Evaluate the Significance of the Boltzmann H theorem in the Foundations of Statistical Mechanics. Describe and Compare the Detailed Nature of the Reversibility and the Recurrence Objections.

Abstract

The recurrence and reversibility objections entail that the Boltzmann equation and H theorem cannot hold for all states at any time and any state for all times respectively and as such limit the applicability of the theorem. The force of these objections are overshadowed by the existence of fluctuation phenomena as only these force us to understand the theorem probabilistically. In virtue of the theorem's success I argue that despite these objections the Boltzmann equation should be understood as a probabilistic coarse-grained dynamical evolution equation that holds for all systems and times of interest. The significance of the H theorem is that it provides a dynamical underpinning to the dynamic-less explanations of the increase in entropy and approach to equilibrium of thermodynamics and statistical mechanics.

1 The Boltzmann Equation as a Coarse Grained Dynamical Evolution Equation

The Boltzmann equation has two general uses. Firstly, it can be used to calculate the macroscopic behaviour of a gas, in particular the viscosity and heat conduction coefficients, from a microphysical model of that gas. Secondly, it can be used to calculate the typical behaviour of *single* particle in regimes where molecular encounters are negligible. Such regimes include the fluid-like regime of a moderately dense gas and the free molecular regime. Technically, the first usage corresponds the the regime in which the mean distance between particle collisions is much less than other scales in the system (e.g. the vessel size) and the second the converse (Cercignani). Furthermore, in both these roles, the Boltzmann equation is not just applicable to the classical gases for which it was originally derived for, rather, when properly generalized, the Boltzmann equation can be used in the study of electron transport in solids and plasmas, neutron transport in nuclear reactors, phonon transport in superfluids, and radiative transfer in planetary and stellar atmospheres (Cercignani). In these roles the Boltzmann equation has been incredibly successful and it is continued to be used on a day to day basis by physicists, engineers and chemists across a range of disciplines from nanotechnology to astrology.

The Boltzmann Equation is derived by considering a gas contained in an isolated fixed vessel. It describes the evolution of the position and velocity distribution of the gas, $f(\mathbf{v}, \mathbf{r}, t)d\mathbf{v}\ d\mathbf{r}$, over time. In Boltzmann's original derivation $f(\mathbf{v},\ \mathbf{r}, t)d\mathbf{v}d\mathbf{r}$ denotes the fraction of molecules in the gas with velocities between $\mathbf{v}$ and $\mathbf{v}+d\mathbf{v}$ and located between r and $\mathbf{r}+d\mathbf{r}$. In modern text books (e.g. Liboff/Springer) $f(\mathbf{v},\ \mathbf{r}, t)d\mathbf{v}d\mathbf{r}$ is generally taken to denote the single particle velocity and position probability distribution in the single particle velocity and position probability distribution. Although conceptually different, it is possible to run the derivations, the structure, and the objections to both cases $f(\mathbf{v},\ \mathbf{r}, t)d\mathbf{v}d\mathbf{r}$ in parallel.

I will set out a derivation of the Boltzmann equation for a dilute, isotropic and spatially homogeneous gas. In this case we assume that: i. the distribution is spatially homogeneous for all times $f(\mathbf{v}, r, t) = f(\mathbf{v}, t)$; ii. the velocity distribution is isotropic for all times $f(\mathbf{v}, t) = f(v, t)$; iii. the gas is sufficiently dilute that we only need to consider binary collisions. These assumptions can be relaxed to derive more general versions of the Boltzmann equation. The derivations are conceptually the same but technically more complex.

Essential to any derivation of the Boltzmann Equation is the assumption that: iv. the molecules only interact at short distances and so only collisions between molecules affect f(v). When $f(\mathbf{v},\ \mathbf{r}, t)$ is taken to represent the fraction of particles with a given probability distribution we also need to assume: v. the number of molecules is large enough that the discrete distribution can be approximated as continuous.

A final assumption is required. It can be formulated in two ways. The SZA is generally accepted (Brown, Uffink, Sklar) to be Boltzmann's original assumption. The assumption of molecular chaos is a modern alternative that fits naturally with the probabilistic understanding of f. In both cases the assumption is formulated by dividing a group of particles that are about to collide into two groups which can be referred to as the '**I**ncoming' and '**T**arget' particles.

SZA: The distributions of the incoming and target particles just before the collision f_BC are uncorrelated:

$$f_{BC}(v_I, v_T, t) = f_{BC}(v_I, t) \times f_{BC}(v_T, t) \tag{1}$$

Molecular Chaos: The fine grained correlations between the velocities of any pair of particles do not affect the evolution of the coarse grained degrees of freedom. Thus we are free to factorise the velocity distribution:

$$f(v_I, v_T, t) \rightarrow f(v_I, t) \times f(v_T, t) \tag{2}$$

The fine grained description includes all dynamical details of the microstate of the system. A coarse grained description of a system is a description of a system in terms of its macroscopic variables, such as temperature and pressure.

