamic limit of f/V is constant. Strictly speaking, finite systems, gravitational or not, are never extensive.

But no matter. It's still a fact that self-gravitating systems, unlike Coulomb systems, have no built-in tendency causing them to be even approximately extensive. This physical difference, although invisible in certain contexts, is no less real for being a matter of degree. Indeed, the cost of the energy and entropy being non-extensive in this sense is quite profound.

5.1 Extensivity in Thermodynamics

Let's concentrate initially on the costs incurred by classical phenomenological thermodynamics. The trouble with doing so is that there is no one canonical way of developing the theory. Some authors take as axioms what others derive, and vice versa. Still, it seems that non-extensivity wreaks havoc with traditional thermodynamics no matter the route by which it reverberates through the subject. Reif [44] and Landsberg [33], for example, take extensivity as axiomatic.

Focusing briefly on one example, note that the first three postulates of Callen's [5] now classic textbook explicitly or implicitly assume the extensivity of the system. Postulate III states that "the entropy of a closed system is additive over the constituent subsystems" (25), so additivity is treated as an axiom. Postulate II requires that entropy is a function of the extensive parameters of the system in equilibrium. In that case, if energy is not extensive then, strictly speaking, entropy is not a function of energy. Even Postulate I, the existence of equilibrium states, tacitly requires extensivity of the entropy and/or energy, for Callen declares an equilibrium state as being completely characterizable by extensive parameters. Of course, there are extensive parameters in addition to energy and entropy, but a formalism based around, say, volume and mole number, would scarcely be recognizable as thermodynamics. With the first three basic postulates impacted by the loss of extensivity, scores of important implications are lost. For instance, one cannot derive:

- that subsystems mutually in equilibrium have equal temperatures (Callen, p. 37)
- that when a system is in equilibrium, its large subsystems also will be in equilibrium

- that the important Gibbs-Duhem equality holds
- that the Second Law holds (e.g., in [37]).

The reader is invited to go through a thermodynamics text and circle all the derivations lost by removing the assumption of extensivity. I'll assume, however, that the above list suffices as an indication of the damage.

5.2 Extensivity in Statistical Mechanics

Moving to statistical mechanics, matters are in many ways more interesting. However, in the end, here too the approximate extensivity of the energy is often taken as axiomatic.

For example, Landau and Lifshitz [32] take the statistical independence of subsystems and the associated additivity as perhaps their most basic assumption, for through a chain of reasoning they use it to ground the "most important conclusion... [that the] values of the additive integrals of motion... completely determine the statistical properties of a closed system" (p. 11). With this, they write, it is possible to replace the "staggeringly large data" of the mechanical approach with a simple statistical distribution, i.e., the microcanonical ensemble, the basis of equilibrium statistical mechanics (p. 11). It is not an exaggeration to say that *additivity of energy*, in their treatment, *is what makes statistical mechanics possible*.

To take another example, consider Ruelle's [47] elegant and influential treatment of statistical mechanics. In his hands, "normal" thermodynamic behavior is possible only when two conditions hold, stability and temperedness. We won't spend time describing these conditions here, but suffice to say, in self-gravitating systems stability is violated [8, 27]. Interestingly, if we make our potential repulsive rather than attractive at large distances, then the temperedness condition is violated [8, p. 11]. So one way or the other, long-range forces spell trouble for thermodynamic behavior as Ruelle understands it.

As in the thermodynamic case, one finds extensivity assumed throughout many treatments of statistical mechanics. Not surprisingly, since statistical mechanics grounds thermodynamics, it's assumed in statistical mechanics in many of the same places it is in thermodynamics. However, let's note a few additional features of interest:

