Cosmology and entropy: in search of further clarity

Gordon McCabe

July 3, 2017

Abstract

The concept of cosmic entropy, and the purported need to explain the initial conditions of Friedmann-Robertson-Walker 'Big Bang cosmology', continue to cause confusion within the scientific community. David Wallace's 2010 paper went some way towards disentangling this confusion, but left a number of significant issues unaddressed. The purpose of this paper is to define and resolve these issues.

The paper begins by making a clear distinction between the entropy density and the entropy of a comoving volume. The different behaviour of these two quantities in Big Bang cosmology is explained and identified. A second distinction is drawn between the different behaviour of radiative entropy and the entropy of matter, and a third distinction is made between actual entropy and maximum possible entropy.

The paper then devotes some attention to the particular issues associated with the entropy of matter, and its relationship to the existence of life and complexity. Wallace's account of Big Bang nucleosynthesis is endorsed, albeit in the context of a more general line of argument, demonstrating that the expansion of the universe generates information in both the radiation content as well as the matter content, and does so without the need for any 'clumping' of matter.

The role of stars and galaxies in the cosmic entropy budget is then expounded, the argument concurring and extending that provided by Wallace. In particular, attention is drawn to the eventual evaporation of gravitationally bound systems.

However, whilst Wallace accepts black-hole entropy as something of a special case, this paper takes a more sceptical approach to black hole thermodynamics, endorsing and extending recent arguments from Dougherty and Callender.

The paper concludes by analysing the role of entropy within inflationary cosmology, identifying the differences and similarities with Big Bang cosmology.

1 Introduction

To set the stage, recall some basic facts about entropy. Entropy is a property possessed by the states of physical systems. From the perspective of thermodynamics, entropy is a property of the equilibrium states of a system. These are defined by macroscopic variables such as pressure, temperature and density.

The entropy of an equilibrium state can be empirically defined, up to an additive constant, by selecting a reference state $\mathscr{O}$, and defining the entropy of a state A as:

$$S = \int_{\mathscr{O}}^{A} \frac{dQ}{T} \,,$$

where Q is the heat energy transfer and T is the temperature.

The integral can be taken over any continuous *reversible* succession of states forming a path between $\mathscr{O}$ and A. Irrespective of the process by which a system actually acquires its entropy, its value is defined to equal the heat-energy per unit temperature transferred into the system by a counterfactual quasi-static process, conducted at a constant temperature whilst the system is held in contact with a heat-bath.

In more intuitive terms, the entropy of a system at temperature T is that portion of the internal energy U of a system, per unit temperature, which is not available for performing work. The free energy F is duly defined to be $F = U - ST$.

From the perspective of classical physics, each physical system possesses a multi-dimensional space of possible states, called the phase space Γ of the system. In classical mechanics, each state is defined by the precise components of position and momentum of all the particles in the system, whilst in classical field theory, each state is defined by the combined configuration of the field in 3-dimensional space, and that of its 'conjugate' momentum field.

The phase space is partitioned into regions $\Gamma_M \subset \Gamma$ called 'macrostates', consisting of states which share macroscopically indistinguishable properties. The exact states, (the points of the phase space), are then referred to as the 'microstates'. A third type of state, dubbed a 'statistical state', is defined by a probability distribution ρ upon Γ.[1]

From the perspective of thermodynamics, entropy is a property of the equilibrium macrostates of a system, whilst from the perspective of statistical mechanics, entropy is a property of either the statistical states or the macrostates. As a property of a statistical states, entropy is defined as:

$$S = k \int_{\Gamma} - \log(\rho(x)) d\rho(x) \,.$$

where k is Boltzmann's constant.

[1]A macrostate can be seen as a special type of statistical state in which the probability distribution is of a constant value $[\mathrm{Vol}(\Gamma_M)]^{-1}$ on Γ_M, and zero elsewhere.

Depending on the whether the context is thermodynamic or information-theoretic, this concept is termed 'Gibbs entropy' or 'Shannon entropy', respectively. In the guise of Gibbs entropy, the probability distribution ρ tends to be interpreted as representing objective frequencies, whilst in the guise of Shannon entropy ρ tends to represent an expression of incomplete knowledge or uncertainty.

Using the epistemic language often employed in the context of Shannon entropy, the negative of the log of $\rho(x^*)$ corresponds to the information which would be gained if the previously unknown microstate of the system were established to be x^*. For example, if $\rho(x^*) = 1/8 = 2^{-3}$, then $-\log_2(1/8) = 3$, and it is said that 3 bits of information would be gained if the microstate were identified to be x^*. The integral is then a type of expectation value, which expresses the average amount of information gained when the microstate of the system is ultimately identified.

In contrast, defined as a property of macrostates, the entropy is:

$$S = k \log(\text{Vol}(\Gamma_M)) = k \log \left(\int_{\Gamma_M} dx \right) .$$

This is often termed the 'Boltzmann entropy'. One can then define entropy as an observable $S(x)$ on phase space by defining the entropy of a microstate as the entropy of the macrostate to which that microstate belongs.

From the perspective of statistical mechanics, the entropy of a macrostate is simply a measure of the size of a macrostate volume in phase space, and the thermodynamic entropy is simply the restriction of that concept to the equilibrium macrostates which exists at different temperatures. The entropy of a microstate then measures how typical that state is within the phase-space.

It is not, however, the purpose of this paper to analyse the conceptual relationships between the various notions of entropy, or to determine whether the concepts are ultimately ontic or epistemic. Although these issues surface when the discussion turns to black-holes, the analysis of Big Bang cosmology below will proceed as if there is no distinction between Boltzmann entropy and Gibbs/Shannon entropy. Rather, the analysis will be devoted to other conceptual distinctions, to which we now turn.

