3

The past hypothesis meets gravity

CRAIG CALLENDER

3.1 Introduction

Why does the universe have a thermodynamic arrow of time? The standard reasoning relies on the truism: no asymmetry in, no asymmetry out. If the fundamental laws of Nature are time symmetric invariant (that is, time reversal), then the origin of the thermodynamic asymmetry in time must lie in temporally asymmetric boundary conditions. However, this conclusion can follow even if the fundamental laws are not time reversal invariant. The more basic question is whether the fundamental laws – whether time symmetric or not – entail the existence of a thermodynamic arrow. If not, then the answer must lie in temporally asymmetric boundary conditions. No asymmetry of the right kind in, no asymmetry out. As it happens, as I understand them, none of the current candidates for fundamental law of Nature entail the thermodynamic arrow. String theory, canonical quantum gravity, quantum field theory, general relativity, and more all admit solutions lacking a thermodynamic arrow. So a first pass at an answer to our initial question is: the universe has a thermodynamic arrow due in part to its temporally asymmetric boundary conditions.

Merely locating the answer in boundary conditions, however, is not to say much. All it does is rule out thermodynamic phenomena being understood as a corollary of the fundamental laws. But that's true of almost all phenomena. Few events or regularities can be explained directly via the fundamental laws. If we are to have a satisfying explanation, we need to get much more specific.

One promising way of doing so is via the explanation Boltzmann initially devised. Roughly put, the idea is as follows. Identify the thermodynamic entropy of a system with the so-called Boltzmann entropy. Then make plausible the claim that if the initial Boltzmann entropy of the system is low, then over 'reasonable' time spans in the future it is highly likely that it will increase. Finally, assume as a boundary condition that the initial Boltzmann entropy of the system was low. With these pieces in place, one can infer that the system will display a thermodynamic arrow over the time spans in question. What is the system to which this applies? Because it is difficult to know how to decouple systems in a non-arbitrary way, Boltzmann took himself to be describing the entire universe. If this is right, we now have a theory explaining why we have a thermodynamic arrow in our universe.[1] And this explanation appeals to a much more specific claim about boundary conditions than the generic reasoning we engaged in above: namely, that the Boltzmann entropy of the entire universe was very low (compared to now) roughly 15 billion years ago. In particular, the entropy of this state was low enough to make subsequent entropy increase likely for many billions of years. Let us follow Albert (2000) in calling this claim the past hypothesis; let us call the state it posits the *past state*. Physicists such as Boltzmann, Einstein, Feynman, Penrose and Schrödinger have all posited the past state in one form or other. To me, if the Boltzmann framework can be defended, then positing the past state in one form or other appears to be the simplest answer to the problem of the direction of time in statistical mechanics.

Simplicity is nice, but truth is better. Is the past hypothesis true? When we look to the early universe, as described by contemporary cosmology, do we observe something resembling the past state? Some authors (e.g. Price, 1996) believe that Boltzmann's prediction is spectacularly well confirmed by cosmology. I agree that if correct, the vindication of Boltzmann's novel retrodiction should count among the great achievements of science. It would be a prediction of the early state of the universe from seemingly independent statistical mechanical arguments. However, if Boltzmann's prediction is right, why is it so unsung? The answer is that we cannot be confident that the prediction is right. The reason for this is that it has never been entirely clear how to apply Boltzmann's statistical mechanical framework in conditions such as those in the early universe.

Bracket all the questions still under debate about the big bang. Let us not worry about cosmic inflation periods, the baryogenesis that allegedly led to the dominance of matter over anti-matter, the spontaneous symmetry breaking that purportedly led to our forces, and so on. The past state does not have to be the 'first' moment. Skip to 10^{-11} seconds into the story when the physics is less speculative. Or skip even further into the future if you are worried about the standard model in particle physics. (And do not even think about dark energy or dark matter.)

Even still, for confirmation of Boltzmann's insight, at the very least one needs to understand statistical mechanics and Boltzmann entropy in generally relativistic spacetimes, the entropy of radiation, how this entropy relates to the entropy of

[1] Many physicists and some philosophers want *more*: they want to explain why the boundary condition is what it is. In Callender (2004) I argue that this is not necessary.

the matter fields, and more. Needless to say, all of this is highly non-trivial.[2] That the physics needed is largely unknown, of course, does not imply that the past hypothesis is false; it merely explains why Boltzmann is not lauded or faulted for his prediction.

What would suggest falsity is if – as a matter of principle – the basics of Boltzmann's framework just cannot be applied in the non-classical theories needed to describe the physics of the early universe. That is Earman's (2006) claim with respect to general relativity. In a sharp attack, Earman claims that the past hypothesis is 'not even false'. The reason for this conclusion is that Earman is unable to define a coherent and non-trivial Boltzmann entropy in general relativity. For the Boltzmann entropy to make sense (as we will see) one needs a well-defined state space for the theory and a measure invariant under dynamical evolution. We do not have this for the space of all solutions to Einstein's field equations. We do have it for some very special cases. Restricted to Friedman–Robertson–Walker metrics with a scalar matter field, one can use the Hawking–Page measure over a two-dimensional reduced phase space or the Holland–Wald measure over a three-dimensional reduced phase space. Earman shows that using either makes nonsense of the past hypothesis.

Earman's result is troubling, but perhaps not fatal to the past hypothesis.[3] The measures he cites, it must be admitted, are developed only for a highly idealized set of solutions to Einstein's field equations. The problem of developing measures on the space of solutions to Einstein's field equation is still in its infancy, e.g., we are very far from claiming that either of the above measures is uniquely invariant with respect to time evolution. Hence we do not have an in principle demonstration that the Boltzmann entropy is indefinable in general relativity. There may be other measures that work. What Earman shows is that given what we know, things do not look good.

Given this situation, a natural question is whether the Boltzmann entropy makes sense even in classical physics when we consider cosmological systems. In particular, since what is causing the present trouble is gravity, one would like to understand an early classical state when the gravitational interactions are included in the system. Such an approach would be deeply limited. As mentioned, to describe anything like our universe one needs general relativity, the expansion of space, strong and weak nuclear forces, and much more. While this claim is no doubt true, there are virtues in beginning simply. For if we have trouble even here, then we know we

[2] Please do not be fooled into thinking that various entropies used in the literature, including the so-called Bekenstein entropy, are a quick fix. It is commonly asserted that the Bekenstein entropy is low in the past and high in the future. Without a clear connection between this entropy and the Boltzmann entropy, however, this claim simply is not relevant to our question.

[3] Earman also launches other attacks on the past hypothesis and the uses to which it has been put, but we only have space to focus on this problem.

have a problem with gravity no matter how the measure-theoretic details work out in general relativity. And if some problems in the classical context can only be solved by adding non-classical elements, then that is still something interesting to learn. Before worrying about general relativistic or quantum gravitational thermodynamics, let us figure out whether classical gravitational thermodynamics works.

As we shall see, even here in the Newtonian context – surprisingly – matters get tremendously complex. Nasty 'paradoxes' threaten the very foundations of gravitational thermodynamics. The point of the present chapter is to introduce these problems and show how they affect the Boltzmann explanation described above.

This chapter has two very modest goals. Firstly, and primarily, I want to demonstrate why even classical gravity is a serious problem for the standard explanation of entropy increase. If the chapter does nothing else, my hope is that it gets the problems induced by gravity the attention they deserve in the foundations of physics. Secondly, I want to outline a possible way out of at least one difficulty. Most of the work here will be in the set-up, both in seeing the exact nature of the problem and in understanding how the work done on the statistical mechanics of stellar systems can be conceived from a foundational perspective. Once framed, I want to make plausible a very weak claim: that there is a well-defined Boltzmann entropy that *can* increase in *some* interesting self-gravitating systems – where I get to define 'interesting'. More work will need to be done to see if this claim really answers the threat to the standard explanation of entropy increase. However, establishing the claim might remove some of the pessimism one might have about the standard explanation in the gravitational context, in addition to suggesting a clear path for future study.

