Scientific
Research
Publishing

A Different Cosmology—Thoughts from Outside the Box

J. C. Botke

Nogales, Arizona, USA
Email: jcbotke(at)gmail.com

Abstract

In this paper, we present a new cosmology based on the idea of a universe dominated by vacuum energy with time-varying curvature. In this model, the universe began with an exponential Plank era inflation before transitioning to a spacetime described by Einstein's equations. While no explicit model of the Plank era is yet known, we do establish a number of properties that the vacuum of that time must have exhibited. In particular, we show that structures came into existence during that inflation that were later responsible for all cosmic structures. A new solution of Einstein's equations incorporating time-varying curvature is presented which predicts that the scaling was initially power law with a parameter of $\gamma = 1/2$ before transitioning to an exponential acceleration of the present-day scaling. A formula relating the curvature to the vacuum energy density is also a part of the solution. A non-conventional model of nucleosynthesis provides a solution for the matter/antimatter asymmetry problem and a non-standard origin of the CMB. The CMB power spectrum is shown to be a consequence of uncertainties embedded during the initial inflation and the existence of superclusters. Using Einstein's equations, we show that so-called dark matter is, in fact, vacuum energy. A number of other issues are discussed.

Keywords

Evolution of the Universe, Inflation, Big Bang Nucleosynthesis, Cosmic Microwave Background, Dark Matter, Dark Energy

1. Introduction

The following **Figure 1** is a preview of three of the many predictions of the model described in this paper. The first shows the acceleration of the scaling that follows from our solution of Einstein's equations. The second shows the predicted
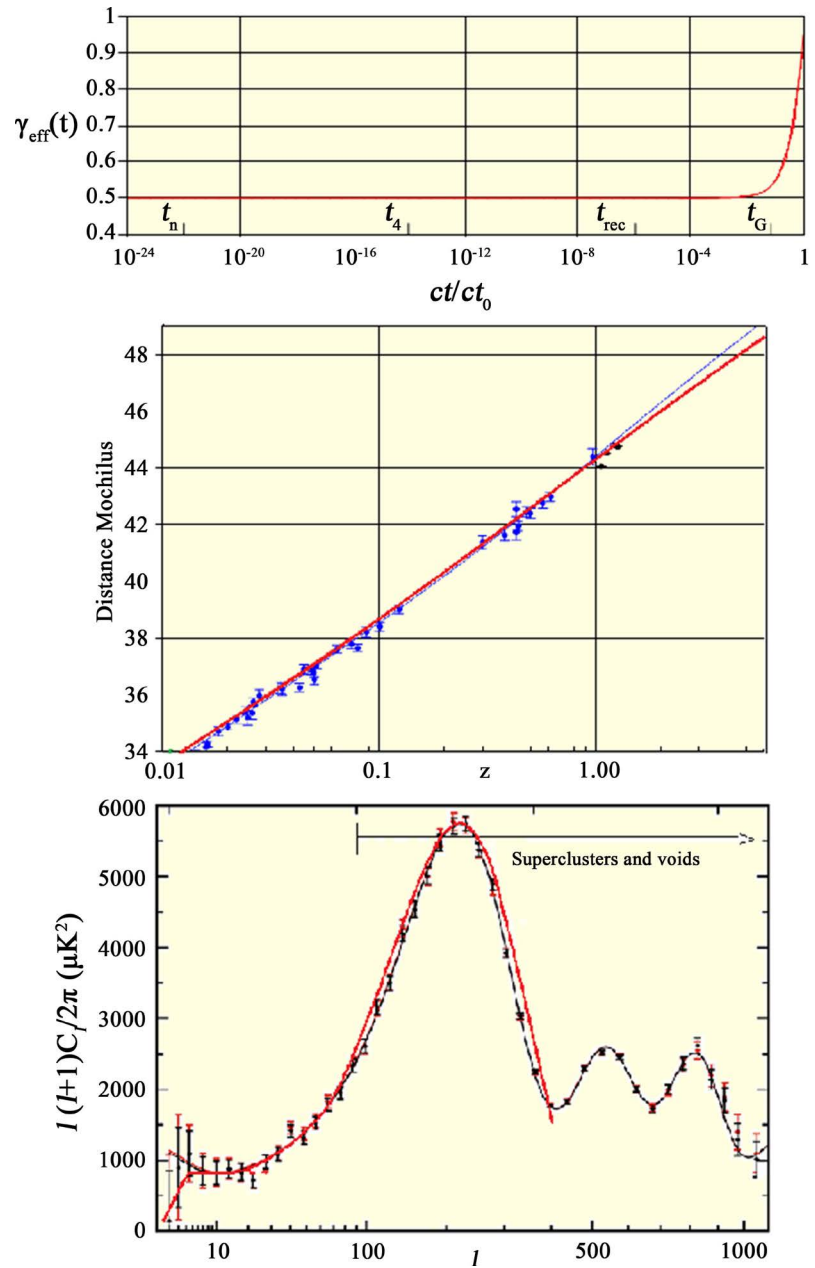
Figure 1. Preview of results.

luminosity distance curve and the third is our prediction of the CMB spectrum whose structure we show originated during an initial Plank era inflation.

Over the past several decades, a tremendous effort has been expended making detailed and careful observations of the cosmos. As a result of this, a great deal is now known about compact objects such as stars, planets, black holes, and galaxies. When it comes to the understanding the large-scale evolution of the universe, however, the situation is very different because, for one, the beginnings of the universe are out of reach of direct observation and, for another and more importantly, observations on the largest scales can only be understood within the context of a cosmological model. To date, the only model that anyone has

seriously considered is the Friedmann-Robertson-Walker (FRW) model and, in fact, this model has such unquestioned acceptance that researchers seem to have forgotten that it is just a model. We are going to assert in this paper that the FRW model and the view of cosmology that follows from it are incorrect.

In place of the FRW model, we are going to advance a new view of cosmology that challenges much of the standard viewpoint. What we will show is that a Plank era inflation of the vacuum leading into the Einstein era can account for all the major features of cosmology provided that the curvature of the vacuum varies with time. This idea results in a consistent story that makes a considerable number of predictions that are in agreement with present-day observations.

Before getting into the details of the model in the body of the paper, we will present some background materials that will help to establish the ideas. No one will argue with that the fact that a truly staggering amount of energy came into existence during the Big Bang. Two questions that immediately arise are where did this energy come from and what form did this energy take? As to the second question, energy is not a substance and so can only exist as a condition or property of something else. In the FRW case, it is assumed that the "something else" was either the field of an exotic meson or radiation or both. In this new model, we will argue that the "something else" is vacuum energy.

We can actually answer a narrow interpretation of that question immediately if, as it is generally assumed, existence began with the Big Bang. The energy did not "come from" anywhere because there was no "from". Existence began with the Big Bang and so our universe defines the totality of existence and any model of the Plank era must reflect that fact. There is no "outside" or "time" beyond our existence so we cannot talk about a period of "time", large or small, that elapsed before the Big Bang. The Big Bang simply happened. The idea that the universe defines existence also precludes the idea of multiple universes that share some sort of simultaneous existence. The keyword is simultaneous since it would be impossible to say whether separate existences occurred before, during, or after our existence and at our location or somewhere else since such distinctions are meaningless without some degree of shared existence.

Subsequent to the Plank era, we are on solid ground because Einstein's equations can be used and here, we present a new solution of the equations based on a metric with time-varying curvature that describes the evolution from the end of the Plank era onwards. The Plank era, on the other hand, presents a real problem because Einstein's equations are not applicable and we do not as yet have an alternative. Not only do we not have an analytical model to describe this era, we don't even have a convincing framework that can be used to talk about it. Nevertheless, as we will show in this paper, we can say quite a lot about the properties of the vacuum that came into existence during that era. These ideas will be developed as we proceed and a summing up will be given at the end of the paper.

As a starting point, consider for a moment the issue of measuring an interval of time. In order to do so, one must have a clock whose ticks are of shorter dura-

tion than the interval to be measured. Carrying this back in time, the ticks must get smaller and smaller until eventually we reach the Plank era. What we are proposing here is that that is as far as one can go. There is an ultimate tick and its value is the Plank time. Similarly, the ultimate length is the Plank length. The consequence is that there is fuzziness limiting the degree to which spacetime points can be specified which leads to a concept of uncertainty in which uncertainty is not a condition of some field but of the coordinates. (There is a large literature under the general heading of non-commutative geometry (NCG) which attempts to formalize this concept but these models fall outside the range of ideas that we are asserting here. We will have more to say about this at the end of the paper). The curvature of spacetime is presumably continuous but we cannot say precisely what the curvature is at "some point" because, in part, we cannot say precisely what we mean by "some point". It follows then that existence did not begin at a point but within a Plank-sized volume with a time uncertainty (from our point of view) equal to the Plank time.

In conventional field theories, one typically works in a spacetime that can be described by a differential manifold. A significant difference between such fields and this new model of spacetime is that, while the former has limitations or uncertainties that limit one's knowledge in some way, no limitation is placed on our ability to distinguish between two points arbitrarily close together either in time or distance, a point that is essential if we are to describe spacetime in terms of a differential manifold. One of the tenets of our new model, on the other hand, is that there *is* such a limitation.

The uncertainty principle requires that the initial vacuum energy was encapsulated within a Plank-sized volume with a magnitude uncertain by an amount given by $\Delta t \Delta E \geq \hbar/2$. Substituting the Plank time, we find, aside from a factor of 2, that ΔE is the Plank energy with a corresponding energy density equal to $\rho_{vac} c^2 = c^7/\hbar G^2$. The manifestation of this energy must have been the curvature of spacetime since there was no other existence. Given this fact, we can go one step further to conclude that the Plank energy density is the maximum possible energy density because a larger energy density would necessarily require a curvature more compact than given by the Plank length which we have just asserted is the smallest possible dimension. In some way then, the vacuum energy's existence was connected with the uncertainties of time and dimension. Another consequence of this uncertainty is that our normal concept of causality is not applicable and, as we will show, this had a number of important consequences.

Even though we believe that uncertainty was a crucial element during the Plank era, we don't believe that quantization had anything to do with this which is one of the points of departure from the NCG models. In fact, we don't believe that it makes any sense to talk about the quantization of gravity at all. Quantized fields describing ordinary matter all share a few general characteristics such as that they can be localized, have identifying properties such as mass and can exert forces on one another. Gravity exhibits none of these characteristics. It cannot be

localized and most importantly, in spite of the fact that it is almost impossible to discuss gravitation without using the word "force", gravity does not actually exert a force on anything. As a result, it is meaningless to wonder at the weakness of the so-called gravitational force compared with actual fundamental forces because there is no such thing as a gravitational force. The gravitational constant, G, after all is the proportionality between energy and curvature and only incidentally when one takes a Newtonian point-of-view does it become a proportionality between mass and force.

Moving on from the Plank era, we know that Einstein's equations are non-linear but, in conventional usage, one might say only in a trivial way. If one distorts spacetime at one point, the Ricci tensor components, which happen to be nonlinear, propagate that distortion to the surrounding spacetime in such a way as to maintain continuous derivatives but spacetime is passive in this process or in other words, spacetime is not acting as its own source. In the new model, we interpret the equations differently to achieve a set of equations that are non-linear in a non-trivial way. If we specify, *a priori*, some distribution of ordinary mass/ energy, it will give rise to some configuration of curvature but more generally, in our new model, instead of ordinary matter being the source of the energy or curvature, the curvature of spacetime itself becomes the source and we derive an exact expression of this idea. In addition to the equations that relate geometry to energy density and pressure, we will find that conservation of energy-momentum demands that the curvature of spacetime at any point is proportional to the sum of the vacuum energy density, pressure and any matter at that same point.

We now wish to consider the generally held belief that on the largest scales, the universe is homogeneous and isotropic. One models such a universe in terms of a sequence of hypersurfaces each of which is homogeneous and isotropic. Expressed in terms of a symmetry of spacetime, this leads directly to the requirement that the spacetime curvature of each hypersurface must be constant. To build a complete model of the universe, however, these hypersurfaces must be strung together in some way and symmetry arguments say nothing about how this is to be done. This brings us to the second and quite independent idea which is how the universe should appear to fundamental observers. When we speak of appearance, we are speaking about light that reaches us from distant objects and of necessity, that light will have passed through a sequence of hyperspaces. In the FRW case, it is *assumed* that that all hypersurfaces have the same constant curvature with the result that the universe will appear homogeneous and isotropic to fundamental observers. In this new model, on the other hand, we assert that the curvature varies with time so while the universe will appear isotropic to fundamental observers, it will *not* appear homogeneous even though each hypersurface on its own is homogeneous and isotropic.

In the first part of this paper, we will consider the inflation during the Plank era. As noted above we do not have a proper model to describe this period but nevertheless we will establish some important facts about the evolution of the

universe by examining a simple model. Having done so, we will then consider in some detail the correct form of the metric for a spacetime with time-varying curvature. The resulting equations are then solved. Two important results that follow are a prediction of a present-day exponential expansion of the universe independent of any parameter adjustments and that the time-varying curvature of spacetime is proportional to the vacuum energy density. Next, we will present a detailed non-conventional model of nucleosynthesis and the origin of the CMB which, incidentally, contains a solution of the so-called Lithium problem. Still later, we show that dark matter is, in fact, vacuum energy. Next, we discuss the origin of cosmic structures and the power spectrum of the CMB from which we discover that all such structures had a common origin in an imprint that was embedded in the vacuum during the Plank era inflation thus bringing us back to our starting point.

2. Plank Era

We will begin with some order of magnitude arguments that connect the initial curvature of spacetime with the total energy of the universe. As we proceed, we will need the values of some basic parameters and while there is some uncertainty about these, there does seem to be some consensus that the following values are reasonable (subscript 0 denotes present-day values) with the age of the universe having the smallest uncertainty.

$$
\begin{aligned}
&a_0 \approx 4.4 \times 10^{26}\,\text{m} \\
&M_U \approx 3.7 \times 10^{54}\,\text{kg} \\
&E_U = M_U c^2 \approx 3.3 \times 10^{71}\,\text{J} \\
&t_0 \approx 13.8 \times 10^9\,\text{y} = 4.36 \times 10^{17}\,\text{s}
\end{aligned}
\tag{2-1}
$$

In order to connect the curvature of spacetime with the energy density, we turn to the time component of Einstein's equations expressed in terms of a perfect fluid with the interpretation that the energy density and pressure are properties of the vacuum. The simplest expression of this idea comes from the FRW metric and reads as follows.

$$
\mathsf{R}_{00} = \left(4\pi G/c^4\right)\left(\rho c^2 + 3p\right).
\tag{2-2}
$$

Making use of the facts that the scaled Ricci tensor has the dimensions of (length)$^{-2}$ and that it embodies the geometry that defines the curvature, we can define a parameter we will call the characteristic radius of curvature or norm of the Ricci tensor

$$
O\left(\mathsf{R}_{00}\right) \sim 1/R_c^2 .
\tag{2-3}
$$

Later, once we have proposed a metric, we will show that the characteristic curvature is equivalent to the Ricci scalar. Ignoring the pressure term and equating these two gives us a connection between the radius of curvature and the vacuum energy,

$$
1/R_c^2 \sim \left(4\pi G/c^4\right)\rho c^2 .
\tag{2-4}
$$

Suppose we assume that the total energy of the universe was packed into a Plank-sized volume. The energy density would then be

$$\rho c^2 = \frac{3.3 \times 10^{71}}{(4\pi/3)(1.6 \times 10^{-35})^3} \, \text{J} \cdot \text{m}^{-3} = 1.9 \times 10^{175} \, \text{J} \cdot \text{m}^{-3} \tag{2-5}$$

which results in a characteristic radius of curvature of

$$R_c = 7.1 \times 10^{-67} \, \text{m}. \tag{2-6}$$

We asserted earlier that no dimension can be smaller than a Plank length and so we find that packing the total energy into a Plank-sized volume is impossible. We next ask what volume is necessary to contain the present-day energy of the universe without exceeding the Plank length limit. Again using (2-4) we have,

$$\frac{1}{(1.6 \times 10^{-35})^2} = \frac{(4\pi G/c^4) \times 3.3 \times 10^{71}}{(4\pi/3) R_M^3} \tag{2-7}$$

which gives $R_M = 1.3 \times 10^{-14} \, \text{m}$ and an energy density of

$$\rho_P C^2 = 3.7 \times 10^{112} \, \text{J} \cdot \text{m}^{-3} \tag{2-8}$$

which is thus the maximum allowed energy density of spacetime. (We note that ρ_P equals $(1/4\pi)\rho_{\text{Plank}}$ where $\rho_{\text{Plank}}c^2$ is the actual Plank energy density). The conclusion we draw from this is that by placing a limit on the minimum possible distance, we place an upper limit on the allowed energy density of spacetime which echoes the arguments we made in the introduction.

3. A Simple Model

We now wish to develop a model that will allow us to probe the initial expansion of the universe. The main problem we have is that, as dimensions approach Plank dimensions, our normal notion of differentiation is no longer applicable and as noted earlier, the concept of causality becomes an issue because of the uncertainties of both time and dimension. In order to build a model, we must first have a mental picture of the process we are trying to understand. One concept that comes immediately to mind is the idea that the universe began as a Plank-sized volume that underwent an exponential inflation and, in fact, this is the model we will develop in what follows. As the development proceeds in this paper, however, we will be forced to recognize that this concept is only a partial solution because, as we will prove, structures developed in the vacuum that were both very smooth and vastly too large to be explained within the constraints of normal causality. If we suppose, on the other hand, that some sort simultaneous beginning over a volume much larger than a Plank volume occurred, we are faced with the even more intractable problem of explained the existence of some influence that coordinated all these simultaneous beginnings. It seems most likely that our first idea is closer to being correct but with causality expressed in terms of an essentially unlimited speed of influence. The word influence is used here to distinguish this idea from radiation which was definitely not a part of

this process. We will build on this idea throughout this paper and finally tie things together at the end of the paper in Section 16.

We will now define a simple model embodying three general constraints that result in an exponential inflation. The significant result that follows from this model is that the total vacuum energy equaled the present-day energy of the universe at the end of the inflation and that the end of the inflation occurred when the uncertainty in time became small relative to the age of the universe. Our contention is that even though the model isn't correct, the physical picture that emerges is valid. Keep in mind that while Einstein's equations had no meaning during the Plank era, the correct theory must approach Einstein's equation asymptotically for times large compared to the Plank time so the use of an Einstein equation model is not entirely unwarranted.

The first of these constraints is that the acceleration of the scaling is dependent on the energy density and the pressure. For the purposes of this argument any metric that embodies the idea that energy density slows the expansion and a negative pressure accelerates the expansion will work. Since the time coordinate Einstein equation of the FRW metric provides the simplest expression of this idea, that is the equation we will use.

$$\ddot{a}/a = -\left(4\pi G/3c^2\right)\rho c^2\left(1-3f\right). \tag{3-1}$$

We have introduced the parameter

$$f = -p/\rho c^2 \tag{3-2}$$

where p is the familiar perfect fluid pressure term. Note that we have introduced a minus sign in the definition of f.

The next idea is that the initial expansion was non-adiabatic. First, there was nothing and later there was something so the expansion was definitely non-adiabatic. For a closed, or adiabatic, system, energy-momentum conservation requires that

$$\nabla_\mu \mathsf{T}^{\mu\nu} = 0. \tag{3-3}$$

Again, using the FRW metric for simplicity, the time component equation becomes

$$\dot{\rho} + \left(3/c^2\right)\left(\dot{a}/a\right)\rho c^2\left(1-f\right) = 0. \tag{3-4}$$

In the previous section, we established that the total energy could not have been simply dumped into a Plank-sized volume so the only alternative was that the energy was realized over a span of time sufficiently long to allow the energy of the universe to reach its present value without the energy density limit being exceeded. The simplest modification that incorporates the idea of a non-adiabatic expansion is to simply add a source to the right-hand side of (3-4)

$$\dot{\rho} + \left(3/c^2\right)\left(\dot{a}/a\right)\rho c^2\left(1-f\right) = \dot{\rho}_s. \tag{3-5}$$

Keep in mind that we are not proposing this as an actual theory but only as a means of modeling a non-adiabatic expansion. We will repeatedly use the term

"source" but this is not a source in the conventional sense but is instead a consequence of trying to confine spacetime within a Plank-sized volume. A genie in a bottle is perhaps a better mental picture.

Another approach to this same problem begins with Einstein's equations which have the general form

$$\mathsf{R}_{\mu\nu} - (\mathbf{R}/2)\mathsf{g}_{\mu\nu} = \kappa\mathsf{T}_{\mu\nu}. \tag{3-6}$$

By construction, the covariant derivative of the left-hand side vanishes but in order to model the introduction of existence, the covariant derivative of the right-hand side must not vanish. To fix things up, we could add a "source" term to the left-hand side,

$$\mathsf{R}_{\mu\nu} - (\mathbf{R}/2)\mathsf{g}_{\mu\nu} + S(t)\mathsf{g}_{\mu\nu} = \kappa\mathsf{T}_{\mu\nu} \tag{3-7}$$

where the source represents vacuum energy that varies with time but not location. In fact, it could not vary with location because during the inflation, there was not yet a well-defined concept of location. Calculating the covariant derivative of both sides results in

$$\nabla_{\mu}\mathsf{T}^{\mu\nu} = \partial_{\mu}S(t)\mathsf{g}^{\mu\nu}/\kappa. \tag{3-8}$$

Such a term would imply that the vacuum is, in fact, its own source which is consistent with the notion that the energy arose within the vacuum as a consequence of uncertainty. Carrying this further, (3-1) and (3-4) become

$$\ddot{a}/a = -(4\pi G/3c^2)\rho c^2(1-3f) + S(t)/3 \tag{3-9a}$$

$$\dot{\rho} + (3/c^2)(\dot{a}/a)\rho c^2(1-f) = -\dot{S}(t). \tag{3-9b}$$

Since we expect the source to lead to an increase in the energy density, from (3-9b) it is apparent that its time derivative must be negative. This in turn requires that initially the source or curvature must have been maximal which is in accord with the arguments given earlier. This model also provides a built-in cutoff of the source corresponding to the time at which $S(t) = 0$.

Finally, we come to the third constraint which is simply that the energy density of the vacuum cannot exceed the Plank energy density, *i.e.* $\rho c^2 \le \rho_P c^2$.

It is important to appreciate that even though we have borrowed two of the FRW equations for our model, the interpretation of these equations is very different from the FRW interpretation. The new model asserts that the universe began as spacetime vacuum with a high degree of curvature that we interpret as energy and that this energy is not related in any way to ordinary matter.

We now wish to solve this set of equations given some initial conditions. Getting to practical matters, there is some arbitrariness in how one chooses to define the coordinates. In this paper, we will choose the radial coordinate to have no units so it ranges in value from 0 to 1. Next because of the huge range of values of the scaling, it is useful to express time and the scaling in terms of Plank dimensions. We define a variable $\tau(\ge 0)$ by

$$t = t_P \mathrm{e}^{\tau} \tag{3-10}$$

where $t_P = \sqrt{3}t_{\text{Plank}}$. This definition of t_P follows from the constants of the field equation as will be shown below. The scaling is defined similarly in terms of a function $\alpha(t)(\geq 0)$

$$a(t) = a_P e^{\alpha(t)} \tag{3-11}$$

where a_P is the actual Plank length ($= 1.6 \times 10^{-35}\,\text{m}$) We also define the function ζ to be the ratio of the energy density to its limiting value.

$$\zeta = \rho(\tau)c^2 / \rho_P c^2. \tag{3-12}$$

Substituting and using the chain rule results in two new equations

$$d^2\alpha/d\tau^2 + (d\alpha/d\tau)^2 - d\alpha/d\tau = -\zeta e^{2\tau}(1-3f) \tag{3-13a}$$

$$d\zeta/d\tau + 3(d\alpha/d\tau)(1-f) = d\zeta_s/d\tau. \tag{3-13b}$$

During the process of making the change of variables, the physical constants combine to form the definition of t_P,

$$t_P \equiv \sqrt{3c^2/(4\pi G \rho_P c^2)} = 9.3 \times 10^{-44}\,\text{s} \tag{3-14}$$

At this point, we need to choose a model for the source and it probably doesn't matter too much what choice we make. In the event, we simply set $d\rho_S/dt$ to be a constant. Using the chain rule and scaling by $\rho_P c^2$ we find

$$d\zeta_s/dt = (d\zeta_s/d\tau)(e^{-\tau}/t_P) \tag{3-15}$$

so

$$d\zeta_s/d\tau = \sigma e^{\tau} \tag{3-16}$$

where σ has a non-zero, constant value for just the short period during which the universe was escaping from the uncertainty constraints. We noted earlier that the $S(t)$ model quite naturally imposes such a cutoff.