2 Big Bang entropy

As a first distinction, applicable inside and outside cosmology, there are two entropic values attributable to a system at any moment in time:

1. The actual entropy S.
2. The maximum possible entropy S^{max}.

Furthermore, in the particular case of cosmology, there are at least three distinct concepts of the entropy of the universe:

1. The entropy of the whole universe.
2. The entropy of a comoving volume S_c.
3. The entropy of a physical volume of unit size, i.e., the entropy density S_d.

Let's analyse these in turn:

2.1 The entropy of the whole universe

If the universe is spatially compact and homogeneous, then the entropy of the universe at each moment of cosmic time, both the actual entropy and the maximum possible entropy, is finite.

If the entropy of a comoving volume is constant in time, then in particular the entropy of the entire spatially compact and homogeneous universe will also be constant in time.

If, however, the universe is spatially non-compact and homogeneous, then the entropy of the universe, (both the actual entropy and the maximum entropy), will be infinite at each moment of cosmic time.

2.2 The entropy of a comoving volume

A comoving volume is typically defined as that lying within a fixed range of a dimensionless radial spatial coordinate r. Typically, one defines spherical polar coordinates (θ, ϕ, r) on a spacelike hypersurface Σ, and in the warped product space-time $I \times_R \Sigma$ of a Friedmann-Robertson-Walker model (see McCabe 2004), these coordinates are automatically inherited by the points of each spatial slice $\Sigma_t = t \times \Sigma$. Given a particular origin to such a coordinate system, one might define a comoving volume as that within the solid ball $r \in [0, 1)$. Such a set of points can be identified in each spacelike hypersurface Σ_t.

Because the scale factor $R(t)$ is increasing, the physical volume of the solid ball defined by $r \in [0, 1)$ is increasing, but each point within the volume retains the same spatial coordinates over time. Hence, the coordinates are said to be 'comoving' with the expansion, and the volume defined by $r \in [0, 1)$ is said to be a 'comoving volume'.

In Big Bang thermodynamics the entropy S_c of a finite comoving volume is finite and constant in time; the expansion is claimed to be 'isentropic'.

The derivation of this typically proceeds as follows: First, it is pointed out that the entropy of the radiation is far greater in order of magnitude than the entropy of the matter. Then, the entropy density s_d^{rad} of the radiation is given by (Coles and Lucchin 1995, p102) :

$$s_d^{rad} = \frac{\rho_r c^2 + p_r}{T_r} \, ,$$

where ρ_r is the radiative energy density, and p_r is the radiative pressure.

The equation-of-state of the radiation is such that $p_r = 1/3 \rho_r c^2$, hence:

$$s_d^{rad} = \frac{(4/3)\rho_r c^2}{T_r} \, .$$

Furthermore, the energy density of the radiation is such that $\rho_r c^2 = \sigma T_r^4$, hence:

$$s_d^{rad} = (4/3)\sigma T_r^3 \, ,$$

where σ is the 'black-body constant', related to the Stefan-Boltzmann constant by $\sigma_{SB} = \sigma c/4$, (ibid. p82).

Now, the temperature of the radiation is inversely proportional to the scale factor, $T_r(t) \propto R(t)^{-1}$, hence the time evolution of the radiative entropy density is:

$$s_d^{rad}(t) \propto R(t)^{-3} \, .$$

The radiative entropy of a comoving volume is then obtained by multiplying the entropy density by the cube of the scale-factor:

$$s_c^{rad}(t) = s_d^{rad}(t) R^3(t) \, .$$

As a result of which it is concluded that the expansion of the universe is isentropic:

$$\frac{ds_c^{rad}(t)}{dt} = 0 \, .$$

This conclusion is often conjoined or conflated with the statement that the 'entropy per baryon' s_b is constant in time, which might lull the unwary reader into believing that the entropy of the baryonic matter, and in particular the entropy of the nuclear matter, in a comoving volume is also constant in time.

However, the entropy per baryon is not a property of baryons *per se*. Having obtained the radiative entropy density s_d^{rad}, this quantity is divided by the number density of baryons n_b to obtain:

$$s_b = \frac{s_d^{rad}}{n_b} \, .$$

The 'entropy per baryon' is therefore most definitely not the baryonic entropy, but, rather, it is the ratio of radiative entropy to the number of baryons. In a comoving volume, the number of baryons remains constant (by definition, only particles with peculiar velocities or 'proper motion' with respect to the expansion, pass through the boundary of a comoving volume), so the ratio of the radiative entropy to the number of baryons remains constant with time in a comoving volume. We shall turn to the entropy of baryonic matter itself at a subsequent stage below.

For the moment, however, let us return to equating the entropy in a comoving volume with the radiative entropy in a comoving volume. As derived above, cosmologists trumpet the fact that this entropy is constant in time. Crucially, however, it is only the *actual* entropy of a comoving volume which is constant.

The maximum possible entropy of a comoving volume is *increasing* because the physical volume of a comoving region increases as the value of the scale factor increases. The maximum possible entropy is proportional to the phase-space volume, and the phase-space volume is proportional to the physical volume, not the comoving volume.

Hence, if the actual entropy of a comoving volume is constant during cosmic expansion, but the maximum possible entropy of a comoving volume increases, the difference between the two increases with time. The difference between the two is commonly defined to be the information content of a system, (Layzer 1990 p138; Layzer 2010 p12). Hence, Big Bang thermodynamics, without any assistance from inflationary cosmology, and before the onset of structure formation (i.e., before the formation of stars, galaxies and galaxy clusters), entails that the information content of the universe increases with time.