The molecular chaos assumption is stronger than the SZA in the sense that it applies to the distribution of all particles at all times in the gas rather than solely the distributions of pairs of particles that are about to collide. However, the assumption is also weaker in the sense that while the SZA is making an absolute claim about this fine grained description, a claim about the actual microstate of all dilute gases at all times, the molecular chaos assumption accepts that there may be correlations but that these fine grained details are dynamically ignorable when interested in the evolution of the coarse grained degrees of freedom. Understood as an absolute claim, molecular chaos would clearly be false. The velocities of a pair of particles who have just collided are correlated given momentum conservation. However, because by the time the particles re-collide, if ever, the correlations will have been washed out and forgotten we can ignore these correlations. Given, this two way pull, there could also be an intermediary version of the SZA, call it SZA_{CG}, that states that any correlations between the about to collide particles are dynamically ignorable.

It is natural to assume molecular chaos rather than the SZA when considering $f(\mathbf{v},\ \mathbf{r}, t)d\mathbf{v}d\mathbf{r}$. This is because given molecular chaos, the probability distribution of each particle in the gas can be treated as independent, and thus the probability distribution for the gas of N particles is simply the tensor product of N single particle probability distributions and the probable number of particles with velocity and positions in the range v ,$v + dv$ is $Nf(\mathbf{v},\ \mathbf{r}, t)d\mathbf{v}d\mathbf{r}$.

Given the molecules act only at short distances, to determine how the distribution evolves with time we need to ascertain how collisions change the distribution. Given the continuity of the distribution, and a little inspection, it is clear that we can express the change in the distribution of the incoming group by the following integrals.

$$\begin{aligned}\Delta f(v_I) = &\int f_{BC}(v_I', v_T', t)C(v_I', v_T' \rightarrow v_I, v_T)dv_I'dv_T'dV_T\Delta t \\ &- \int f_{BC}(v_I, v_T, t)C(v_I, v_T \rightarrow v_I', v_T')dv_I'dv_T'dv_T\Delta t\end{aligned} \tag{3}$$

The primed coordinates are the velocities of the molecules before the collision, the unprimed the velocities after. The first integral gives the number of molecules added to range of v to $v + dv^3$, the second the number taken out. $C(v_I', v_T' \rightarrow v_I, v_T)$ gives the fraction of collisions that take a pair of particles from initial velocities v_I' and v_T' respectively to the final velocities v_I and v_T. Physically, $C(v_I', v_T' \rightarrow v_I, v_T) = |v_I - v_T|\sigma(v_I'v_T'v_Iv_T)$ where σ is the collision cross section. The exact form of the collision cross-section is calculated for a particular system using collision dynamics.

The time symmetry of classical mechanics ensures that $C(v_I', v_T' \rightarrow v_I, v_T) = C(v_I, v_T \rightarrow v_I', v_T')$. Given this, the SZA, and the validity of the continuous limit, we can rearrange Eq. 3 to give the Boltzmann Equation:

$$\frac{df(v)}{dt} = \int [f(v_I', t)f(v_T', t) - f(v_I, t)f(v_T, t)]C(v_I', v_T' \rightarrow v_I, v_T)dv_I'dv_T'dv_T \tag{4}$$

This equation gives the rate of change in the distribution of speeds of the gas molecules over time.

Having derived the Boltzmann equation, it directly follows, via a few mathematical manipulations, that the functional:

$$H[f(v, t)] = \int f(v, t)ln[f(v, t)]d^3v \tag{5}$$

Is non-increasing for all times;

$$\frac{dH}{dt} \leq 0 \tag{6}$$

Furthermore, iff $\frac{dH}{dt} = 0$, the distribution is the Maxwell-Boltzmann distribution: $f(v, t) = Ae^{-Bv^2}$.

If all the stated assumptions are valid then we have established that: The H of that system decreases monotonically until it reaches the minimum value of H at equilibrium. Then, it remains at that minimum value of H for all subsequent times.