We will reduce the equations further by introducing the variable $\beta = d\alpha/d\tau$ so finally we have

$$d\alpha/d\tau = \beta \tag{3-17a}$$

$$d\beta/d\tau = \beta(1-\beta) - \zeta e^{2\tau}(1-3f) \tag{3-17b}$$

$$d\zeta/d\tau = -3\zeta\beta(1-f) + \sigma e^{\tau}. \tag{3-17c}$$

The next step is to fix the source cutoff which we will accomplish by making the connection between the total energy of the universe and the scaling parameter. The total energy at any time is

$$E_U = \int dV \rho c^2 \tag{3-18}$$

and since the density is assumed to be a function of time only, this becomes

$$E_U = \rho c^2 \int dV = \zeta \rho_P C^2 (4\pi/3) a^3 \tag{3-19}$$

Setting E_U to its current value and taking the logarithm of both sides, we obtain

$$\ln \zeta + 3\alpha = 144 \qquad (3\text{-}20)$$

and since ζ will be unity at the time the limit is reached, we find an estimate of the cutoff,

$$\alpha_I \approx 48. \qquad (3\text{-}21)$$

This value is only a first estimate, however, because in addition to adding energy directly to the curvature, the source also generates pressure. After the source is cut off, the pressure does not immediately vanish but instead decays over a period of time during which it acts as a source adding more energy to the curvature. This energy must be included in the final total and consequently the actual source cutoff must occur at a somewhat smaller value of α than indicated by (3-21).

This is an initial value problem in which we begin with a Plank-sized volume and the source so we have $\alpha(0) = 0$ and $\sigma(0) > 0$. The remaining parameters that must be specified are $\beta(0)$, $\zeta(0)$, and $f(0)$ which can be done in several ways. We considered several possibilities but found that the end result is much the same no matter what assumptions are made. In all cases, the evolution divides into 3 phases. The first, which we will call the inflationary phase, is the period during which the source is non-zero and the energy density is at its maximal value. Said another way, it is during this period that the covariant derivative of the energy tensor is non-zero. The second phase, which we will call the transition phase, is the period during which *f* decays to zero and the third is the Einstein era. The inflationary phase can be further subdivided into an initialization period which lasted for 3 Plank times or less followed by the actual inflationary period.

We are now in a position to examine some results. The numerical integrations were performed using the standard 4$^{\text{th}}$ order Runge-Kutta method.

4. Inflation

What we found was that the model results are for the most part insensitive to the details of the initial conditions or the transition pressure decay. Either the scaling curve had the shape shown in **Figure 2** or there is no solution at all.

The detailed evolution during the initialization period varies considerably depending on the initial conditions chosen but it never lasts more than 2 or 3 Plank times and once the inflation begins, the differences cease to matter because the inflation is essentially the same in all cases. The end of the inflation generally occurs
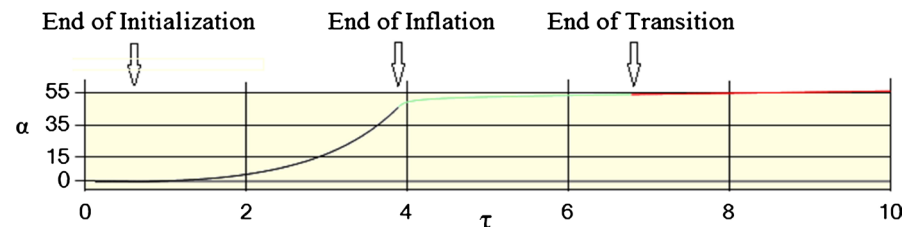
Figure 2. Typical initial evolution of the universe.

at a value of $\tau_I \approx 3.8 \text{-} 3.9$ or about 46 - 49 Plank times. Following the inflation is the transition era which eventually ends at a value of τ somewhere between 6 and 8 or a time between 400 - 3000 Plank times.

No matter how we start things off, at some point, the energy density reaches its Plank limit at which point $d\zeta/d\tau = 0$. From that point on, the pressure ratio is fixed by the (3-17c) so we have

$$f = 1 - \frac{\sigma e^\tau}{3\beta}. \tag{4-1}$$

If we substitute this into (3-17b) with $\zeta = 1$, we find that

$$\beta(\tau) = be^\tau \tag{4-2a}$$

$$b^2 + \sigma/b = 2 \tag{4-2b}$$

$$f = 1 - \sigma/(3b) \tag{4-2c}$$

where b is a constant. The significant point is that β, and hence α, is an exponential in τ which means that there is an exponential inflation of the scaling,

$$a(t) \propto e^{t/t_p}. \tag{4-3}$$

Returning to (3-9b), if we set $d\zeta/d\tau = 0$ as before, assume $-\dot{S}(t)$ is a positive constant, ignoring f results in a differential equation for $a(t)$ which has the solution given by (4-3). Thus, in either model, we get an initial exponential inflation of the universe.

What we can now see is that the Plank limits fit very nicely with the model's prediction of the evolution. It was only during the Plank era that the non-adiabatic condition existed which suggests that the nature of the source is connected with the fuzziness limiting the degree to which spacetime points can be specified. As we mentioned earlier, an uncertainty of time equal to the Plank time implies an uncertainty in the energy density equal to the Plank energy density. The end of the inflation occurred at about 46 Plank times which means that the source cutoff, based on the energy argument, happened at about the time that the overall time scale was beginning to be large compared to the Plank time and the corresponding energy uncertainty would have become small.

We started by specifying the total energy but now turning this around, we now see that that model is really saying that the *present-day energy and size of the universe was fixed by the condition that the uncertainty of the vacuum energy had become negligible* and our contention is that this result is a statement of fact rather than a result limited to this particular model. This is the principal result of this portion of the paper. Existence began as a vacuum with Plank uncertainties of time and distance that then became realized when time and distance became large compared to Plank dimensions.

With the ending of the inflation, the evolution entered the transition period which was not only a transition from an exponential expansion to perhaps a power law expansion but also a transition from the Plank era to the Einstein era.

The model predicts that the end of the transition occurred at precisely the point in time at which Einstein's equations would be expected to have become valid. The "creation" of vacuum energy had long since ended and the granularity of the coordinates was by then a very small fraction of the current age so one would then expect a differential manifold description to be a reasonable approximation.

Initially, during the transition phase, the energy density and pressure dropped very rapidly and the constraint given by (4-1) was no longer valid. Eventually, f vanishes but we cannot just set $f = 0$ at the end of the inflation because the solution fails if we do so. It is necessary then to postulate a decay model for f and while the equations do not give us an explicit expression for the decay, they do impose a constraint on the decay rate. Various decay models were tried (linear, exponential, Gaussian) and it turned out that the results are not sensitive to the particular choice made. The exponential model seemed to be the most reasonable since the other quantities are exponential so that is the model used to obtain the results shown below. The formula we used was

$$f(\tau) = f_I e^{-\left(\frac{\tau - \tau_I}{\tau_2 - \tau_I}\right)}$$

(4-4)

where the subscript "I" denotes values at the time of the source cutoff and τ_2 is an adjustable parameter.

After the source cutoff, (3-5) can be rewritten as

$$\dot{\rho} + \left(3/c^2\right)\left(\dot{a}/a\right)\rho c^2 = \left(3/c^2\right)\left(\dot{a}/a\right)\rho c^2 f$$

(4-5)

which shows that the pressure does act as a source during the transition period.

After running a number of simulations, we determined that the minimum value of the decay parameter τ_2 is about 4 and that the effect of increasing τ_2 is to postpone the end of the transition to some degree. Each time we adjusted the decay parameter, we also had to adjust the source cut off so that the final energy matched the present-day energy of the universe.

Eventually, the pressure ratio vanishes, the total energy reaches its final, constant value and we enter the Einstein era. Since the total energy is proportional to $a^3 \rho c^2$, setting $dE_U/dt = 0$ gives

$$\dot{\rho} + \left(3/c^2\right)\left(\dot{a}/a\right)\rho c^2 = 0$$

(4-6)

which agrees with the previous equation with $f = 0$. We can now easily obtain an exact solution. We define a scaling parameter γ by

$$a(t) = a_T \left(t/t_T\right)^\gamma.$$

(4-7)

where t_T is the time at the end of the transition period. Substituting into (4-6) gives the energy density

$$\rho c^2 = \rho_T c^2 \left(t_T/t\right)^{3\gamma}$$

(4-8)

and substituting both into (3-1) gives

$$\frac{\gamma(\gamma - 1)}{t^2} = -\left(4\pi G/3c^2\right)\rho_T c^2 \left(t_T/t\right)^{3\gamma}$$

(4-9)

which has the solution

$$\gamma = 2/3 \tag{4-10a}$$

$$\rho_T c^2 = 8.2 \times 10^{111} \left(t_P / t_T \right)^2 \text{ J} \cdot \text{m}^{-3}. \tag{4-10b}$$

Substitution back into (4-8), we find for $t > t_T$

$$\rho(t) c^2 = 8.2 \times 10^{111} \left(t_P / t \right)^2 \text{ J} \cdot \text{m}^{-3}. \tag{4-11}$$

The point at which $\beta = 2/3$ is the point at which the total energy ceased to change.

In order to test the sensitivity of the model, we tried a number of different scenarios to determine if the initial conditions were important. What we found was that the solution is insensitive to the initial conditions. The arguments based on uncertainty suggest that it makes the most sense to assume that the universe began with the maximum possible curvature and a positive expansion rate for which the starting conditions are $\alpha(0) = 0$, $\zeta(0) = 1$, $\sigma = 1$, and $\beta(0) > 0$. Any value greater that about $\beta(0) = 0.7$ will work. Note that this value is actually quite small given that the value of β at the end of the inflation is about 46.

We also examined the sensitivity to the source strength. We won't show the results but it turns out the similar results are obtained for any σ in the range $0.01 < \sigma < 1.5$. For larger values of σ the solution becomes erratic and by $\sigma = 2.0$, the pressure is no longer large enough to prevent the collapse induced by the energy density. For smaller values of σ, a solution still exists but with a decreasing source cutoff time and an increasingly long transition period.

After running many simulations, we found that the following set of parameters form a reasonable picture of the model up to this point. Starting with $\beta(0) = 1.0$, $\sigma = 1$, $\alpha_c = 45$, and $\tau_2 = 4.4$ we find

$$\tau_I = 3.83 \Rightarrow t_I = 4.3 \times 10^{-42} \text{s} = 46 t_P$$

$$\alpha_I = 45 \Rightarrow a_I = 6.2 \times 10^{-16} \text{m}$$

$$\tau_T = 7.36 \Rightarrow t_T = 1.47 \times 10^{-40} \text{s} = 1570 t_P \tag{4-12}$$

$$\alpha_T = 53.4 \Rightarrow a_T = 2.49 \times 10^{-12} \text{m}$$

$$\zeta_T = 9.1 \times 10^{-8} \Rightarrow \rho_T c^2 = 3.34 \times 10^{105} \text{J} \cdot \text{m}^{-3}$$

$$E_U = 3.2 \times 10^{71} \text{J}$$

$$a_0 = a_T \left(t_0 / t_T \right)^{2/3} = 5.3 \times 10^{26} \text{m}$$

The total energy is in reasonable agreement with the value of (2-1) and we also see that value of a_0 is only a little larger than the value given in (2-1).

For times greater that t_T, we have

$$a(t) = 2.49 \times 10^{-12} \left(t / t_T \right)^{2/3} \text{m} \tag{4-13a}$$

$$\rho(t) c^2 = 3.34 \times 10^{105} \left(t_T / t \right)^2 \text{ J} \cdot \text{m}^{-3} = 8.2 \times 10^{111} \left(t_P / t \right)^2 \text{ J} \cdot \text{m}^{-3}. \tag{4-13b}$$

In conclusion, what we found was that we either get a Big Bang or we don't and when we do, the scaling and energy density are much the same no matter what values of the adjustable parameters are used.

Looking back at (3-1) and (3-5), we see that they are linear in the energy density. This is a consequence of the fact that the equations do not include self-interactions of the field. With self-interactions, on the other hand, the equations will not be linear in the energy density and further, these self-interactions will tend to slow the expansion with the result that the end of transition period will occur at a somewhat later time than given by $\gamma = 2/3$ criterion which, after all, is simply the asymptotic limit of the particular simple model we used. In fact, as we will show below, observations and the requirements imposed by the existence of the CMB require that the scaling parameter during the post-Plank era had a value of $\gamma \approx 0.5$, a value which characterized the expansion up until about the time of galaxy formation. Subsequently, the exact solution of the scaling which we have found shows that the scaling began an exponential acceleration.

Evidence in support of this view follows from observations of the Hubble parameter. By definition,

$$H(t) \equiv \dot{a}(t)/a(t) \qquad (4\text{-}14)$$

which for power-law scaling has the value,

$$H(t)^{-1} = t/\gamma. \qquad (4\text{-}15)$$

The actual value of the Hubble constant is still a matter of debate. For the purposes of this paper, we will use a value of $H_0^{-1} \approx 4.6 \times 10^{17}\,\text{s}$ ($H_0 \approx 67.3$) which corresponds to an effective power-law scaling of

$$\gamma_0 = 0.95 \qquad (4\text{-}16)$$

with the understanding that an adjustment will be required later. It is important to note that these results follow directly from our model of the vacuum and have nothing to do with ordinary matter which didn't appear on the scene for another 10^{38} ticks of our Plank clock. The expansion of the universe was and is controlled by the vacuum energy from start to finish.

5. Curvature

One of the principal tenets of the new model is that the curvature must vary with time and consequently that the FRW field equations do not correctly describe the evolution of the universe. By dimensional arguments, the FRW curvature K must be related to the curvature parameter k by $K \propto k/a^2$ and since $K \sim 1/R_c^2$, we have

$$1/R_c^2 \sim k/a^2. \qquad (5\text{-}1)$$

During the inflation, the energy density and hence the radius of curvature would have been constant so from (5-1), we see that initially k would have increased exponentially with time. During the transition period, the time dependence of k would have been more complicated but eventually it entered a slow

decay dictated by the scaling. Substituting (2-4) into (5-1) gives

$$k = \left(a/c\right)^2 \left(4\pi G/c^2\right)\rho c^2 . \tag{5-2}$$

With the assumption that k is proportional to a linear combination of the energy density and pressure, and using the fact that the coupling between energy and curvature must include a factor of G, we find that the only combination of variables that has the correct units is $a^2 G\left(\rho c^2\right)/c^4$ which leads us to expect a relation of the form

$$k = \left(a/c\right)^2 \left(4\pi G/c^2\right)\left(k_1 \rho c^2 + k_2 p\right) . \tag{5-3}$$

Substituting (4-13) into (5-2) gives

$$k(t) = \left(3/2\right)\left(a_T/c\right)^2 \left(8\pi G/3c^2\right)\rho_T c^2 \left(t_T/t\right)^{2/3} = 2.2\times10^{39}\left(t_T/t\right)^{2/3} \tag{5-4}$$

and, with this result, the radius of curvature is

$$R_c(t) = a(t)/\sqrt{k(t)} = 5.3\times10^{-32}\left(t/t_T\right)\ \mathrm{m} \tag{5-5}$$

which we see varies linearly with time. Later we will find that the linear dependence is, in fact, an exact result independent of the scaling. At the transition time, R_C was three orders of magnitude larger than the Plank distance. This value is not unreasonable from the point of view of just having exited the Plank era but it is widely different from the curvature of the essentially flat universe of the present day. Equation (5-4) suggests that the curvature varies as $t^{-2/3}$ but, as we will show in Section 8, this is a special case of the exact result that $k(t) \propto \left(a(t)/t\right)^2$.

6. Radiation

In some versions of the FRW model of the Big Bang, it is posited that energy was dumped into the nascent universe in the form of radiation. We have already demonstrated that the evolution of the universe can be understood without reference to radiation but we now will go even further and argue that radiation during the initial evolution of the universe was not even possible.

The initial inflation ended at a time around $t_I = 4.3\times10^{-42}\ \mathrm{s}$ (46 Plank times) and thus the maximum distance any radiation could possibly have traveled in that time was on the order of 1.3×10^{-33} m. But that isn't the whole story because any radiation would have been restricted to the geodesics of the metric (assuming that such a concept had any meaning during the Plank era). Since the radius of curvature at that time was given by the Plank length, any extant radiation would be turned back onto itself in a volume also given by the Plank length. This being the case, instead of going somewhere, the radiation would be confined to the minimal possible physical dimension. At the same time that the radiation wasn't going anywhere, the scaling was increasing by a factor of 10^{19} to a value on the order of 10^{-16} m so even if there had been some form of radiation present, it would have been impossible for a signal to propagate from any one point in the universe to any other point. We will give a more formal statement of this re-

sult in Sec. 8 where we show that what we will call the horizon distance is, in fact, equal to the radius of curvature and during the inflation, the radius of curvature was fixed at one Plank length.

We also know that any extant radiation must have been deposited by the initial energy source and so by the end of the inflation, all the radiation that would have existed in the early universe must have, by then, been there. But since radiation could not have existed during the inflation, it cannot have been around shortly afterwards either.

In the standard model, the radiation is supposed to have somehow come into existence at the termination of a period of inflation although the model does not actually explain how that happened. The time at which this was supposed to happen would have been long after the end of the Plank era on a logarithmic scale. In this new model, on the other hand, it is asserted that the radiation came into existence at a still much later time during nucleosynthesis and there was never a period during which the expansion was dominated by radiation.

7. Homogeneity and Isotropy

In the standard model, the curvature is assumed to be fixed and a consequence of that is that at an early enough time, it would have been possible for any point in the universe to be within the horizon of any other. That being the case, it is assumed that any initial anisotropies would have been smoothed out. That is, in itself, a big "if" since smoothing requires sufficient time for mixing to occur and the time scale involved is limited to only about 10^{-35} s. The next problem was how to propagate that uniformity to the present-day without large in homogeneities developing via the interactions of different regions of space time. The solution was to imagine an inflation in which the spatial dimensions outran the signaling distance thus preserving the initial uniformity.

Another assumption made is that the conventional inflation was adiabatic. A consequence of this is that the entropy at the end of the inflation would have been the same as at the beginning when it is assumed, it is was of $O(1)$. The present-day entropy, on the other hand, is thought to be on the order of

$$S = O\left(10^{90}\right) \tag{7-1}$$

and in the conventional model, this huge increase is assumed to have happened during a period immediately after the end of inflation when the energy of the inflation mesons was converted into the radiation plasma. The conventional model assumes that the inflation was driven by the action of an exotic meson but makes no attempt to explain the origin of the exotic meson field, which itself would have been a non-adiabatic event.

In the new model, the situation is quite different. First, the horizon distance was fixed at the Plank length during the inflation so there was *never* a time during the Plank era when all points or even a few points in the universe could communicate in a conventional manner. As the inflation progressed, more and more Plank-sized regions came into existence but each was isolated from all

others. (These last two statements will require amendment later). Since each such region had a fixed energy equal to the Plank energy, the inflation was *non-adiabatic* to the tune of all the energy of the present-day universe. At the end of the inflation, the total number of Plank regions would have been

$$\left(a\left(t_I\right)/a_P\right)^3 = e^{3\alpha_I}. \tag{7-2}$$

At that time, $\alpha_I \geq 45$ so the number of independent regions and hence the entropy was

$$S\left(t_f\right) = O\left(10^{60}\right). \tag{7-3}$$

The entropy is quite sensitive to the value of τ at the end of the inflation, however, so this value will require some adjustment. Since the actual initial expansion of the universe would have been slower than $\gamma = 2/3$, the inflation would of necessity extended for a slightly longer period of time in order to match the present size of the universe. We will return to this point in Sec. 8.

At the end of the inflation with the source cut off, the radius of curvature began to increase faster than the scaling. Remember that the defining condition of the inflation was that the energy density would be at its maximum possible value. The consequence of this would have been a universe that was homogeneous to a high degree because any departure from homogeneity would imply that the density at one point was different from the density somewhere else. Since the energy density is directly related to the curvature and the scaling, these too would have been homogenous. Subsequent to the inflation, the universe would have remained homogeneous on large scales because there was no mechanism by which the homogeneity could have been disrupted. Each small region of the universe evolved without communication with any other region and since the physics was the same everywhere, the regions evolved in lockstep. The universe remained homogeneous precisely because of the *lack* of communication.

Well, *almost* homogeneous.

Almost, because we must allow for differences in the energy that would have resulted from the uncertainties at the time the inflation ended. Initially, the energy of each Plank-sized region was nominally the Plank energy with an uncertainty equal to that same value. By the end of the inflation, however, the uncertainty would have been reduced by the ratio of the finish time to the Plank time so we have

$$\delta\rho/\rho = e^{-\tau_I} \tag{7-4}$$

with $\tau_I = 4$ or a little more. At that point, these variations in the energies of the Plank-sized regions would have become locked in because, with the source cut off, the expansion became adiabatic and the energy content of each region would have become fixed. The result was that there were small fluctuations in the scaling, curvature and so on at all scales larger than the size of these regions.

In order to determine the spatial characteristics, we examine the expectation value of the density, or any other parameter, at two different points,

$$\left\langle \left(\delta\rho(r_1)/\rho(r_1) - \delta\rho(r_2)/\rho(r_2) \right)^2 \right\rangle$$
$$= \left\langle \left(\delta\rho(r_1)/\rho(r_1) \right)^2 \right\rangle - 2\left\langle \delta\rho(r_1)/\rho(r_1)\,\delta\rho(r_2)/\rho(r_2) \right\rangle + \left\langle \left(\delta\rho(r_2)/\rho(r_2) \right)^2 \right\rangle. \tag{7-5}$$

Because the fluctuations were random, the middle term vanishes for separations greater than a Plank length so the result is

$$\left\langle ... \right\rangle = 0 \ \text{ for } \left| r_1 - r_2 \right| < a_{\text{Plank}} \tag{7-6}$$

$$\left\langle ... \right\rangle = 2\left\langle \left(\delta\rho(r)/\rho(r) \right)^2 \right\rangle = \text{const. for } \left| r_1 - r_2 \right| > a_{\text{Plank}}.$$

Thus, we see that the distribution is scale-invariant for points further apart than one Plank length. What that means is that the fluctuations of the universe as a whole do not get smoothed out as the communication regions expand even though the internal fluctuations within each such region will, to some extent, get smoothed. In fact, we will see in a later section of this paper that it is this variance that responsible for the large angle CMB power spectrum.

You will note that we have not said anything about the smooth structures that we have been insisting must also have existed. In fact, as we will see, these were of very small amplitude and so their existence does not alter the general picture presented here.

The next significant step in the evolution was the creation of ordinary matter but before we get to that, we will present the full metric along with its solution and examine more carefully how a homogeneous and isotropic universe can be reconciled with time-varying curvature.

8. Time-Varying Curvature

This section is devoted to the problem of understanding the evolution of a universe in which the curvature varies with time. One of the significant consequences of time-varying curvature is that we must distinguish between the universe as it actually is and as it is perceived by an observer.

When we speak of the universe as it actually is, we are speaking about such characteristics as the curvature and scaling of a sequence of spacelike hypersurfaces which are described by the 3-space portion of the metric and which exist outside the context of Einstein's equations. When we speak of the perceptions of an observer, on the other hand, we are speaking about the capture of signals that originated at some point in spacetime and then passed through a sequence of such hypersurfaces to reach an observer at some later point in time. It is these signals that are described by Einstein's equations. In an FRW universe where the curvature is time-invariant, one can, for the most part, ignore this distinction. With time-varying curvature, however, this distinction is important and has numerous consequences that must be considered.