We have, then, a concept of the information content of a comoving volume:

$$I_c = S_c^{max} - S_c \,.$$

Furthermore, this comoving information content is increasing with time:

$$\frac{dI_c}{dt} > 0 \,.$$

2.3 The entropy of a physical volume

The entropy of a physical volume is the same thing as the entropy density S_d. i.e., it is the entropy per unit size of physical volume. The physical volume is proportional to the cube of the scale factor, hence the entropy density is proportional to the inverse cube of the scale factor:

$$S_d(t) = \frac{S_c(t)}{R(t)^3} \,.$$

If the actual entropy of a comoving volume is constant, then because the physical volume is increasing it follows that the actual entropy density is decreasing. We have already derived this fact above in the case of the radiative entropy density, which diminishes as the inverse-cube of the scale factor, $s_d^{rad}(t) \propto R(t)^{-3}$. Hence,

$$\frac{dS_d(t)}{dt} < 0 \,.$$

Thus, whilst Big Bang thermodynamics is claimed to be isentropic, the actual entropy of a physical volume decreases with time.

For the case of radiation, the maximum possible entropy per unit of physical volume is constant in time because the maximum possible entropy is defined by phase-space volume, which is constant for a volume of constant physical size if the number of degrees of freedom is also constant. Assuming there is no net destruction or creation of radiative particles, the number of radiative degrees of freedom will be constant.

Hence, just as we have a concept of the information content of a comoving volume, we also have a concept of the information content of a unit physical volume, i.e., the information density:

$$I_d = S_d^{max} - S_d \,.$$

Furthermore, like the comoving information content, the information density is increasing with time, this time because the actual entropy density decreases:

$$\frac{dI_d}{dt} > 0 \,.$$

Either way, Big Bang thermodynamics increases the information content of the universe. Moreover, to repeat, it does so without need for the exponential expansion of inflationary cosmology, and before the onset of structure formation.

2.4 The entropy of matter

Whether Big Bang thermodynamics is consistent with the 2nd Law of thermodynamics depends upon whether, and how, the 2nd Law can be generalised to the case where the geometry of space is expanding. If the 2nd Law generalises to a statement that the actual comoving entropy never decreases,

$$S_c(t) \geq 0 \,,$$

then this statement is consistent with Big Bang thermodynamics.

However, if the 2nd Law generalises to the statement that the actual entropy density never decreases,

$$S_d(t) \geq 0 \,,$$

then this is clearly *inconsistent* with Big Bang thermodynamics. The entropy density *does* decrease.

One possible escape clause is offered by the condition that the 2nd Law only applies to successions of equilibrium states. One might assert that an expanding universe simply cannot provide a succession of equilibrium states. Or, to put it in specific geometry terms, the Friedmann-Robertson-Walker models do not possess timelike Killing vector fields. Equivalently, it is said that the geometry is not 'stationary', hence the mass-energy within cannot possess an equilibrium state. (In black-hole thermodynamics, a stationary spacetime geometry is equated with a state of thermal equilibrium). Kolb and Turner (1990, p70) raise this possibility, only to relent on the basis that "For practical purposes...the Universe has for much of its history been very nearly in thermal equilibrium."

As Layzer (2010, p4) puts it, "particle reaction rates exceed the cosmic expansion rate at sufficiently early times. Consequently, local thermal equilibrium prevails at the instantaneous and rapidly changing values of temperature and mass density."

However, as Wallace (2010) and Layzer (ibid.) both emphasise, the early universe fell out of equilibrium when the temperature and density dropped to the point at which certain particle reaction timescales were long compared with the rate of expansion. "As the universe expanded, both its rate of expansion and the rates of particle encounters decreased, but the latter decreased faster than the former. As a result, some kinds of equilibrium ceased to prevail...matter and radiation decoupled when their joint temperature fell low enough for hydrogen to recombine. Thereafter, the matter temperature and the radiation temperature declined at different rates. So while matter-radiation interactions tended to equalize the matter and radiation temperatures, generating entropy in the process, the cosmic expansion drove the two temperatures farther apart, generating information. Earlier, the relative concentrations of nuclides were frozen in when the rates of nuclear reactions became too slow relative to the rate of cosmic expansion to maintain their equilibrium values. Again, competition between the cosmic expansion and nuclear reactions generated both entropy and information," (Layzer ibid., p14).

Wallace points out that when $T \gg 10^{11}$K, the average kinetic energies of the hydrogen nuclei (i.e., protons) were such that it was entropically favourable for matter to reside in the state of a hydrogen plasma. During this period, matter maintained a state of instantaneous local thermal equilibrium. Fusion results in the loss of degrees of freedom as free particles become bound systems, hence it was only when the temperature declined to $T \sim 10^{11}$K that the heat energy released by fusion compensated for this loss, in terms of its contribution to entropy.

At this point, the decline in temperature increased the maximum possible entropy of matter, and the actual entropy became less than the maximum possible entropy. Hence, the matter was no longer in a state of instantaneous local thermal equilibrium. Some nuclear fusion then occurred ('primordial nucleosynthesis'), and this process increased the *actual* entropy of the matter. However, with the declining temperatures and densities, the reaction-rate of such fusion reactions was slow compared to the rate of expansion, and when the universe was approximately 3 minutes old, at a temperature of 8×10^8K, the primordial nucleosynthesis ceased (Schneider p201), with only $\sim 25\%$ of the nucleons fused into helium. From that point on, the matter in the universe was locked out of thermal equilibrium until the formation of stars could generate the pressures and temperatures at which fusion could resume.

There are two principal conclusions to derive here:

1. The expansion generates information in the matter content of the universe, defined as the deficit between actual matter entropy and the maximum matter entropy.

2. The story for matter is rather different to the story for radiation: whilst the actual radiative entropy in a comoving volume remains constant, the actual entropy of the matter in a comoving volume *increases* when $T \sim 10^{11}$K. It might be insignificant to the total entropic budget of the

universe, which is dominated by the entropy of the radiation, but it is *very* significant to the subsequent evolution of life and complexity. At this point, one can see how misleading it is to conflate the entropy possessed by baryons with the constant radiative 'entropy per baryon'.