- The justification of the canonical ensemble is considered parasitic upon the justification of the microcanonical ensemble. Large research programs are devoted to justifying the choice of a uniform measure for a closed system. Once that is in hand, or even assumed, then one gets the justification of the more widely used canonical ensemble for free. That is, one justifies the canonical ensemble by assuming the system of interest is a small subsystem of a much larger system itself described by the microcanonical ensemble. With the added assumption of additivity of energies, one can then derive the canonical ensemble as an appropriate probability distribution for the subsystem. Without this additivity, one expects strange non-canonical behavior to emerge from non-extensive systems in contact with heat reservoirs.
- Non-extensive systems have no conventional thermodynamic limit. The thermodynamic limit is one where we normalize the extensive variables by the volume V or

particle number N, e.g., E/N, thereby making the ratios intensive; then we let N, $V \to \infty$ while keeping these intensive ratios finite. For a rigorous discussion of this problem see [42] and references therein. Intuitively, this fact is expected. The existence of the thermodynamic limit depends on making the contribution of surface effects go to zero as $N, V \to \infty$. But with a non-additive system, we saw that the surface effects aren't going to get smaller as N and V increase. If we have a self-gravitating system in a box, as we increase the size of the box, the total kinetic energy will scale with the mass, which in turn scales with the size of the box: $E_{kin} \propto M \propto R^3$. However, the potential energy E_{pot} will grow much faster: $E_{pot} \propto M^{5/3}$. This makes the specific gravitational potential energy of the system, the total potential energy divided by N, increase indefinitely as N increases (see [21]). Another way the lack of a thermodynamic limit is unsurprising is that the connection between it and extensivity is sometimes especially direct. Beck and Schloegl [2], like many others, state that a system is said to possess a thermodynamic limit only when its total entropy and energy are extensive quantities (in the rigorous sense described earlier). And adding one more link to the chain, Ruelle proves that the thermodynamic limit exists only for interactions that are thermodynamically stable, whereas non-extensive systems aren't stable. Ironically, the thermodynamic limit doesn't apply to very "large" systems.

- There can be negative specific heats in self-gravitating systems [38]. The specific heat capacity C is a measure of the amount of heat energy required to raise the temperature of a material one degree Kelvin. In classical thermodynamics, C is always positive. For such a system, adding energy makes the temperature increase and removing energy makes the temperature decrease. Because it is so closely tied to the stability of a thermodynamic system, some texts treat positive C as a basic principle of thermodynamics (e.g., [52]). We'll return to this topic in Sect. 7. Suffice to say, negative specific heat is quite counter-intuitive; furthermore, it is connected to the onset of phase transitions and ensemble inequivalence, two fascinating topics.

6 Extensivity Regained?

The physics community's reaction to non-extensivity divides into three broad camps. One group seeks to extend/modify equilibrium statistical mechanics to cover cases of non-extensive energy and entropy. Another uses the problems mentioned as a reason to ditch equilibrium statistical mechanics in favor of non-equilibrium techniques. But probably the majority of physicists pursue a conservative approach. Instead of modifying or abandoning equilibrium statistical mechanics, they continue to use it but seek approximations, limiting regimes, and so on, wherein extensivity is regained. In this section I'll focus on this last response. Since the other two groups pursue agendas outside the scope of this paper, I'll ignore them apart from a little discussion of the first group in Sect. 7.

Divide our N-body system up into two cells, 1 and 2. Due to the gravitational correlation energy U_{int} between cells, we've seen that the internal energy isn't extensive: $U = U_1 + U_2 + U_{int}$. Yet we can still ask whether there are systems for whom

$U_{int} \ll U_1 + U_2$. If so then extensivity is a good approximation and ordinary equilibrium statistical mechanical techniques are used again. Let's investigate two ways of regaining extensivity.

First, here is a quick and loose argument one sometimes sees to show that entropy can be extensive for self-gravitating systems. Define the entropy via the standard

$$\frac{1}{T} = \frac{\partial S}{\partial E} \tag{7}$$

where E is the total energy and T the temperature. Then make two assumptions: (a) that the system is virialized, and (b) that the temperature is proportional to the mean kinetic energy. The first assumption holds if the system is stably bound. Then the virial theorem states that

$$\langle K \rangle_\tau = \frac{1}{2} \langle V \rangle_\tau \tag{8}$$

where τ is the time over which the averages are taken, and K, V, the kinetic and potential energy, respectively. The second assumption is that

$$T \propto \frac{\langle K \rangle}{k_B(3N - 6)} \tag{9}$$

where k_B is Boltzmann's constant. As one can quickly inspect, assuming that $\langle K \rangle_\tau \approx \langle K \rangle$, it then follows that both E and T are (similarly) N-dependent. Hence, the entropy, defined in terms of these notions, must also be N-dependent and therefore extensive.