3.2 The past hypothesis

Classical phenomenological thermodynamics is a system of functional relationships among various macroscopic variables, e.g., volume, temperature, pressure. It tells us that some macro-states M covary or evolve into others, e.g., $M_{t_1} \rightarrow M_{t_2}$. One of these relationships is the famous second law of thermodynamics. It tells us that an extensive state function S, the entropy, defined at equilibrium, is such that changes in it are either positive or zero, i.e. entropy does not decrease. For realistic cases, it seems to imply that in the spontaneous evolution of thermally closed systems, the entropy increases and attains its maximum value at equilibrium. Actually, there is controversy whether the spontaneous movement from non-equilibrium to equilibrium strictly follows from the second law; but even if it does not, there is no controversy that this spontaneous movement occurs and is a central feature of thermodynamics. This feature describes many of the temporally directed aspects of our world, e.g., heat going from hot to cold, gases spontaneously expanding throughout their available volumes.

Why, from a mechanical perspective, do these temporally directed generalizations hold? Let us restrict ourselves to classical statistical mechanics, and in particular, the Boltzmannian interpretation of statistical mechanics. I find the Boltzmannian view of statistical mechanics provides a more 'physical' description of what is going on from a foundational perspective than the rival Gibbsian perspective.[4]

The first step in understanding the Boltzmannian explanation of the approach to equilibrium is distinguishing the macroscopic from microscopic description of the system. The exact microscopic description of an unconstrained classical system of n particles is given by a point $X \in \Gamma$, where $X = (\mathbf{q}_1, \mathbf{p}_1 \ldots \mathbf{q}_n, \mathbf{p}_n)$ and Γ is a $6n$-dimensional abstract space spanned by the possible locations and momenta of each particle. X evolves with time via Hamilton's equations of motion. Since energy is conserved, this evolution is restricted to a $6n - 1$ dimensional hypersurface of Γ.

The same system described by X can also be described in the macro-language by certain macroscopic variables (volume, pressure, temperature, etc.). This characterization picks out the system's macro-state M. Notice that many other micro-states will also give rise to the same macro-state M. If we consider all the $X \in \Gamma$ that give the same values for macroscopic variables as M gives, this will pick out a volume Γ_M. The set of all such volumes partitions is the energy hypersurface of Γ.

A quick word about the volume. A continuous infinity of micro-states will give rise to any particular macro-state, so one requires the resources of measure theory. The $(6n-1)$-dimensional energy hypersurface of Γ has a Lebesgue measure naturally associated with it. From this measure one creates a probability measure, and one assumes or hopes to prove that the probability of finding a system in region Γ_M of the energy hypersurface of Γ is proportional to the volume of Γ_M, $|\Gamma_M|$, within $|\Gamma|$.

We can now define the entropy of a macro-state M. The *Boltzmann entropy* of a system X that realizes M is defined by

$$S = k \log |\Gamma_M(X)|$$

where k is Boltzmann's constant and $||$ indicates volume with respect to Lebesgue measure. Notice that this entropy is defined in and out of equilibrium. In equilibrium, it will take the same value as the Gibbs fine-grained entropy if n is large.

[4] I am not alone. Lavis (2005: 246) writes, 'When confronted with the question of what is "actually going on" in a gas of particles (say) when it is in equilibrium, or when it is coming to equilibrium, many physicists are quite prepared to desert the Gibbsian approach entirely and to embrace a Boltzmannian view'. See Lavis for a description of the Gibbsian view.

Outside equilibrium, the entropy can take different values and will exist so long as a well-defined macro-state exists.

Why should Boltzmann entropy increase? The answer to this is controversial, and we do not have space to discuss it fully here. The hope is that one will be able to show that *typical* micro-states underlying a non-equilibrium macro-state subsequently head for equilibrium. One way to understand this is as follows.[5] The Boltzmann equation describes the evolution of the distribution function $f(\mathbf{x}, \mathbf{v})$, over a certain span of time, and this evolution is one toward equilibrium. Let $\Gamma_\delta \subset \Gamma$ be the set of all particle configurations X that have distance δ, $\delta > 0$, from $f(\mathbf{x}, \mathbf{v})$. A *good* point $X \in \Gamma_\delta$ is one whose solution (a curve $t \rightarrow X(t)$) for some reasonable span of time stays close to the solution of the Boltzmann equation (a curve $t \rightarrow f_t(\mathbf{x}, \mathbf{v})$). A *bad* point $X \in \Gamma_\delta$ is one that departs from the solution to the Boltzmann equation. The claim that typical micro-states underlying a non-equilibrium macro-state subsequently head for equilibrium is the statement that, measure theoretically, most points $X \in \Gamma_\delta$ are good. The expectation – proven only in limited cases – is that the weight of good points grows as n increases. The Boltzmannian wants to understand this as providing warrant for the belief that the micro-state underlying any non-equilibrium macro-state one observes is almost certainly one subsequently heading toward equilibrium. As mentioned, the desired conclusion does hang on highly non-trivial claims, in particular, the claim that the solution to Hamilton's equations of motion for typical points follows the solution to the Boltzmann equation.

Here is a loose bottom-to-top way of picturing matters that will come in handy later (see DeRoeck, Maes and Netočný, 2006). We know at the macroscopic level that non-equilibrium macro-states evolve over short periods of time into closer-to-equilibrium macro-states. That is, M_1 at t_1 will evolve by some time t_2 into a closer-to-equilibrium macro-state M_2. Call $\Gamma_{M_1 t_1}$ the set of states in Γ corresponding to M_1 at t_1, $\Gamma_{M_2 t_2}$ the set corresponding to M_2 at t_2, and $\phi_{t_2 - t_1}\Gamma_{Mt_1}$ the time evolved image of the original set M_1. Then, if our picture is right, the second law is telling us that $\phi_{t_2 - t_1}\Gamma_{Mt_1}$ is virtually a proper subset of Γ_{Mt_2}. That is, almost all of the points originally in M_1 have evolved into the set corresponding to M_2. Liouville's theorem states that a set of points retains its size through Hamiltonian evolution. Hence the volume of $\phi_{t_2 - t_1}\Gamma_{Mt_1}$ is equal to the volume of $\Gamma_{M_1 t_1}$. Since the former is virtually a proper subset of Γ_{Mt_2}, that means that $|\Gamma_{Mt_1}| \leq |\Gamma_{Mt_2}|$. From the definition of entropy it follows that $S(M_{t_2}) \geq S(M_{t_1})$.

The problem of the direction of time is simple to see. Nowhere in the above argument did I say whether t_2 is before or after t_1. Given a non-equilibrium state at

[5] For a general discussion, see Goldstein (2002) and references therein. For the specific formulation here, see Spohn (1991: 151). And for some of the challenges this approach faces, see Frigg (2009).

t_1, the above reasoning shows that it is very likely that it will subsequently evolve to a later higher entropy state at t_2, where t_1 is earlier than t_2. However, it is also true that the reasoning shows that most likely the state at t_1 evolved from an earlier higher entropy state, in this case where t_2 is earlier than t_1. There is nothing in the time reversible dynamics nor in the above reasoning to rule out entropy increase in both temporal directions from the non-equilibrium present. The famous recurrence and reversibility challenges to Boltzmann point out that even good points X will go bad if given enough time (recurrence) or allowed to go in the wrong temporal direction (reversibility).