We will begin with a review of the formalism defining homogeneous and isotropic hypersurfaces. This will be mostly familiar ground but since everything that follows is dependent on these ideas it will be useful to make sure we have a

common starting point. Referring to e.g. [1], chapter 5, we find that on any such hypersurface, symmetry arguments require that the spatial portion of the Riemann tensor must have the following form

$$^{(3)}\mathsf{R}_{abcd} = K h_{c[a} h_{b]d} \qquad (8\text{-}1)$$

where the curvature K is a constant (on *that* hypersurface). Given this fact, it then follows that the spatial portion of the correct metric must have the following form (see e.g. [2], chapter 14.)

$$\mathrm{d}\sigma^2 = \mathrm{d}r^2 / \left(1 - kr^2\right) + r^2 \left(\mathrm{d}\vartheta^2 + \sin^2 \vartheta \mathrm{d}\phi^2\right) = \mathrm{d}r^2 / \left(1 - kr^2\right) + r^2 \mathrm{d}\Omega^2. \quad (8\text{-}2)$$

We emphasize this is a statement about *each* spacelike hypersurface and that it follows from symmetry arguments alone and has nothing to do with Einstein's equations. It is also important that even though this expression involves the radial coordinate, r, there is *no* notion of a preferred origin. All points in the hyperspace are equivalent.

So far, so good. The problem comes when we set about stitching these hypersurfaces together to form the complete spacetime. The symmetry argument tells us nothing about how to go about doing this so an additional assumption must be made. In the FRW case, the additional assumption is implicitly made that the universe must not only be composed of homogeneous and isotropic hypersurfaces, it must also *appear* homogeneous and isotropic to fundamental observers. In order for this to be true, it turns out that all hypersurfaces must have the same constant curvature. It is one of the main contentions of the work, however, that the curvature does vary with time and from that, it follows that the curvature is *not* the same for all hypersurfaces so the 3-space line element must instead have the form

$$\mathrm{d}\sigma^2 = \mathrm{d}r^2 / \left(1 - k\left(ct\right)r^2\right) + r^2 \mathrm{d}\Omega^2. \qquad (8\text{-}3)$$

(Henceforth, we will define the time coordinate to be ct which has the units of length).

The next step is to specify the general form of the complete metric. If we just set $k = k\left(ct\right)$ in the FRW metric and calculate the Ricci tensor, we will find that there is an off-diagonal component proportional to the time derivative of the curvature. This demands that the metric must also have an off-diagonal term so the simplest generalization of the FRW metric must have the form

$$\mathrm{d}s^2 = -\tilde{q}\left(ct,r\right)\left(c\mathrm{d}t\right)^2 + \tilde{h}\left(ct,r\right)\left(c\mathrm{d}t\right)\mathrm{d}r + a^2\left(ct\right)\left(\mathrm{d}r^2 / \left(1 - k\left(ct\right)r^2\right) + r^2 \mathrm{d}\Omega^2\right). \quad (8\text{-}4)$$

The scaling, like the curvature, is a property of the 3-space so by our assumption of homogeneity and isotropy, it, like the curvature, must also depend only on the time.

Without providing the proof, it follows from the equations that to avoid singularities at $r = 0$, $\tilde{h}$ must be proportional to r and $\tilde{q}\left(ct,r\right)$ must have the form $\tilde{q}\left(ct,r\right) = 1 - r^2 q\left(ct,r\right)$. A redefinition of time was used to fix $\tilde{q}\left(ct,0\right) = 1$. Next, the work of [3] allows us to replace $q\left(ct,r\right)$ with the form

$$q(ct,r) = \left(1 - k(ct)r^2\right)\left(h(ct,r)/a(ct)\right)^2 \tag{8-5}$$

where we have defined $\tilde{h}(ct,r) = rh(ct,r)$. To avoid another singularity, it happens that the radial derivative of h must also vanish at $r = 0$.

We now note that redefining the radial coordinate as minus itself leaves the metric unchanged from which we can conclude that h is an even function of r, or in other words, a function of r^2 rather than just r. The first derivative then automatically vanishes at $r = 0$ thus satisfying the various conditions. The same argument applied to the energy-momentum tensor shows that the energy density and pressure are also functions of r^2 so the first derivative of the pressure also automatically vanishes at $r = 0$.

The final metric is then

$$\begin{aligned}
\mathrm{d}s^2 = {} & \left(-1 + \frac{r^2 h(ct,r)^2}{a(ct)^2}\left(1 - k(ct)r^2\right)\right)(c\mathrm{d}t)^2 \\
& + 2h(ct,r)(c\mathrm{d}t)r\mathrm{d}r + a^2(ct)\left(\frac{\mathrm{d}r^2}{1 - k(ct)r^2} + r^2\mathrm{d}\Omega^2\right)
\end{aligned} \tag{8-6}$$

With this metric, we have reached the point where the necessary calculations are truly beyond the capabilities of a human working by hand. Aside from the time required, the vast number of calculations simply cannot be completed without error. In our case, we choose to use the symbolic capabilities of Mathematica © Wolfram Research, Inc. to do the heavy lifting.

Developing the field equations is straightforward. The energy-momentum tensor is

$$\mathsf{T}^{\mu\nu} = \left(\rho c^2(ct,r) + p(ct,r)\right)\delta_0^\mu \delta_0^\nu + p(ct,r)\mathsf{g}^{\mu\nu} \tag{8-7}$$

since the spacetime is at rest. From this metric, the field equations have the form,

$$\mathsf{R}_{00}(ct,r)/S_{00} = \kappa\left(\rho c^2 + \left(\frac{3a^2 - r^2 h^2\left(1 - kr^2\right)}{a^2 - r^2 h^2\left(1 - kr^2\right)}\right)p\right) \tag{8-8a}$$

$$\mathsf{R}_{01}(ct,r)/S_{01} = \kappa\left(\rho c^2 + \left(\frac{3a^2 - r^2 h^2\left(1 - kr^2\right)}{a^2 - r^2 h^2\left(1 - kr^2\right)}\right)p\right) \tag{8-8b}$$

$$\mathsf{R}_{11}(ct,r)/S_{11} = \kappa\left(\rho c^2 - \left(\frac{a^2 - r^2 h^2\left(1 - kr^2\right)}{a^2 + r^2 h^2\left(1 - kr^2\right)}\right)p\right) \tag{8-8c}$$

$$\mathsf{R}_{22}(ct,r)/S_{22} = \kappa\left(\rho c^2 - \left(\frac{a^2 + r^2 h^2\left(1 - kr^2\right)}{a^2 - r^2 h^2\left(1 - kr^2\right)}\right)p\right) \tag{8-8d}$$

where $\kappa = 8\pi G/c^4$. The equation for R_{33} is identical to that for R_{22}. For brevity, the arguments were suppressed on the RHS of the equations. Unlike the FRW case, the Ricci tensor components of (8-8c) and (8-8d) are different func-

tions of the metric components. In all cases, some simplification was achieved by the scaling of each equation by the coefficients of the energy density.

$$S_{00}(ct,r) = \frac{\left(a^2 - r^2 h^2 \left(1-kr^2\right)\right)^2}{2a^4} \tag{8-9a}$$

$$S_{01}(ct,r) = -\frac{rh}{2}\left(1 - \frac{r^2 h^2 \left(1-kr^2\right)}{a^2}\right) \tag{8-9b}$$

$$S_{11}(ct,r) = \frac{a^2 + r^2 h^2 \left(1-kr^2\right)}{2\left(1-kr^2\right)} \tag{8-9c}$$

$$S_{22}(ct,r) = \frac{r^2 a^2}{2} - r^2 h^2 \left(1-kr^2\right) \tag{8-9d}$$

The expansions of the Ricci tensor components are long so we won't write them out here. (By long, we mean that the some of the expansions contain well over 300 terms).

In addition to these equations, we have the two equations that follow from the conservation condition, (3-3). Like the Ricci tensor components, these are also rather long. Symbolically these are

$$\nabla_\mu \mathsf{T}^{\mu ct}(ct,r) = 0 \tag{8-10a}$$

$$\nabla_\mu \mathsf{T}^{\mu r}(ct,r) = 0. \tag{8-10b}$$

As we noted earlier, there is a difference between the "real" universe as a sequence of hyperspaces and the universe as perceived by any observer. At any moment of cosmic time, the universe consists of one, single hyperspace which is characterized by its curvature and scaling. There is no notion of time or location on this hyperspace because time is the same everywhere and all points are equivalent, *i.e.* there is no preferred origin. Also, because a hyperspace exists at a single moment of time, signals within a hyperspace are impossible and thus, an observer placed on such a hyperspace would not be able to say anything about that hyperspace because his or her hyperspace is unobservable.

Observer's do receive signals, of course, but what they are observing are signals arriving from previous hyperspaces. It is these signals that constitute the observer's perception of the universe and it is these signals that Einstein's equations describe. In other words, Einstein's equations describe *any observer's perception of the universe in terms of his or her time and radial coordinates*. A different observer would have a different perception, even though they exist in a single universe, and the relationship between these is also fixed by the equations. We have then a "real" universe consisting of a sequence of homogeneous hypersurfaces which is overlaid by non-homogeneous observer perceptions which are unique to each observer. Since the equations are dependent on a metric, it follows that any conclusions drawn from *observations* about the expansion of the universe, for example, are totally dependent on the choice of metric since any observer's perception is dependent on the all the intervening spacetime between

the observer and the observed object.

A key point is that with time-varying curvature, these field equations are functions not of just time but also of the radial coordinate, *i.e.* $R_{00} = R_{00}(ct, r)$, etc. The question, then, is how do we interpret these field equations, which certainly do have reference to an origin ($r = 0$) and which do describe signals, in such a way that they describe hyperspaces which have neither. The resolution of this dichotomy comes when we realize that each hypersurface is just the set of all possible observer origins and since all such locations are equivalent, any one observer's field equations will comprise the field equations of the hyperspace as a whole *when evaluated at that observer's origin*. Thus, the field equations that replace the FRW field equations follow not from equations which are free of the radial coordinate, as is the case with the FRW metric, but from the $r = 0$ limit of the more general field equations which *are* dependent on the radial coordinate. We can conclude then that Einstein's equations, which are concerned with signals, make contact with an observer's hyperspace only in the limit of signals of zero extent.

The curvature is a property of the hypersurface and so must relate to the energy density and pressure of that hypersurface so (5-3) now becomes

$$k(ct) = \kappa a^2(ct)\left(k_1 \rho c^2(ct, 0) + k_2 p(ct, 0)\right) \tag{8-11}$$

where the quantities on the RHS are evaluated at $r = 0$ and k_1, k_2 are constants yet to be determined. As we noted earlier, this equation is, in a sense, a replacement for the "equation of state" of the FRW model but with the difference that, in this case, it falls out as part of the solution rather than being introduced as an *ad hoc* assumption.

Starting with the above equations and then taking the limit of $r \to 0$, we obtain the following equations. (Note that we are switching to the Mathematica notation which is more compact that the standard notation for partial derivatives. The notation $g^{(i,j)}[ct, r]$ denoted the ith derivative with respect to first listed coordinate which happens to be the time coordinate, ct, and j^{th} derivative with respect to 2nd coordinate. Equations (8-12a) 0 (8-12c) follow directly from (8-8a) 0 (8-8c). The equation that follows from (8-8d) is identical to (8-12c). Equation (8-12d) follows from (8-10a). Equation (8-10b) becomes $p^{(0,1)}[ct, 0] = 0$ which is satisfied identically because the pressure is a function of r^2.

$$-\frac{6h[ct, 0]^2}{a[ct]^4} - \kappa\left(3p[ct, 0] + \rho c^2[ct, 0]\right) - \frac{6a''[ct]}{a[ct]} + \frac{6h^{(1,0)}[ct, 0]}{a[ct]^2} = 0 \tag{8-12a}$$

$$-\frac{6h[ct, 0]^2}{a[ct]^4} - \kappa\left(3p[ct, 0] + \rho c^2[ct, 0]\right) + \frac{8h[ct, 0]a'[ct]}{a[ct]^3}$$
$$-\frac{4a'[ct]^2}{a[ct]^2} - \frac{2k'[ct]}{h[ct, 0]} - \frac{2a''[ct]}{a[ct]} + \frac{2h^{(1,0)}[ct, 0]}{a[ct]^2} = 0 \tag{8-12b}$$

$$\frac{6h[ct,0]^2}{a[ct]^4}+\kappa\left(p[ct,0]-\rho c^2[ct,0]\right)+\frac{4k[ct]}{a[ct]^2}-\frac{8h[ct,0]a'[ct]}{a[ct]^3}$$
$$+\frac{4a'[ct]^2}{a[ct]^2}+\frac{2a''[ct]}{a[ct]}-\frac{2h^{(1,0)}[ct,0]}{a[ct]^2}=0$$

(8-12c)

$$\left(p[ct,0]+\rho c^2[ct,0]\right)\frac{3a'[ct]}{a[ct]}+\left(\rho c^2\right)^{(1,0)}[ct,0]=0.$$

(8-12d)

Our original six equations have thus been reduced to four. An important simplification has occurred because none of these contain spatial derivatives.

We will now set about solving these equations. We first subtract the 2nd equation from the 1st and solve for $h^{(1,0)}[ct,0]$. At the same time, we introduce a new function $\bar{h}$ defined by

$$h[ct,r]=\frac{a[ct]^2\,k'[ct]}{2k[ct]}\bar{h}[ct,r]$$

(8-13)

After rearranging, we have

$$\bar{h}^{(1,0)}[ct,0]=-\frac{2k[ct]^2}{a[ct]\bar{h}[ct,0]k'[ct]}-\frac{2k[ct]a'[ct]^2}{k'[ct]a[ct]^2}+\frac{\bar{h}[ct,0]k'[ct]}{k[ct]}$$
$$+\frac{2k[ct]a''[ct]}{a[ct]k'[ct]}-\frac{\bar{h}[ct,0]k''[ct]}{k'[ct]}$$

(8-14)

We next solve (8-12a) and (8-12c) for $\rho c^2[ct,0]$ and p [ct, 0],

$$\rho c^2[ct,0]$$
$$=\frac{3}{\kappa a[ct]^4}\left(h[ct,0]^2-2a[ct]a'[ct]h[ct,0]+a[ct]^2\left(k[ct]+a'[ct]^2\right)\right)$$

(8-15a)

$$p[ct,0]=-\frac{1}{\kappa a[ct]^4 h[ct,0]}\left(3h[ct,0]^3-6a[ct]a'[ct]h[ct,0]^2\right.$$
$$\left.+a[ct]^2 h[ct,0]\left(k[ct]+3a'[ct]^2\right)+a[ct]^4 k'[ct]\right)$$

(8-15b)

We now substitute into (8-12d), solve for h and then use (8-13) to obtain

$$\bar{h}[ct,0]=\left(1-\frac{\kappa\left(\rho c^2[ct,0]+p[ct,0]\right)}{2k[ct]/a[ct]^2}\right)^{-1}.$$

(8-16)

We now define a parameter γ_h by

$$\bar{h}[ct,0]=-\frac{\gamma_h}{1-\gamma_h}$$

(8-17)

and solve for $k[ct]$ to find

$$k[ct]=\frac{1}{2}\gamma_h a[ct]^2\kappa\left(\rho c^2[ct,0]+p[ct,0]\right).$$

(8-18)

This constitutes the proof that the relationship between the curvature and the energy density and pressure that we have been hinting at is an exact conse-

quence of time-varying curvature. Comparing with (8-11), we find that $k_1 = k_2 = \gamma_h/2$ which means that the curvature is dependent only on the *sum* of the vacuum energy and the pressure. In fact, we will later discover that *all* physical quantities depend only on that sum. (We note that no such relationship exists in the standard model because it does not contain the necessary off-diagonal component in the metric; *i.e.* in the FRW model, $\gamma_h = 0$). As a matter of terminology, instead of always writing out "vacuum energy density plus the pressure", we will often use the shorter "vacuum energy density" which we intend to mean the same thing.

Based on our original contention that $k_{1,2}$ were constants, it follows that γ_h will also be a constant. That, however, is still an assumption. Nevertheless, proceeding with that assumption, we will find a complete solution of the equations which demands that indeed, γ_h is a constant.

That begin the case, $\bar{h}[ct, 0]$ is also a constant and it follows that the RHS of (8-14) vanishes. We now assert that a substitution of

$$k[ct] = \bar{k}_0 \left(\frac{a[ct]}{ct} \right)^2 \tag{8-19}$$

where

$$\bar{k}_0 = k_0 \left(\frac{ct_0}{a_0} \right)^2 = 0.0884 k_0 \tag{8-20}$$

reduces the equations to a single non-linear differential equation for $a[ct]$. This equation has a quadratic form given by

$$\frac{A_1[ct] \pm \sqrt{A_1[ct]^2 - 4\bar{k}_0 a[ct]^2 A_2[ct]}}{2A_2[ct]} = -\frac{\gamma_h}{1-\gamma_h} \tag{8-21a}$$

$$A_1[ct] = -(ct)^2 \left(\dot{a}[ct]^2 - a[ct]\ddot{a}[ct] \right) \tag{8-21b}$$

$$A_2[ct] = a[ct]^2 - (ct) \left(\dot{a}[ct]^2 - a[ct]\ddot{a}[ct] \right) \tag{8-21c}$$

which has the solution,

$$a[ct] = \mathbb{C}_2 (ct)^{\gamma_h + \bar{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h}} e^{\mathbb{C}_1 ct} \tag{8-22}$$

where $\mathbb{C}_1, \mathbb{C}_2$ are constants. (The second solution happens to be a special case of the solution shown). Expressing this in terms of a reference time of t_0, this becomes,

$$a[ct] = \left(a_0 e^{-c_1} \right) \left(\frac{ct}{ct_0} \right)^{\gamma_h + \bar{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h}} e^{\frac{ct}{ct_0} c_1}. \tag{8-23}$$

Here c_1 is also a constant but its value is dependent on the choice of the reference time, *i.e.* $c_1 = \mathbb{C}_1 / (ct_0)^2$.

What we find is that with time-varying curvature, there *must* be an acceleration of the *scaling*. We emphasize that this is a *prediction* of the model. In con-

trast, the standard model does not actually make any prediction at all. Instead, the so-called prediction of an accelerated scaling results from curve fitting rather than from any fundamental constraint imposed by the structure of the model. Put another way, the standard model claims an accelerated scaling *after the fact* of the luminosity distance observations whereas the new model predicts an accelerated scaling without any reference to luminosity distance or any other observation.

The energy density and pressure are now

$$\rho c^2 [ct, 0] = \frac{3}{\kappa (ct_0)^2} \left(\frac{c_1^2}{(1-\gamma_h)^2} + \left(\frac{2c_1 \overline{k}_0}{\gamma_h} \right) \frac{ct_0}{ct} + \overline{k}_0 \left(1 + \overline{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h^2} \right) \frac{(ct_0)^2}{(ct)^2} \right) \quad \text{(8-24a)}$$

$$p[ct, 0] = \frac{-3}{\kappa (ct_0)^2} \left(\frac{c_1^2}{(1-\gamma_h)^2} + \left(\frac{2c_1 \overline{k}_0}{\gamma_h} \right) \frac{ct_0}{ct} + \overline{k}_0 \left(1 - \frac{2}{3\gamma_h} + \overline{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h^2} \right) \frac{(ct_0)^2}{(ct)^2} \right) \quad \text{(8-24b)}$$

$$\rho c^2 [ct, 0] + p[ct, 0] = \frac{2\overline{k}_0}{\kappa (ct_0)^2 \gamma_h} \frac{(ct_0)^2}{(ct)^2} = \frac{2k_0}{\kappa a_0^2 \gamma_h} \frac{(ct_0)^2}{(ct)^2}. \quad \text{(8-24c)}$$

The Ricci scalar is

$$R[ct, 0] = \frac{12}{(ct_0)^2} \left(\frac{c_1^2}{(1-\gamma_h)^2} + \left(\frac{2c_1 \overline{k}_0}{\gamma_h} \right) \frac{ct_0}{ct} + \overline{k}_0 \left(1 - \frac{1}{2\gamma_h} + \overline{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h^2} \right) \frac{(ct_0)^2}{(ct)^2} \right). \quad \text{(8-25)}$$

Comparing with (8-19), we see that the Ricci scalar varies as $k[ct]/a[ct]^2$ which with (5-1), ties the characteristic curvature of (2-3) to the Ricci scalar as promised earlier.

We now wish to fix the unknown constants. It will be useful to make two additional definitions. First, we define the constants

$$\gamma_* = \gamma_h + \overline{k}_0 \frac{(1-\gamma_h)^2}{\gamma_h} \quad \text{(8-26a)}$$

$$a_* = a_0 e^{-c_1}. \quad \text{(8-26b)}$$

In terms of these, we have

$$a[ct] = a_* \left(\frac{ct}{ct_0} \right)^{\gamma_*} e^{\frac{ct}{ct_0} c_1}. \quad \text{(8-27)}$$

We see that the scaling is power law for $ct/ct_0 \ll 1$ and exponential for $ct/ct_0 \geq 1$.

At this point, aside from the value of k_0, we now have exact results for $k(ct)$ and $a(ct)$ so we could substitute these back into the original equations leaving us with a set of equations for $h(ct, r)$, $\rho c^2(ct, r)$, and $p(ct, r)$. It happens that subtracting the 2nd equation from the 1st leaves us with a non-linear PDE for $h(ct, r)$ *alone* as it did for the $r \to 0$ limit. (There are other linear combinations that also result in equations for $h(ct, r)$ alone but these are far more complicated). Once a solution for $h(ct, r)$ is known, subtracting (8-8d) from (8-8c) yields a formula for $p(ct, r)$ and lastly substituting again into

(8-8c), say, determines $\rho c^2 (ct, r)$ and the solution is complete. For now, we will set aside the problem of determining $h(ct, r)$ with the promise to return to it later after developing further context.

For any power law scaling with a constant scaling parameter, we have $a'[ct]/a[ct] = \gamma/t$ and it will be useful to emulate this result in the general case by defining an effective scaling parameter by

$$\gamma_{eff}[ct] = ct\frac{a'[ct]}{a[ct]} = \gamma_* + c_1 \frac{ct}{ct_0}. \tag{8-28}$$

To fix the parameters, we use the fact that, as shown in Sec. 4, the effective scaling at the present time is $\gamma_{eff}[ct_0] = 0.95$. Also, as we will see in Sec. 10, the existence of the CMB requires that, at the time of nucleosynthesis, the effective scaling must have been $\gamma_{eff}[ct_n] \approx 0.5$. Making the assumption that the scaling parameter is exactly 0.5, we then have $\gamma_* = 0.5$ and $c_1 = 0.45$ with γ_h given by

$$\gamma_h = \frac{\overline{k}_0 + \dfrac{\gamma_*}{2}\left(1 + \sqrt{1 - 4\overline{k}_0 \dfrac{1 - \gamma_*}{\gamma_*^2}}\right)}{1 + \overline{k}_0}. \tag{8-29}$$

For $\overline{k}_0 = 1/8$, which is a value we will explain shortly, we have $k_0 = 1.414$ and $\gamma_h = 1/3$. We note that this result supports our earlier assumption that γ_h is a constant.

The resulting curves for the scaling parameter and the scaling are shown in **Figure 3**. Also indicated are certain milestone times. The time t_n marks the beginning of nucleosynthesis and as we will see later, the CMB temperature at t_n is about a factor of 10 smaller than that of standard model so the time of recombination is correspondingly earlier. There is no change in the time of galaxy formation because it is not dependent on the temperature.

The effective scaling parameter is essentially constant up until about 1% of the present age of the universe and then gradually approaches an exponential with increasing time. The middle chart shows the actual scaling and the lower shows the last two decades in more detail. We see that even though the effective scaling parameter is increasing rapidly, the actual scaling does not differ greatly from 2/3rds scaling over that time range.

Next, we show the Hubble parameter in **Figure 4**. We see first that the Hubble parameter increases with increased look-back time (or redshift.) It is a constant power law curve for times earlier than $ct/ct_0 < 0.1$ but is very non-linear for more recent times.