This argument is problematic at every step. The virial theorem is a theorem of mechanics stating that *if* the average time derivative of a function G known as the scalar virial is zero, then (8) holds. In astrophysics the antecedent is often ignored if the system is stably bound for a long time. The reasoning is that if G is bound between G_{min} and G_{max} for a long time, as happens in orbits, then its average is zero. For the time scales of interest, this assumption is usually perfectly innocuous. However, a priori there is no reason to think these time scales are equal to the relevant thermodynamic time scales. Being in "dynamical equilibrium" doesn't imply that one is in thermal equilibrium (or vice versa). That the system is stably bound for a time period τ doesn't in the least imply that we can assign it a thermodynamic temperature. Furthermore, even if we do assume the system is in thermal equilibrium, it doesn't follow that its temperature is proportional to kinetic energy. There are many systems—and in fact some gravitating systems—where this is not so. The quick argument needs more justification before it can be employed in general. Nonetheless, it does at least show that if these assumptions can be consistently employed, then the energy and entropy can be extensive even if the system is self-gravitational.

The more common method of regaining extensivity is by going to the mean field limit. First, one rescales (2) by what is known as the "Kac" prescription [26]. That is, one chooses the spatial and temporal parameters so that the coupling parameter $c = -Gm^2$ becomes $c = \pm 1/N$. Hence (2) becomes

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2N} \sum_{i,j=1}^{N} V(\mathbf{q}_i - \mathbf{q}_j) \tag{10}$$

Now, keeping all parameters fixed, and in particular the volume, take the $N \to \infty$ limit. This differs from the usual thermodynamic limit, for in that limit the volume V also goes to infinity. The resulting limit has a number of nice features, for example, from it one can obtain the collisionless Boltzmann equation, which is quite useful in astrophysics. For our purposes, however, what is important is that the energy H is proportional to N at this limit, and hence, the energy is extensive. In the literature it is commonly asserted that the $N, V \to \infty$ limit is appropriate for short range systems but the $N \to \infty$ limit is what is appropriate for long range systems. (In a similar development, de Vega and Sánchez [14] argue that in the so-called "dilute limit" where $N, V \to \infty$ while $N/V^{1/3}$ stays fixed, the internal energy, free energy and entropy become extensive; however, this claim has been disputed [30].) We'll return to these limits in the next section.

Finally, I should mention two further regimes in which extensivity can be regained, but which I won't describe in detail here. They can be found in [49], who focuses especially on infinite N systems and the canonical ensemble. Although the first justification departs from our Newtonian N-body situation, it is too good to ignore: Saslaw [49] shows with a neat argument that the expansion of the universe effectively cancels out the long-range part of the gravitational force. Expansion limits the bulk effects of gravity to certain length scales. So there is a gravitational kind of screening after all. And at these length scales the energy is approximately additive. The second justification focuses on features of the so-called correlation length used in astrophysics, e.g., its steepness, and shows that in some cases it too will leave the interaction energy negligible for certain spatial and temporal scales. In sum, there are some regimes where extensivity is regained.

7 Foundational Questions

Gravitational thermodynamics stimulates many foundational questions in statistical mechanics. We have already broached some of them—e.g., what is equilibrium, whether systems are strictly extensive, the status of divergences, the problem of negative specific heat—and no doubt each can be discussed in far greater detail. Although there is much one could discuss, let me finish by concentrating on four philosophical questions little discussed in the literature.

7.1 Statistical Mechanics Without Thermodynamic Limits—and Which Limits?

The lack of a well-defined conventional thermodynamic limit is prima facie disturbing. Many approaches in the foundations of statistical mechanics crucially rely upon this limit. It is also said to provide the resolution of various "paradoxes" in statistical mechanics [50]. I want to suggest one general heretical claim, and then raise another.