All manner of answers to this problem have been proposed – appeals to time asymmetric environmental perturbations, ignorance, electromagnetism, and more. In my opinion, where these proposals have merit, they eventually reduce to an appeal to temporally asymmetric boundary conditions. Ultimately we need to assert that in the direction we call 'earlier' entropy was in fact very low compared to now. As mentioned at the outset, the specific form of this claim in the present context is that the past hypothesis is true; that is, that the Boltzmann entropy of the universe was extremely low roughly 15 billion years ago.

3.3 The past hypothesis meets gravity

No sooner is the past state posited than it is immediately challenged with a bit of a problem: it seems to be manifestly false. When we look to cosmology for information about the actual past state, we find early cosmological states that appear to be states of very high entropy, not very low entropy. Cosmology tells us that the early universe is an almost homogeneous isotropic state of approximately uniform temperature, i.e. a very high entropy state, not a low-entropy state as mandated by the past hypothesis. Here is the physicist Wald:

> The above claim that the entropy of the very early universe must have been extremely low might appear to blatantly contradict the 'standard model' of cosmology: there is overwhelmingly strong reason to believe that in the early universe matter was (very nearly) uniformly distributed and (very nearly) in thermal equilibrium at uniform temperature. Does not this correspond to a state of (very nearly) maximum entropy, not a state of low entropy?
>
> *(Wald, 2006: 395)*

If we consider point particles interacting without gravity, then the answer certainly seems to be in the affirmative.

Once the problem is stated, however, authors quickly reassure us that it is only apparent. We forgot to include gravity, we are told, and yet by including gravity the 'situation changes dramatically' (Wald, 2006: 395). Gravity saves the past hypothesis. This claim is made with equal frequency and force by scores of physicists and philosophers of physics.

How does gravity save the past hypothesis? Here is a (too) simple expression of the idea. If we think of a normal terrestrial gas in a box, as a result of repulsive forces and collisions, its 'natural tendency' is to spontaneously spread throughout its available volume into a homogeneous state. If this is right, then when we add an attractive force like gravity the reasoning should reverse. For it is the 'natural tendency' of a gravitating system to spontaneously move toward more clumped states. Masses attract one another, and both in theory and computer simulation self-gravitating Newtonian systems get more and more clumpy with time. With gravity, inhomogeneity is the new homogeneity. Since low-to-high entropy transitions express the natural tendency of systems, it ought to be that in gravitating systems clumped states are of high entropy and spread out ones of low entropy. The cosmic background radiation shows that the universe was more homogeneous in the past. Hence the past state is vindicated. In fact, one might go so far as to say not only that it does not falsify the standard explanation of entropy increase, but that it is a stunningly accurate prediction made by the standard explanation.

Of course, this simple idea leaves out the momentum sector of phase space. There is no 'natural tendency' toward *spatial* homogeneity or inhomogeneity in either gravitating or non-gravitating systems. The oil and vinegar separating in your container of salad dressing is an entropy increasing process. Many spatial inhomogeneities grow in perfectly normal entropy increasing situations, and presumably homogeneities can develop in gravitational situations. The idea must be, then, that the increasing concentration in the configuration sector of phase space is compensated by a greater decreasing concentration in the momentum sector of phase space. As we go forward in time, one – not implausibly – imagines the velocity vectors as becoming increasingly chaotic.

Assume that the total entropy can be expressed as a simple sum of the configuration sector entropy $|\Gamma_{Mq}|$ and the momentum sector entropy $|\Gamma_{Mp}|$. Then it is easy to see that it is possible that entropy increase or decrease with time. When gravity is the dominant force then presumably $|\Gamma_{Mq}|$ will decrease as time passes if the initial state is originally very dispersed. Although the system may develop into various quasi-stable configurations, in the long run we might expect it to become more concentrated in space. On the other hand, we might expect particles gradually to be 'slingshot' far away, so that the system evaporates and becomes very dispersed in space. Similarly, it is possible that $|\Gamma_{Mp}|$ grows as particles' velocities become increasingly randomly distributed with time; but it's also possible that the velocities become more aligned as time passes. What is needed is that $\log|\Gamma_M|$ increases with time and this can be achieved in a variety of ways.

We know in the normal non-gravitational case that entropy can go up or down. There are, as described before, good and bad initial states, the bad ones leading to subsequent entropy decrease. Fortunately, with respect to the Lebesgue measure, most of the states are believed to be good ones. So what is of interest is not whether entropy can go up or down when gravity is turned on – of course it can – but whether for *most* initial states entropy increases.

A really pressing question then is whether the standard probability distribution crafted from the Lebesgue measure is *empirically adequate* when gravitational interactions are included. Can we see the motion of the stars, and so on, as the movement to an equilibrium state, where equilibrium is understood as the largest, 'most probable' macro-state according to the Lebesgue measure? A priori, the applicability of the Boltzmann framework, and in particular, the empirical adequacy of the Maxwell–Boltzmann probability distribution, is not guaranteed. New physics presents new challenges. Indeed, when one states this hypothesis one realizes that the standard explanation of the direction of time – which assumes this framework works with gravity – tries to explain with one stroke two possibly quite distinct processes. It tries to account for ordinary thermodynamics *and* the rise of structure in the cosmos. The first is a primarily non-gravitational process and the second is a primarily gravitational process. The past hypothesis is thus a tremendously ambitious claim, and if successful, the result would be a major unification in physics. But we should be clear that it is ambitious and that it's not obvious that the two processes can be given the same explanation.[6]

At this point the natural thing to do would be to calculate the Boltzmann entropy, with gravity included, of some toy gravitational systems and see if entropy increases. Then one would like to compare the results there with our actual cosmological history. However, for various reasons to be discussed, we are stymied in this attempt.

3.4 The gravitational paradoxes

Statistical mechanics and thermodynamics work flawlessly in some gravitational contexts. In terrestrial cases where we can approximate the gravitational field as

[6] I should not give the impression that no one else is aware of potential difficulties with the usual response besides Earman. Wald (2006), for instance, comments that statistical thermodynamics is usually justified via ergodicity, and yet ergodicity will not obtain in a general relativistic universe (the universe might be open, and it is not time translation invariant in the right way). He also warns that the real story will include discussion of black hole entropy and quantum gravity. As mentioned above, I think that unless one shows that the black hole entropy is connected to the Boltzmann entropy, then the black hole entropy will not be relevant to our explanation. The first worry may also be irrelevant, as stated, since the Boltzmannian hopes his or her explanation uses requirements on the dynamics that are weaker than ergodicity. But the spirit of Wald's point is right: once the Boltzmannian is clear about the necessary dynamics, it will be a good question whether they obtain in generally relativistic spacetimes.

uniform, there is simply no problem. Thermodynamics obviously works in such cases, and the extension of statistical mechanics to systems with external uniform fields does not require any major modification (see, e.g., Landau and Lifshitz, 1969: 72; Rowlinson, 1993). However, we are interested in the more general question, the thermodynamic and statistical mechanical properties of a system self-interacting via time-varying gravitational forces. We will operate under the idealization that gravity is the only force obtaining between the particles and we will restrict ourselves to classical gravitation theory. Therefore, we investigate the thermodynamic properties of the famous classical *N*-body problem in gravitation theory.