The molecular chaos assumption behaves as a correlation-discard coarse graining method (Wallace's terminology). In effect the molecular chaos assumption amounts to a Markov assumption in that we assume that any correlations between particles, say resulting from collisions, that could affect the evolution of macroscopic variables are rapidly 'forgotten'. A similar correlation discard method is used in the derivation of the quantum Markovian master equation. In this case, we assume that the density operator of the system and environment can be written as a product state for all times $\rho_{S,E} = \rho_E \otimes \rho_S$ if we are interested solely in the system's evolution. For this assumption to be justified we need the time scale on which the system and environment correlations are forgotten (quantified by the rate of decay of the system environment correlation functions) to be much quicker than that of the time scale on which system evolves (Breuer 2013). Discarding correlations is only one example of a coarse graining method; other include diagonalising the density operator of the system in a particular basis or smoothing the probability distribution over certain cells in phase space.

The structure of the derivation of the Boltzmann equation, in particular what it means to say the molecular chaos assumption should be understood as a correlation discard coarse graining method and the implications of disregarding the correlations, can be clarified by a quantum analogue to the Boltzmann H theorem (variant on Peres). Switching to the quantum picture is insightful both because the correlation discard method is more familiar in the context quantum mechanics and because the familiar abstract formalism provides an introduction to seeing the Boltzmann equation as part of a more general framework.

Call the combined density state of the target and incoming particles $\rho_{I'T'}$. To make the SZA_2 we disregard any correlations between the incoming and target particle states by replacing the combined density operator with the tensor product of their two reduced density operators. [1]

$$\rho_{I'T'} \rightarrow Tr_{T'}[\rho_{I'T'}] \otimes Tr_{I'}[\rho_{I'T'}] \tag{7}$$

Let the operator U represent the effect of colliding the particles. After the collision the state of the combined system is thus:

$$\rho_{IT} = U(Tr_{T'}[\rho_{I'T'}] \otimes Tr_{I'}[\rho_{I'T'}])U^{\dagger} \tag{8}$$

We now consider the function H' during this process:

$$H'(\rho) = Tr[\rho ln(\rho)] \tag{9}$$

I make no claims here about what H' represents, in the same way I have yet to make any claims about what H represents. All that matters is that it is a definable function that, as we will see, we can prove is non increasing with time.

H' clearly remains constant under the unitary evolution of the system $H'(\rho_{IT}) = H'(\rho_{I'T'})$ as unitary evolution is reversible.

However H' decreases during the correlation discard process if the initial state $\rho_{I'T'}$ cannot be written as the product state $\rho_{I'} \otimes \rho_{T'}$. This follows from the easily proven sub-additivity inequality [2]:

$$H'(Tr_A[\rho_{AB}]) \otimes Tr_B[\rho_{AB}]) = H'(Tr_A[\rho_{AB}]) + H'(Tr_B[\rho_{AB}]) \leq H'(\rho_{AB}) \tag{10}$$

From which it directly follows that:

$$H'(\rho_{IT}) = H'(Tr_{T'}[\rho_{I'T'}] \otimes Tr_{I'}[\rho_{I'T'}]) \leq H'(\rho_{I'T'}) \tag{11}$$

---

[1] When Peres sets out this derivation he assumes that there are initially no correlations between the particles and so starts with the initial state $\rho'_I \otimes \rho'_T$. He then takes the partial trace of the two particles after they have collided $\rho_{IT} \rightarrow Tr_T[\rho_{IT}] \otimes Tr_I[\rho_{IT}]$ on the presumption that these correlations are rapidly forgotten and so will not affect the future evolution of the macroscopic variables. While operationally this assumption amounts to the same result, it is a conceptually stronger assumption than the SZA. I assume that the correlations between any pair particles are lost by the time that pair next collide, not that there are no correlations at all subsequent times.