We now turn to the energy density and pressure. The first thing we note is that there is a constant contribution to both reminiscent of a cosmological constant with a value of 6.8×10^{-10} J·m^{-3}. This contribution, however, has no physical significance and just amounts to a redefinition of what we mean by zero energy and pressure. A constant energy or pressure everywhere is the same as no energy or pressure at all. In any case, we could simply eliminate it by adding a
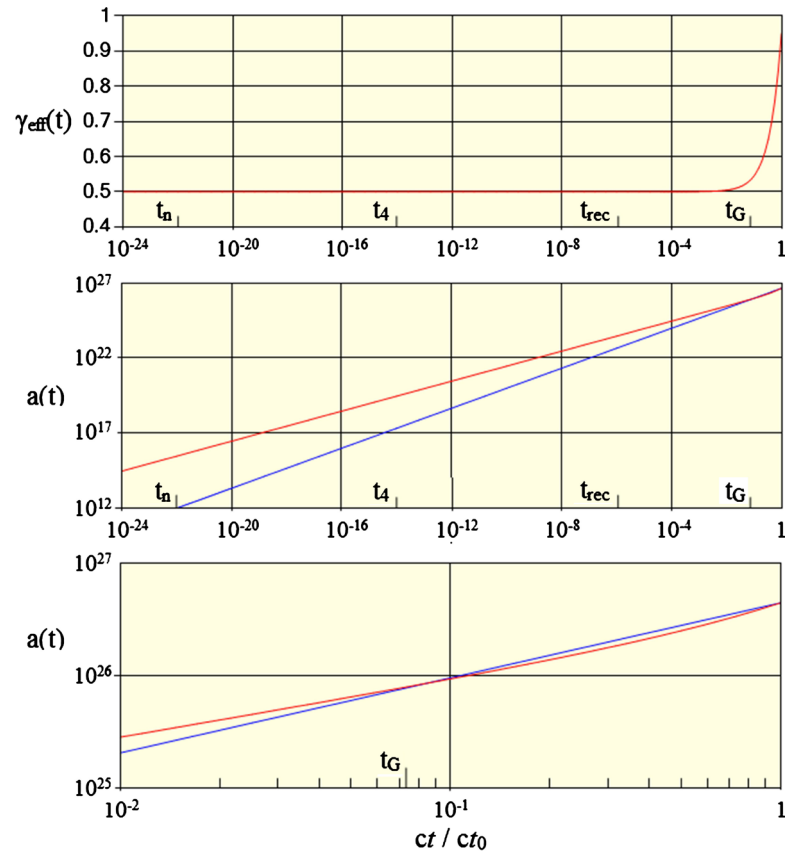
Figure 3. Time-varying curvature predictions in red. For comparison, the curves for $2/3^{rds}$ scaling are shown in blue. The indicated times are: t_n = time of neutron formation to be explained below, t_4 = end of nucleosynthesis, t_{rec} = recombination, and t_G = galaxy formation.
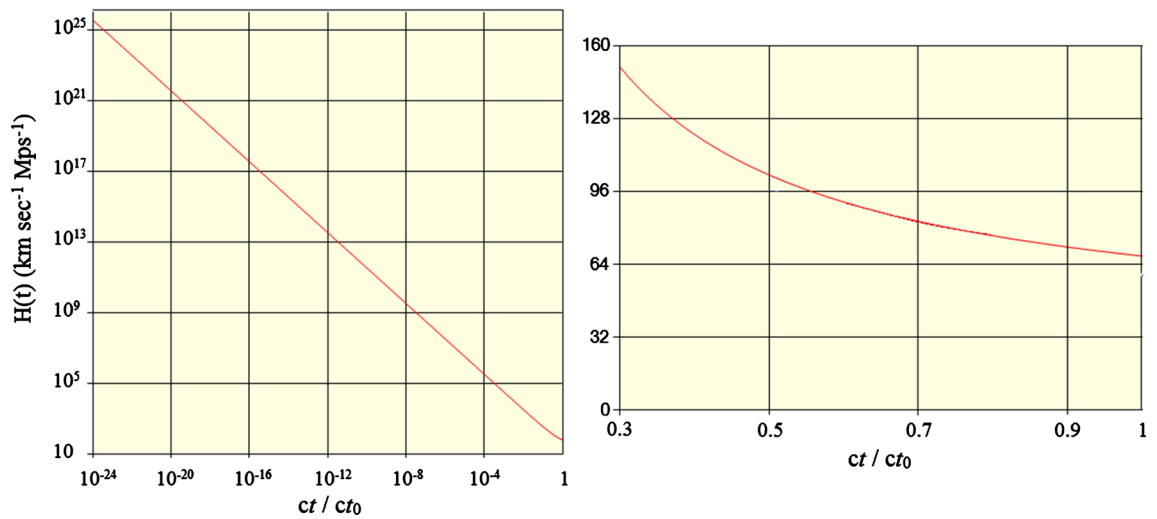
Figure 4. Hubble parameter.

cancelling "cosmological" constant to the equations. More importantly, however, the parameters of physical interest, e.g. the scaling and the curvature, *are dependent only on the sum* of the energy density and the pressure and the constant

contribution cancels in that sum.

Ignoring this constant value and using the scaling parameters just determined along with a value of $k_0 = 1.41$, we calculate the energy density and pressure shown in Figure 5. The time-varying curvature thus predicts that there is pressure and that both it and the energy density vary as t^{-2} up until shortly before the time of galaxy formation at which point, the pressure begins a rapid decline and eventually becomes negative.

Returning to the point about physical quantities, we now want to consider the motion of a test particle with 4-velocity $u^\mu = \left(u^t, u^r, u^\theta, u^\varphi\right)$. The particle's geodetic equations are given by

$$\frac{du^\mu}{d\tau} + \Gamma^\mu_{\nu\sigma} u^\nu u^\sigma = 0. \tag{8-30}$$

The significant point here is that the connection coefficients are dependent only on the metric functions and because $h(ct, r)$ is the solution of a differential equation that does not contain either the vacuum energy or pressure separately, it, like the curvature, is a function only of the sum of the energy and the pressure. The result is that the motion of the test particle is, in turn, only dependent on that sum. Thus, while we can talk about the vacuum energy and pressure as distinct quantities, only their *sum* is of physical significance.
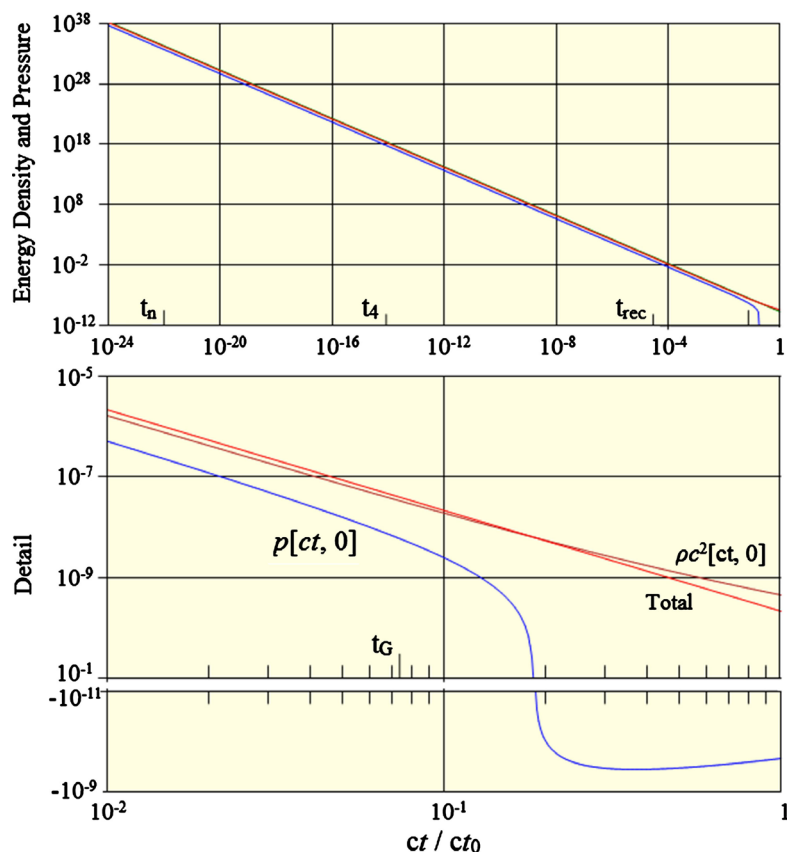
Figure 5. Energy density and pressure for $k_0 = 1.41$. The energy density is shown in red, the pressure in blue and their sum in firebrick.

If we actually work out the geodesic equations, we find that the *ct* and *r* equations are rather long. The angle equations, however, are short. The θ equation is

$$\frac{\mathrm{d}u^\theta}{\mathrm{d}\tau} + \frac{2u^r u^\theta}{r} - \left(u^\varphi\right)^2 \sin(\theta)\cos(\theta) + 2u^t u^\theta \frac{\dot{a}(ct)}{a(ct)} = 0 \tag{8-31}$$

with a similar equation holding for u^φ. We see now that if at any point on the test particle's trajectory, $u^\theta = u^\varphi = 0$, the corresponding velocity derivatives vanish so on any such trajectory, the angles are constant which is a reflection of the lack of off-diagonal metric components connecting the angle and time coordinates.

At this point, we will pause to compare two predictions with currently accepted values. First, we note that the predicted present-day value of the sum is

$$\rho c^2 (ct_0, 0) + p(ct_0, 0) = 2.1 \times 10^{-10}\,\mathrm{J} \cdot \mathrm{m}^{-3} \tag{8-32}$$

which differs from the currently accepted dark energy density (6.3×10^{-10} J·m^{-3}) by *no more than a factor of* 3. We can also compute the total energy to find, $E_{\text{total}} = 7.5 \times 10^{70}\,\mathrm{J}$ which is smaller than the value in (2-1) by a factor of about 4. We thus find that the vacuum energy as determined by the exact solution of Einstein's equations can account for two of the properties of spacetime that are considered mysteries in the standard model.

The radius of curvature, defined earlier in (5-1), is

$$R_c(ct) = \frac{a(ct, 0)}{\sqrt{k(ct)}} = \frac{a_0}{\sqrt{k_0}}\frac{ct}{ct_0} = 2.85ct \tag{8-33}$$

which varies linearly with time as we determined earlier.

At this point, we need to raise an issue concerning ordinary matter. The solution presented is correct and in particular, the predicted scaling is correct. What is missing, however, an understanding of the contribution from the ordinary matter. This potentially becomes significant during the latter stages of the evolution of the universe but we need further development before we can address this issue.

We will now establish an upper limit on k_0. Referring back to (8-26a) and (8-27), we see that for any γ_*, there is maximum value of $\bar{k}_0$ above which there is no value of γ_h that can realize that value of γ_*. This condition is expressed by the requirement that (8-29) must yield a positive, real number and thus the limiting value of $\bar{k}_0$ is given by the vanishing of the radical which with $\gamma_* = 0.5$ yields

$$\bar{k}_0 = \frac{\gamma_*^2}{4(1-\gamma_*)} = 1/8 \tag{8-34}$$

which explains the value we have been using. Later, in fact, we will give observational evidence that $\bar{k}_0$, and hence, k_0, *always* has this maximal value and thus is not an adjustable parameter but instead is a prediction.

We next wish to make contact with our inflation model. Earlier, we asserted

that the self-interaction of the curvature would result in a slower than $2/3^{rds}$ expansion which we have now found to be the case. We can now extrapolate backwards to the inflation to determine the change in the cutoff necessary to account for the different scaling. Without going through the details, by re-running the inflation simulation, we determine that the end of the inflation occurred at a value of $\tau_I \approx 4.2$. Also, by comparing the energy density at the end of the transition with (8-24c), we obtain a value of $k_0 \approx 6.9$. Even though this is larger than the allowed limit, this is really a remarkable result because we are tying together the two ends of the evolution of the universe. At the end of the inflation, the energy density was equal to the Plank energy density and during the transition phase, the density dropped by a factor of about 10^{-6} and yet this simple model of the transition yields a curvature that exceeds the upper limit by less than a factor of 5. The adjusted inflation is shown in Figure 6.

This result also suggests that as a general principle, the curvature always has its maximum possible value and we will later find evidence for this when we examine the luminosity distance data in the next section.

Recall that the entropy is determined by the number of Plank cells and in this case, the adjusted α_I has a value of $\alpha_I \approx 62$ which results in an entropy of $O\left(10^{80}\right)$.

The next topic we will discuss is the interpretation of the radial coordinate. In order to do so, however, we need to develop the equation that relates time and the radial coordinate along the path of a photon emitted by a source at time t_e and received sometime later at time t_r by an observer so we are now concerned with an observer's perceptions and Einstein's equations. Starting with the metric, setting $\mathrm{d}s^2$ and also $\mathrm{d}\Omega = 0$ results in a quadratic equation for $c\mathrm{d}t$ which has the solution,

$$c\mathrm{d}t = \frac{a(ct)\mathrm{d}r}{\sqrt{1-k(ct)r^2}}F(ct,r) \tag{8-35a}$$

where

$$F(ct,r) = \frac{\sqrt{1+r^2\hat{h}^2(ct,r)}+r\hat{h}(ct,r)}{\sqrt{1-q(ct,r)r^2}} \tag{8-35b}$$

$$\hat{h}(ct,r) = \frac{h(ct,r)\sqrt{1-k(ct)r^2}}{2a(ct)\sqrt{1-q(ct,r)r^2}}. \tag{8-35c}$$
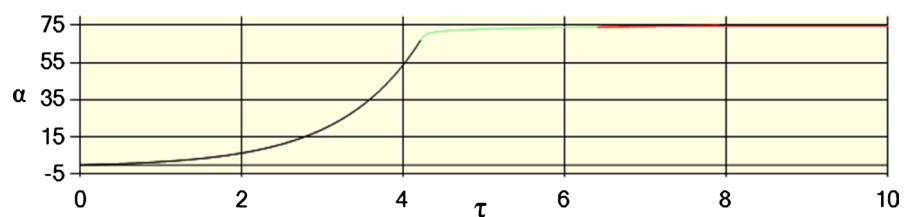
Figure 6. Inflation adjusted to exact solution.

(The absence of off-diagonal elements connecting the time and angular coordinates means that photons travel along lines of constant angular coordinate so it is meaningful to speak of a single value for the angles). Rearranging we have

$$r(ct_r, ct_e) = \int_{ct_e}^{ct_r} c\mathrm{d}t' \frac{\sqrt{1 - k(ct')r^2}}{a(ct')F(ct', r)} \qquad (8\text{-}36)$$

which is the integral form of a nonlinear differential equation for r. In this case, r is defined with respect to the source so $r(ct_e, ct_e) = 0$. The present-day redshift is given by

$$z = \frac{\partial_t r(ct, ct_e)\big|_{ct=ct_e}}{\partial_t r(ct, ct_e)\big|_{ct=ct_0}} - 1. \qquad (8\text{-}37)$$

Note that these reduce to the FRW formulas when $h = 0$ and $k = \text{constant}$.

You will recall that according to the convention we adopted, the radial coordinate is dimensionless and specifies any location as a fraction of the scaling. It thus has the limits of $0 \leq r \leq 1$. We also see what appears to be a singularity in the metric, (8-6), at $r = 1/\sqrt{k(ct)}$ whenever $k \geq 1$ which with the maximal curvature will always be the case. We can now understand the nature of this apparent singularity. From (8-35a), it appears that the time interval corresponding to an infinitesimal increase in the radial coordinate becomes infinite at that value of r. In other words, for sources at or beyond that coordinate limit, photons would require an infinite amount of time to reach the observer which means that they are not visible. Thus, although any observer would know that there must be a universe lying beyond this horizon, the field equations describing the observer's perception of the universe only retain validity out to the limiting value of

$$r_h(ct) = 1/\sqrt{k(ct)}. \qquad (8\text{-}38)$$

The corresponding actual proper distance would be

$$R_h(ct) = a(ct, 0)/\sqrt{k(ct)} \qquad (8\text{-}39)$$

since the appropriate scaling is that of the hyperspace at time coordinate, ct. We will refer to this as the horizon distance to avoid confusion with other definitions of related concepts. This brings us full circle back to the radius of curvature of (8-33); *the horizon distance and the radius of curvature are the same thing*. The meaning of this distance is that it is the proper distance between a source and observer (at time t) that are just beyond the limit of being able to influence each other assuming that each emitted a signal at time $t = 0$. As we will see shortly, however, this result is an oversimplification and the actual limit on our ability to detect distance sources is slightly less. The horizon distance is a different concept than the limit on communication since the latter requires an exchange of multiple signals within a meaningful period of time and so is much smaller.

We noted earlier that because the metric components are functions of both t and r, the universe will not appear homogeneous to an observer even though

each hyperspace is homogeneous. It is natural then to ask to what degree and in what manner will the universe not appear homogeneous. To answer this question, we will calculate the radial coordinate and redshift using the above equations. Rearranging (8-35a) and introducing the dimensionless time variable, $\xi = ct/ct_0$, we find

$$\frac{\mathrm{d}r}{\mathrm{d}\xi} = \frac{ct_0}{a(\xi)} \frac{\sqrt{1-k(\xi)r^2}}{F(\xi,r)} \tag{8-40}$$

which can be solved using the 4th order Runge-Kutta method. Working from the point of view of the source, the initial condition is $r(\xi_e,\xi_e) = 0$. At this point, we do not yet have a solution for $\bar{h}(ct,r)$ so for the moment, we will assume that it has the constant value given by (8-17). For $\bar{k}_0 = 1/8$, this has the value $\bar{h}[ct,0] = -0.5$.

In the following figures, we wish to compare with standard model results. Because in the latter case, a redefinition of the radial coordinate is usually done, we cannot easily compare with results in that formulation. Instead, we just compute the curves with k set to a constant value of $k = 1$. The time-varying curvature solutions are shown in red and the constant curvature solutions in black.

In **Figure 7**, the curves are the locus of the radial coordinates of sources that emitted signals at the indicated time that were later received by an observer at the present time.
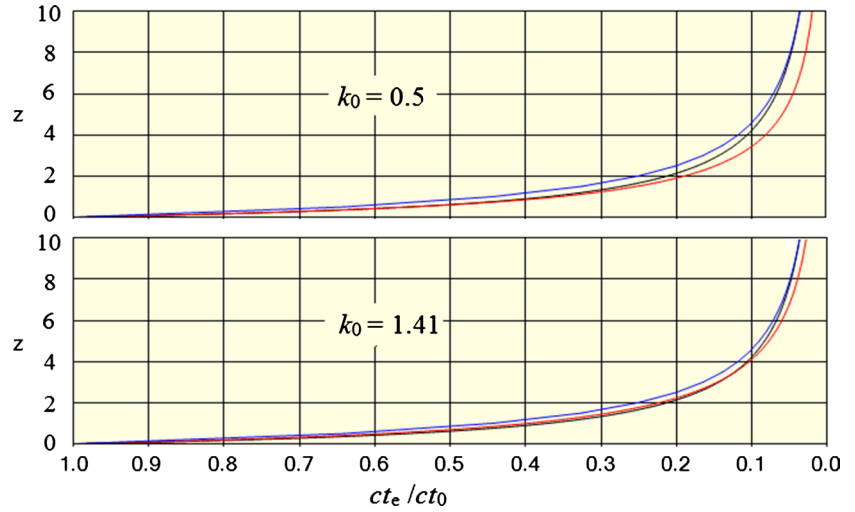
In **Figure 8**, we show the computed redshifts for the same set of parameters. Also shown in blue are the redshifts calculated using the FRW lookback time. ([2], page 409).

$$1-\xi = (t_0 H_0)^{-1} \int_0^z \mathrm{d}z'(1+z')^{-1}\left[(1+z')^2(1+\Omega_m z')-z'(2+z')\Omega_\Lambda\right]^{-1/2}. \tag{8-41}$$

where we have used $\Omega_k = 1-\Omega_m-\Omega_\Lambda$. In this case, we used the preferred values of $(\Omega_m,\Omega_\Lambda) = (0.24,0.76)$.
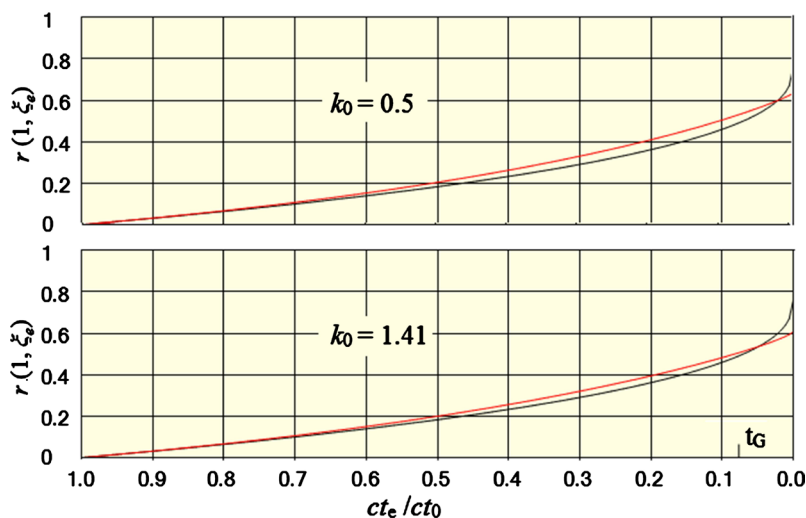
Figure 7. $r(1,\xi_e)$ vs ξ_e for two values of k_0 .

Figure 8. Redshift vs ξ_e for two values of k_0. Time-varying curvature in red, constant curvature in black and the FRW lookback time result in blue.

Starting with **Figure 7**, we see that while the curves are not far apart for a given lookback time, if we ask for the lookback time corresponding to a particular radial coordinate, we see that there can differences on the order of 75% between the two model predictions. We also see that the new model predicts an upper limit on the radial coordinate of visible sources unlike the standard model which shows no such limit. This means that we can only see sources with a radial coordinate less than about 0.62 no matter how early the source emitted its signal.

Considering now **Figure 8**, for time ratios greater than about 0.2, the same arguments just made apply to the potential error when interpreting the redshift of a source. For smaller value of the time ratio, the curve becomes steep and the opposite condition applies, namely that there will be small error in the time determined from a known redshift but a large error in the redshift for a particular time ratio. These curves serve to show that there are differences between the standard and time-varying curvature predictions that could be significant when interpreting observations.

Figure 9 shows the Hubble parameter as a function of the redshift. The exact and constant curvature results were obtained by plotting the exact Hubble parameter of **Figure 4** as a function of the redshift using the curves of **Figure 8** to make the conversions. We see the magnification of the effect of the apparent small difference between the curves of **Figure 8** which we touched upon earlier.

For comparison, we also show the FRW formula

$$H_{\text{FRW}}(z) = H_0\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0} + \Omega_{k,0}(1+z)^2} \tag{8-42}$$

for two values of $(\Omega_m, \Omega_\Lambda)$. The Hubble constant was set to a value of $H_0 = 67.3$ to match the exact curve at $t = t_0$. What we find is that there is a considerable difference between the exact and FRW results.