The Hamiltonian describing the system for *N* gravitationally interacting particles of mass $\boldsymbol{m}$ is

$$H(q, p) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} - \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i} \frac{Gm_i m_j}{r_{ij}} \tag{3.1}$$

where $r_{ij} = \left\| q_i - q_j \right\|$. Although this system is ideal, some globular star clusters ($N = 10^5 - 10^6$), galaxies ($N = 10^6 - 10^{12}$), open clusters ($N = 10^2 - 10^4$), and planetary systems are decent instantiations of this ideal system. That is, the salient features of these systems over large time and space scales are due to gravity. Collisions, close encounters and other behaviours where other forces are relevant are rare, so ignoring inelastic interactions does not cause great harm. The question is then whether the stars in such systems or even the galaxies themselves, when idealized as point particles, admit a thermodynamic description. Can we think of the stars as the point particles in a thermodynamic gas?

In the literature on classical gravitational thermodynamics,[7] most papers mention some subset of five obstacles facing any such theory: non-extensivity, ultraviolet divergence, infrared divergence, lack of equilibrium, and negative heat capacity. The first problem is that the energy and entropy of systems evolving according to (3.1) can be non-extensive, even though in thermodynamics these quantities are extensive. The second problem arises from the infinite range of the gravitational potential and the lack of gravitational shielding; together they imply that the integral over the density of states can diverge. The third problem arises instead from the short-range nature of the potential. 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 fourth problem comes in many forms, some linked to

[7] For entries into this literature, see Padmanabhan (1990); Saslaw (2000); Dauxois *et al.* (2002); Heggie and Hut (2003).

the divergence problems. But in general there are many problems with defining an equilibrium state for a system evolving via (3.1). Finally, the fifth problem, which is not really a paradox but merely extremely counterintuitive, is that the heat capacity for systems evolving via (3.1) can often be negative, whereas in classical thermodynamics it is always positive.

To get an intuitive feel for how gravity causes trouble, focus on just one issue, the non-extensivity problem. Intuitively put, extensive quantities are those that depend upon the amount of material or size of the system, whereas intensive quantities are those that do not. The mass, internal energy, entropy, volume and various thermodynamic potentials (e.g., F, G, H) are examples of extensive variables. The density, temperature, and pressure are examples of intensive variables. Mathematically, the most common expression for extensivity is the definition that 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

$$f(aU, aV, aN) = af(U, V, N) \tag{3.2}$$

for all positive numbers a. 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 $S(2U, 2V, 2N)$ represents the joint system, and Equation (3.2) says that this is the same as two times the individual entropies of the partitioned component systems. Extensive functions are also assumed to be additive, and with a slight assumption, they are. A function – using our example, entropy – is additive if

$$S(U_1 + U_2, V_1 + V_2, N_1 + N_2) = S(U_1 + V_1 + N_1) + S(U_2 + V_2 + N_2)$$

With minimal assumptions homogeneity of degree one implies additivity.[8] With these definitions in hand, let us turn to statistical mechanics, the theory that explains why thermodynamical relationships hold for mechanical systems.

Perhaps the most basic assumption of thermodynamics / statistical mechanics is that the total energy of any thermodynamic system is approximately equal to the sum of the energies of that system's subsystems. If we have a large gas in a box, and we conceptually divide it into two subsystems, we expect the total energy to be the sum of the two subsystem energies – so long as the subsystems are still macroscopic systems. In many influential treatments of the theory, this assumption is regarded as the most basic of all, e.g., Landau and Lifshitz's (1969) classic treatment begins with essentially this assumption.

One of the features that makes this assumption plausible is that in terrestrial cases we are usually dealing with short-range potentials. At a certain scale matter

[8] See Dunning-Davies (1983) and Touchette (2002) for useful discussions of extensivity and additivity. Because of their close connections for realistic systems, I will use the two more or less interchangeably.

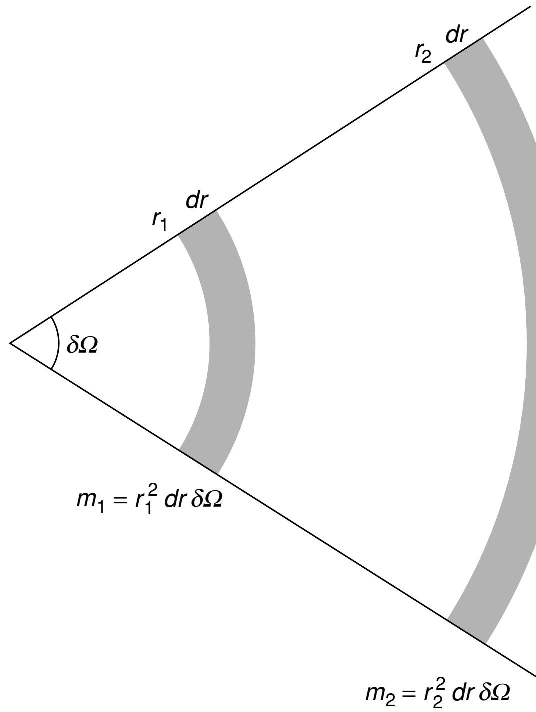

Fig. 3.1. Fashioned after Binney and Tremaine (1987: 188)

is electrically neutral and gravity is so weak as to be insignificant. If the potential is short range and our subsystems are not too small, then the subsystems will interact with one another only at or in the neighbourhood 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 (length)^3 and the latter as (length)^2. The basic assumption is then justified.

However, if gas molecules are replaced by stars – that is, short-range potentials replaced by long-range potentials – this reasoning does not work. Consider a star at the apex of a cone (Binney and Tremain, 1987: 187–8) and the force by which the stars in the cone attract the star (Fig. 3.1). Suppose the other stars are distributed with a uniform density. The force between this star and any other falls off as r^{-2}, but the number of such stars increases along the length of the cone as r^2. Thus any two equal lengths of the cone will attract the target star with equal force. If the density of stars is perfectly homogeneous and isotropic, the star will not feel any force. But if not homogeneous – even if not homogeneous only at great distances – the star will feel a net force. For this reason the force on any particular star is typically determined more by the gross distribution of matter in the galaxy than by the stars close to it. Collisions do not play as large a role as they do in a typical gas in a box on Earth. As is sometimes said, terrestrial gas molecules tend to lead violent lives determined in large part by sudden disputes with their neighbours; stars tend to lead comparatively peaceful lives because they are in harmony with the overall universe.

Returning to energy, we see that 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 comes from distant 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:

$$U \propto \int_0^R 4\pi r^2\, dr\, \rho r^{-3-\varepsilon} \propto \int_0^R dr\, r^{-1-\varepsilon}$$

One can then verify that with $\varepsilon > 0$ ('short-range potentials'), the significant contribution to the integral comes from near the particle's origin, whereas with $\varepsilon < 0$ ('long-range potentials'), the contribution comes from far from the origin (Padmanabhan, 1990).

Consider again a chamber of gas divided into two equal boxes, A and B. 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. As a result of the interaction, it is easy to devise scenarios whereby $E_{\mathrm{A}} = E_{\mathrm{B}} = -a$, where $a > 0$, yet where the energy of the combined system $E = 0$, not $-2a$. The energy might not be even approximately additive.