At the time that primordial nucleosynthesis ceased, the energy density, and therefore the expansion rate of the universe was dominated by 'relativistic particles' (i.e., photons and neutrinos). The matter in the universe remained in a plasma state, with Thomson scattering between photons and free electrons maintaining thermal equilibrium between the matter and radiation. Hence, although the matter was locked out of thermal equilibrium itself, it remained in equilibrium with the radiation.

After the temperature fell to $T \sim 3,000$K, when the universe was approximately $400,000$ years old, the protons and electrons combined to make a gas of hydrogen atoms (so-called 'recombination'), and Thomson scattering between photons and electrons effectively ceased. However, it is not quite true to say that the radiation effectively decoupled from the matter at this point. In fact, the equilibrium between atoms and radiation was maintained for some time by a population of residual free electrons. It was only when the universe was approximately 10 million years old that densities dropped to the point at which the equilibrium between matter and radiation broke down. Thereafter, the temperature of the radiation and gas followed different profiles, the temperature of the gas dropping more rapidly than that of the radiation during the so-called cosmic 'Dark Ages', the period before the first stars began to form after approximately 100 million years, (Loeb 2006). During this period, the cosmic gas was a net absorber of radiation.

The Zeroth Law of thermodynamics asserts that every part of a system in thermal equilibrium is at the same temperature, hence once the matter and radiation, as the different components of the energy content of the universe, acquired different temperatures, they fell out of thermal equilibrium with each other.

Thus, the expansion of the universe not only reduces the entropy density of radiation and matter; it also knocks them out of thermal equilibrium with each other, and drops the matter out of equilibrium itself. Moreover, it does so when the universe is still statistically homogeneous, before galaxies and stars have formed. It is to such 'structure formation' that we now turn.

3 Structure formation and complexity

Within cosmology and physics, something of an industry has been built around the belief that the initial conditions of the universe had to possess a state of low entropy, otherwise complexity and life would not have been possible.

Charles Lineweaver provides a typical recent expression of this belief:

> *"The entropy of the very early universe had to have some initially low value $S_{initial}$, where 'low' means low enough compared*

to the maximum possible entropy S_{max} so that the entropy gap $\Delta S (= S_{max} - S_{uni}(t))$ was large and could produce and support irreversible processes, such as stars and life forms", (2014, p419).

As we saw above, however, this statement is false for two reasons: (i) The universe is expanding, and the density of matter and radiation is falling, hence the entropy density is decreasing, and the information content of the universe is increasing. (ii) The expansion knocked the matter content of the universe out of equilibrium, and increased its maximum possible entropy.

As Layzer (1990, p144) points out, the universe began with a rapid decompression, not an bang or explosion. Whilst an explosion is a process far from equilibrium, a sudden decompression takes a system which is initially at equilibrium, and then throws it out of equilibrium.

So the existence of information, (as defined), does not require the initial actual entropy to be less than the initial maximum possible entropy, as often claimed. On the contrary, the initial actual entropy can be equal to the maximum possible entropy. To be precise, even if the difference $[s_d^{max} - s_d(t)] \rightarrow 0$ as $t \rightarrow 0$, the expansion itself is capable of creating information $I_d > 0$ at $t > 0$.

As alluded to in the last section, the information content of the universe increases before the onset of structure formation, and before the development of the complexity permitted by such structure (e.g. geological and meteorological systems, ecosystems, and cognitive systems). Hence, the growth of information, as defined, is not the same thing as the growth of complexity.

The evolution of life and complexity is dependent upon two conditions:

1. Primordial nucleosynthesis is incomplete.

2. The subsequent rate of expansion declines to the point that, for a period of time, it becomes entropically favourable for the nuclear and atomic matter to form gravitationally bound systems.

Only if both conditions are satisfied will planets form around stars which burn with the power of nuclear fusion, thereby providing the radiation flux necessary for the generation of life and complexity. If primordial nucleosynthesis continues until the point at which all the nuclear matter is converted into iron, then even if stars and planets form later in the history of such a universe, the stars will merely glow with the conversion of gravitational potential energy into heat energy. Conversely, if there is a supply of low entropy nuclear matter (i.e., hydrogen and helium), but the subsequent expansion is too rapid, then it will be entropically favourable for the matter to remain homogeneously dispersed for all time. Once again, there will be no complexity or life.

There is, however, a widespread and misconceived belief that a 'clumpy', inhomogeneous distribution of matter, such as that exhibited by stars and galaxies, possesses high entropy. This belief is often found in conjunction with the claim that there is an as-yet undiscovered definition of gravitational entropy, and that it is only with the introduction of such a notion that the clumping of gravitational systems can be rendered consistent with the 2nd Law of thermodynamics. A prime example of this belief-system is expressed by general relativist George Ellis:

> "*If you take the statements about entropy in almost every elementary textbook, and indeed most advanced ones, they are contradicted when the gravitational field is turned on and is significant. For example, in the famous case of the gas container split into two halves by a barrier, with all the gas initially on one side, the standard statement is that the gas then spreads out to uniformly fill the whole container when the barrier is removed, with the entropy correspondingly increasing. But when gravitation is turned on, the final state is with all the matter clumped in a blob somewhere in the container, rather than being uniformly spread out...The question then is whether there is a definition of entropy for the gravitational field itself (as distinct from the matter filling space-time), and if so if the second law of thermodynamics applies to the system when this gravitational entropy is taken into account...*
>
> "*This issue remains one of the most significant unsolved problems in classical gravitational theory, for as explained above, even though this is not usually made explicit, it underlies the spontaneous formation of structure in the universe - the ability of the universe to act as a 'self-organizing' system where ever more complex structures evolve by natural processes, starting off with structure formed by the action of the gravitational field," (Ellis 2002).*

On the contrary, as Wallace (2010) meticulously explains, the coalescence of stars and galaxies, and the formation of gravitationally bound systems in general, increases the total entropy of the matter and radiation in the universe, and is perfectly consistent with the 2nd Law. In summary, Wallace's exposition is as follows:

If the initial total energy of a system is positive (i.e., if the sum of the gravitational potential energy U and the kinetic energy K, is greater than zero, $E = U + K > 0$), then the entropy is maximised by the expansion of the system. In such a system the typical velocity of the constituent particles exceeds the gravitational escape velocity of the system. If, however, $E = U + K < 0$, then the system is said to be gravitationally bound, and in this case it may actually be contraction which maximises the entropy. This, then, immediately demonstrates the complacency of the assumption that gravitational contraction is an entropy-lowering process.