[2] Proof of Sub-Additivity (Peres):
First, consider the quantity $S(\rho_a|\rho_b) = Tr[\rho_a(ln(\rho_a) - ln(\rho_b))]$. This quantity can be shown to be always non negative as follows.
Let $\rho_a = \sum_m a_m \left|a_m\right\rangle \left\langle a_m\right|$ and $\rho_b = \sum_n b_n \left|b_n\right\rangle \left\langle b_n\right|$.
In the $\left|a_m\right\rangle$ basis the diagonal elements of $ln(\rho_b)$ are: $ln(\rho_b)_{mm} = \sum_n ln(b_n)| \left\langle b_n|a_m\right\rangle |^2$.
$| \left\langle b_n|a_m\right\rangle |^2$ is a matrix, call it M_{nm}, with the property that $\sum_m M_{nm} = \sum_n M_{nm} = 1$. Making use of this property, $Tr[\rho_a(ln(\rho_a) - ln(\rho_b))] = \sum_{mn} a_m M_{nm} ln(\frac{a_m}{b_n})$
Now, $ln(x) \geq 1 - x^{-1}$ with equality holding for x = 1.
Thus, $S(\rho_a|\rho_b) \geq 0$ and the equality holds for $\rho_a = \rho_b$.
Let $\rho = \rho_{I'T'}$. Let $\rho_1 = Tr'_T[\rho_{I'T'}]$ and $\rho_2 = Tr'_I[\rho_{I'T'}]$ and have eigenvalues ω_{1i} and ω_{2j} respectively. We know from the above that $S(\rho_1 \otimes \rho_2|\rho) := Tr[\rho(ln(\rho - ln(\rho_1 \otimes \rho_2) \leq 0$.
Now, $H(\rho_1) + H(\rho_2) = \sum_i \omega_{1i} ln(\omega_{1i}) + \sum_j \omega_{2j} ln(\omega_{2j}) = \sum_{ij} \omega_{1i}\omega_{2j} ln(\omega_{1i}\omega_{2j}) = H(\rho_1 \otimes \rho_2)$
Furthermore, $Tr(\rho ln(\rho_1) = \sum_{mpnq} \rho_{mp,nq} ln(\rho_1)_{nm} \delta_{pq} = Tr(\rho_1 ln(\rho_1)$. (And similarly for ρ_2.)
Thus, substituting these back in, we have proved sub-additivity: $H(\rho) \geq H(\rho_1) + H(\rho_2)$ where the equality only holds if $\rho = \rho_1 \otimes \rho_2$.

The change in H' occurs during the step at which correlations are discarded, not during the evolution. Thus this demonstrates it is the fact that we chose to focus on coarse grained details, rather than the underlying fine grained dynamics, that results in a decrease in H'.

It is possible to run an analogous derivation in the original classical case in terms of abstract evolution and projection maps. The sub-additivity of H can also be derived in classical information theory (Penrose and Percival).

The process set out above of alternating between evolving a system forward in time and then coarse graining the evolution equation can be iterated. This generates a coarse grained evolution equation for a system.

Let C denote the coarse graining map of a state. A coarse graining map is one such that $C^2 = C$ (an already coarse grained state cannot be coarse grained further). Let L denote the evolution map of the system. In the case of my coarse graining process set out above $C\rho_{AB} = Tr_B\rho_{AB} \otimes Tr_B\rho_{AB}$ and $L = U\rho U^\dagger$.(I have assumed here for simplicity that the evolution operator is time independent. The generalisation is straightforward).

For an appropriately chosen coarse graining method and evolution, the evolution of the system will be insensitive to the exact time δt that the system is evolved for between course grain projections. In the case of the quantum Boltzmann equation, it does not matter exactly when we forget the correlations, as long as they are forgotten.

We can thus write:

Classically: $f(t) = LCL....CL\ f(0)$ Quantum Mechanically: $\rho(t) = LCL....CL\rho(0)$

In the classical case f(t) could either represent a single microstate, a fractional distribution, or a probability distribution. In the quantum case ρ could be either a pure density operator or a mixed state density operator. The probabilistic formulation is more common in physics. However, a non probabilistic interpretation is compatible with my set out of the H theorem thus far.

In order for the coarse grained equation to work in a time period [0,t] the microstate/distribution/density operator generated by alternately evolving and coarse graining the microstate/distribution/density operator must have the same macroscopic properties as the microstate/distribution/density operator would have had evolving for the entire time period under just the evolution operator. In other words, We require, for an initial state $f(0)$, that $f(t) = LCL...CLf(0)$ and $f'(t) = Lf(0)$ have the same macroscopic properties.

A large body of equations can be thought to have this general structure including heat transport in oceans, environment induced decoherence, the BBGKY hierarchy, as well as the Boltzmann equation.

2 Reversibility, Recurrence and Fluctuations

In this section of the essay I start by setting and comparing out the reversibility, recurrence and then fluctuation objections. I argue these objections rule out an absolute understanding of the H theorem but that instead we should understand the theorem probabilistically.

Our best theories of microphysics, in particular classical mechanics and unitary quantum dynamics, are time symmetric in the sense that the fundamental laws of motion are form invariant under time reversal. The reversibility objection highlights that the derivation of the H theorem cannot involve only mechanical assumptions because it is impossible, almost as a matter of logic, for the H theorem to establish an asymmetric dynamical law from entirely time symmetric premises. Thus it raises the challenge of if/how the non mechanical assumptions that introduce the asymmetry to the theorem can be justified.