First the heretical statement: the thermodynamic limit is neither necessary nor sufficient for thermodynamic behavior. That it is not necessary for thermodynamic behavior is exemplified by gravitational systems that display thermodynamic behavior, but also all the many, many systems that do not meet the exacting conditions for having a thermodynamic limit. That it is not sufficient can be seen as follows. Suppose we do have a proof, for instance, that in the thermodynamic limit a system is

exactly additive: doubling the E, V, and N, say, doubles the entropy. Does it automatically follow from the existence of this limit that real finite N systems are additive or even approximately additive? If the limit is not smooth, not much may follow about the properties of large N systems. All the usual "short-run" versus "long-run" worries can be raised. The existence of thermodynamic property P obtaining in the thermodynamic limit is no guarantee by itself that real systems exemplify some approximation of P.

For this reason, I regard the lack of a conventional limit in the gravitational case as more edifying than problematic. The emphasis on the thermodynamic limit in the foundations of the subject often arises from a misplaced desire to show that thermodynamics holds *exactly* of statistical mechanical systems. However the goal of statistical mechanics is to explain the thermodynamic patterns of observables, not the thermodynamic formalism [6]. The thermodynamic limit is sometimes useful in this explanatory enterprise, but it should not replace it.

Second, the specific question: what should we make of the use of unconventional limits? In the previous section we saw that extensivity can be regained by using an unconventional thermodynamic limit. In fact, it's not uncommon for physicists to use modified thermodynamic limits to achieve their ends. What should we make of this practice? Is one limit the "right" physical limit and the other the "wrong" one, given the goals? What makes a limiting procedure correct? Is a pragmatic justification, i.e., that the limit succeeds in getting the desired results, enough? Or do we require a more physically based justification? One naturally wants to be on guard against reverse engineering: assuming that extensivity holds and then taking the limits needed to get there. This worry arises in the present case. Laliena [30, 31], for instance, charges de Vega and Sánchez [15] with imposing "ad hoc" conditions to obtain a well-defined "dilute" thermodynamic limit. This criticism presupposes that some restrictions are legitimate and others not. By contrast, while not exactly denying this, in a side remark Bouchet and Barre [4] approach the topic by demanding a limit that gives physics largely independent of how large N is. That is, they want to find the physics of effectively intensive variables. But in the context of the foundational question of whether one can regain extensive versus intensive variables, this motivation would seem to beg the question. It says: take whatever limit is needed to recover the 'intensive versus extensive' split. Obviously a criterion of some kind is needed to determine what limits (with what restrictions) justify what physical claims. What is it?

7.2 Ensemble Nonequivalence: Will the True Ensemble Please Stand?

In mainstream Gibbsian statistical mechanics, the choice of probability distribution or ensemble is widely deemed a matter of convenience. One can choose the microcanonical, canonical, or grand canonical ensemble. They are supposed to be merely different windows onto the same physics, each corresponding merely to the fixing of different thermodynamic parameters. The physics isn't supposed to hang essentially on our choices modeling. This belief is justified in textbooks by gesturing at facts like the following. Focusing on the canonical and the microcanonical ensembles, when N goes to infinity, the energies of most configurations in the canonical ensemble are equal to the average energy; hence in the limit one gets systems all of

the same energy, i.e., the microcanonical ensemble. Reasoning like this leads to the common belief that the three ensembles are equivalent in the conventional thermodynamic limit.

Of course, matters are not so simple. More than a century ago Gibbs [19] pointed out inequivalence for systems undergoing phase transitions, something treated as a loophole by some textbooks. And more careful treatments acknowledge that the so-called "problem of equivalence" has been a foundational problem for a long time.