Nevertheless, such contraction will only take a system to at most half of its initial radius, and fails to explain the formation of stars from interstellar gas clouds. Instead, the formation of stars is dependent upon the existence of a mechanism for removing heat (and therefore energy) from the contracting system.

Suppose the initial state is one in which thermal pressure balances the gravitational attraction, but suppose that there is then some heat flow out of the system. The kinetic energy of the system K, and therefore its thermal pressure, will reduce, as will the total energy E. As a result of the reduced thermal pressure, the system will contract, reducing the gravitational potential energy U of the system, and converting it into a kinetic energy K which is greater than the initial kinetic energy. Thus, the removal of heat from such a gravitationally bound system will actually increase its temperature, (such a system is said to possess a negative heat capacity). This, in turn, will create a greater temperature gradient between the system and its surroundings, leading to more heat flow out of the system, and to further contraction. The process will continue until the pressure becomes sufficiently great at the centre of the contracting mass that nuclear fusion is triggered. As long as the heat produced by nuclear fusion is able to balance the heat flow out of the system, the thermal pressure will balance the gravitational force, and the contraction will cease.

Whilst this process successfully explains star formation, it is a mechanism which requires heat flow out of the system, and this is an entropy-decreasing effect. Given that the reduction in spatial volume also constitutes an entropy-decreasing effect, it can be safely concluded that the entropy of the *matter* in a gravitationally contracting stellar system decreases. However, contrary to the suggestions made by Ellis and others, the entropy of the gravitational field does not need to be defined and invoked to reconcile this fact with the second Law of thermodynamics.

As a gravitationally-bound system contracts, the frequency of the collisions between the constituent particles increases, and a certain fraction of those interactions will be so-called inelastic collisions, in which the atoms or molecules are raised into excited energy states. Those excited states decay via the emission of photons, and this electromagnetic radiation is then lost to the surroundings. It is this radiative emission which is the most effective means by which heat is transferred from the contracting body to its lower temperature surroundings. And crucially, the entropy of this radiation is sufficiently huge that it easily compensates, and then some, for the lower entropy of the contracting matter. The total entropy of a contracting gravitational system therefore increases, as long as one counts the contribution from the electromagnetic radiation.

Hence, there is no need to invoke an as-yet undefined notion of gravitational entropy to explain why gravitational systems become 'clumpy'. In addition, it follows that an inhomogeneous initial distribution of matter would not necessarily provide a high entropy beginning to the universe at all.[2]

Once gravitationally bound systems containing out-of-equilibrium nuclear matter have formed, nuclear fusion then provides the means by which the nuclear matter can return to equilibrium. Whilst the number of degrees of freedom is

[2] Note, however, Penrose's argument (2016, p269-275) that if black holes possess huge entropies, and if white holes can be interpreted as time-reverse equivalents with the same entropy, then initial conditions in which the universe is perforated by a population of white holes would possess much greater entropy than the homogeneous initial conditions of a Friedmann-Robertson-Walker universe. We shall address the purported black-hole entropy below.

reduced by the formation of larger nuclei, fusion reduces the potential energy of the nuclear matter, and generates large amounts of heat radiation, increasing the actual entropy of the matter.

Ultimately however, and independently of the return of the nuclear matter to equilibrium, the gravitationally bound systems will evaporate.

> "*The course of...long term galactic dynamical evolution is dictated by two generalized competing processes. First, in an isolated physical system containing any type of dissipative mechanism (for example, gravitational radiation, or extremely close inelastic encounters between individual stars), the system must evolve toward a state of lower energy while simultaneously conserving angular momentum. The net outgrowth of this process is a configuration in which most of the mass is concentrated in the center and most of the angular momentum is carried by small parcels at large radii. (The present day solar system presents a good example of this process at work.) Alternatively, a second competing trend occurs when collisionless relaxation processes are viable. In a galaxy, distant encounters between individual stars are effectively collisionless. Over time, stars tend to be evaporated from the system, the end product of this process is a tightly bound agglomeration (perhaps a massive black hole) in the center, containing only a fairly small fraction of the total mass,*" (Adams and Laughlin, 1997).

Having cooled down and released all the radiation from fusion, it is entropically favourable for the nuclear matter in galaxies to maximise their configuration space degrees of freedom, and evaporate. If the average cosmic density continues to decline, the concentration gradient between a bound system and its surroundings increases. This drives the evaporation-rate of the gravitationally bound systems until statistical homogeneity returns.

This does not entail, however, that the phase-space of the universe expands as the universe expands. Penrose warns against the "misunderstanding that there are comparatively few degrees of freedom available to the universe when it is 'small', providing some kind of low 'ceiling' to possible entropy values, and more available degrees of freedom when the universe gets larger, giving a higher 'ceiling', thereby allowing higher entropies. As the universe expands, this allowable maximum would increase, so the actual entropy of the universe could increase also...This cannot be the correct explanation for the entropy increase...This phase space is just 'there' and it does not in any sense 'grow with time'," (2004, p701, quoted in a different context by Wallace 2010).