There are two, interrelated, ways to bringing out the force of this objection. Firstly, the reversibility argument provides a mechanism for generating counter examples to the Boltzmann H theorem. Say I have a box gas that I open in a room at t = 0. The gas doubles in volume by time t and H decreases accordingly. Now consider this resultant lower H state at t but with all the velocities reversed. For a classical gas, the time symmetry entails that $(\mathbf{q}_t, \mathbf{p}_t) = L(\Delta t)(\mathbf{q}_0, \mathbf{p}_0)$ iff $L(\Delta t)(\mathbf{q}_t, -\mathbf{p}_t) = (\mathbf{q}_0, -\mathbf{p}_0)$ where $\mathbf{q}_t$ and $\mathbf{p}_t$ are the position and momentum vectors respectively of a particle at time t. Thus, such a state will evolve back to the initial high H state with all the particles in smaller volume of the box. There is nothing in the laws of classical mechanics that rules out such an initial state and so this acts as a counter example to the H theorem when understood as making claims about all systems.

Secondly, in practice, the counter examples are correlated; they all correspond to taking the time inverse of an evolving system. As such, the Boltzmann equation does not work for retrodicting the evolution of a system. In the above example, the Boltzmann equation could predict that the gas would have doubled in volume by t. However, the Boltzmann equations time inverse at t would retrodict that at t = 0 the gas would be more expanded still. The prediction made at t = 0 and retrodiction made at t contradict one another and the retrodiction contradicts what we know to have in fact occurred. We need to explain what it is about the world and the nature of the Boltzmann equation that makes the equation fail abjectly to correctly determine the past evolutions while being so successful at explaining forward evolutions.

Poincare's recurrence theorem loosely states that almost all initial states of an energetically isolated system that is confined to a finite volume will return arbitrarily close to their initial state in a finite time. The intuition behind why this the case, and the intuition that drives proofs of recurrence theorems (Wallace 2013), is that a state confined to a finite region of phase space has a limited number of points in phase space that it can visit and thus it will eventually start revisiting ones.

More formally (Brown 2015), the system's fine grained microstate is given by a point in its 6N dimensional phase space. Say we have a dynamical system represented by $\langle \Gamma, A, \mu, \Phi_t \rangle$ where Γ is the state space of the system, A is a σ algebra of measurable subsets of Γ, μ is measure on the measure space $\langle \Gamma, A \rangle$ and Φ_r is a one parameter family of mappings: $\Gamma \Rightarrow \Gamma$ such that $\Phi_t(\Gamma) = \Gamma$.

Recurrence Theorem: Consider an arbitrary measurable set of points g in A and some finite number τ. The subset g* of g that contains all the points p in g such that $\Phi_t(p)$ is not a member of g, for any finite $t \geq \tau$, has measure zero: $\mu(g*) = 0$.

We should not postulate that our world is one of the special measure zero sets because this exception does not appear in the quantum recurrence theorem. The reason for this is that the Hilbert space norm, which acts as a natural distance measure in quantum dynamics, is preserved. Thus, if all non measure zero points recur, then all points, including measure zero points, recur. Given this, it cannot be the case that the measure zero clause allows that actual world to avoid the force of the recurrence theorem.

The recurrence objection provides a counter example to the claim that H is non-increasing for any system (with non stationary H) for all times. Given the recurrence theorem, there is some time t_r, such that the microstate of the system at this time is arbitrarily close to the initial microstate of the system: $x(t_r) \simeq x(t_0)$. At this time, $H(x(t_r)) \simeq H(x(t_0))$. As such, if the H of a system changes at all, it must increase as much as it decreases in a complete cycle.

The recurrence objection, to a greater extent than the reversibility objection, is purely a conceptual worry. The theorem does not apply to any real system because the phase space of the actual universe is infinite and it does not apply to any subsystem of the universe because no subsystem is completely energetically isolated. Furthermore, recurrence times of realistic macroscopic systems are longer than the age of the universe. Thus, we do not have to worry about any actual system recurring, but the idealised systems that the H theorem is derived for, would theoretically recur.

Thermal fluctuations are an empirically well confirmed phenomena. They have significant implications for phase transitions and chemical kinetics. Furthermore, we have developed empirically successful formulae (such as the fluctuation-dissipation theorem) to quantify them. Given the existence of fluctuations, we know that, for any finite system, H is not non-increasing for all times, rather there will be small fluctuations when H increases. In particular, such increases in H will not be confined to times close to recurrence. Thus, the existence of fluctuations provides a counter example to the statement that the H theorem is non-increasing for any system for all times up to a finite time τ where τ is close to recurrence time t_r. Furthermore, in that fluctuations are empirically observed, this objection has practical as well as conceptual force.