With self-gravitating systems, it seems that this problem is all the more pressing. Without a well-defined conventional thermodynamic limit, we cannot prove equivalence of ensembles. In self-gravitating systems, the three ensembles can describe very different physics. In fact, in toy self-gravitating systems, one can demonstrate strikingly divergent physical behaviors for self-gravitating systems between the microcanonical and canonical ensembles. The former, for instance, allows negative specific heats whereas the latter cannot. The situation has now changed: it's not just that we don't know (i.e., haven't yet proved) that the three ensembles are equivalent for some systems, excepting phase transitions; it's that now we know that they're not equivalent for some systems. On its face, this situation seems dire. Gallavotti [18] writes that the problem of equivalence "is a fundamental problem because it would be a serious setback for the whole theory if there were different orthodic ensembles predicting different thermodynamics for the same system" (69). And Huang [22] writes, "From a physical point of view, a microcanonical ensemble must be equivalent to a canonical ensemble, otherwise we would seriously doubt the utility of either" (148).

What this nonequivalence means for the entire Gibbsian program has yet to be considered from a foundational perspective, but it would seem to call for serious study. A number of questions face us. Do we treat one ensemble as distinguished? Much of the gravitational thermodynamics literature does do this—Gross [20] most vocally. They understand the real physics as being described by the microcanonical ensemble. And since one needs additivity to derive the canonical from microcanonical ensembles, one cannot derive the canonical from the microcanonical even if they aren't equivalent in the thermodynamic limit. Thus Combes and Robert [11] write that as a result of non-extensivity "the canonical formalism of Gibbs is no longer justified". Others, by contrast, stick with the canonical description as basic. It is, after all, the easiest to use, and so there is real incentive to standing by it. Still others consider moving to the use of new ensembles. Costeniuc et al. [12] define a new "generalized canonical" ensemble by modifying the Lebesgue measure from which the canonical ensemble is crafted. They then show that this 'generalized canonical ensemble' is equivalent to the microcanonical measure in the conventional thermodynamic limit. They also consider the little-discussed Gaussian ensemble, which is the ensemble said in the first instance to be appropriate to a system in contact with a finite heat reservoir.

For myself, I don't think inequivalence is so surprising. After all, a system with an infinite heat bath attached to it and one without are quite different physical systems! What's surprising is that for some systems this difference *didn't matter* to *some* thermodynamic observables. In any case, the phenomenon of ensemble nonequivalence and its connection to features such as non-concavity of the entropy, phase transitions, and negative specific heat have been the subject of an awful lot of research lately (for an introduction, see [54]) and it begs for philosophical appraisal.

7.3 Negative Specific Heats

The specific heat capacity C is a measure of the amount of heat energy required to raise the temperature of a material one degree Kelvin. For constant volume, it is defined as

$$C_V = \frac{\partial U}{\partial T} \tag{11}$$

where U is the internal energy and T the temperature. In classical thermodynamics, C is always positive. For such a system, adding energy makes the temperature increase and removing energy makes the temperature decrease. But as mentioned earlier, in the gravitational literature systems with negative heat capacity are a commonplace. Eddington long ago mentioned that stars and star clusters would cool down if energy is added. Then in the 60's and 70's many astrophysical models appeared that displayed negative specific heat. As Lynden-Bell [38] describes matters, negative specific heat was a "paradoxical" result, but the paradox was dispelled by Thirring [51], who showed that although the specific heat is always positive in the canonical ensemble it can be negative in the microcanonical ensemble.

If there really are negative specific heats, however, they raise many questions worthy of further study. To name one, can systems of negative specific heat really be in equilibrium? Suppose we say that it's essential to equilibrium that it be transitive. Then on its face it seems that systems that are otherwise in equilibrium do not satisfy transitivity. This claim relies on two premises (see [16]):

Premise 1: Two negative C_v systems cannot reach equilibrium if put in thermal contact with one another.

Suppose system A loses energy to system B. Then because of the negative C_v, the one losing energy gets hotter and the one gaining energy gets cooler.[6]

Premise 2: a negative C_v system can be in equilibrium with a positive C_v system.

Suppose we place a negative C_v system A into thermal contact with a positive C_v system B. Assume the negative C_v system begins with a slighter hotter temperature than the positive C_v system, so that heat flows from A to B. But that means that A will get hotter as it loses energy. But B, given its positive C_v, will also get hotter as it increases energy. And in fact it will get hotter quicker than A will, for A's—C_v—is less than B's—C_v—because A started off hotter. So B's temperature can catch up to A's temperature. They can therefore reach an equilibrium where they are both hotter than they were before!