In the late universe, it is not the number degrees of freedom which increase with time, but the range of unconstrained configuration-space values available to each particle, which increase as the density of the universe decreases.

The fact that evaporation, and the ultimate return to homogeneity is entropically favoured in the late future of the universe again undermines the belief that a clumpy distribution of matter possesses high entropy. As Wallace (2010) points out when discussing the formation and lifespan of a star which culminates with a Type Ia supernova, the process begins with a low density cloud, and ends with a low density cloud. Entropy has increased because numerous protons and neutrons have become locked into nuclei with a greater binding energy per nucleon, and a substantial amount of radiation has been released, but gravitational collapse merely generated the temperatures and pressures which provided the trigger for the process.

4 Black holes and entropy

It is not the purpose of this paper to provide a detailed analysis of black-hole thermodynamics. Nevertheless, given the significant putative role which black-hole entropy is believed to play in the overall cosmic entropy budget, it is necessary to provide at least a cursory account of the subject.

It is a widespread belief amongst modern physicists that black holes, or their horizons, possess temperature and entropy. The putative black-hole temperature is inversely proportional to the surface area of the horizon, while the entropy is proportional to the surface area. In natural units, the entropy of a non-rotating black hole is (Penrose 2016, p271):

$$S_{bh} = \frac{1}{4}A = 4\pi M^2 \,,$$

where A is the area and M is the mass.

The concept that a black hole could be the bearer of entropy is often justified by claiming that the black-hole entropy compensates for the 'loss of information', or the 'lost degrees of freedom', associated with matter and radiation falling into the black hole, never to be seen again. Bekenstein's original argument went as follows:

> "*Suppose that a body carrying entropy S goes down a black hole...The S is the uncertainty in one's knowledge of the internal configuration of the body...once the body has fallen in...the information about the internal configuration of the body becomes truly inaccessible. We thus expect the black hole entropy, as the measure of the inaccessible information, to increase by an amount S*," (Bekenstein 1973).

Presumably, the idea is that one loses both the actual entropy and the maximum possible entropy associated with these extinguished dimensions of phase-space. However, as Dougherty and Callender (2016) point out, Bekenstein-type arguments express an epistemic and operationalistic interpretation of entropy. They rightly complain that "The system itself doesn't vanish; indeed, it had better not because its mass is needed to drive area increase...there is no reason to believe that a body slipping past an event horizon would lose its entropy...no compensation is necessary...we could observe the entropy of steam engines and the like that fall behind event horizons. Just jump in with them!"

We can make the objection more precise in general relativistic terms. For example, take the Oppenheimer-Snyder spacetime for a star collapsing to a black hole, or the Schwarzschild spacetime for a black hole itself. In each case, the spacetime is globally hyperbolic, hence it can be foliated by a one-parameter family of spacelike Cauchy hypersurfaces Σ_t, and the entire spacetime is diffeomorphic to $\mathbb{R} \times \Sigma$.

Each Cauchy surface is a complete and boundaryless 3-dimensional Riemannian manifold. There is no sense in which any Cauchy surface intersects the singularity. Each Cauchy surface which contains a region inside the event horizon also contains a region outside the horizon. Moreover, every inextendible causal curve in a globally hyperbolic spacetime $\mathbb{R} \times \Sigma$ intersects each Cauchy surface Σ_t once and only once. Particles follow causal curves, hence because each particle will intersect each Cauchy surface exactly once, assuming that none of those particles form bound systems, it follows that no degrees of freedom are lost. The future may well be finite inside the event horizon, but that doesn't entail that any degrees of freedom are lost from the universe.

The entropy of one part of the universe can decrease, just as the entropy of a volume of water decreases when it transfers heat to some ice cubes immersed within it. Similarly, if a material system possessing entropy falls into a black hole, whilst the region of the universe exterior to the black hole loses entropy, the total entropy does not decrease from one spacelike Cauchy hypersurface to the next. To echo Dougherty and Callender, there is no reason for the event horizon of a black hole to possess entropy; there is simply no loss to compensate for.

Penrose, however, argues that "the enormous entropy that black holes possess is to be expected from...the remarkable fact that the structure of a stationary black hole needs only a very few parameters [mass, charge and angular momentum] to characterize its state. Since there must be a vast volume of phase space corresponding to any particular set of values of these parameters, Boltzmann's formula suggests a very large entropy," (2010, p179).

This appeal to the 'no-hair' theorem of black holes is based upon a sleight of hand: it is the space-time geometry of the stationary, asymptotically flat, *vacuum* solutions which are classified by just three parameters. Such vacuum solutions are useful idealisations for studying the behaviour of test particles in a black hole spacetime, but they do not represent the history of actual black-holes.

The spacetime of an actual black-hole contains the mass-energy which collapses to form the black hole, and any mass-energy which falls into the black-hole thereafter, including swirling accretion disks of matter and so forth. Hence, actual black holes are represented by variations of the Oppenheimer-Snyder spacetime, not the Schwarzschild space-time. As Dafermos and Rodnianski comment, "It is traditional in general relativity to 'think' Oppenheimer-Snyder but 'write' maximally-extended Schwarzschild," (2013, p18).

Whilst the exterior region of a collapse solution is isometric to an exterior region of the vacuum solution, the difference in the interior solution makes all the difference in the world. Spacetimes which represent collapse to a black-hole are not classified by just three parameters; on the contrary, they are classified by a large number of parameters, characterising the specifics of the collapsing matter, including its entropy. The entropy of such black-hole spacetimes is possessed, not by the geometry of the black-hole horizon, but by the infalling mass-energy, just as it should be, (see Figure 1).