How do the objections compare? Given how I have set out the objections above, the reversibility, recurrence and fluctuation objections form a hierarchy. The reversibility objection shows that the H theorem cannot hold for all initial states of a system at any time but leaves open the option of claiming that there are no non-ignorable correlations in the velocities of particles before a collision is simply a contingent fact of the universe. It is only the recurrence argument that shows on pain of contradicting classical mechanics that this cannot be the case for all times for any closed isolated system. This point it recognised by Brown et al. However, if this style of argument is extended to include fluctuations it is clear both the recurrence and reversibility objections are overshadowed by the impact of fluctuations. This is because while the recurrence and reversibility objections show that there must be certain microstates that could in effect allow for fluctuations in that they evolve into higher H states, it could be a contingent fact that these states do not occur in the universe except going backwards in time or at times close to recurrence. It is only the existence of fluctuation phenomena that in effect rules out this option. Furthermore, in doing so, as I will later argue, the existence of fluctuation phenomena, force us to interpret the H theorem, at best, as probabilistically. As it is only the existence of fluctuation phenomena that forces this conceptual shift, I take it that its importance over shadows that of either the reversibility or recurrence objections.

The counter examples provided by the reversibility, recurrence and fluctuation objections do rule out understanding the H theorem as applying to all systems at all times. However, it should be clear that this is not the only way in which we can understand the H theorem. It is entirely consistent with these objections that H is typically non-increasing for all systems of interest and all times of interest.

Why should we understand the H theorem in this weaker way? We should care because it would be mad to disregard the theorem entirely in virtue of these objections given the incredible empirical success that the Boltzmann equation has enjoyed. It would be especially mad because the reversibility, recurrence and fluctuation phenomena are not unique to the Boltzmann Equation but rather affect any coarse grained dynamical evolution equation. As I emphasised earlier, there are a large body of such equations and these equations have also been incredibly successful. This success across a broad spectrum of types of systems, with different microphysical make ups, using similarly structured equations of motion, begs explanation. The simplest explanation, one that is generally adopted by scientific realists in other contexts and that we should also accept now in the absence of any alternative, is that these equations have latched onto some fact about reality. This fact, I take it, is that for systems of interest and times of interest certain microscopic fine-grained details are typically dynamically ignorable.

In order to defend my claim that the H theorem establishes this weaker claim, I will now proceed to explain: firstly, why the theorem is limited in scope and yet is successful otherwise and secondly, why and how the theorem should, in the light of thermal fluctuations, be understood probabilistically.

The correct response to the existence of fluctuation phenomena is to accept that the theorem should be understood probabilistically: Boltzmann's original interpretation of $f(v, x, t)$ as the fractional number of particles in the volume of gas with velocities and positions in the regions v and $v + dv$ and x and $x + dx$ is inadequate, rather we should understand f as the probability distribution of a single particle. The H theorem would then refer to the average value of H over the distribution. Given that H is now an average it is clear that the theorem would allow for fluctuations away from the average.

That the Boltzmann equation should be understood probabilistically is unsurprising when we consider its second general usage (mentioned on pg. 2). The Boltzmann equation can be used to calculate the typical behaviour of a single particle in regimes where molecular encounters are negligible. That the equation can be successfully used in this context where it is to be understood as a single particle probability distribution suggests that it makes sense to consider the equation probabilistically.

It is the SZA that introduces the asymmetry to the account. As seen for the quantum analogue, H is constant during the unitary evolution process but non-increasing when the system correlations are traced out. The same would also hold if we ran the process back in time: if we start with the final state and take the trace of the combined system this decreases H. (Note, the end state generated by the retrodiction will not be the same as the actual one at t_0.) The asymmetry is introduced by the fact that we chose to apply the coarse graining in only one temporal direction.

We can justify applying the coarse graining in only one temporal objection and in doing so explain why the Boltzmann equation works so well forwards for all times of interest, but not backwards and not at distant future times, by appealing to the special initial condition of all systems of interest. Consider again my box of gas. If we know the gas ends up (back at the beginning of the time frame we are considering) in the box, then it is entirely expectant that there will be unusually complex correlations holding between the velocities of about-to-collide particles that we cannot ignore if we want to obtain the correct macrodynamics. In this way we can explain why we cannot take the state of the system half way through its expansion and apply the Boltzmann equation backwards. This special initial condition can also be seen as a special end condition which explains the fact that there will later (for an idealised energetically isolated finite volumed system), close to recurrence times, also be special complex correlations between the velocities of the particles that cannot be ignored.

Similar response to these objections run for other coarse-grained dynamical equations. Any time asymmetric equation that is derived from classical mechanics or unitary quantum mechanics will have to appeal to a special initial state of the system that allows coarse graining to work only one direction in time. Similarly, fluctuation phenomena, require that these equations, as they in fact generally are in most textbooks, to be treated probabilistically.