When the additivity and extensivity go, so do large portions of equilibrium thermodynamics and statistical mechanics. For example, when a system is in equilibrium, its large subsystems also will be in equilibrium. This is no longer necessarily the case. And additivity is a requirement for the equilibrium second law in thermodynamics (see Lieb and Yngvason, 1998). Moreover, in statistical mechanics it is built into the heart of the theory. The famous Boltzmannian probability W of a macro-state is assumed equal to the product of the probabilities of the subsystem macro-states, i.e. $W_{\text{total}} = W_a W_b$. Boltzmann's definition of entropy as $S = k \ln W$ straightforwardly implies that $S_{\text{total}} = S_a + S_b$. And there is no conventional thermodynamic limit for non-extensive systems. For a rigorous discussion of this see Padmanabhan (1990), Lévy-Leblond (1969) and Hertel *et al.* (1972). This last fact should not be surprising. The existence of the thermodynamic limit depends on making the contribution of surface effects go to zero as N, V go to infinity. In a non-additive system, we saw that the surface effects are not going to get smaller as N and V increase. Ironically, the thermodynamic limit does not apply to very large systems if one includes the force primarily relevant to the dynamics of those systems.[9]

3.5 The problem

As interesting as these problems are, they are – at first glance – orthogonal to our main worry. The problems described are problems for *equilibrium* thermodynamics and statistical mechanics, but we are interested in *non-equilibrium* statistical mechanics.

The reason this is so is because the past state is surely not an equilibrium state, yet arguably it still has a Boltzmann entropy. Why is the past state a non-equilibrium state? It is the global state of the universe, and the very reason it is posited is that subsequent evolution will spontaneously take the global state to regions corresponding to higher entropy. But if it is in equilibrium, the system will not change unless an external constraint is removed; yet since the system is the global state, there is no external constraint to remove. The unavoidable conclusion is that the past state must not be an equilibrium state. Indeed, no one expects the past state to stay that way. It is expected, under the attractive force of gravity, to begin clumping. The past state, therefore, simply does not have an ordinary equilibrium entropy corresponding to equilibrium thermodynamics or equilibrium statistical mechanics. But it does have a Boltzmann entropy. The definability of the Boltzmann entropy in systems outside equilibrium is touted as perhaps its greatest virtue. Since we are restricting ourselves to the classical phase space and assuming a Lebesgue measure upon it, we do not have the potentially in-principle problems Earman worries about with general relativity. All one really needs for a Boltzmann entropy to exist is a well-defined macro-state and a well-defined notion of volume in phase space. If the earliest cosmological times do not correspond to a macro-state for some reason, then the past hypothesis picks out the 'first' state that does. This macro-state will correspond to a particular volume $|\Gamma_M|$, and hence it has an entropy. The problems with equilibrium statistical mechanics in the presence of

[9] Before concluding this section I should point out that there is a large research program devoted to the statistics of non-extensive systems that I am here bracketing aside. This is the approach of Tsallis statistics. The Tsallis school develops 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. 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. Long-range force systems like self-gravitating systems are supposed to be one example. The debate between the Tsallis school and others believing Boltzmann–Gibbs suffices is often very heated. For the purposes of this chapter I want to stay conservative and remain within the Boltzmann framework – though for some criticism of the Tsallis school, see Nauenberg (2003). That said, we ought to acknowledge that one way of responding to the above worries is to change frameworks and go outside the normal Boltzmann–Gibbs picture.

gravity are worrying, but so far not directly relevant to the increase of Boltzmann entropy.

Or are they? One cannot completely divorce non-equilibrium theory from equilibrium theory. Think of the issue as follows. The Boltzmann entropy for the gravitating system described by the past state will exist, but what will it do? What one wants is not merely the existence of entropy but also the functional relationships that are usually entailed by a system having an entropy. Why think, for instance, that a system in the past state will increase its entropy? And more generally, granted that the past state picks out some volume in phase space, what gives this volume its physical significance?

As Boltzmann famously showed, in the case of the dilute gas we have everything we could want from the Boltzmann entropy. Recall that the argument goes as follows. The H-theorem shows that the entropy $S(f(x, v))$

$$S(f) = -N \int f \ln f \, d^3\mathbf{x} \, d^3\mathbf{v} \tag{3.3}$$

increases monotonically with time when f is evolving via the (independently motivated) Boltzmann equation. Here f is the distribution function defined over six-dimensional μ-space when partitioned into a finite number of cells. $S(f)$ is shown to increase with time except for when f is a local Maxwellian, whereupon $S(f)$ is stationary. Since Maxwell had already shown that his distribution corresponded to equilibrium, the idea of $S(f)$ playing the role of entropy is naturally suggested. Note that so far none of this bears on the Boltzmann entropy. The crucial link is provided by the detour via six-dimensional μ-space. By making a number of assumptions appropriate to the dilute gas – but certainly not to gases with strong interactions – Boltzmann is able to 'translate' distributions f into hypervolumes in Γ. In particular, he is able to show via the famous 'combinatorial argument' that the distribution f corresponding to the Maxwell distribution occupies far and away the greatest proportion of volume in Γ. Via this translation Boltzmann shows that all the desirable properties true of $S(f)$ are true of the Boltzmann entropy too in the case of the dilute gas. Doing so motivates the entire picture of micro-states most likely evolving into the dominant equilibrium sections of Γ. (See Uffink, 2007 for more discussion.)

It is important to stress that it is the above connections to the H-function and the Boltzmann equation that gives the volume in Γ any claim to be physically significant. After all, there are other volumes calculated in other bases, e.g., energy, which do not have this feature.

Now we immediately see at least one big problem for providing the Boltzmann entropy physical significance in the gravitational case. Boltzmann's argument can plausibly be extended to some systems for which it was not originally intended, and new arguments mimicking Boltzmann's can show that the Boltzmann entropy for some non-dilute gases has physical significance (e.g, Garrido *et al.*, 2004; Goldstein and Lebowitz, 2004). However, in the gravitational case we know we in general cannot use Boltzmann's argument and there is not much reason to hope anything like it will help.

For instance, consider an important property we need to know of our system: the macro-state $f(\mathbf{x}, \mathbf{v})$ that has maximum volume in Γ. One can hope to find this via the combinatorial argument only if one can translate between μ-space and Γ-space – and one can only do this because the gas is dilute and interactions are effectively turned off. What one does is maximize $f(\mathbf{x}, \mathbf{v})$ subject to two constraints. One constraint is associated with particle number, but the other is more directly relevant to us:

$$\int_V d\mathbf{x} \int_{R^3} d\mathbf{v} \, \frac{1}{2} m \mathbf{v}^2 f(\mathbf{x}, \mathbf{v}) = E \tag{3.4}$$

where E is the total energy. In other words, one is maximizing conditional on the claim that the total energy is the sum of kinetic energies. If this is so, the Maxwell equilibrium distribution is the macro-state with the maximum volume in Γ. In fact, as n goes to infinity departures from the equilibrium macro-state go to zero. This step warrants the additional geometric interpretation the Boltzmannian asserts. The picture of typical non-equilibrium states moving to equilibrium states because there are vastly more of the latter than former is not justified except by this procedure.

In the present case, however, we are directly challenged by the total energy not being approximately the sum of independent individual energies. Equation (3.4) is manifestly false and not even approximately true for many self-gravitating systems. The gravitational potential energy contributes to the overall energy of the system. Without (3.4) one cannot show that the largest macro-state in phase space is the equilibrium state; absent this, one cannot make plausible that typical initial states go to equilibrium. So although the loss of a maximization constraint may seem like a quibble, an awful lot hangs on it. In fact, the very terms by which we conceived the original question depend on this; unless the energy factorizes there is no reason to think the entropy $S(f)$ is a simple sum of a configurational contribution and momenta contribution, so the intuitive reasoning we engaged in earlier does not hold. And if this were not bad enough, we are also lacking a gravitational version of the Boltzmann equation for which one can prove an H-theorem (more on this later). I hope this discussion adequately displays the problem: although the gravitational system has a well-defined Boltzmann entropy, that by itself does not imply any particular subsequent behaviour.