In **Figure 10**, we compare the scaled angular distance from the two models. In the FRW case, the angular distance is given by
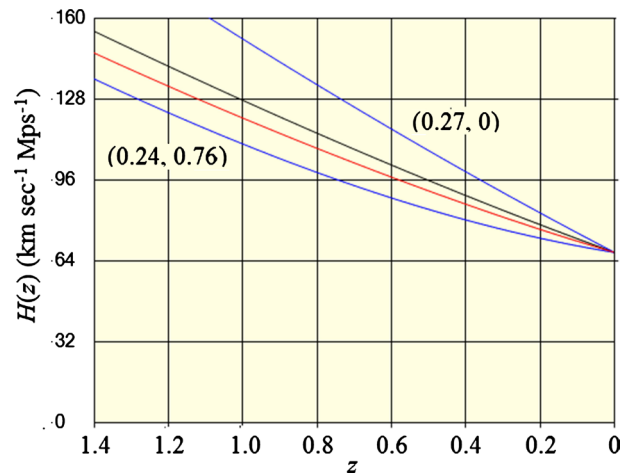
Figure 9. Hubble parameter vs redshift. Time-varying curvature in red, constant curvature in black, and the FRW results for two values of the densities in blue.
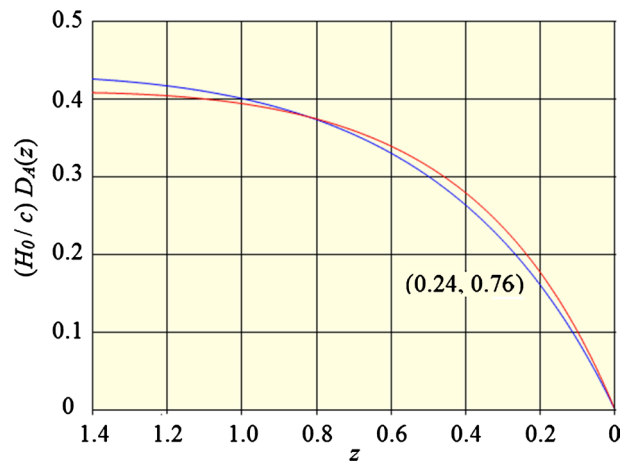
Figure 10. Scaled Angular Distance vs Redshift. Exact solution in red and the FRW curve in blue.

$$D_A(z) = \left(\frac{c}{H_0}\right)\frac{1}{1+z}|\Omega_k|^{-1/2}\sinh\left\{|\Omega_k|^{1/2}I(z)\right\} \qquad (8\text{-}43a)$$

where

$$I(z) = \int_0^z dz'\left[(1+z')^2(1+\Omega_m z') - z'(2+z')\Omega_\Lambda\right]^{-1/2}. \qquad (8\text{-}43b)$$

In the exact case, it is given by

$$D_A(z) = a(z)r(z) \qquad (8\text{-}44)$$

where the scaling is given in (8-27) and the coordinate distance and redshirt are shown in **Figure 7** and **Figure 8**. We again see that there are differences which very likely have a bearing on the current difficulties in trying to fix the Hubble constant.

These results were calculated assuming a constant value for $\overline{h}(ct, r)$ instead of the actual solution. We now wish to establish that these results have some va-

lidity which requires that we have at least some idea about the r dependence of $\bar{h}$ and its effect on the calculated curves. Developing the appropriate equation for $\bar{h}(ct,r)$ is straightforward. We start with the difference between (8-8a) and (8-8b) and perform a series of transformations to get the result into its final form. First, we make replacements using (8-13), (8-19), (8-27) and their derivatives. Next, we make a change of variable to η defined by $ct/ct_0 = \mathrm{e}^\eta$ followed by a change of radial coordinate to a scaled coordinate defined by $r_s = r/r_h(ct)$ which ranges from 0 to 1. To finish, we substitute numerical values for the various constants, e.g. a_0, ct_0, γ_h, etc. The result is shown in (8-45) and (8-46).

$$0.0891 + \bar{h}[\eta, r_s]\Big(f_1[\eta, r_s] + f_2[\eta, r_s]\bar{h}[\eta, r_s]$$
$$+ f_3[\eta, r_s]\bar{h}^{(0,1)}[\eta, r_s] + f_4[\eta, r_s]\bar{h}^{(1,0)}[\eta, r_s] \Big) = 0 \tag{8-45a}$$

where

$$f_1[\eta, r_s] = 0.356 + \left(-0.534 + 0.144\mathrm{e}^{2\eta}\right)r_s^2 \tag{8-45b}$$

$$f_2[\eta, r_s] = 0.356 + \left(-1.247 + 0.641\mathrm{e}^{\eta} - 0.144\mathrm{e}^{2\eta}\right)r_s^2$$
$$+ \left(0.891 - 0.641\mathrm{e}^{\eta} + 0.144\mathrm{e}^{2\eta}\right)r_s^4 \tag{8-45c}$$

$$f_3[\eta, r_s] = r_s\Big[\left(0.178 - 0.144\mathrm{e}^{2\eta}\right) + \left(-0.356 + 0.289\mathrm{e}^{2\eta}\right)r_s^2$$
$$+ \left(0.178 - 0.144\mathrm{e}^{2\eta}\right)r_s^4 \Big] \tag{8-45d}$$

$$f_4[\eta, r_s] = \left(-0.356 + 0.321\mathrm{e}^{\eta}\right) + \left(0.712 - 0.641\mathrm{e}^{\eta}\right)r_s^2$$
$$+ \left(-0.356 + 0.321\mathrm{e}^{\eta}\right)r_s^4 \tag{8-45e}$$

and

$$r_h(\eta) = \frac{1.319}{\sqrt{\mathrm{e}^{0.9\mathrm{e}^{\eta} - \eta}}}. \tag{8-46}$$

The coefficients are shown in **Figure 11**. At this point, the obvious next step would be to specify boundary conditions and ask Mathematica to grind out the result. Unfortunately, while Mathematica can deal with some nonlinear PDEs, those must be quasilinear which this equation is not. The technical reason for this limitation is that Mathematica applies Newton's method to a linearized version of the equations. In this case, that procedure becomes highly unstable and no solution can be found. In so happens, however, that we can learn enough to answer the question concerning the importance of the r dependence on the curves of **Figure 7** and **Figure 8** without having an exact solution.

The first result is found by evaluating the equation at $r_s = 1$. The solution of the resulting equation is

$$\bar{h}[\eta, 1] = \frac{1}{2 - 1.62\mathrm{e}^{2\eta}}. \tag{8-47}$$

The significant point is that this is positive for all $\eta \leq 0$. We know, on the other hand, that $\bar{h}[\eta, 0] = -0.5$ so the two together imply that there must exist a curve $r_s[\eta]$ such that $\bar{h}[\eta, r_s[\eta]] = 0$ for all η. But this is not possible
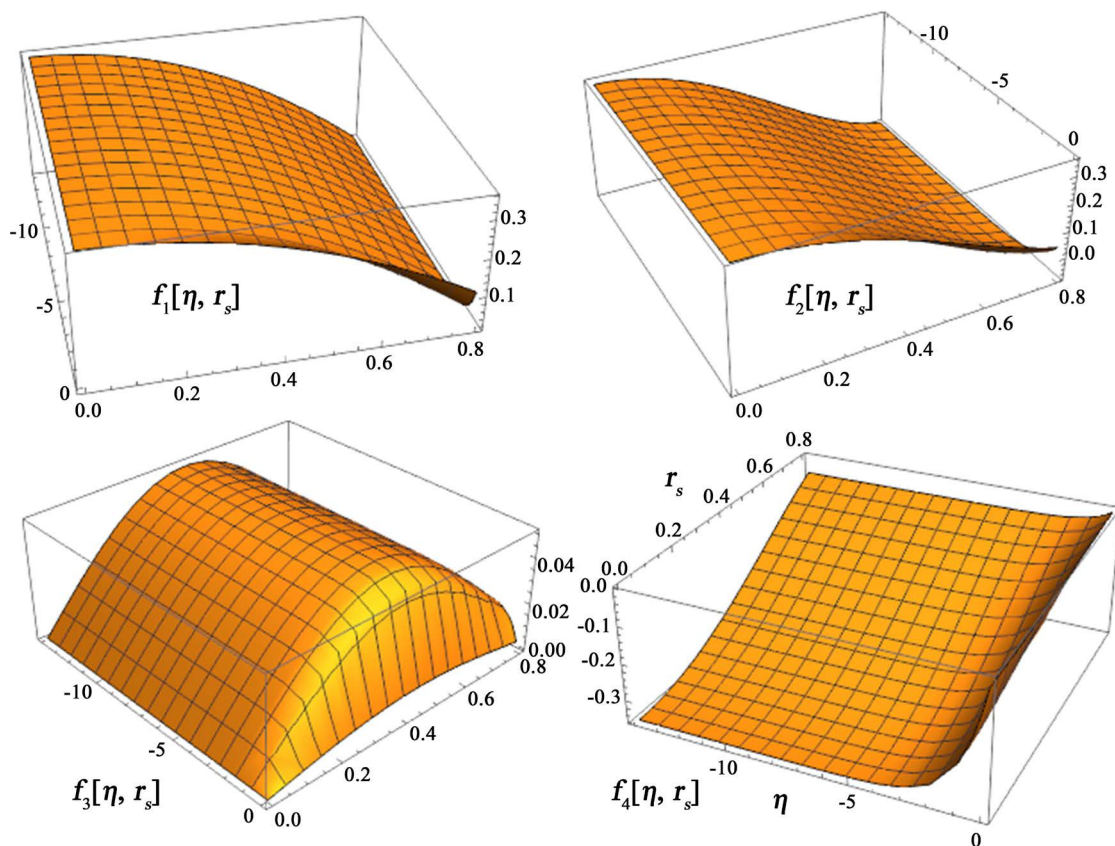
Figure 11. $\bar{h}[\eta, r_s]$ coefficients for $-14 \leq \eta \leq 0$ and $0 \leq r_s \leq 0.8$. ($\eta_{trec} = -13.7$).

because such a value would not be a solution of (8-45a). The conclusion is that there must exist a singularity in the equation along some curve, $r_s[\eta] < 1$. This implies an upper limit on the radial coordinates of visible sources.

We now turn to small values of r_s. We noted earlier that we must have $h^{(0,1)}[ct,0] = 0$ in order to avoid a singularity. We will now examine the equation for small r_s to show that this is indeed the case.

First, we write

$$\bar{h} \approx -0.5 + h_1 r_s + h_2 r_s^2 \tag{8-48}$$

which we substitute into the sum of the first 3 terms of (8-45a). The result is

$$-0.108\left(0.412 - 1.481\mathrm{e}^{\eta} + \mathrm{e}^{2\eta} - 3.29 h_1^2\right) r_s^2. \tag{8-49}$$

Since we are asserting that the derivative vanishes at $r_s = 0$, we can safely drop the last term. We next substitute into the 4th term but this time, we include only the constant term of our approximate solution because the 4th term already contains a factor of r_s. The result is

$$-0.5\left(0.178 - 0.144\mathrm{e}^{2\eta}\right) r_s \bar{h}^{(0,1)}[\eta, r_s]. \tag{8-50}$$

Finally, after dropping the 5th term because the time derivative vanishes at $r_s = 0$ and is generally small elsewhere. We now solve for $\bar{h}^{(0,1)}[\eta, r_s]$ and integrating the result to obtain

$$\bar{h}[\eta, r_s] = -0.5 + \frac{0.75(0.412 - 1.481e^{\eta} + e^{2\eta})}{1.235 - e^{2\eta}} r_s^2. \tag{8-51}$$

Figure 12 shows the result. We see that at least in this approximation that the solution varies slowly for small r_s (notice the vertical scale) and that, indeed, $h^{(0,1)}[ct, 0] = 0$ as asserted. As a check, we substitute (8-51) back into the original equation with the result shown in Figure 13. In addition to the 3D plot, we have shown 3 slices at the indicated values of η.

Comparing these curves while using the constant term of (8-45a) as a magnitude reference, we see that this approximation does a reasonable job of satisfying the equation even out to values of $r_s > 0.2$.

Finally, we will now return to the original question concerning the validity of the curves of Figure 7 and Figure 8. We recalculated the curves (for $k_0 = 1.41$ only) with a set of trial functions with the same small r_s behavior as the approximate result. The result was that the radial coordinate curves were indistinguishable from the plotted curve in Figure 7 from which we can conclude that to
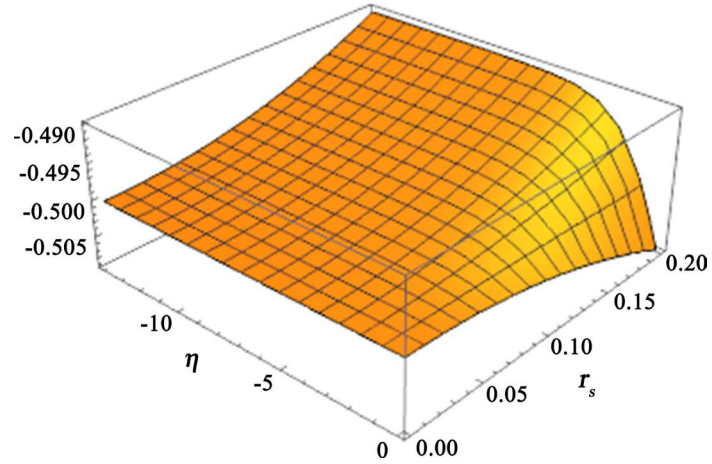
Figure 12. Solution of (8-45) for small r_s.
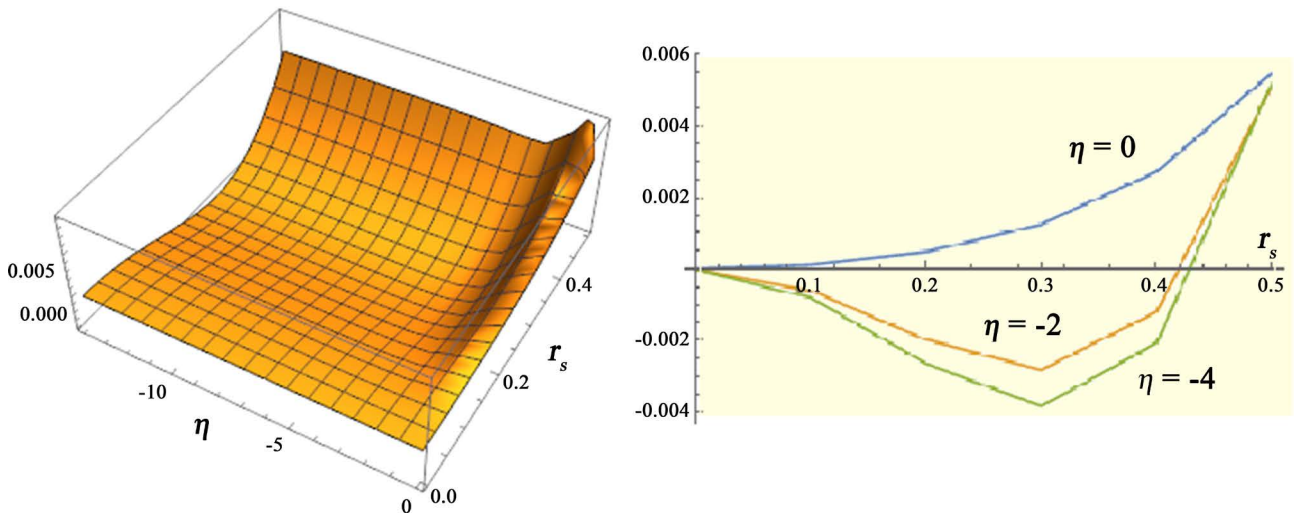
Figure 13. Evaluation of (8-45) using (8-51).

a reasonable degree of accuracy, the result shown in Figure 7 is correct. The same process was applied to the redshift. For values of $ct_e/ct_0 > 0.1$, the curves were again indistinguishable but for smaller values of ct_e/ct_0, there were small variations with magnitudes of about twice the width of the plotted lines. We can conclude that plotted curve of Figure 8 also gives a reasonable approximation to the actual redshift but with somewhat less confidence that the result of Figure 7.

The solutions for $\alpha(ct)$, $k(ct)$, $\rho c^2(ct,0)$, and $p(ct,0)$ are exact. Given that fact, we then have a problem with the expansion in the future. The issue is that the scaling is predicted to be exponential whereas the energy density, for example, is varying as t^{-2}. The result is that the total energy of the universe would be predicted to eventually increase exponentially. The standard model, of course, has the same problem, if indeed it is a problem. Einstein's equations already enforce energy conservation locally so it may not be even meaningful to be concerned about the total energy. Nevertheless, it does leave one to wonder if it would be possible to introduce an additional constraint on the model that would preserve local energy conservation while at the same time forcing energy conservation globally. Such a constraint would not have a large effect on the results from the Plank era to the present because the total energy predicted by the current solution does not vary greatly during that period but such a condition have a significant impact on the future evolution as a result of suppressing either or both of the scaling and the vacuum energy density in such a matter as to keep the total energy constant.

To get a sense of when these effects will begin to be important, the curvature has been decreasing since the initial inflation but at the time that effective scaling equals unity, the scaling will begin to outrun the influence of the Big Bang and the curvature will start to increase. The point at which that will happens is $t = 1.1t_0$ which on cosmic times scales is just around the corner.

In summary, we have presented a new model of the expansion of the universe that provides a good match to observations. The only parameters that appear are t_0, a_0, $\gamma_{eff}(t_0)$, and $\gamma_{eff}(t_n)$ which are fixed by observation. The universe that is described by this model is open while at the same time, the curvature is always positive which is diametrically opposed to the FRW model. We have also shown that predictions based on the FRW model should yield reasonable results when working backwards from the present to about a redshirt of unity or so but for earlier times, that is not the case at all. This means that the error estimates on the values of various physical parameters obtained from earlier observations are suspect because they were obtained using an invalid "ruler".

9. Luminosity Distance

We will turn to the issue of the ongoing observations of the luminosity "distance". In 1998, Riess, *et al.*, [4], reported observations of type 1a supernovae that, when interpreted in the context of the FRW model, suggest that there was an observable acceleration of the scaling for values of $z \approx 1$ which in turn sug-

gested the existence of a cosmological constant. Later, in 2016, Nielsen, *et al.*, [5], published a new analysis of a much larger data set that cast doubt on the original conclusions. In this section, we will review the data and its FRW interpretation and then consider the situation in light of the new model.

We begin with the definition of the luminosity "distance" of some source. We put "distance" in quotes because luminosity "distance" is not a distance at all but instead is a *model dependent* construct that happens to have the dimension of length. Its usefulness is that it can be both measured and calculated thus allowing theory and observation to be compared. The definition is

$$D_L \equiv \left(\frac{L}{4\pi F} \right)^{1/2} \tag{9-1}$$

where L is the absolute luminosity of the source and F is the energy flux arriving at the Earth. Observationally, this quantity is determined by measuring the flux received from a multitude of sources at different distances that are known to have the same absolute luminosity. To calculate this quantity, we start with the formula for the arriving flux which is

$$F = \frac{L}{A(1+z)^2} \tag{9-2}$$

where A is the area of the sphere centered at the source. In this formula, there are two factors of $(1+z)$. One of these is the result of the photons being redshifted because of the expansion of the universe and the other is a consequence of the fact that the arrival rate of the photons is also reduced by the expansion. Substituting into (9-1) gives

$$D_L = \frac{\sqrt{A}(1+z)}{2\sqrt{\pi}} \tag{9-3}$$

Note that the absolute luminosity cancels when calculating the distance. So far, this formula is valid for any metric. It is when computing the area that the metric becomes involved. For the metric of (8-6) the area is $A = 4\pi \left(a(t_0) r(z) \right)^2$, where $r(z)$ is the solution to (8-40) that is shown in **Figure 7**. Thus,

$$D_L = a(t_0)(1+z)r(z). \tag{9-4}$$

In the FRW case, the area is $A = 4\pi \left(a \sinn(\chi) \right)^2$ where sinn() is sinh() in an open universe and sin() in a closed universe,

$$D_{L,\text{FRW}} = a(t_0)(1+z)\sinn(\chi(z)) \tag{9-5}$$

In its most general form, the FRW model allows for ordinary matter, radiation, and a cosmological constant. It is standard practice, however, to neglect the radiation component in which case the coordinate is given by the following formula (see [2], page 411)

$$a(t_0)\chi(z) = \left(\frac{c}{H_0} \right) I(z) \tag{9-6}$$

where $I(z)$ is given in (8-43b). Substituting yields for $\Omega_k > 0$,

$$D_{L,\text{FRW}}(z) = \left(\frac{c}{H_0}\right)(1+z)|\Omega_k|^{-1/2}\sinh\left\{|\Omega_k|^{1/2}I(z)\right\} \tag{9-7}$$

and when $\Omega_k = 0$,

$$D_{L,\text{FRW}} = \left(\frac{c}{H_0}\right)(1+z)I(z). \tag{9-8}$$

Turning now to the data, luminosity distance data is generally presented in the form of a Hubble diagram in which the distance modulus, defined by

$$\mu_P = 5\log_{10}(D_L) + 25 \tag{9-9}$$

is plotted against the redshift. In this formula, D_L is measured in units of Megaparsecs.

The Riess *et al.* observations are shown in **Figure 14** that follows. These results by themselves don't actually tell us much because nothing can be said about the relationship between observations and hyperspaces outside the context of a metric. In this case, the authors used the FRW metric and by doing a best fit they came to the conclusion that there must be a cosmological constant which in turn implies a pressure term in (3-1) that results in an acceleration of the scaling. This conclusion is, however, based almost entirely on the single data point at the largest redshift and even more importantly, on the assumption of the FRW metric.

We now jump ahead to the data set complied by Nielsen *et al.* which is shown in **Figure 15**. This analysis includes a much larger data set than does the Riess analysis and more specifically includes four data points with redshifts greater that the last point of the Riess data set. Because the authors chose to use a linear redshift scale, it is difficult to compare this figure with the previous results. To facilitate a comparison, we have combined the two sets in **Figure 16**. Instead of duplicating the mass of data points of the Nielsen graph, we just show the general trend of the data by the black line with only the last four data points plotted
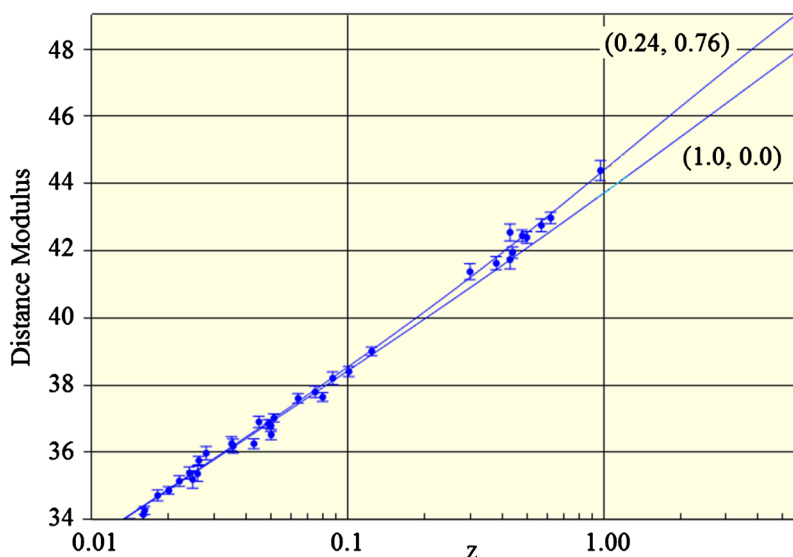
Figure 14. Hubble Diagram from [4].
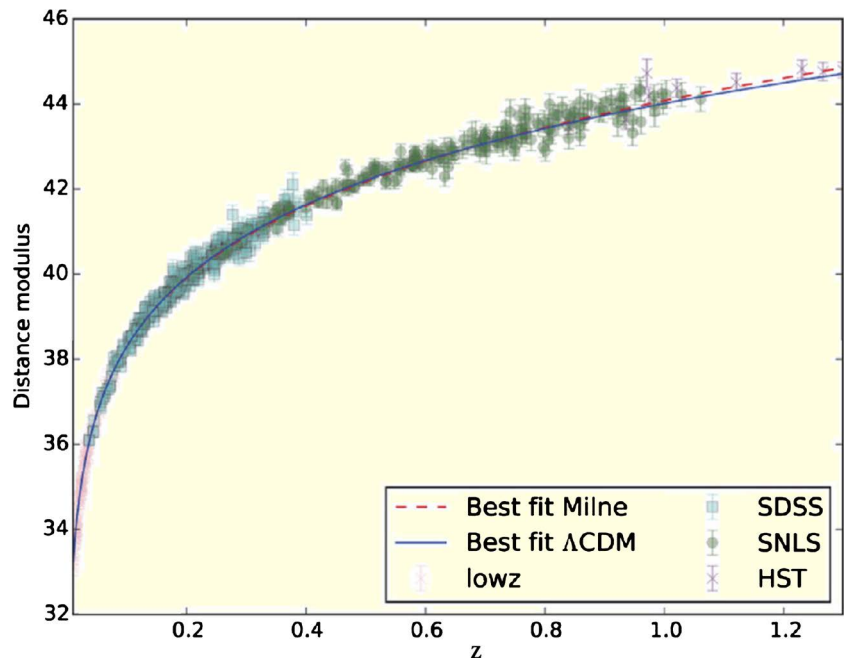
Figure 15. Summary of luminosity distance observations from [5].
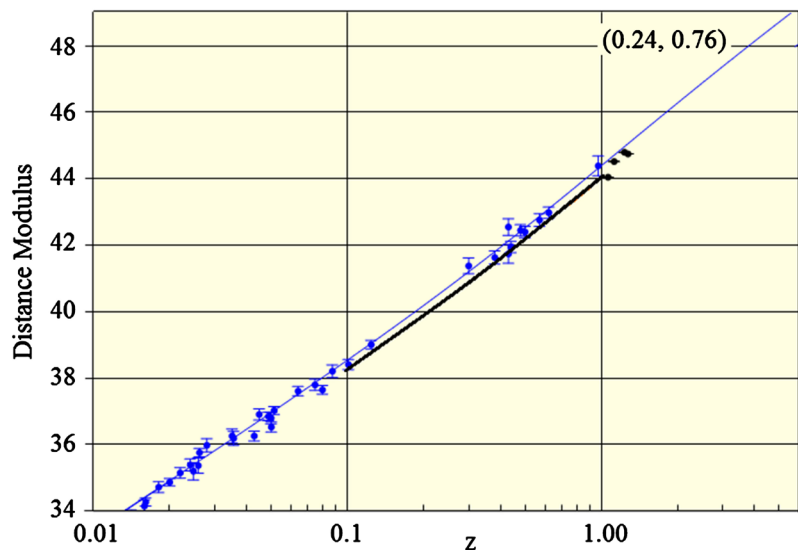
Figure 16. (Important note: The black line and data points where scaled off a printed copy of **Figure 15** and must not be considered as accurate representations of the data.)

separately.