Naturally, there are still conceptual issues to address. Firstly, it not clear what exactly it is about the underlying microdynamics that means certain correlations are ignorable. Sklar is especially concerned by this final point. Put another way, molecular chaos just assumes that these details are ignorable, this claim has not be proven and so the theorem have not been rigorously proven. (Nor has any interesting probabilistic variant on the Boltzmann equation been rigorously proven: Lanfords theorem is conceptually different and only holds for 1/5th of the time between particle collisions). Secondly, it is not clear what type of probabilities we are using here. The distribution could be an ignorance distribution, a distribution over an ensemble of systems, or as emerging from the probabilities intrinsic to quantum mechanics.

In response to the first issue, it is tempting to point out that the details must be ignorable because of the equations are successful; however, if we want to explain the success of these equations on the basis that these details are ignorable then we are in danger of arguing circularly. Nonetheless, I do not think a lack of current understanding of certain dynamical details is a reason reject this account of the Boltzmann equation and other coarse grained dynamical equations. It is heuristically plausible that it is just an objective fact of the world tat certain fine grained details do not affect the dynamics irrespective of whether we yet know the dynamic details grounding this fact. This objection thus highlights a worthwhile area of research. How rigorous a proof of a theorem is required to accept a theorem depends on one's approach to physics and philosophy of physics. The derivation of the Boltzmann equation is not a rigorous mathematical proof but then little is in physics. Take quantisation for example, there is no mathematically rigorous way to quantise a classical operator as a quantum mechanical operator given the physical significance of the complex phase in quantum mechanics. Ultimately it is experiment that determines whether a particular means of proof is successful. Similarly, that a probabilistic version has not been proven should not prevent us accepting that it is the correct way of understanding the theorem unless a convincing alternative explanation can be given for both the empirical success of the Boltzmann equation and fluctuation phenomena. Furthermore, we can always think of the Boltzmann equation as applying over a coarse grained time scale as a stop gap until a more complete probabilistic understanding it found.

To take the probabilistic concern, I will only point out that the philosophy of probability is far from clear cut in most contexts and so this it is unclear how we should interpret probabilities in the context of the H theorem is unsurprising.

I have tried to demonstrate that it is reasonable to accept that despite the reversibility , recurrence and fluctuation phenomena, H is typically non-increasing for all systems of interest and times of interest. In establishing that something, H, typically increases, (irrespective of whether -H is taken to refer to thermodynamic entropy) the H theorem has foundational significance in that it establishes why certain processes involving dilute gases appear irreversible.

3 The Broader Significance of The H Theorem Within Statistical Mechanics

I will now argue that the Boltzmann equation can also be understood as providing an explanation of the second law of thermodynamics for that range of systems that the Boltzmann equation applies to.

The Second Law of Thermodynamics: The change in thermodynamical entropy between any two subsequent equilibrium states of a total system is always non negative.

Thermodynamic Entropy (S): $S_A - S_B = \int_A^B \frac{dQ}{T}$ where dQ is the heat transfer during the reversible process from macrostate A to B at temperate T.

Thermodynamics, in consisting of a series of empirical principles, is a principle rather than a constructive theory. However, constructive theories, which build a more complex picture out of a relatively simple formal scheme, are more explanatory. The paradigm example of a constructive theory is kinetic theory where properties such as pressure and temperature are derived from a microphysical model of a gas. Statistical mechanics has historically been seen as in some sense a constructive theory; an attempt to explain the results of thermodynamics using a simple formal scheme. The oddity about statistical mechanics is that this formal scheme is built around the mathematical tool of the phase space rather than from the constituents of the actual dynamical systems. The Boltzmann equation is not part of kinetic theory as it depends on the SZA; however, nonetheless, it clearly is derived directly from a microphysical picture of the nature of gases. The Boltzmann Equation in this way has stronger credentials to explain thermodynamics.

The Boltzmannian phase space argument, subsequent to the H theorem and historically seen as superior, which purports to provide the explanatory background to the second law of thermodynamics, runs as follows.

Heuristically: The region of phase space corresponding to a system at equilibrium takes up the majority of the phase space. As such a system initially outside of this equilibrium will typically, within a short time, wander into equilibrium and then stay in there. The reversibility objection is avoided via the claim that if the initial state of a system was in a small macrostate far from equilibrium, it is entirely expected that any states backwards evolution will be away from equilibrium towards its initial far from equilibrium state.

This explanation is very general in that it applies to a range of microphysical models but at the cost of remaining completely silent about the dynamics involved. In particular it does not provide the dynamics that cause a system to evolve towards the equilibrium nor the dynamics that distinguish the typical, highly probable, initial microstates from the rare non typical ones that do not evolve towards equilibrium.