Perhaps we can look at the glass as half full? We already knew the above problems for the Boltzmann explanation. Many critics of Boltzmann (e.g. Schrödinger, 1948 [1989]) point out that it works rigorously only for the case of dilute gases, yet most systems are of course not dilute gases. The Boltzmannian can deflate some of these worries by showing how many systems are approximately like dilute gases, how numerical simulations of cases that are not dilute gases vindicate the Boltzmannian claims, and so on. But there are of course many systems that do not fit this mould, and the strongly self-gravitational system is one of them. All we have done is highlight the existing problem by displaying a class of systems that are especially far from being treated as dilute gases. And we could have made this argument with plenty of non-gravitational systems too, e.g., some types of plasmas. Maybe, perversely, this is good news to the Boltzmannian. The problem gravitational interaction presents to the standard story the Boltzmannian tells is *as bad as* but *not obviously worse than* the problem other systems already cause the Boltzmannian.

It would be nice if we could view the problem as simply a new version of the same old one already challenging Boltzmann. But it's not clear that even this is the case. In astrophysics researchers often make assumptions about the stars that warrant a description of the system via f on μ-space, not the full Γ-space. That is, they often work with the 'one-particle' distribution function on six-dimensional μ-space just as Boltzmann did in his work on dilute gases. This restriction on f is typically justified in the astrophysics literature by the fact that gravitational systems are essentially collisionless for long periods of time. So what we are doing now is restricting ourselves to a regime wherein some of the usual Boltzmann apparatus can be salvaged. The entropy is defined as (3.3) above.

Under these restrictions, let us now search for the state of maximum entropy, which will be our equilibrium state. Even here, however, we run into problems. To find out what equilibrium looks like for self-gravitating systems, therefore, we can find the distribution $f(\mathbf{x}, \mathbf{v})$ that maximizes the equilibrium entropy (3.3). However, if one looks for the distribution that maximizes S for a given mass M and energy E, then it is a major result in the field that S is extremized iff $f(\mathbf{x}, \mathbf{v})$ is the distribution function of the isothermal sphere (Ogorodnikov, 1965; Lynden-Bell, 1967; Lynden-Bell and Wood, 1968). The isothermal sphere is an infinite n self-gravitating ideal gas. That is, there is *no* distribution function that maximizes S while keeping M and E finite. Maintaining finite M and E, one can obtain arbitrarily large entropies by rearranging the configuration of stars, as Binney and Tremaine (1987) show. There is no $f(\mathbf{x}, \mathbf{v})$ that maximizes the entropy (3.3) for finite M and E. (Binney and Tremaine (1987) take this result to show that galaxies and presumably other typical stellar configurations are *not* the result of long-term thermal equilibrium.

The quest in the astrophysics literature is to associate typical stellar configurations with quasi-stationary states, not true equilibrium states.[10])

The Boltzmannian may reply that this problem is an artefact of the simplification, that with the 'true physics' on Γ (instead of on μ) the problem will go away. That puts the Boltzmannian in an awkward position, however. The Boltzmannian cannot show this is the case because then she meets the gravitational version of Schrödinger's worry: that one cannot prove much outside the simple case. In the non-gravitational case the Boltzmannian replies to Schrödinger by pointing out all the success she had with dilute gases, toy models, computer simulations, and so on. Now in the gravitational case it looks like the Boltzmannian needs to solve the hard case to help answer problems with the allegedly easy case.

Obviously more study is needed of this problem. Perhaps there is still a way the Boltzmannian can by-pass these difficulties. Right now, however, gravity seems to have pulled the Boltzmannian into a serious thicket of problems.

3.6 A way forward?

That was the bad news. Let us conclude, however, with some good news.

In Section 3.3 we learned that the problem we have is giving the Boltzmann entropy of a gravitating system physical meaning. In the case of a non-gravitating dilute gas we saw that the Boltzmann equation, the *H*-theorem and the connection to the *H*-function provided the Boltzmann entropy physical significance. Can we do this for other systems, in particular, gravitational systems?

In some regimes, yes.

To understand the general idea, recall again what Boltzmann does for dilute gases. Boltzmann considers a distribution $f(\mathbf{x}, \mathbf{v}, t)$ that evolves according to the Boltzmann equation, an equation independently motivated on physical grounds. His famous *H*-theorem shows that a function of f, $S(f)$, increases with time except when $f(\mathbf{x}, \mathbf{v})$ is a local Maxwellian. The reason why any of this is interesting is that $S(f)$ is shown to be roughly equal to the Bolztmann entropy $k \log|\Gamma_M(X)|$ and the Maxwellian distribution is the distribution corresponding to the maximum Boltzmann entropy. Thanks to this connection, we know that so long as the system is appropriately modelled via the Boltzmann equation, the genuine Boltzmann entropy will increase with time until reaching equilibrium.

There are physical regimes, of course, where the Boltzmann equation is not a good approximation, and in those cases different macroscopic kinetic equations often apply. Are there macroscopic kinetic equations that are good approximations

[10] See, for instance, Chavanis (2005).

of real systems wherein the dominant interaction is gravitational? Can we get an H-theorem for these regimes? And can we show that this H-function $H = -S(f)$ is roughly equal to the Boltzmann entropy? The answer in each case is yes.

In astrophysics the full n-body problem is often too hard to study, even in computer simulations in many instances, and hence one considers regimes wherein various kinetic equations are appropriate. Astrophysics is filled with macroscopic equations of motion for distributions. Indeed, since plasmas interact via long-range forces too, many of the same kinetic equations used in plasma physics often work in astrophysics too, so an awful lot is known about many equations.

Consider a galaxy of n identical stars with characteristic radius r. The time it takes any star to cross the galaxy is r/v, where v is the typical speed of a star (determined by G, n, r, mass); this is called the *crossing time* of the star. Suppose that the star evolves in a background wherein the mass is perfectly smoothly distributed, not clumped up into individual stars. Call this its mean trajectory. When would the difference between the star evolving in this background versus a more realistic background show up in its velocity (where by 'show up' we mean the velocity changes by order of itself)? Leaving the details to textbooks on galactic dynamics, the answer is that the star is deflected from its mean trajectory over order $0.1\, n/\ln n$ crossing times. Hence one concludes that for systems that are less than $0.1\, n/\ln n$ crossing times old, individual stellar encounters are more or less unimportant. Many galaxies, with $n \approx 10^{11}$ stars and a few hundred crossing times old, are examples. For these systems, typically where the forces are long range and weak, a natural move is to replace the actual force by its spatial average. Many self-gravitating systems enjoy large space and time scales where this approximation is justified.

A major equation of study in galactic dynamics is therefore the Vlasov equation, or the collisionless Boltzmann equation. The equation is for a density of particles subject to an average force field:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = 0 \qquad (3.5)$$

where $f = f(\mathbf{x}, \mathbf{v}, t)\, d^3\mathbf{x}\, d^3\mathbf{v}$ and ϕ is a smooth gravitational potential. Equation (3.5) is essentially a special case of the Liouville equation (for a derivation of (3.5), see Kandrup, ms; 1981). Despite its simplicity, the Vlasov equation is described in textbooks as the fundamental equation of stellar dynamics. What is nice for us is that if we define an entropy via this f, $S(f)$, then one can show that it is proportional to the Boltzmann entropy. What is not so nice, however, is that we cannot show that entropy increases for distributions evolving according to the Vlasov equation.[11]

[11] In terms of the conjecture mentioned in Section 3.2, DeRoeck *et al.* (2006) make clear that not all macroscopic equations will produce an H-theorem. In particular, and skipping the details, they explain that if *every* microstate X is typical of the macroscopic equation, then the argument does not go through. For Equation (3.5),

This is not at all surprising, since the Vlasov equation is more or less the Liouville equation.[12]