What we see is that the last four data points indicate that the upward curvature of the data for large redshifts in less pronounced than indicated by the single Riess data point thus casting doubt on the conclusion concerning a cosmological constant. It is also apparent from this graph and the previous one that the deviation of the single Riess data point from the "no acceleration" line is not unlike the scatter in some of the other data points at smaller redshifts; for example, at $z = 0.43$. Keep in mind too that a larger distance modulus just indicates that the radiation is dimmer than expected and so could be the result of some un-

identified mechanism that absorbed or scattered the light along the light of sight to that particular source.

Figure 17 shows the model predictions for two values of k_0. What we find is that the time-varying curvature solution with *no* adjustable parameters other than k_0 provides a very close fit to the data. (A slightly larger value of H_0 gives an even better fit). The predicted curves show the upward trend of the data but not to the extent of the cosmological constant model. This is a rather ironic result because it shows that while there is an acceleration of the scaling which the time-varying curvature result explicitly expresses, the luminosity "distance" observations within the range of observed redshifts do not provide any clear evidence for that acceleration and certainly do not provide evidence for a cosmological constant.

Comparing the two model results, we see that the larger k_0 curve gives a noticeably closer fit to the data and from this we conclude that the curvature is a large as it can be so from this point on, we will take it as a *general principle of the model* that the curvature *always* takes on its maximum possible value. Thus, $\bar{k}_0 = 1/8$ and $k_0 = 1.41$ which is now a prediction leaving us with *no* adjustable parameters. It follows also that the energy density and pressure are also as large as they can be. This result is actually a continuation of the situation existing during and immediately after the original inflation when those quantities had maximal values set by the Plank dimensions.

We will now turn to the problem of the origin of ordinary matter and the CMB.

10. Asymmetry, Ordinary Matter, and the CMB

With respect to the nucleosynthesis era, it is important to differentiate between what is known from what is conjecture. Nucleosynthesis proper consists of the binding of an initial population of neutrons and protons into light elements via
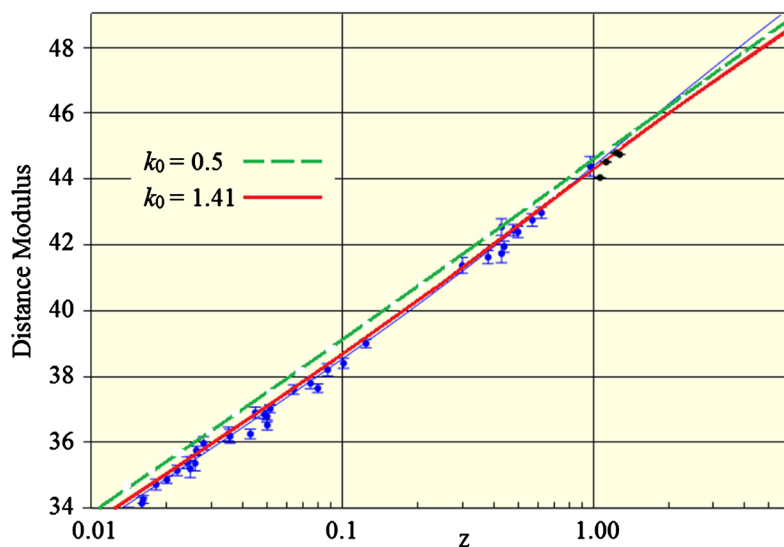
Figure 17. Time-varying curvature prediction of the luminosity distance.

well-known reactions. Because all the important reactions have been studied in the laboratory, it is a straight-forward problem to calculate the final densities of the light elements. Validation of these results have come from measurements of light element densities in young galaxies and while such validations are indirect because of the long time span between the end of nucleosynthesis and the formation of galaxies, it is felt that the physics of the intervening time period is sufficiently well-understood to consider the validations as significant. With the final densities known, working backwards to the beginning of nucleosynthesis proper allows us to be also reasonably confident about the initial densities of neutrons and protons required to make it all work.

That, however, is as far as one can go with respect to observations. What this means is that, at least with respect to nucleosynthesis, everything leading up to that initial population of neutrons and protons is conjecture. In other words, there is *no* observational evidence that the standard inflation/ quantum field theory model of the pre-nucleosynthesis period actually happened.

In this and the next section, we will present an alternative model that leads to the same nucleosynthesis starting point and in addition, accounts for the matter/ antimatter asymmetry of the universe.

We will first establish some basic parameters that will give us a framework for the arguments that follow. The matter in the universe, as is well known, exists in long relatively thin filaments of galaxies which contain about 94% of the mass of the universe and which surround voids that make up about 80% of the volume of the universe and contain the remaining 6% of the total mass. Simulation results are shown in Figure 18. As an aside, we will show in Sec. 16 that this filament structure is a consequence of vacuum energy structures that originated during the initial inflation.

First, we must separate the 20% of the volume that contains 96% of the matter from the voids. In the latter, the density is
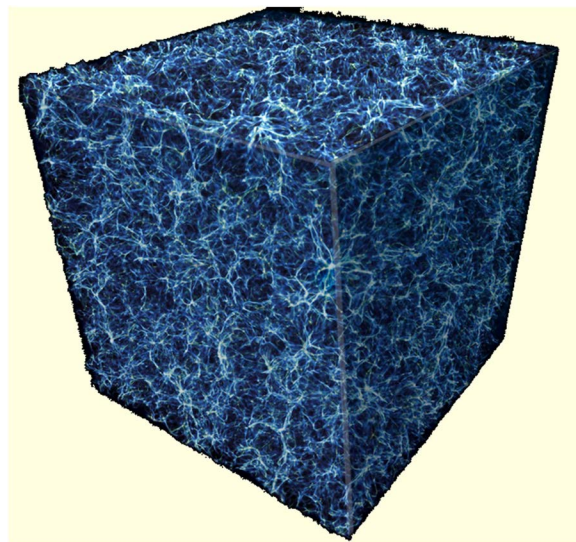
Figure 18. Cosmic web, Nasa [6].

$$n_{void} = \frac{0.06N}{0.8V} = 0.075\frac{N}{V} \tag{10-1}$$

and in the material portion,

$$n_m = \frac{0.94N}{0.2V} = 4.7\frac{N}{V} \tag{10-2}$$

Observations suggest that in the material regions the present-day average density is about 1 m^{-3} (see, e.g. [7] [8]) which gives an overall average value of $N/V = 0.21$. Thus, the number density in the voids must be on the order of

$$n_{void} = 0.016 \text{ m}^{-3}. \tag{10-3}$$

The average in the material regions was the value just given but in the subregions in which most of the nucleosynthesis actually occurred, the density was 2 - 3 times larger as determined by the nucleosynthesis process itself. This latter value is in agreement with the value determined by counting stars. The general consensus is that there exist 10^{22} - 10^{24} stars and using the Sun as an average mass, the equivalent present-day number density of hadrons is on the order of 0.034 - 3.4 m^{-3} so the two are roughly the same.

In order to create these particles, the energy density of the source must have been at least as great as their rest mass which, using a neutron as the architype, is 1.35×10^{44} J·m^{-3}. Equating this to the vacuum energy density, (8-24) will allow us to fix the point in time, denoted by t_n, at which the primary particle creation must have ceased; that is, provided we know the scaling parameter. This, however, we can determine from the energy density of the CMB.

The temperature of the CMB varies with time according to

$$T(t_n) = T(t_0)\frac{a_0}{a(t_n)} \tag{10-4}$$

and, assuming a black-body spectrum, the corresponding energy density was

$$\rho_\gamma c^2(t_n) = a_B T^4(t_n). \tag{10-5}$$

Now, if we assume a trial value of $\gamma_{eff}(t_{rec}) = 0.6$, we find that $t_n = 5.2 \times 10^{-5}$ s and $\rho_\gamma c^2(t_n) = 6.9 \times 10^{39}$ J·m^{-3} but we also have $\rho_{vac} c^2(t_n) = 2.1 \times 10^{34}$ J·m^{-3} so we immediately see that this value isn't going to work because the necessary CMB energy density would then be vastly larger than the total energy of the universe.

Turning the problem around, we can instead ask what value of the scaling is necessary to bring the radiation energy density into line with the vacuum energy density? The result is a value a little larger than 0.5. This value, however, cannot be correct either because the vacuum energy accounts for most of present-day energy of the universe so the radiation energy must be much less. Making a jump, we will henceforth suppose that $\gamma_{eff}(t_n) = 0.5$. This happens to be the same value as that of a radiation dominated universe and it also has the virtue that the ratio $\rho_\gamma c^2(t)/\rho_{vac} c^2(t)$ is a constant up until a time somewhat later than t_{rec} when the acceleration of the scaling began to be significant. With this value, we have

$t_n = 4.3 \times 10^{-5}$ s which is not much different from the earlier result.

We can now calculate the various quantities of interest assuming a present-day particle density of 2 m^{-3} in those regions where nucleosynthesis was significant. The results are shown in Table 1. Of course, we haven't created the particles or radiation yet but these will be their densities when we do.

Looking at these numbers, we see that the radiation energy is about 0.1% of the vacuum energy density and that the particle energy density is vastly smaller even when their rest mass is included. This clearly reinforces the idea presented earlier that the scaling of the universe is entirely a consequence of the time-varying vacuum energy density. We also see that the temperature is about a factor of 10 less that the standard model temperature and that the ratio of particles to photons, $n_{part}(t_n)/n_\gamma(t_n) = 5.1 \times 10^{-9}$ is about a factor of 10 larger. Finally, we have listed both the horizon distance and ct_n. The former defined the greatest distance to a source that emits a signal at time $t = 0$ that can be received by an observer. The latter, on the other hand, is a measure of the distance over which a source and observer can communicate. A source emitting a signal at time $t_n = 4.3 \times 10^{-5}$ s will be received by an observer at a distance of ct_n at a time $t = 2t_n = 8.6 \times 10^{-5}$ s .

We now want to characterize the possible scenarios leading up to the starting point of nucleosynthesis. The problem is to not only to account for the values just discussed, but to account for the matter/antimatter asymmetry of the universe. Since we start with a vacuum and end up with both particles and radiation, there are three possibilities as shown in Figure 19. What we will show is

Table 1. Various quantities at the time $t_n = 4.3 \times 10^{-5}$ s .

a (m)	k	R_h (m)	ct (m)	T (K)	kT (MeV)
2.8×10^{15}	5.8×10^{21}	3.7×10^4	1.3×10^4	4.2×10^{11}	36.5

Energy Density (j·m^{-3})				Num Density (m^{-3})	
Vacuum	Particle thermal	Particle mass	Photons	Particles	Photons
2.1×10^{34}	4.5×10^{22}	1.2×10^{24}	2.4×10^{31}	7.7×10^{33}	1.5×10^{42}
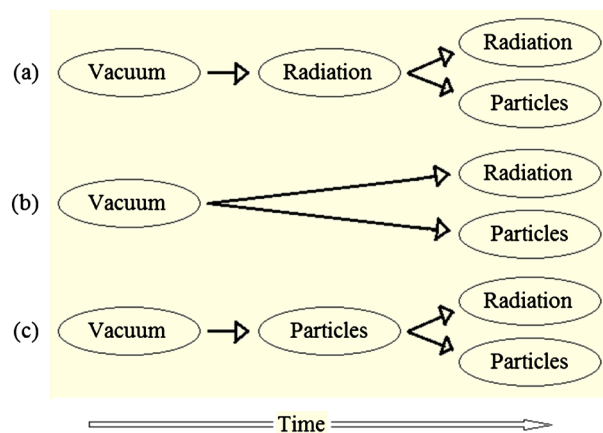
Figure 19. Possible nucleosynthesis scenarios.

that a scenario of type (a) cannot explain the matter/antimatter asymmetry. Scenarios of type (b) could explain the asymmetry but suffer from a number of problems that render such a scenario as very unlikely. This leaves the last type as the one most likely to be correct. We want to emphasize that the big jump is to go from vacuum to matter or, in other words, from nothing to something. Whether the something is radiation or particles is really a secondary issue since we have no idea of how the vacuum could accomplish either. We can only say with certainty that it did happen.

Scenario (a)

The standard model is an example of this type in which it is assumed that vacuum energy underwent a transition into radiation that eventually transitioned into the mix of particles and radiation via processes described by quantum field theory.

The main point in this case is that photons are matter/antimatter neutral so even if the vacuum had an asymmetry, such an asymmetry could not have been imprinted on the radiation. Likewise, quantum field theory is also matter/antimatter neutral, at least at a level that can be detected via experiments, so it follows that there is no mechanism by which an asymmetry with a single "sign" could have been created on a large scale.

We might imagine, however, that locally some asymmetry could have been introduced via random fluctuations. Here now is an essential point; because of the finite speed of light, there was no communication over distances larger than 10^4 m at time t_n and thus correlations of matter vs antimatter could not have extended over any region larger than that dimension. Further, the state of each such cell would have been random not just with respect to its "sign" but also with respect to its percentage of asymmetry.

We can consider two limiting cases. In the first, let us assume that each entire cell was either matter or antimatter. Soon after their formation, nearest neighbors would have begun a process of annihilation. The total number of such cells would have been $N_{\text{cell}}(t_n) = \left(a(t_n)/R_h(t_n)\right)^3 = 4.3 \times 10^{32}$ and after the annihilations were complete, the excess of matter cells over antimatter cells could not have exceeded $\sqrt{N_{\text{cells}}} = 2.1 \times 10^{16} \text{ m}^{-3}$. If we assume that the initial energy density of the particles was the same as that of the radiation, we would have started with a particle density of roughly 10^{41} m^{-3} so the final density would have been no greater than 10^{25} m^{-3} which is vastly smaller than the value of 10^{33} m^{-3} indicated by the present-day particle density.

The other limiting case is that in which matter and antimatter particles were created at random. In that case, annihilation would immediately have reduced the density within each cell to a value no greater than $\sqrt{10^{41}} = 3.2 \times 10^{20} \text{ m}^{-3}$ of either particles or antiparticles. Following that process, this random mix of matter and antimatter cells would have then undergone a subsequent annihilation so the final count of either particles or antiparticles density would be reduced even further.

The conclusion is that no scenario such as the standard model that begins with radiation will be able to explain the asymmetry.

We also feel that the field theory model has additional problems. For one, it is just too complicated. It is supposed that the request neutrons and protons were the result of a scenario in which radiation evolved into quarks and gluons and then into baryons and leptons all in a time period of less than 10^{-5} s. It wasn't until a time of 10^{-24} s, for example, that information could have traveled across the dimension of neutron which places severe limitations on any sort of cooperative interaction. Another problem with the quark plasma idea is that such a process would require three-body reactions which are notoriously slow. The strong force is short range so in this case, 3 relativistic quarks of the correct type would have had to simultaneously occupy a volume no larger than a neutron and with relative velocities small enough that a reaction could take place. With random distributions and velocities, such a condition is extremely unlikely so the rate of binding into hadrons would have been extremely small. There is also the problem of explaining how the required numbers of each quark type could have randomly formed out of the radiation with no quarks left over.

Scenario (b)

The second scenario assumes that the particles and radiation coalesced simultaneously directed out of the vacuum at a time at or near $t = t_n$. The asymmetry problem can be solved in this case but it suffers from the lack of a mechanism that could account for any particular mix of protons, neutrons, and photons necessary for the subsequent nucleosynthesis. In other words, it is too complicated to be correct.

Scenario (c)

In this case, it is assumed that particles coalesced out of the vacuum without any initial accompanying radiation. It further simplifies matters considerably if only a single particle type was created with the obvious candidates being neutrons and/or antineutrons.

Suppose for the moment that spacetime had the property that it could only form neutrons or antineutrons but not both. The asymmetry problem is then solved by fiat and it is also possible to account for the radiation as being the result of some initial kinetic energy of the neutrons being converted into photons during the early phase prior to nucleosynthesis proper via the reactions $np \rightarrow \gamma d$ followed by the breakup reaction $nd \rightarrow nnp$. The problem with that idea is that such a model implies the creation of far too much matter because in order to account for the energy density of the CMB, the number of particles would need to have been on the order of 1×10^{42} m^{-3} which is too large by a factor of 10^8.

A second option is that both neutrons and antineutrons were created in nearly equal numbers. In this case, the source of the CMB radiation was annihilation. Initially, each such photon would have had an energy equal to 939 MeV but these would have evolved into a thermal spectrum as a result of scattering off the

charged particles that soon came into existence. In order to account for the radiation energy density, the initial number of original particles must have been $n_m(t_n) = 1.6 \times 10^{41} \text{ m}^{-3}$ counting both neutrons and antineutrons. As far as the asymmetry problem goes, however, we have the same two possibilities we discussed under the first scenario. After the annihilation, in both cases, the final particle densities would have been vastly too small.

What we learn from all this is that no symmetric random process can account for the present-day particle density of matter so we conclude that the process that initiated the existence of matter must have been a *biased* random process and the only agent that could have been responsible for that is the vacuum. Going further, the action of this bias must have manifested during the creation process of the primary particles because all the subsequent reactions are matter/antimatter neutral.

Let us assume that in the creation process, the probability of creating a neutron is *p* and an antineutron is *q*. From the theory of a biased random walk (see e.g. [9]), the mean densities of neutrons and antineutrons created would then be $n_{\text{total}}(t_n)p$ and $n_{\text{total}}(t_n)q$ respectively where, in this case, $n_{\text{total}}(t_n) = 1.6 \times 10^{41} \text{ m}^{-3}$. After annihilation, the mean number of remaining neutrons (or antineutrons) would be $n_m(t) = n_{\text{total}}(t_n)(p-q)$ with a variance about this mean given by $\sigma = \sqrt{n_{\text{total}} pq}$. *p* and *q* are probabilities so we also have $p + q = 1$. Solving for the probabilities and assuming a present-day particle density of 2 m^{-3}, we find,

$$p = \frac{1}{2} + 2.4 \times 10^{-8} \tag{10-6a}$$

$$q = \frac{1}{2} - 2.4 \times 10^{-8} \tag{10-6b}$$

and

$$\sigma = 2 \times 10^{20}. \tag{10-6c}$$

To be clear about this, the bias must have been the same, or nearly the same, everywhere in order for the end result to have been either all matter or all antimatter rather than a mix. We also see from the very small size of the variance relative to the number of particles that all the cells would have finished up with the same number of particles. This is significant because nucleosynthesis proper is sensitive to the initial particle densities. W*hat we find is that a very small asymmetry in the "fabric" of the vacuum can account for the necessary matter/antimatter asymmetry* and further, there does not appear to be any other mechanism that can account for it. This is the first indication that the structure of the vacuum is far more complex than is generally thought.

Going back to the standard model, now that we know the magnitude of the bias, we can ask if there could there be such a bias in quantum field theory of scenario (a)? The answer to that is no because, although the bias is small, it is not so small that it would have escaped notice in present-day experiments. Further, in order for such a bias to manifest itself, the scenario would have to follow

along the lines of the "only neutrons" model. A small bias in the field theory cannot directly account for the very small particle/radiation ratio since it would require a huge bias to create nothing but matter. Thus, the initial radiation would have had to have first converted almost entirely into matter and antimatter with populations reflecting the small bias followed by subsequent annihilations that would have rebuilt the radiation and further, this small bias would have to be the same everywhere.
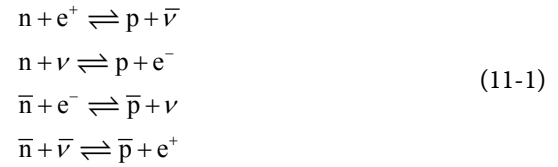
Having proposed that a slightly bias spacetime can account for both the present-day density of particles and the CMB, we next need to demonstrate that an all-neutron/antineutron beginning can account for the subsequent formation of the light elements.

11. Neutron Nucleosynthesis

In this section, we will examine a model of nucleosynthesis based on the idea that neutrons and antineutrons formed directly out of the vacuum energy of spacetime. Surprisingly, there is actually a hint that this idea has merit from the results of experiments conducted over the last 25 years that are attempting to nail down the lifetime of free neutrons. The article by Greene and Geltenbort, [10], provides a concise review of the situation. These experiments are of two types. One is known as the "Bottle" approach and the other as the "Beam" approach. The "Bottle" approach measures the lifetime by counting the number of neutrons remaining in a "Bottle" as a function of time. This approach makes no attempt to identify the decay products or even the mechanism of the decay. The "Beam" approach, on the other hand, counts the protons that result from the expected β decay of the neutrons. What is known as the neutron enigma is the fact that neutron lifetime measured by the "Bottle" approach (878.5 s) is a bit shorter than that measured by the "Beam" approach (887.7 s) which indicates that there is some as yet unknown decay path that allows roughly 1% of the neutrons to simply disappear without leaving behind a proton. Since a neutron cannot decay into any other baryon, it would seem that the decay violates the conservation of baryon number along with a few other conservation laws. But a violation of the conservation of baryon number is exactly what is needed to account for the bias that is needed to explain the matter/antimatter asymmetry.

In this new model, the particles and nucleosynthesis reactions are, of course, the same as those of the standard model but the initialization process was quite different. We also must recognize that the standard model seems to give a reasonable account of the final particle distributions. This means that the new model must account for a similar distribution of particles and radiation going into nucleosynthesis proper.