These gaps can be filled by the Boltzmann equation. The state of a gas approaches equilibrium at a coarse grained level because of the large numbers of collisions between gas molecules leading to a redistribution of energy and diffusion. Entropy increases because we focus on the evolution of coarse grained degrees of freedom. The typical states are the ones that SZA_2 holds for, ones where the fine grained correlations between the particles velocities make no difference to the evolution of the macroscopic degrees of freedom.

Excluding more subtle concerns that would require delving into a more precise account of reference, I can see no major obstacle to prevent us taking H to refer to thermodynamical entropy and Boltzmann Entropy.

The relation between H and thermodynamical entropy (ST) and Boltzmann entropy (SB) cannot and need not be exact. Firstly, H is defined for gases (and other entities that can broadly be modeled as gases) where as both Boltzmann entropy and Thermodynamic entropy are more general. This is entirely unsurprising as H's main advantage is that it is defined for a particular dynamical model. Secondly, thermodynamic entropy is only defined for equilibrium systems where as both H and statistical mechanical entropy are defined for non equilibrium systems. Given this, H should be considered a generalisation of thermodynamical entropy to non equilibrium evolutions but restricted to dilute gas models.

Once these caveats are set it, it would indeed seem that, - H and ST and SB play the same quantitative and heuristic roles in their retrospective theories. Quantitatively, it is readily shown using the Maxwell Boltzmann distribution and the phase space of a dilute gas respectively that the values of H and ST and SB agree (up to an arbitrary constant) at equilibrium. Heuristically, they are all properties that typically increase when we consider any thermodynamic process (accepting here that the second law of thermodynamics ought to also be understood probabilistically). As long as the heat sinks and reservoirs are gaseous then, the H theorem can apply to all the heat engine and heat pump cases that thermodynamics focuses on.

Thus, taking H as a generalisation of thermodynamical entropy to non equilibrium evolutions but restricted to dilute gas (and other similar) models, the Boltzmann equation provides a dynamical basis for the approach to equilibrium and increase in entropy.

4 Conclusion

The Boltzmann Equation should be understood as a coarse grained dynamical evolution equation comparable to a large number of other such equations in non equilibrium statistical mechanics. The reversibility and recurrence objections and the existence of fluctuation phenomena form a hierarchy first restricting the applicability and then forcing a probabilistic interpretation of the H theorem. Despite not being on entirely stable conceptual grounds, the empirical success of the Boltzmann equation and other similarly structured equations, make it eminently reasonable to accept that the H theorem establishes that the functional H is typically non-increasing for all systems of interest and all times of interest. Furthermore, as there appears no serious obstacle to accepting that -H refers to entropy, the H theorem would appear to have the credentials to provide a dynamical underpinning to the dynamic-less explanations of the increase in entropy and approach to equilibrium of thermodynamics and Boltzmannian statistical mechanics.

References

[1] Brown, H. R.; Myrvold, W. and Uffink, J. (2009), 'Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics', Studies in History and Philosophy of Modern Physics 40.

[2] Brown, H. R. (2015), Basic Concepts in Statistical Mechanics. (Teaching Note).

[3] Breuer, H-P.; Petruccione F. (2002). The Theory of Open Quantum Systems. Oxford: Oxford University Press.

[4] Cercignani, C. (1988). The Boltzmann Equation and its Applications. New York: Springer-Verlag.

[5] Frigg, R. (2007), A Field Guide to Recent Work on the Foundations of Thermodynamics and Statistical Mechanics, in Dean Rickles, ed.,'The Ashgate Companion to the New Philosophy of Physics', Ashgate, London.

Holger, A. (2013), Theoretical Terms in Science, Available online at: ¡http://plato.stanford.edu/archives/sum2013/entries/theoretical-terms-science/¿.

[6] Liboff, R.L. (2003), Kinetic theory: classical, quantum, and relativistic descriptions, Springer.

[7] Peres, A. (1995), Quantum Theory, Springer, London.

[8] Sklar, L. (1993), Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics, Cambridge University Press, Cambridge.

[9] Uffink, J. (2007), Compendium of the foundations of classical statistical physics, Available online at http://philsci-archive.pitt.edu/archive/00002691/.

[10] Wallace, D. (2011), The Logic of the Past Hypothesis, Available online at http:// philsci-archive.pitt.edu/8894/.

[11] Wallace D. (2013), What Statistical Mechanics Actually Does, Available online at http://philsci-archive.pitt.edu/9846/

[12] Wallace D. (2013), Recurrence Theorems: a Unified Account, Available online at http://arxiv.org/abs/1306.3925

[13] Wallace D. (2001), "Implications of quantum theory in the foundations of statistical mechanics", Available online at: http://philsci-archive.pitt.edu/410.