Nonetheless, there is a lot more to stellar dynamics than the Vlasov equation. Many systems are such that stellar encounters have played a major part in their development. Globular clusters, open clusters, galactic nuclei and clusters of galaxies all have n, crossing times and lifetimes making the collisionless regime inappropriate to describe them. Outside this Vlasov regime, kinetic equations other than (3.5) are required, equations including some effect of collisions and close encounters. There are scores of kinetic equations used in the subject, but for concreteness let me mention two, namely, the Fokker–Planck equation and the essentially equivalent Landau equation. (The Landau equation is a symmetric form of the Fokker–Planck equation.) These are equations of form

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = C[f] \tag{3.6}$$

where $C[f]$ denotes the rate of change of f due to encounters and collisions. The Fokker–Planck and Landau equations are of form (3.6) with specific collision terms derived for small-angle grazing collisions. The equations are derived by expanding the Boltzmann equation about small-angle grazing collisions. For the exact form of $C[f]$ see Balescu (1963), Spohn (1991) or Heggie and Hut (2003). Both Fokker–Planck and Landau are useful for gas/fluid systems that are weakly coupled, and they are particular useful for stellar systems in which collisions are rare and interactions weak. In astrophysics, the Fokker–Planck equation is advertised as the 'most accurate model of a stellar system, short of the N-body model' (Heggie and Hut, 2003: 87).

What I want to point out is that for the Fokker–Planck equation, one of the most successful kinetic equations in astrophysics, one can get everything one wants. In particular, for many broad classes of collision terms $C[f]$ one can prove an H-theorem for (3.6). One can show that this H-function is related to the Boltzmann entropy in the same way Boltzmann does for the dilute gases. And one can show that the stationary or equilibrium distribution of (3.6) is equivalent to the solution one obtains from maximizing the Boltzmann entropy in the presence of an external potential. Since the Fokker–Planck equation has been extensively studied, and these results are relatively well known, I will not prove any of it here. I simply will refer the reader to the relevant literature for proofs and discussions of these assertions (see, e.g., Green, 1952; Balescu, 1963: 170ff; Liboff and Fedele, 1967; Spohn, 1991: 83;

every X is typical: every solution of Hamilton's equations will follow solutions of (3.5) for f. We will not, therefore, get an H-theorem.

[12] A coarse-grained entropy might increase, however. In gravitational dynamics physicists speak of non-collisional 'phase mixing' as another means of a system moving to equilibrium. See Heggie and Hut (2003: 93) and Chavanis (1998).

van Kampen, 1981; Risken, 1989).[13] I note in addition that many of these results have recently been extended to the non-linear Fokker–Planck equation too (e.g., Frank (2005) and references therein). Of course, complete vindication of my claim will hang on demonstrating the match between particular astrophysical systems and the assumptions (boundary conditions and so on) used in any particular *H*-thoerem.

There are scores of other kinetic equations used in astrophysics and for many of these one will also find an *H*-theorem in the literature. And for those that do not readily admit of an *H*-theorem, one may also try employing the conjecture of De Roeck *et al.* to find an '*H*-theorem' of sorts. Recall that in Section 3.2 I described a top–down way of thinking about entropy increase and *H*-theorems. Imagine we have some deterministic macroscopic equation of motion, one that tells us that macro-states such as M_1 at t_1 will evolve by time t_2 into macro-state M_2. We saw that Liouville's theorem and the claim that (effectively) $\phi_{t_2-t_1}\Gamma_{M_1 t_1} \subset \Gamma_{M_2 t_2}$ implies that almost all of the points originally in M_1 have evolved into the set corresponding to M_2. From this it follows that $\Gamma_{M_1 t_1} \leq \Gamma_{M_2 t_2}$ and we therefore have a kind of *H*-theorem. Whether this strategy is defensible and whether it works with certain equations in astrophysics are questions that require study. I will not argue for either here. Presently I merely wish to point out that with the plethora of macroscopic kinetic equations in the field, there will be many opportunities to try to employ this strategy.

The picture we have developed, then, is this. We have not calculated the Boltzmann entropy including strong gravitational coupling directly, so we do not know whether it increases or decreases from an initial state like the early cosmological state. For the reasons discussed in Section 3.3, unless we simplify our system considerably we cannot show what the Boltzmann entropy of such a state will do. We have no answer to the main question of this chapter; indeed, displaying this problem is the main point of the chapter. As mentioned, however, the news is not all bad. We know that when some large self-gravitating structures in the universe reach a certain stage of development it becomes appropriate to idealize them as obeying a gravitational kinetic equation. For some of these equations, and in fact for some very accurate ones, we can show that the Boltzmann entropy increases. I have not shown this here, as it is implicit in the literature.[14] Moreover, I have

[13] Please bear in mind that often these works are not written from the perspective of the Boltzmann viewpoint used here. To complete all the links mentioned, one sometimes will need to use, for example, the fact that the Boltzmann entropy is close in value to the Gibbs entropy at equilibrium for large systems, as well as results from Boltzmann's original derivation.

[14] It may be worth pointing out that the diffusion coefficient in the Fokker–Planck equation causes dispersal in velocity space. So if we think back to Section 3.3, where we wanted to know what was happening in momentum space in such systems, we see that these kinetic equations are describing systems whose momenta are getting more dispersed as time goes on, just as we hoped.

pointed to the vast range of gravitational kinetic equations in use as a place to investigate this question further.

To what extent the Boltzmannian program is ultimately successful in the face of gravity depends on what we hope for and on the empirical facts. The original past hypothesis covered the entire universe, but this theory will not be vindicated by the current very limited result. The current move only yields an increasing Boltzmann entropy in regimes appropriately described by a gravitational kinetic equation. For instance, the Fokker–Planck regime only lasts when the system is weakly coupled. The whole universe is certainly not such a regime. If one hopes for a Boltzmann entropy for the universe, this avenue cannot meet this goal. Also, if one wanted to tackle the problems of extensivity *et al.* head-on, we have not done that here either. By going to a regime where a mean force is used, even where close encounters are considered to some extent, we may be accused of ignoring the problem of extensivity rather than addressing it.

Yet if one has more modest expectations, one has encouraging news. What is perhaps the best kinetic equation incorporating gravitational effects generates the increase of Boltzmann entropy. The natural reconstruction of the past hypothesis is as the claim that the early states of (e.g.) Fokker–Planck regimes are of very low Boltzmann entropy compared to now. The pressing empirical question for this approach is whether we are in such a regime and if so how big it is and how many there are.[15]

This picture, it must be said, bears some similarity to the 'branch' systems approach to statistical thermodynamics. Reichenbach's (1999) branch hypothesis is the claim that thermodynamics applies only to quasi-isolated macroscopic 'branch' systems. Thermodynamics does not apply to the universe as a whole on this view, but only to certain systems when they become sufficiently isolated from the rest of the world. Historically, one objection to this picture is that it is not at all clear what 'sufficiently isolated' could possibly mean. See Albert (2000: 88–89) for a forceful statement of this objection, among others. Here I just want to note that the proposal under review is not guilty of *this* mistake, at least on one reading. The criterion of whether a system fits the assumptions underlying the use of a Fokker–Planck equation is quite clear. The identification of branches can proceed without too much difficulty. The larger problem, also mentioned by Albert, of whether one has any

[15] Actually, probably more of the action will come in looking at the level of detail – i.e. the choice of macro-states – than simply the size of the system. For instance, our galaxy, the Milky Way, has approximately $N = 10^{10}$ stars in it and a 'crossing time' of 10^8 yr, making stellar close encounters a relatively unimportant part of its evolution. This means the Vlasov equation is a good description of our galaxy. This equation, recall, provides no entropy increase. However, that does not mean that if one wants to look at more fine-grained structure in our galaxy one cannot use the Fokker–Planck equation, an equation from which one can derive entropy increase. And that does not mean that one cannot also enlarge the scale and use the Fokker–Planck equation to describe the dynamics of clusters of galaxies, with $N = 10^3$, which may include the Milky Way.

right to impose a uniform probability distribution over the 'first' such state when we know it has evolved from previous states lingers, however, and demands further thought (for a little in this direction, see Callender, 2010).