We start with neither radiation nor protons so the first problem is to account for their existence. We have already asserted that the radiation was the result of annihilation but we must also account for a significant number of protons. The solution lies in the neutrino reactions listed below.

$$n + e^+ \rightleftharpoons p + \bar{\nu}$$
$$n + \nu \rightleftharpoons p + e^-$$
$$\bar{n} + e^- \rightleftharpoons \bar{p} + \nu$$
$$\bar{n} + \bar{\nu} \rightleftharpoons \bar{p} + e^+$$

(11-1)

By assumption, we are starting with almost equal, very dense populations of both neutrons and antineutrons. Almost immediately, annihilation reactions would have begun creating very energetic photons. Simultaneously a few of the neutrons and antineutrons would have begun to decay initiating a cascade of neutrino reactions as shown below in **Figure 20** with a corresponding cascade beginning with the antineutrons. (Given the initial density of neutrons/antineutrons, the interparticle spacing was 4×10^{-23} s). Clearly, this process would have resulted in a large number of protons and antiprotons being created very rapidly. It is important to note that this cascade is completely dependent on the initial existence of both neutrons and antineutrons. The cascade would have continued until the density of protons became comparable to the density of neutrons at which point, the inverse reactions would have become significant. Eventually, an equilibrium would have been reached in which the inverse reactions were in balance with the forward reactions. At the same time, radiation was being created which would have brought the whole ensemble into thermal equilibrium via scattering. The final equilibrium densities of the neutrons and protons would then have been

$$\frac{n_n}{n_p} = e^{-\left(m_n c^2 - m_p c^2\right)/kT}$$

(11-2)

From here on out, nucleosynthesis would have proceeded along the lines described by the standard model. Because the starting conditions temperature and densities are different from those of the standard model and also because we wished study the importance of non-thermal particles, a numerical model has been developed to study this problem. The reaction equations are simple enough to write down. The basic rate equation for any particle can be written as

$$dN_i/dt = \sum R_i^{(+)} - \sum R_i^{(-)}$$

(11-3)

where the terms on the right are sums over the reaction rates that increase and decrease the count of particle "*i*" respectively. For two-body thermal reactions in which both particles have mass, the number of reactions per unit time in a volume *V* is
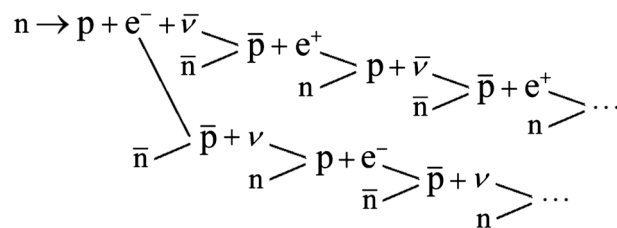
Figure 20. Neutron-neutrino interaction cascade.

$$R_{ij} = V\rho_i\rho_j I\left(\sigma_{ij}, T\right) \tag{11-4a}$$

$$I_{ij}\left(\sigma, T\right) = \sqrt{8/\pi\mu_{ij}\left(kT\right)^3} \int_0^\infty dE'\, E'\sigma_{ij}\left(E'\right)e^{-E'/kT} \tag{11-4b}$$

where $\rho_{i,j}$ are the reactant densities, $\sigma_{ij}\left(E\right)$ is the cross section and μ_{ij} is the reduced mass of the reactants. There is a similar formula for the case in which one of the reactants is a photon but we won't show it here because thermal photons play no direct role in nucleosynthesis once the deuteron bottleneck is passed. It is important to appreciate that this definition of the reaction rate concerns the number of reactions per unit time *in an expanding volume containing a fixed number of particles*. It is *not* the number of reactions per unit volume per unit time.

Starting with, for example, a cubic meter of spacetime at t_n, the numerical simulation tracks that expanding volume of particles so the number of baryons under consideration remains constant. This condition actually provides a sensitive test for errors in the simulation software.

Along with the standard thermal equilibrium model of nucleosynthesis, we were also interested in studying the importance of what we will denote as "fast" particles that acquire their energies from the various exothermic reactions (we are including the energetic photons in this designation). These particles are potentially important because they are continuously being produced during nucleosynthesis and thus, their initial energies do not decrease with time as a result of the expansion of the universe as do those of the thermal particles.

One of the nice features in the thermal case is that the reaction rate formula provides a clean separation between the lab and CM reference frames. For non-thermal particles, things are not quite so tidy. Our starting point is the usual reaction formulation with a Maxwell-Boltzmann distribution for the thermal particle and an unknown distribution function for the fast particle,

$$R_{Tf} = V\rho_T\rho_f \int_0^\infty d^3v_T \int_0^\infty d^3v_f\, \sigma_{Tf}\left(v\right)v\left(m_T/\left(2\pi kT\right)\right)^{3/2}\exp\left(-m_T v_T^2\right)F\left(v_f\right). \tag{11-5}$$

Here, v is the magnitude of the relative velocity of the reactants. The first simplification we make is to ignore reactions between two fast particles. The population of fast particles will generally be smaller than the density of thermal particles so this is a reasonable approximation especially considering the fact that very few of the reactions could involve two fast reactants. With this restriction, since the fast particle velocity will always be much larger than the thermal velocity, we can drop the thermal contribution to the relative velocity which allows the integrations to be separated. With a change of variable, the rate becomes

$$R_{Tf} = V\rho_T\rho_f c\sqrt{2/mc^2}\,\frac{\displaystyle\int_0^\infty dE'\,\sigma_{Tf}\left(E'\right)E'f\left(E'\right)}{\displaystyle\int_0^\infty dE'\,\sqrt{E'}f\left(E'\right)} \tag{11-6}$$

where we have made explicit the normalization of the fast particle distribution function.

The next step is to establish some sort of model for the fast particles. Obviously, we can't track the actual velocities of the particles so instead we divide the energy range of each type of particle into a number of bins and consider each bin to represent a single particle type that has a single fixed energy. This is equivalent to assigning to each type of fast particle. a distribution function which is constant within its bin and zero everywhere else. As nucleosynthesis proceeds, the numbers of each type of fast particle will change but their energies will not. With this approximation, we have

$$R_{Tf} = V \rho_T c \sqrt{2/mc^2} \sum_i \rho_f(E_i) \frac{\int\limits_{E_i}^{E_i+\Delta E} dE' \sigma_{Tf}(E') E' f(E')}{\int\limits_{E_i}^{E_i+\Delta E} dE' \sqrt{E'} f(E')} \tag{11-7}$$

where $\rho_f(E_i)$ is the number density of fast particle i. The rate equation for each particle type is then

$$R_{Ti} = V \rho_T \rho_i(\bar{E}_i) c \sqrt{2\bar{E}_i/mc^2} \bar{\sigma}_{Ti}(\bar{E}_i) \tag{11-8}$$

where $\bar{\sigma}_{Ti}(\bar{E}_i)$ is the bin-averaged cross section. For photons, a similar argument yields

$$R_{Ti} = V \rho_T \rho_i(\bar{E}_i) c \bar{\sigma}_{Ti}(\bar{E}_i). \tag{11-9}$$

The principal difficulty with this model that there is no clear separation between the lab and CM energies as there is when both particles have Maxwell-Boltzmann distributions. We will generally think of the nominal bin energies as CM energies and try to adjust to lab energies when possible but this cannot be easily done with any rigor because such a transformation for the energies of the outgoing particles would then be angle dependent which in turn would mean that we could not assign the outgoing particles to a single bin. The consequence of ignoring this issue is that the cross sections will be evaluated at energies that might differ by as much as a factor of 2 from the "correct" energy but since the cross sections vary slowly on a logarithmic scale and also since the definition of each particle is no better than the width of its bin, such an energy shift will not have a significant effect on the results. When a fast particle is one of the inputs to a reaction, we calculate the input CM energy of the reaction by assuming that the bin energy is the lab energy and using the normal kinematics based on the particle masses. For reactions with two output particles, we determine the output energies using the normal two-particle CM kinematics and then allocate each particle to the bin corresponding to its energy. With three output particles, we calculate the maximum possible energy that each particle could have and then allocate that particle to the bins assuming that each particle has a uniform spread of energy from its maximum value down to zero. Of course, when we speak of a particle, we are actually talking about a huge number of particles of

any given type.

As a practical matter, considering the Q values of the reactions and then allowing for the input kinetic energies of the fast particles and energetic photons, we found that fast neutron and proton energies would reach 18 MeV, that alpha particle energies would reach 10 MeV, and that photon energies would reach 35 MeV. The number of bins for each type is somewhat arbitrary. Enough are needed to give reasonable distributions but not so many as to create excessive numerical work or place too great a strain on our lab/CM energy blurring. We found after a few trials that 12 bins each for the neutrons, protons, and alpha particles and 16 bins for the photons seemed to be a reasonable compromise.

We set the low end of the fast particle energy range to be 0.3 MeV and decreed that any particle whose energy dropped below that value was henceforth a thermal particle. The results are not sensitive to the exact value as long as it is not zero. Finally, we tried two models for the bin widths. In one case we used a linear scale so the bins had equal energy widths and in the other case, we used a log scale so the bins had equal widths when plotted on any of the log-log cross section plots. Trials showed that the results were not particularly sensitive to the choice but since a logarithmic pattern better matches the cross section data, that was the option we chose to use.

Having dealt with the model, we will now turn to the cross section data. In Table 2 that follows, we list the reactions that were included in this model along with their Q values. Note that we have not included any particles with atomic numbers greater than 7. Although we attempted to locate as many cross sections as possible, in many cases, it was necessary to use reaction rate formulas (replacements for (11-4b)) directly. This was not a restriction for the thermal simulations but was a serious hinderance for the "fast" particle simulations because those require knowledge of the cross sections. References to the original sources of the rate formulas are generally given although in 3 cases, we were not able to access the original source so instead took the formula directly from the BBN code.

The ID in the first column is just a reference number that will allow us to refer to any particular reaction. The "Refs" column lists the references to the cross section and rate formula data. The CS and RF columns indicate whether or not we had cross section and/or rate formula data and the last column indicates whether or not the reaction is included in the standard BBN simulation. The reactions in which we had both cross section and rate formula data allowed us to verify our calculations of the reaction rates. The results are generally in good agreement although in some cases, we did find some differences in detail.

Because of the large reaction rates and the fact that number densities of the different particle types vary by many orders of magnitude, the equations are stiff and cannot be solved by using the standard Runge-Kutta methods. Instead, we used a predictor-corrector solver known as "Lsoda." This solver was developed over a period of time at the Lawrence Livermore Laboratory several decades ago.

Table 2. Model reactions.

ID	Reaction	Q (MeV)	Refs	CS	RF	BBN
	p Production					
1	$n \rightarrow p + e^- + \bar{\nu}$					y
	d Production					
2	$n + p \rightarrow d + \gamma$	2.2	[11] [12]	y	y	y
	^{3}H & ^{3}He Production					
3	$d + d \rightarrow {}^3H + p$	4.0	[12] [13] [14]	y	y	y
4	$d + d \rightarrow {}^3He + n$	3.3	[11] [13] [14]	y	y	y
5	$d + n \rightarrow {}^3H + \gamma$	6.3	[15] [16]	y	y	y
6	$d + p \rightarrow {}^3He + \gamma$	5.5	[13] [14] [17]	y	y	y
	^{4}He Production					
7	$d + d \rightarrow {}^4He + \gamma$	23.0	[13] [14]	y	y	
8	${}^3H + d \rightarrow {}^4He + n$	17.6	[11] [13] [14]	y	y	y
9	${}^3H + p \rightarrow {}^4He + \gamma$	19.8	[13] [14]	y	y	y
10	${}^3He + d \rightarrow {}^4He + p$	18.4	[11] [13] [14]	y	y	y
11	${}^3He + n \rightarrow {}^4He + \gamma$	20.6	[15]	y	y	y
12	${}^3H + {}^3H \rightarrow {}^4He + n + n$	11.3	[13] [14]	y	y	y
13	${}^3He + {}^3H \rightarrow {}^4He + d$	14.3	[13] [14]	y	y	y
14	${}^3He + {}^3H \rightarrow {}^4He + n + p$	12.1	[13] [14]	y	y	y
15	${}^3He + {}^3He \rightarrow {}^4He + p + p$	12.9	[13] [14] [18]	y	y	y
	Exchange Reactions					
16	${}^3He + n \rightarrow {}^3H + p$	0.8	[11] [14]	y	y	y
17	${}^3H + p \rightarrow {}^3He + n$	−0.8	[13] [14]	y	y	y
	Breakup Reactions					
18	$d + n \rightarrow n + n + p$	−2.2	[19]	y		
19	$d + p \rightarrow n + p + p$	−2.2	[13] [14]	y	y	
20	$d + \gamma \rightarrow n + p$	−2.2	[12] [20]	y	y	y
	^{6}Li, ^{7}Li, ^{7}Be Production					
21	${}^4He + {}^3H \rightarrow {}^7Li + \gamma$	1.6	[11] [13] [14]	y	y	y
22	${}^4He + {}^3He \rightarrow {}^7Be + \gamma$	1.6	[11] [13] [14]	y	y	y
23	${}^4He + d \rightarrow {}^6Li + \gamma$	0.7	[14]		y	y
24	${}^3He + {}^3H \rightarrow {}^6Li + \gamma$	15.0	[21]		y	y
25	${}^4He + {}^3H \rightarrow {}^6Li + n$	−4.8	[14]		y	
	^{6}Li, ^{7}Li, ^{7}Be Exchange					
26	${}^6Li + n \rightarrow {}^7Li + \gamma$	7.2	[22]		y	y

Continued

|							
27	$^6\text{Li} + \text{p} \rightarrow {}^7\text{Be} + \gamma$	6.4	[14]			y	y
28	$^6\text{Li} + \text{d} \rightarrow {}^7\text{Be} + \text{n}$	4.2	[22]			y	y
29	$^6\text{Li} + \text{d} \rightarrow {}^7\text{Li} + \text{p}$	5.0	[22]			y	y
30	$^6\text{Li} + {}^3\text{H} \rightarrow {}^7\text{Be} + \text{n} + \text{n}$	−2.1	[13]	y			
31	$^6\text{Li} + {}^3\text{H} \rightarrow {}^7\text{Li} + \text{d}$	1.0	[13]	y			
32	$^6\text{Li} + {}^3\text{He} \rightarrow {}^7\text{Be} + \text{d}$	0.9	[13]	y			
33	$^7\text{Li} + \text{p} \rightarrow {}^7\text{Be} + \text{n}$	−0.8	[14]	y	y		
34	$^7\text{Li} + {}^3\text{He} \rightarrow {}^6\text{Li} + {}^4\text{He}$	13.4	[13]	y			
35	$^7\text{Li} + {}^3\text{He} \rightarrow {}^7\text{Be} + {}^3\text{H}$	−0.02	[13]	y			
36	$^7\text{Be} + \text{n} \rightarrow {}^7\text{Li} + \text{p}$	0.8	[15] [23]	y	y		y
	^{6}Li,^{7}Li,^{7}Be Breakup						
37	$^6\text{Li} + \text{n} \rightarrow {}^4\text{He} + {}^3\text{H}$	5.6	[14]			y	y
38	$^6\text{Li} + \text{p} \rightarrow {}^4\text{He} + {}^3\text{He}$	4.8	[14]			y	y
39	$^6\text{Li} + \text{d} \rightarrow {}^4\text{He} + {}^4\text{He}$	23.2	[13]	y			
40	$^7\text{Li} + \text{p} \rightarrow {}^4\text{He} + {}^4\text{He}$	18.2	[11] [13] [14]	y	y		y
41	$^7\text{Li} + \text{p} \rightarrow {}^4\text{He} + {}^4\text{He} + \gamma$	17.3	[14]	y			
42	$^7\text{Li} + \text{d} \rightarrow {}^4\text{He} + {}^4\text{He} + \text{n}$	16.0	[14]			y	y
43	$^7\text{Be} + \text{d} \rightarrow {}^4\text{He} + {}^4\text{He} + \text{p}$	16.8	[14]			y	y
44	$^7\text{Be} + \text{n} \rightarrow {}^4\text{He} + {}^4\text{He}$	19.0	[24]			y	y

It was originally written in Fortran but later was ported to the "C" language and both of these versions can be found on the internet. For our purposes, we ported it again to the Microsoft VB.Net platform. This solver has a number of essential features. It automatically switches between Adams-Bashford and Gear Stiff equation methods and automatically adjusts the step size and method order at each step. Each type of particle requires an equation so with the bin choices discussed earlier, we end up with 60 simultaneous equations when the fast particles are included. There are no equations reflecting a dependence of the scaling on the radiation or particle densities because, unlike the standard model, the scaling is entirely determined by the vacuum energy density.

The critical reactions that regulate the initiation of nucleosynthesis proper are reactions 2 and 20 and the process could not begin until the reaction rates for the two become approximately equal. The cutoff for the breakup reaction is at an energy of 2.2 MeV. Equating this energy to kT gives a time of $t_2 = 1.2 \times 10^{-2}$ s but because of the very small particle/photon ratio, the actual beginning of nucleosynthesis occurs somewhat later. Once the thermal photons dropped below this cutoff, they ceased to have any effect on nucleosynthesis.

In **Figure 21**, we show the results obtained with thermal particles only and with a present-day particle density of $n_{part}(t_0) = 2 \text{ m}^{-3}$. All the reactions in the
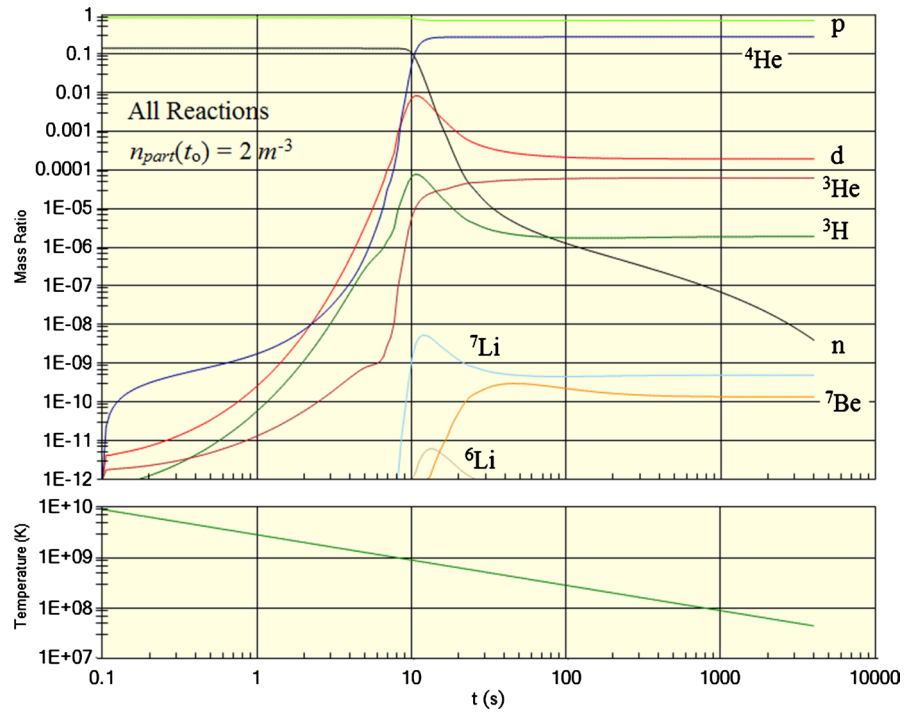
Figure 21. Thermal nucleosynthesis, $n_{part}(t_0) = 2\,\text{m}^{-3}$ with all reactions included.
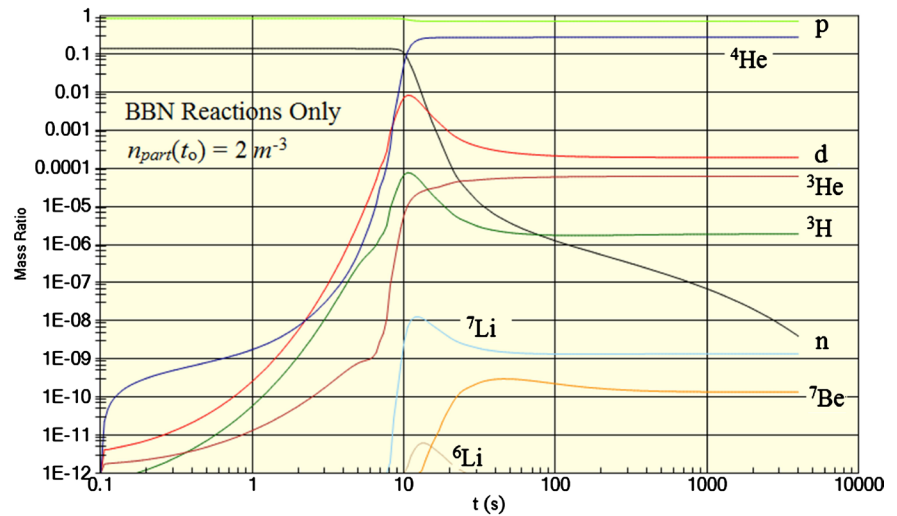
Figure 22. Thermal nucleosynthesis, $n_{part}(t_0) = 2\,\text{m}^{-3}$ with only the BBN reactions included.

table are included. Relative to the standard model, we see that the starting time is earlier by a factor of about 25 and the duration is compressed by about the same factor as well. Nevertheless, the end results are much the same.

In the next figure, **Figure 22**, we show the results obtained by limiting the reactions to those included in the BBN simulation model. All the other parameters are the same. Comparing we see that the results are the same with one exception, namely that the BBN results indicate a significantly larger density of ^{7}Li. *The ratio is* 2.8 *which agrees with the known disparity between the BBN results*

and observation. *This is a strong indication that the so-called lithium problem is simply a matter of not including a number of known lithium reactions in the simulation.*

We ran simulations for a range of values of the present-day particle density and the best results seem to be obtained with the density in the range, $n_{part}(t_0) = 2 - 3\,\mathrm{m}^{-3}$. **Figure 23** shows the results for $n_{part}(t_0) = 3\,\mathrm{m}^{-3}$.

In **Figure 24**, we show the results for nucleosynthesis in the voids. The average particle density is much lower than in the material regions but it is not zero. Using a present-day density of 0.016 m⁻³, we find that in the voids, the protons make up essentially all of the total with the percentage of ^{4}He is less by a factor of about 10 relative to the higher density results. The fractions of the other particle types are generally somewhat larger although still very small.
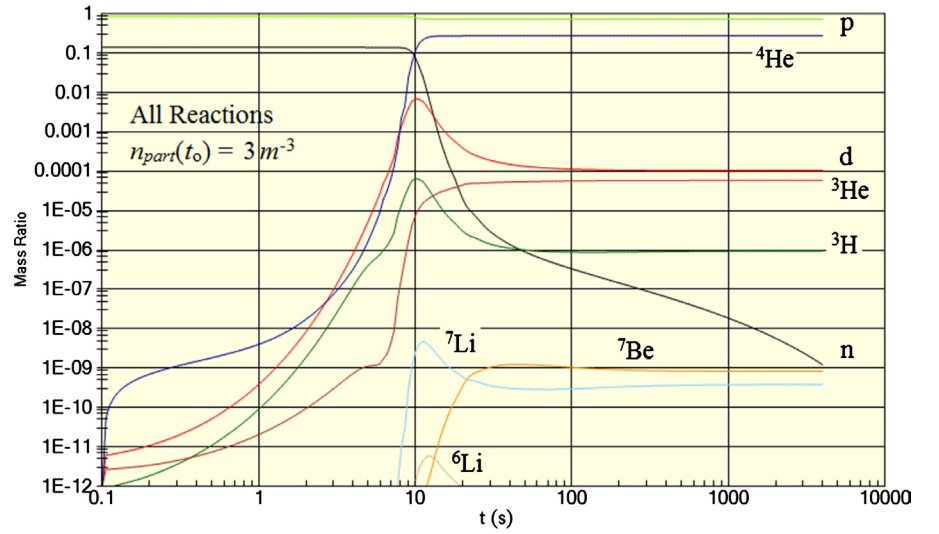
**Figure 23.** Thermal nucleosynthesis, $n_{part}(t_0) = 3\,m^{-3}$ with all reactions included.
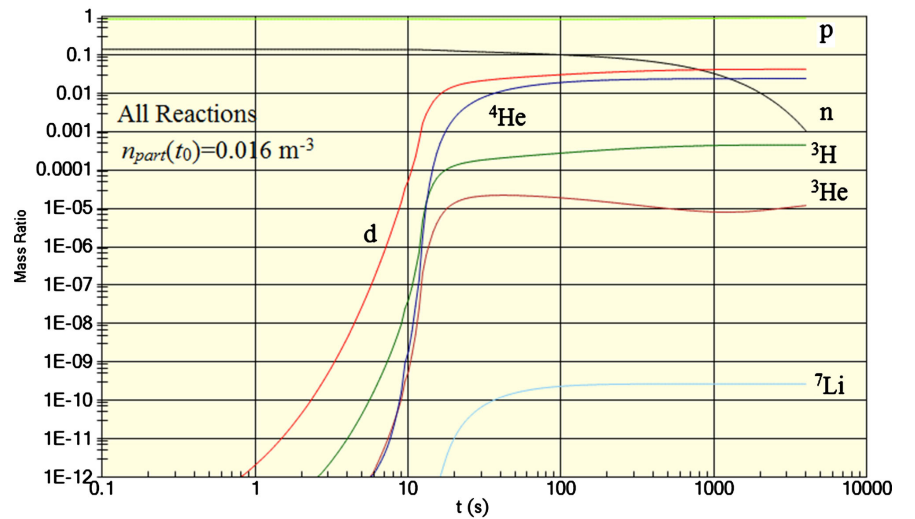
Figure 24. Void thermal nucleosynthesis, $n_{part}(t_0) = 0.016\,\mathrm{m}^{-3}$ with all reactions included.