With these two premises, transitivity of equilibrium is easily violated. Let systems A and C have negative C_v and system B have positive C_v. Then A can be in equilibrium with B, B in equilibrium with C, but never A in equilibrium with C.

This argument is only one possible clash with ordinary thermodynamics. Landau and Lifshitz [32] and Dunning and Davies [17] also describe other problems with stability. So are negative specific heats genuinely holding in equilibrium systems? Are they instead the result of idealizations? Or do they place peculiar constraints on

---

[6]Notice that this is connected with an earlier peculiarity, non-extensivity. This fact implies that we cannot take a system with negative C_v and subdivide it into two subsystems each also with negative C_v. An isolated system with negative C_v must be non-extensive.

how can combine thermodynamic macrostates? And as Chavanis [10] asks, what does the specific heat even mean in statistical mechanics—in the microcanonical ensemble U cannot vary and in the canonical ensemble T does not vary?

7.4 Tsallis Statistics

In response to the problems described here, there is a research program organized around non-extensive statistics. In 1988 Tsallis developed a generalization of the Boltzmann and Gibbs entropies, namely, the Tsallis entropy. The Tsallis entropy reduces to the Boltzmann and Gibbs entropies when the system is extensive, but is different otherwise. Since then a whole school devoted to non-extensive statistics has arisen based around this new entropy formula. The motivation behind the program is to show that the Tsallis entropy works well in situations where the Boltzmann and Gibbs entropies allegedly break down. Normal statistical mechanics is supposed to correspond only to the case where a system is approximately extensive. Long-range force systems like self-gravitating systems are supposed to be one prominent example where Tsallis statistics are needed. The debate in the literature over the necessity of Tsallis statistics is sometimes very heated. Proponents of Tsallis statistics write papers displaying systems for which they claim traditional mechanical techniques come up short; defenders of statistical mechanics respond with criticism of the new techniques and try to show how the Boltzmann-Gibbs framework suffices to account for the physics. However, before one even engages in this back-and-forth, one might inquire from a foundational perspective whether the Tsallis entropy really makes sense. The paper by Nauenberg [41] is quite critical of Tsallis entropy, but there is room for more debate. Those interested in the foundations of statistical mechanics may wish to weigh in on this matter.

8 Conclusion

As advertised, this discussion has raised more questions than it's answered. We are left with puzzles about the thermodynamic limit, ensemble non-equivalence, negative specific heats, the status of Tsallis statistics, and more. That said, some answers have been sketched, and we have learned something about how a thermodynamic description of the world arises. If there is a general lesson, I believe it is that we sometimes have too narrow a vision of thermodynamics. In his beautiful review, Thompson [52] writes that "to show that thermodynamics exists for a given system" we must (a) "prove... the existence of the thermodynamic limit" and (b) "show that the resulting thermodynamics is stable", i.e., prove that specific heat is positive. By these criteria, self-gravitating systems badly fail as thermodynamic systems. Yet thermodynamic techniques sometimes have proven successful when applied to self-gravitating systems. How do we reconcile these two facts? To a large extent I think the conflict arises from thinking that we need to reproduce exactly the generalities of thermodynamics in statistical mechanics. Most real macroscopic systems don't have strictly stationary equilibria, don't have well-defined thermodynamic limits, aren't perfectly extensive, and so on. Still, some of them nevertheless display thermodynamic behavior. Focusing on self-gravitating systems helps us see this point.