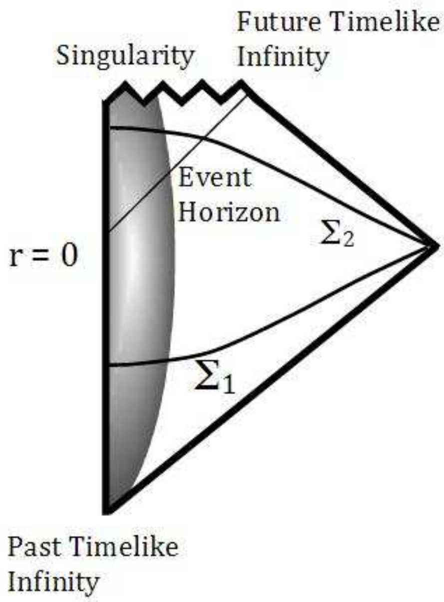

Figure 1: Conformal diagram of a black hole, including a pair of Cauchy surfaces, Σ_1 and Σ_2. The shaded region represents the infalling matter; the thin diagonal line represents the event horizon; and the jagged line represents the singularity. Cauchy surface Σ_2 possesses a region inside the event horizon, and a region outside the event horizon. The shaded region possesses entropy; the horizon doesn't. (From Maudlin 2017)

Dougherty and Callender also draw attention to a number of conceptual contradictions associated with the notion that black-hole horizons possess entropy and temperature. For example:

1. The increase in the area of a black-hole horizon, and therefore its purported entropy, is proportional to the mass-energy of the material which falls into the black-hole. Hence, if a massive object with a small entropy falls into the hole, it produces a large increase in black-hole entropy, whilst if a small object with a large entropy falls in, it produces a small increase in black-hole entropy.

2. Entropy is an 'extensive' thermodynamic property, meaning that it is proportional to the volume of a system. In contrast, black-hole entropy is

proportional to the area of the black-hole.

3. Temperature is an 'intensive' thermodynamic property, meaning it is independent to the size of an object, yet if the size of a black-hole is increased, its temperature decreases.

4. There is no 'equilibrium with' relationship in black-hole thermodynamics. Individual black-holes can be in equilibrium in the sense that the spacetime is stationary, but one black-hole cannot be in equilibrium with another.

5. If two black-holes of the same area, and therefore with the same purported temperature, coalesce, then the area of the merged black-hole is greater than each of its progenitors, hence the purported entropy increases. In contrast, thermodynamics dictates that the coalescence of two entities at the same temperature is an isentropic process.

Even if it is accepted that black holes, or their horizons, possess entropy, a belief in black hole entropy is typically twinned with a belief in the eventual evaporation of black holes. For example, Penrose (2010, p191) asserts that black holes will evaporate by Hawking radiation after the cosmic background radiation cools to a lower temperatures than the temperature of the holes. In this scenario, all the entropy in the universe eventually becomes radiation entropy. Hence, once again, it seems that the clumping of matter is nothing more than an intermediate state. If black holes can evaporate, then black holes are clearly not the ultimate means by which entropy is maximised.

An alternative scenario suggests that large black holes will not evaporate because there is a fundamental lower limit to the temperature of the cosmological radiation field, and this temperature is greater than the possible temperature of large black holes. The belief in such a lower limit is based upon the fact that a universe with a positive cosmological constant $\Lambda > 0$, such as ours currently appears to be, possesses a spacelike future conformal boundary, and the past light cone of each point on this future boundary defines an event horizon. It is then suggested that this event horizon possesses a temperature and an entropy, just as much as the event horizon of a black hole.

However, the reasons for believing that a cosmological event horizon possesses temperature and entropy are much weaker than those for believing a black hole possesses thermodynamic properties. The cosmological event horizon is entirely observer dependent, unlike the case of a black hole event horizon. Moreover, the region rendered unobservable by an event horizon is the region to the future of the event horizon, and in the case of the cosmological event horizon this is the region to the exterior of the past light cone. (In contrast, the region to the future of the event horizon of a black-hole is the interior of the black hole). Penrose (2016, p278-279) points out that the region to the exterior will be of infinite volume if the universe is spatially non-compact, hence its entropy will also be infinite. It therefore makes no sense to interpret the (finite) entropy of a cosmological event horizon as the entropy/information of all the matter and radiation 'lost' beyond that horizon.

5 Inflation and entropy

Inflationary cosmology postulates that the universe underwent a period of exponential expansion in its early history due to the existence of a scalar field ϕ, called the 'inflaton', which possessed an equation of state $p = -\rho$, (where ρ is the energy density and p is the pressure).

The inflationary scenarios postulate that there was at least some patch of the early universe in which this scalar field did not reside at the minimum of its potential energy function $V(\phi)$, and in which the energy density of the universe was dominated by this potential energy of the inflaton, $\rho = V(\phi)$.

Given the equation of state, this value of the scalar field corresponds to a state of negative pressure, in which gravity is effectively repulsive. A region of space in this so-called 'false vacuum' state undergoes exponential expansion until the scalar field eventually falls into the minimum of its potential. After a period of inflation, the false vacuum energy is converted into the energy density of more conventional matter and radiation, and the region of space which underwent inflation subsequently expands in accordance with a Friedmann-Robertson-Walker (FRW) model.

During inflation, the energy density of a comoving volume remains constant whilst the scale factor increases exponentially. As a consequence, the total energy of a comoving volume increases enormously. However, because the energy is provided by the vacuum state of a field, the heat energy itself does not increase during inflation. It is only when inflation terminates that the vacuum energy is transformed into heat energy.

According to Guth's original proposal, inflation occurs when the temperature of the universe drops below the critical temperature T_c of a phase transition, but undergoes supercooling in the process. The phase transition finally occurs at a temperature $T_s \ll T_c$, at which point the universe would be 'reheated' to a temperature $T_r \sim T_c$, and the entropy density would increase by a factor $(T_r/Ts)^3$, (Guth 1981, p350). This factor seems to be such that the entropy density after the cessation of inflation would be comparable to the entropy density prior to inflation. Hence, given the enormous increase in the scale factor, the total entropy of a comoving volume would increase enormously. This is in stark contrast to the thermodynamics of Big Bang cosmology, where the entropy of a comoving volume is constant (at least in terms of the dominating radiation contribution).