In sum, I hope to have shown how the inclusion of gravity into the Boltzmannian account of the direction of time is highly ambitious but also non-trivial. After sketching the serious problems with gravity, I made plausible a sketch of how one can obtain an increasing Boltzmann entropy in self-gravitating systems described by certain types of gravitational kinetic equations. Further work is needed to judge whether this kind of approach is best, but I do hope it removes some of the pessimism one might (reasonably) have about Boltzmann's non-equilibrium framework in the presence of gravitation.[16]

References

Albert, D. Z. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.

Balescu, R. (1963). *Statistical Mechanics of Charged Particles*. New York: Interscience Publishers.

Binney, J. J. and Tremaine, S. (1987). *Galactic Dynamics*. Princeton, NJ: Princeton University Press.

Callender, C. (2004). Measures, explanations and the past: should 'special' initial conditions be explained? *The British Journal for the Philosophy of Science*, **55**(2), 195–217.

Callender, C. (2010). The past histories of molecules. In *Probabilities in Physics*, ed. C. Beisbart and S. Hartmann. Oxford: Oxford University Press.

Chavanis, P. H. (1998). On the 'coarse-grained' evolution of collisionless stellar systems. *Monthly Notices of the Royal Astronomical Society*, **300**, 981–991.

Chavanis, P. H. (2005). On the lifetime of metastable states in self-gravitating systems. *Astronomy and Astrophysics*, **432**, 117–138.

Dauxois, T., Ruffo, S., Arimondo, E. and Wilkens, M. (eds.) (2002). *Dynamics and Thermodynamics of Systems with Long Range Interactions*. Berlin: Springer-Verlag.

DeRoeck, W., Maes, C. and Netočný, K. (2006). H-theorems from macroscopic autonomous equations. *Journal of Statistical Physics*, **123**(3), 571–584.

Dunning-Davies, J. (1983). On the Meaning of Extensivity. *Physics Letters,* **94A**, 346–348.

Earman, J. (2006). The past hypothesis: not even false. *Studies in History and Philosophy of Modern Physics*, **37**(3), 399–430.

Frank, T. (2005). *Nonlinear Fokker–Planck Equations*. Berlin: Springer-Verlag.

Frigg, R. (2009). Typicality and the approach to equilibrium in Boltzmannian Statistical Mechanics. *Philosophy of Science (Suppl.)*.

Garrido, P. L., Goldstein, S. and Lebowitz, J. L. (2004). Boltzmann entropy for dense fluids not in local equilibrium. *Physical Review Letters*, **92**, 50602-1–50602-4.

[16] Thanks to Jonathan Cohen, Roman Frigg, Carl Hoefer, Tarun Menon, Ioan Muntean and Allan Walstad for comments and help, as well as audiences at the 2006 Philosophy of Science Association Meeting, the 2008 Reduction, Emergence and Physics Workshop in Tilburg and the 2008 Annual Meeting of the British Society for Philosophy of Science.

Goldstein, S. (2002). Boltzmann's approach to statistical mechanics. In *Chance in Physics, Foundations and Perspectives*, ed. J. Bricmont, D. Durr, M. C. Galavotti *et al.*, Berlin: Springer-Verlag.

Goldstein, S. and Lebowitz, J. L. (2004). On the (Boltzmann) entropy of non-equilibrium systems. *Physica D: Nonlinear Phenomena*, **193**, 53–66.

Green, M. S. (1952). Markoff random processes and the statistical mechanics of time dependent phenomena. *Journal of Chemical Physics*, **20**, 1281.

Heggie, D. C. and Hut, P. (2003). *The Gravitational Million-body Problem*. Cambridge: Cambridge University Press.

Hertel, P., Narnhofer, H. and Thirring, W. (1972). Thermodynamic functions for fermions with gravostatic and electrostatic interactions. *Communications in Mathematical Physics*, **28**, 159–176.

Kandrup, H. (1981). Generalized Landau equation for a system with a self-consistent mean field–Derivation from an *N*-particle Liouville equation. *The Astrophysical Journal*, **244**, 316.

Kandrup, H. Theoretical techniques in modern galactic dynamics, unpublished class notes available at http://www.astro.ufl.edu/∼galaxy/papers/.

Landau, L. D. and Lifshitz, E. (1969). *Statistical Physics*, part 1. Oxford: Pergamon.

Lavis, D. A. (2005). Boltzmann and Gibbs: an attempted reconciliation. *Studies in History and Philosophy of Modern Physics*, **36**(2), 245–273.

Lévy-Leblond, J.-M. (1969). Nonsaturation of gravitational forces. *Journal of Mathematical Physics*, **10**(5), 806–812.

Liboff, R. L. and Fedele, J. B. (1967). Properties of the Fokker–Planck equation. *Physics of Fluids*, **10**, 1391–1402.

Lieb, E. H. and Yngvason, J. (1998). A guide to entropy and the second law of thermodynamics. *Notices of the American Mathematical Society* **45**, 571–581.

Lynden-Bell, D. (1967). Statistical mechanics of violent relaxation in stellar systems. *Monthly Notices of the Royal Astronomical Society*, **136**, 101–121.

Lynden-Bell, D. and Wood, R. (1968). The gravo-thermal catastrophe in isothermal spheres and the onset of red-giant structure for stellar systems. *Monthly Notices of the Royal Astronomical Society*, **138**, 495.

Nauenberg, M. (2003). Critique of Q-entropy for thermal statistics. *Physical Review E*, **67**(3), 036114.

Ogorodnikov, K. F. (1965). *Dynamics of Stellar Systems*. New York: Pergamon.

Padmanabhan, T. (1990). Statistical mechanics of gravitating systems. *Physics Reports*, **188**(5), 285–362.

Price, H. (1996). *Time's Arrow and Archimedes' Point.* New York: Oxford University Press.

Reichenbach, H. (1999). *The Direction of Time.* Dover, NY: Mineola.

Risken, H. (1989). *The Fokker–Planck Equation: Methods of Solution and Applications*. Berlin: Springer-Verlag.

Rowlinson, J. S. (1993). Thermodynamics of inhomogeneous systems. *Pure and Applied Chemistry*, **65**, 873.

Saslaw, W. C. (2000). *The Distribution of the Galaxies: Gravitational Clustering in Cosmology*. Cambridge: Cambridge University Press.

Schrödinger, E. (1989 [1948]). *Statistical Thermodynamics*. Dover reprint.

Spohn, H. (1991). *Large Scale Dynamics of Interfacing Particles*. Berlin: Springer.

Touchette, H. (2002). When is a quantity additive, and when is it extensive? *Physica A: Statistical Mechanics and its Applications*, **305**(1–2), 84–88.

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In *Philosophy of Physics, Handbook of the Philosophy of Science*, eds. J. Butterfield and J. Earman. Amsterdam: North-Holland, pp. 923–1047.

Van Kampen (1981). *Stochastic Processes in Physics and Chemistry*. Amsterdam North-Holland.

Wald, R. M. (2006). The arrow of time and the initial conditions of the Universe. *Studies in History and Philosophy of Modern Physics,* **37**(3), 394–398.