Does thermodynamics hold of self-gravitating systems? The answer is unexpectedly unclear. It hangs on our evaluation of the means of regaining extensivity surveyed in Sect. 6. For instance, we saw that extensivity could be regained in the mean field limit. In this limit some but not all thermodynamic relationships will hold. So even if we find this limit to be unproblematic, our answer depends on a prior choice of what constitutes "thermodynamics." Putting this semantical quibble aside, should we really say that self-gravitating systems obey thermodynamics even if they only partially do in the mean field limit? We began with a very clear question and proceeded to muddy it up. Sometimes—and I hope this is one of those times—this is a sign of progress.[7]

References

1. Antonov, V.A.: Most probable phase distribution in spherical star systems and conditions for its existence. In: Goodman, J., Hut, P. (eds.) Dynamics of Star Clusters. Dordrecht, Reidel (1962)
2. Beck, C., Schloegl, F.: Thermodynamics of Chaotic Systems: An Introduction. Cambridge University Press, Cambridge (1993)
3. Binney, J.J., Tremaine, S.: Galactic Dynamics. Princeton University Press, Princeton (1987)
4. Bouchet, F., Barre, J.: Statistical mechanics and long range interactions. C. R. Phys. **7**, 414–421 (2006)
5. Callen, H.B.: Thermodynamics. Wiley, New York (1960)
6. Callender, C.: Taking thermodynamics too seriously. Stud. Hist. Philos. Mod. Phys. **32**, 539–553 (2001)
7. Callender, C.: The past hypothesis meets gravity. In: Ernst, G., Hüttemann, A. (eds.) Time, Chance and Reduction: Philosophical Aspects of Statistical Mechanics, pp. 34–58. Cambridge University Press, Cambridge (2010)
8. Campa, A., Giansanti, A., Morigi, G., Labini, F.S. (eds.): Dynamics and Thermodynamics of Systems with Long Range Interactions: Theory and Experiments. AIP Conference Proceedings, p. 970 (2008)
9. Campa, A., Dauxois, T., Ruffo, S.: Statistical mechanics and dynamics of solvable models with long-range interactions. Phys. Rep. **480**, 57–159 (2009)
10. Chavanis, P.H.: Phase transitions in self-gravitating systems. Int. J. Mod. Phys. B **20**, 3113–3198 (2006)
11. Combes, F., Robert, R.: Comment on the thematic issue 'Statistical mechanics of non-extensive systems'. C. R. Phys. **8**, 85 (2007)
12. Costeniuc, M., Ellis, R.S., Touchette, H., Turkington, B.: Generalized canonical ensembles and ensemble equivalence. Phys. Rev. E **73**, 026105 (2006)
13. Dauxois, T., Ruffo, S., Arimondo, E., Wilkens, M. (eds.): Dynamics and Thermodynamics of Systems with Long Range Interactions. Springer, Berlin (2002)
14. de Vega, H.J., Sánchez, N.: Statistical mechanics of the self-gravitating gas: I. Thermodynamic limit and phase diagrams. Nucl. Phys. B **625**, 409–459 (2002)
15. de Vega, H.J., Sánchez, N.: The cluster expansion for the self-gravitating gas and the thermodynamic limit. Nucl. Phys. B **711**, 604–620 (2005)
16. Dunning-Davies, J.: On the meaning of extensivity. Phys. Lett. A **94**, 346–348 (1983)
17. Dunning-Davies, J.: Concise Thermodynamics. Albion, Westgate (1996)
18. Gallavotti, G.: Statistical Mechanics: A Short Treatise. Springer, Berlin (1999)

---

[7]I'm grateful for extremely helpful comments from the anonymous referees, colleagues at UCSD and UCI, and especially Urs Hoffmann and P.H. Chavanis.