The period of exponential expansion postulated in inflation has the consequence that the presently observable universe came from a region sufficiently small that it would have been able to reach homogeneity and thermal equilibrium by means of causal processes before the onset of inflation. Inflation thereby solves the so-called 'horizon problem' of FRW cosmology. Regions of the microwave background sky on opposite sides of our observable universe have the same average temperature and density, even though, in a FRW model, they lay beyond each other's past light cones at the time the background radiation decoupled from matter. In contrast, under inflation the observable universe comes from a region which would have been able to reach a homogeneous state at this time, despite starting from a possibly heterogeneous initial condition.

This is equivalent to the claim that the actual entropy of the comoving volume which has evolved into the observable universe, was equal to its maximum possible entropy before the onset of inflation. However, this hypothetical 'thermalisation' occurs against the backdrop of a vacuum state which, as the temperature of the universe declines, is elevated above its eventual global minimum. In effect, it is a vacuum state which acquires a high potential energy, and a state of high potential energy is a state in which the actual entropy is less than the maximum possible entropy.

In this respect, the inflationary scenario resembles the thermodynamics of the conventional Big Bang cosmology: the expansion and cooling of the universe is capable of increasing the maximum possible entropy, displacing the matter and radiation content of the universe out of thermal equilibrium, and driving a subsequent increase of actual entropy.

6 Conclusions

- In Big Bang cosmology, the actual radiative entropy of a comoving volume is constant, but the actual radiative entropy density is decreasing.
- The maximum possible entropy of a comoving volume increases because of the increase in the scale factor. Hence, the difference between the maximum radiative entropy and actual radiative entropy of a comoving volume increases with time. This increases the radiative information content of the universe.
- Whilst the actual radiative entropy in a comoving volume remains constant, the actual entropy of the matter in a comoving volume *increases*.
- The expansion generates information content in the matter and radiation content of the universe. Hence, the existence of information does not require the initial actual entropy to be less than the initial maximum possible entropy.
- The expansion of the universe knocks matter and radiation out of thermal equilibrium with each other, and drops the matter out of equilibrium itself. Moreover, it does so when the universe is still homogeneous, before galaxies and stars have formed.
- The coalescence of stars and galaxies, and the formation of gravitationally bound systems, increases the total entropy of the matter and radiation in the universe.
- There is no reason to believe that either degrees of freedom or information is lost when matter or radiation falls into a black hole.
- In contrast with Big Bang thermodynamics, the inflationary scenario entails that the entropy of a comoving volume increases enormously.

- The inflationary scenario resembles the thermodynamics of standard Big Bang cosmology, in as much as the expansion and cooling of the universe is capable of increasing the maximum possible entropy, displacing the matter and radiation content of the universe out of thermal equilibrium, and driving the subsequent increase of actual entropy.

References

[1] Adams, F.C. and Laughlin, G. *A dying universe: The Long Term Fate and Evolution of Astrophysical Objects* Rev.Mod.Phys.69, pp337-372.

[2] Bekenstein, J.D. (1973). *Black holes and entropy*, Physical Review D.7, pp2333-2346.

[3] Blau, S.K. and Guth, A.H. (1987). Inflationary Cosmology. In S.Hawking and W.Israel (eds.), *300 Years of Gravitation*, (pp524-603). Cambridge: Cambridge University Press.

[4] Coles, P. and Lucchin, F. (1995). *Cosmology: the origin and evolution of cosmic structure*, Chichester: John Wiley and Sons.

[5] Dafermos, M. and Rodnianski, I. (2013). *Lectures on black holes and linear waves.* Clay Mathematics Proceedings, Vol. 17, Amer. Math. Soc. pp97-205.

[6] Dougherty, J. and Callender, C. (2016) *Black hole thermodynamics: More than an analogy?*, Forthcoming.

[7] Ellis, G.F.R. (2002). *Cosmology and local physics.* Int.J.Mod.Phys.A17:2667-2672.

[8] Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems, *Physical Review* D23, pp347-356.

[9] Kolb, E.W. and Turner, M.S. (1990). *The early universe*, Redwood City Calif: Addison-Wesley.

[10] Layzer, D. (1990). *Cosmogenesis: the growth of order in the universe*, New York-Oxford: Oxford University Press.

[11] Layzer, D. (2010). *Cosmology, initial conditions, and the measurement problem*, arXiv:1008.1229v1.

[12] Lineweaver, C.H. (2014). The entropy of the universe and the maximum entropy production principle, in *Beyond the Second Law*, R.C.Dewar *et al* (eds). Springer.

[13] Loeb, A. (2006). *The dark ages of universe*, Scientific American, November, pp47-53.

[14] Maudlin, T. (2017). *(Information) Paradox Lost.* arXiv:1705.03541v1.

[15] McCabe, G. (2004). The Structure and Interpretation of Cosmology: Part I - General Relativistic Cosmology. *Studies in History and Philosophy of Modern Physics*, Volume 36, Issue 1.

[16] Penrose, R. (2004). *The Road to Reality: A complete guide to the laws of the universe.* Jonathan Cape.

[17] Penrose, R. (2010). *Cycles of time*. The Bodley Head.

[18] Penrose, R. (2016). *Fashion, faith and fantasy in the New Physics of the Universe*. Princeton University Press.

[19] Schneider, P. (2015). *Extragalactic astronomy and cosmology: An introduction*. Second Edition. Springer.

[20] Wallace, D. (2010). *Gravity, Entropy, and Cosmology: In Search of Clarity*, British Journal for Philosophy of Science 61, pp. 513-54.