19. Gibbs, J.W.: Elementary Principles of Statistical Mechanics. Yale University Press, Yale (1902)
20. Gross, D.: Microcanonical Thermodynamics: Phase Transitions in Small Systems. Lecture Notes in Physics, vol. 66. World Scientific, Singapore (2001)
21. Heggie, D., Hut, P.: The Gravitational Million-Body Problem. Cambridge University Press, Cambridge (2003)
22. Huang, K.: Statistical Mechanics. Wiley, New York (1987)
23. Hut, P.: Gravitational thermodynamics. Complexity 3, 38–45 (1997)
24. Ipser, J.R.: On using entropy arguments to study the evolution and secular stability of spherical stellar-dynamical systems. Astrophys. J. **193**, 463–470 (1974)
25. Johal, R.: Mutual thermal equilibrium in long-range Ising model using nonadditive entropy. Physica A **365**, 155–161 (2006)
26. Kac, M., Uhlenbeck, G.E., Hemmer, P.C.: On the van der Waals theory of the vapor-liquid equilibrium. I. Discussion of a one-dimensional model. J. Math. Phys. **4**, 216–228 (1963)
27. Katz, J.: Thermodynamics and self-gravitating systems. Found. Phys. **33**, 223–269 (2003)
28. Kiessling, M.K.-H.: On the equilibrium statistical mechanics of isothermal classical self-gravitating matter. J. Stat. Phys. **55**, 203–257 (1989)
29. Kiessling, M.K.-H.: Statistical mechanics of gravitational condensation and the formation of galaxies. In: Merritt, D., Valluri, M., Sellwood, J.A. (eds.) Galaxy Dynamics. ASP Conference Series, vol. 182, pp. 545–546 (1999)
30. Laliena, V.: On the thermodynamical limit of self-gravitating systems. Nucl. Phys. B **668**, 403–411 (2003)
31. Laliena, V.: Comment on "The cluster expansion for the self-gravitating gas and the thermodynamic limit", by de Vega and Sanchez. Nucl. Phys. B **720**, 439–442 (2005)
32. Landau, L.D., Lifshitz, E.M.: Statistical Physics, 2nd edn. Pergamon, Oxford (1969)
33. Landsberg, P.T.: Thermodynamics with Quantum Statistical Illustrations. Interscience, New York (1961)
34. Landsberg, P.T.: Is equilibrium always an entropy maximum? J. Stat. Phys. **35**, 159–169 (1984)
35. Lavis, D.A.: Boltzmann, Gibbs, and the concept of equilibrium. Philos. Sci. **75**, 682–696 (2008)
36. Lebowitz, J.: Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. Physica A **194**, 1–27 (1993)
37. Lieb, E.H., Yngvason, J.: The physics and mathematics of the second law of thermodynamics. Phys. Rep. **310**, 1–96 (1999)
38. Lynden-Bell, D.: Negative specific heat in astronomy, physics and chemistry. Physica A **263**, 293–304 (1999)
39. Lynden-Bell, D., Wood, R.: The gravo-thermal catastrophe in isothermal spheres and the onset of red giant structure for stellar systems. Mon. Not. R. Astron. Soc. **138**, 495–525 (1968)
40. Ma, S.-K.: Statistical Mechanics. World Scientific, New York (1985)
41. Nauenberg, M.: Critique of q-entropy for thermal statistics. Phys. Rev. E **67**, 036114–0036120 (2003)
42. Padmanabhan, T.: Statistical mechanics of gravitating systems. Phys. Rep. **188**, 285 (1990)
43. Popper, K.: Reply to Hill and Grünbaum. Nature **179**, 1297 (1957)
44. Reif, F.: Fundamentals of Statistical and Thermal Physics. McGraw–Hill, New York (1965)
45. Robert, R.: On the gravitational collapse of stellar systems. Class. Quantum Gravity **15**, 3827–3840 (1998)
46. Rowlinson, J.S.: Thermodynamics of inhomogeneous systems. Pure Appl. Chem. **65**, 873–882 (1993)
47. Ruelle, D.: Statistical Mechanics: Rigorous Results. Addison–Wesley, Reading (1969)
48. Rybicki, G.B.: Exact statistical mechanics of a one-dimensional self-gravitating system. Astrophys. Space Sci. **14**, 56–72 (1971)
49. Saslaw, W.: The Distribution of the Galaxies. Cambridge University Press, Cambridge (2000)
50. Styer, D.: What good is the thermodynamic limit? Am. J. Phys. **74**, 25–29 (2004)
51. Thirring, W.: Systems with negative specific heat. Z. Phys. Hadrons Nucl. **235**, 339–352 (1970)
52. Thompson, C.: Mathematical Statistical Mechanics. Princeton University Press, Princeton (1970)
53. Touchette, H.: When is a quantity additive, and when is it extensive? Physica A **305**, 84 (2002)
54. Touchette, H., Ellis, R.S., Turkington, B.: An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles. Physica A **340**, 138–146 (2004)