We will now turn to the problem of the "fast" particles. We originally developed the "fast" particle model to study an "Only neutron" model (no antineutrons). Starting with only neutrons, it is not possible to get anything like reasonable results with just thermal particles. By including the "fast" particles, on the other hand, it is possible to get final particle densities something like the observed values. The problem with this model is that it is impossible to account for the CMB without at the same time ending up with a final total particle density vastly too large.

There is no doubt about the creation of such particles so the primary question is how fast do the "fast" particles thermalize? We won't be able to get anything like definitive results in large part because we are lacking the necessary cross section data. For many reactions, we don't have such data and for others, large extrapolations were necessary. Also, with "fast" particles, the inverse of many of the forward reactions become significant.

Nevertheless, we show the outcome in **Figure 25** with first, no "fast" particle attenuation and second with 90% attenuation. In the latter case, the results are
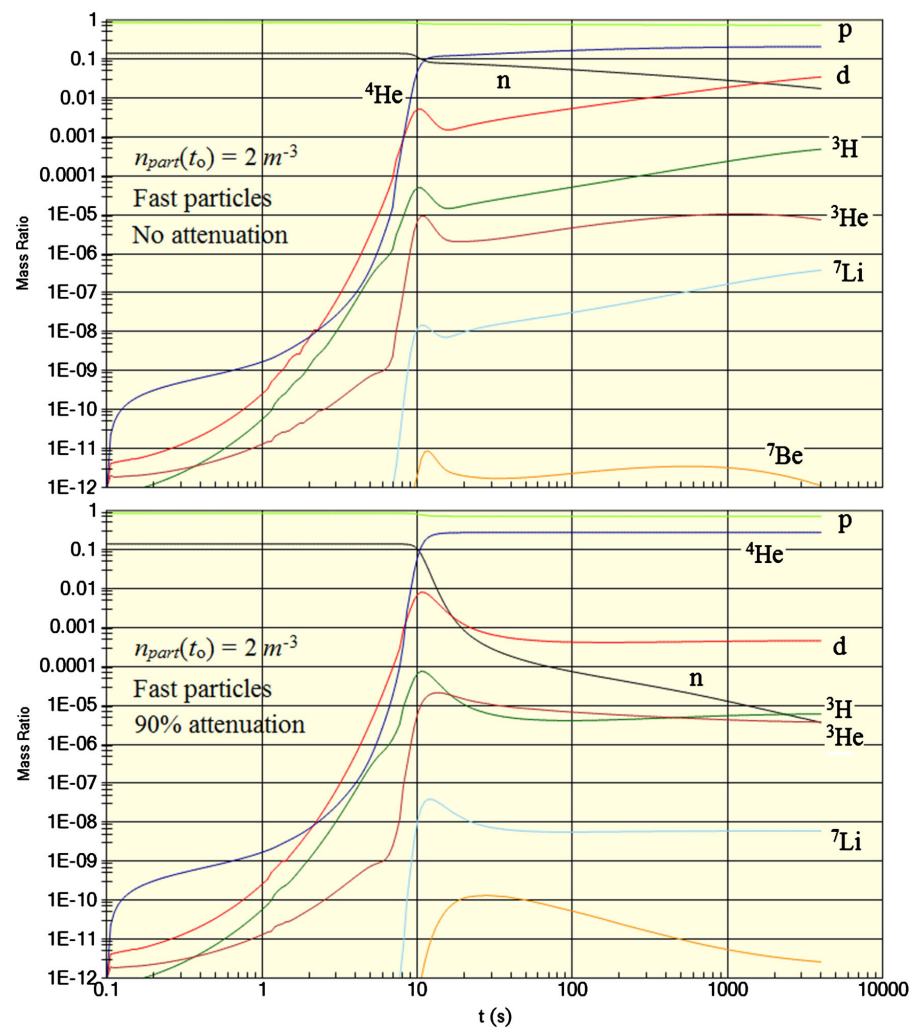
Figure 25. Fast particle nucleosynthesis.

starting to look something like the standard model but there are still significant differences.

What these results do indicate is that the existence of the "fast" particles could have a significant effect on nucleosynthesis. The fact that these results don't agree with the standard model seems to indicates that thermalization is, in fact, very rapid but why that is true is not so obvious. First, a high percentage of the protons, neutrons, and [4]He are "fast" so scattering between these would not lead to rapid thermalization and second, there is a significant density of energetic photons also retarding the thermalization process. Our purpose in showing these results is to indicate that "fast" particles should be considered and that their importance should not be dismissed out of hand. A more careful study of the thermalization process would be needed to settle the question.

A final point concerning nucleosynthesis is that the initial particle density was not the same everywhere so the observed mass ratios are the result of an ensemble average over a spectrum of initial densities which should be incorporated into the model.

12. Solution Revisited

We made a point of saying during the development of Sec. 8 that the solution was correct but incomplete. The issue is the mass density of ordinary matter. We established that the present-day particle density is $n_m(t_0) = 2\,\text{m}^{-3}$ which corresponds to a mass density of $\rho_m c^2(t_0) = 3.0 \times 10^{-10}\,\text{J} \cdot \text{m}^{-3}$ while at the same time, the total vacuum energy density (energy plus pressure) is $\rho_{vac} c^2(t_0) = 2.1 \times 10^{-10}\,\text{J} \cdot \text{m}^{-3}$. The particle energy density is apparently larger than the vacuum energy density which indicates that it cannot simply be ignored.

To include this contribution in the equations, we need to add the particle density to (8-7) which then becomes

$$\mathsf{T}^{\mu\nu} = \left(\rho_{vac} c^2(ct,r) + \rho_m c^2(ct,r) + p(ct,r) \right) \delta_0^\mu \delta_0^\nu + p(ct,r) \mathsf{g}^{\mu\nu} \quad (12\text{-}1)$$

The next step would be to solve the resulting equations but if we think about the solution given earlier, the physical quantities such as the scaling, the curvature, and the motion of test particles are functions of just the *sum of the energy density and pressure* and hence, adding the particle contribution to the sum does not change the solution for the physical quantities since the *sum* is fixed by Einstein's equations.

It is only when we set about separating the contributions the energies and pressure that the contribution of the particle energy becomes apparent. The pressure remains unchanged but what we previously called the vacuum energy density at any point is in reality, the sum of the actual vacuum energy density (plus the pressure) and the particle mass energy density.

Calculating this separation is easy because we know that the particle number density varies according to $a(t)^{-3}$. The result is shown in **Figure 26** which now replaces **Figure 5**. The "Total" and the pressure curves are unchanged but we see that in order to accommodate the particle mass density, the vacuum energy density
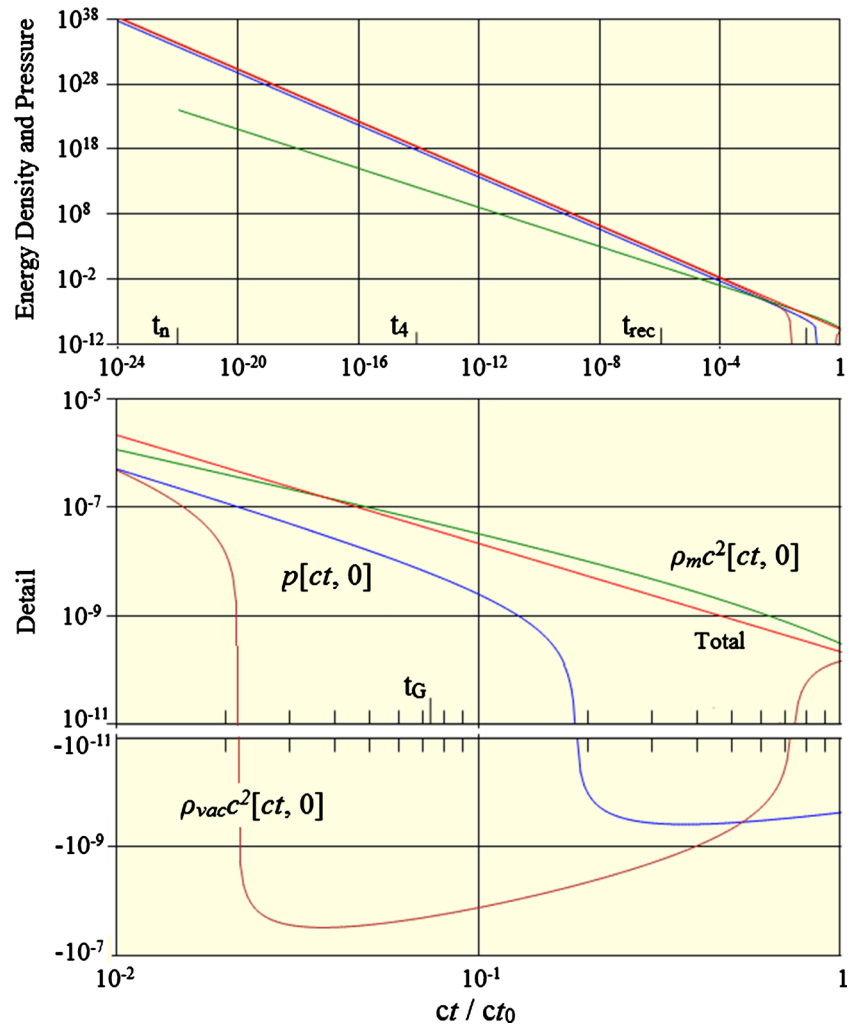
Figure 26. Revised energy densities and pressures. Present-day particle density $n_m(t_0) = 2\,\mathrm{m}^{-3}$.

actually becomes negative for a period of time although it does return to positive values just prior to the present time. This result is sensitive to the assumed particle density and for comparison, we next show the result for $n_m(t_0) = 1\,\mathrm{m}^{-3}$ in **Figure 27**.

We are now in a position to refute the idea that the formation of galaxies began with small particle density fluctuations, random or otherwise, in an otherwise uniform distribution that is generally assumed to have existed subsequent to nucleosynthesis. The facts are that, as just mentioned, the motion of particles is dependent on only the *sum* of energy densities and that sum is fixed by Einstein's equations and is independent of the scaling. Thus, any small variation in the particle density in some region will result in an immediate change in the vacuum energy density sufficient to keep the total constant. The result is that the particles in any region will each experience a uniform gravitational field regardless of any particle density variations and hence will not undergo any sort of accumulation. Thus, *the accretion model of galaxy formation initiated by small*
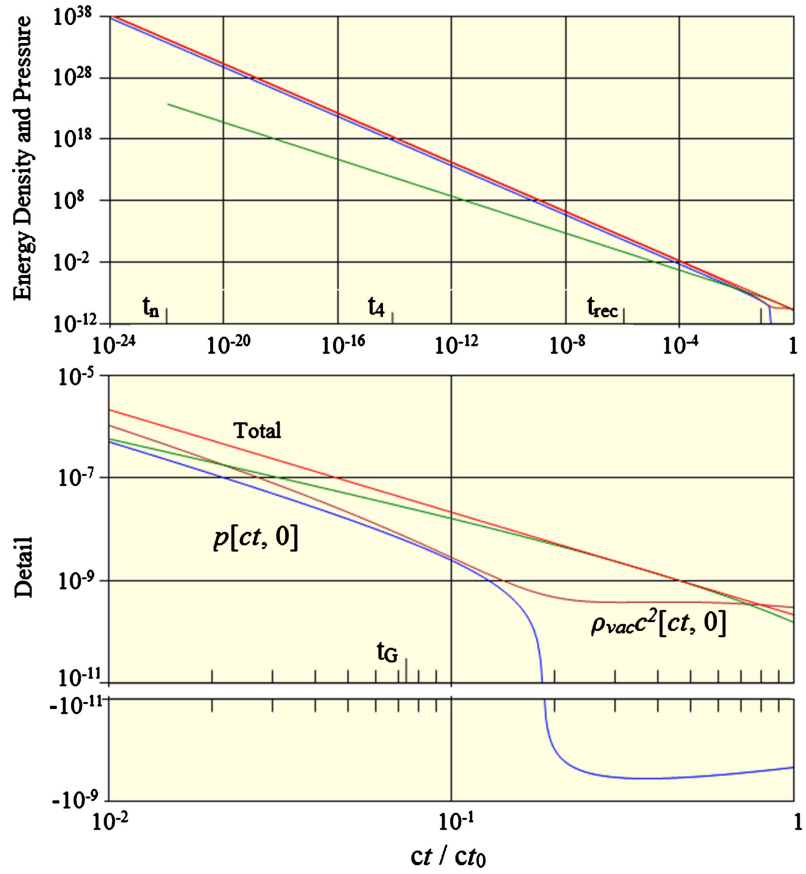
Figure 27. Revised energy densities and pressures. Present-day particle density $n_m(t_0) = 1\,\mathrm{m}^{-3}$.

matter density fluctuations is impossible.

Nevertheless, at some level accretion must have taken place but not nearly to the extent that is generally supposed. The accretion involved not just the particles, but the vacuum energy as well and the focal points of the accretion were the result of large-scale variances in the vacuum energy. We will have more to say about this later in Sec. 16 after an examination of the CMB spectrum.

13. Summary of Parameters

For the remainder of this development, it will be useful to have a summary of the various quantities we have been discussing. The scaling is given by (8-24)-(8-27) with the following parameters

$$
\begin{aligned}
a_0 &= 4.4 \times 10^{26}\ \mathrm{m} \\
t_0 &= 4.36 \times 10^{17}\ \mathrm{s} \\
\gamma_* &= 0.5 \\
k_0 &= 1.414 \\
c_1 &= 0.45 \\
\gamma_h &= 1/3
\end{aligned}
\tag{13-1}
$$

The scaling curves are shown in **Figure 3**.

Table 3. Contains a summary of the various quantities for a number of different times. Typical dimensions for galactic clusters and superclusters are given in the middle portion of the table and corresponding masses are presented in the lower portion.

	t (s)	a (m)	k	R_h (m)	ct (m)	D_{cell} (m)	T (K)
t_n	4.3×10^{-5}	2.8×10^{15}	5.8×10^{21}	3.7×10^{4}	1.3×10^{4}	3.7×10^{4}	4.2×10^{11}
	1	4.3×10^{17}	2.5×10^{17}	8.5×10^{8}	3.0×10^{8}	5.6×10^{6}	2.8×10^{9}
	4000	2.7×10^{19}	2.5×10^{13}	3.4×10^{12}	1.2×10^{12}	3.5×10^{8}	4.4×10^{7}
t_{rec}	4.9×10^{11}	3.0×10^{23}	5.2×10^{5}	4.1×10^{20}	1.5×10^{20}	3.9×10^{12}	4.0×10^{3}
t_G	3.2×10^{16}	7.9×10^{25}	8.4	2.7×10^{25}	9.6×10^{24}	1.0×10^{15}	15.1
t_0	4.4×10^{17}	4.4×10^{26}	1.4	3.7×10^{26}	1.3×10^{26}	5.8×10^{15}	2.71

		Energy Density (J·m^{-3})				Num Density (m^{-3})	
	t (s)	Vacuum + Particles	Particles thermal	Particles mass	Photons	Particles	Photons
t_n	4.3×10^{-5}	2.1×10^{34}	4.5×10^{22} 3.6×10^{20}	1.2×10^{24} 9.2×10^{21}	2.4×10^{31}	7.6×10^{33} 6.1×10^{31}	1.5×10^{42}
	1	4.0×10^{25}	8.5×10^{13} 6.8×10^{11}	3.3×10^{17} 2.7×10^{15}	4.6×10^{22}	2.2×10^{27} 1.8×10^{25}	4.4×10^{35}
	4000	2.5×10^{18}	5.3×10^{6} 4.3×10^{4}	1.3×10^{12} 1.1×10^{10}	2.9×10^{15}	8.7×10^{21} 7.0×10^{19}	1.8×10^{30}
t_{rec}	4.9×10^{11}	1.7×10^{2}	3.6×10^{-10} 2.9×10^{-12}	0.97 7.8×10^{-3}	0.20	6.5×10^{9} 5.2×10^{7}	1.3×10^{18}
t_G	3.2×10^{16}	3.9×10^{-8}		5.3×10^{-8} 4.2×10^{-10}	4.0×10^{-11}	3.5×10^{2} 2.8	7.0×10^{10}
t_0	4.4×10^{17}	2.1×10^{-10}		3.0×10^{-10} 2.4×10^{-12}	4.1×10^{-14}	2 1.6×10^{-2}	4.0×10^{8}

	t (s)	Globular Cluster (m)	Dwarf Galaxy (m)	Milky Way (m)	Spiral Galaxy (m)
t_n	4.3×10^{-5}	$(1.6 - 2.9) \times 10^{6}$	$(1.3 - 13) \times 10^{7}$	6.1×10^{9}	$(1.0 - 20) \times 10^{9}$
	1	$(2.4 - 4.4) \times 10^{8}$	$(2.0 - 20) \times 10^{9}$	9.3×10^{11}	$(1.5 - 30) \times 10^{11}$
	4000	$(1.5 - 2.8) \times 10^{10}$	$(1.2 - 12) \times 10^{11}$	5.8×10^{13}	$(0.9 - 19) \times 10^{13}$
t_{rec}	4.9×10^{11}	$(1.6 - 3.1) \times 10^{14}$	$(1.4 - 14) \times 10^{15}$	6.4×10^{17}	$(1.0 - 21) \times 10^{17}$
t_G	3.2×10^{16}	$(4.5 - 8.2) \times 10^{16}$	$(3.6 - 36) \times 10^{17}$	1.7×10^{20}	$(2.7 - 5.4) \times 10^{19}$
t_0	4.4×10^{17}	$(2.5 - 4.5) \times 10^{17}$	$(2 - 20) \times 10^{18}$	9.5×10^{20}	$(1.5 - 31) \times 10^{20}$

	t (s)	Group (m)	Cluster (m)	Supercluster (m)
t_n	4.3×10^{-5}	$(2 - 4) \times 10^{11}$	$(4 - 19) \times 10^{11}$	$(6 - 58) \times 10^{12}$
	1	$(3 - 6) \times 10^{13}$	$(6 - 29) \times 10^{13}$	$(1 - 9) \times 10^{15}$
	4000	$(2 - 4) \times 10^{15}$	$(4 - 18) \times 10^{15}$	$(6 - 55) \times 10^{16}$
t_{rec}	4.9×10^{11}	$(2 - 4) \times 10^{19}$	$(4 - 20) \times 10^{19}$	$(7 - 61) \times 10^{20}$
t_G	3.2×10^{16}	$(5 - 11) \times 10^{21}$	$(1 - 5) \times 10^{22}$	$(2 - 16) \times 10^{23}$
t_0	4.4×10^{17}	$(3 - 6) \times 10^{22}$	$(6 - 30) \times 10^{22}$	$(1 - 9) \times 10^{24}$

	Mass (kg)
"t_n" Cell	2.7×10^{21}
Globular Cluster	$(2 - 20) \times 10^{35}$
Dwarf galaxy	$(2 - 20) \times 10^{38}$
Milky Way	1.3×10^{41}
Spiral Galaxy	$2 \times 10^{39} - 2 \times 10^{42}$
Cluster	$(2 - 20) \times 10^{44}$
Supercluster	$(1 - 5) \times 10^{46}$

The pair of values under the "Particles" are the values corresponding to the dense and void regions (upper and lower) respectively.

The quantity $D_{\text{Cell}}(t)$ is defined by

$$D_{Cell}(t) = R_h(t_n)\left(\frac{a(t)}{a(t_n)}\right) \tag{13-2}$$

and is simply the size of cell, defined by the horizon distance at the time of neutron formation, scaled by the expansion of the universe. These cells, which we will call "t_n" cells, provide a convenient unit for handling various calculations involving sizes and masses.

The next topic we will consider is the nature of so-called dark matter.

14. Dark Matter

Dark matter was originally proposed to explain the motions of stars and galaxies which cannot be understood solely on the basis of the gravitational field induced by the visible matter. Since that time, dark matter has become something of catch-all for any cosmic phenomena that can't be otherwise explained. In this section, we will show that the vacuum energy we have been discussing can account for these motions thus obviating the need for dark matter as a separate material entity. In another guise, the belief that dark matter is responsible for the filament structure of the cosmos has become popular. Later in Sec. 16, we will show that again, it is vacuum energy that is responsible. We can sum things up with the following statement, dark matter $\leftrightarrow$ vacuum energy .

To make a beginning, we will consider the dynamics of spiral galaxies. In this manifestation of dark matter, the problem to be solved is the disparity between the observed velocity distribution of the stars (see, e.g. [25]) making up the galaxy and the motions calculated on the basis of the distribution of those stars. Figure 28 (adapted from [26]) illustrates the problem. Curve A is the velocity distribution calculated on the basis of the gravitational interactions of the visible matter and curve B is the observed rate. The generally accepted solution of this problem has been to postulate the existence of a halo of dark matter surrounding the galaxy with, in the case of the Milky Way, a total mass of about 5 times the mass of the galaxy's ordinary matter.

There are a number of problems with this model, however. First is the problem of explaining the dynamics of the dark matter halo since such a halo would act just like a halo of stars with the lights turned off so their velocity distribution should match the curve A instead of curve B. Another more general problem is that the dark matter model does not explain why there always seems to be a close association between dark and ordinary matter with the bulk of the dark matter hovering just outside the distribution of ordinary matter. Yet another problem is that the standard model makes no attempt to explain the origin of dark matter and it is completely ignored in the standard model's development of nucleosynthesis.

We get a hint towards the solution to this problem if we subtract the two curves yielding the curve C shown in **Figure 29**. This suggests that the observed velocity distribution can be understood in terms of normal gravitational motion being carried along by a *rotating spacetime*.

With this idea in mind, we will now turn to Einstein's equations. Given the distribution of matter and the motion of a spiral galaxy, it is reasonable to model such as galaxy with a stationary axisymmetric metric. The most general form is

$$
\begin{aligned}
\mathrm{d}s^2 &= -A(c\mathrm{d}t)^2 + B(\mathrm{d}\phi - \omega\mathrm{d}t)^2 + C\mathrm{d}r^2 + D\mathrm{d}\psi^2 \\
&= -\left(A - \frac{B\omega^2}{c^2}\right)(c\mathrm{d}t)^2 - 2\frac{B\omega}{c}\mathrm{d}\psi(c\mathrm{d}t) + B\mathrm{d}\phi^2 + C\mathrm{d}r^2 + D\mathrm{d}\psi^2
\end{aligned} \tag{14-1}
$$

with an energy-momentum tensor of the form
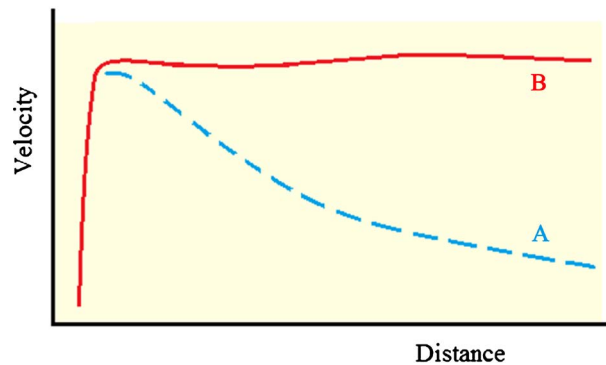
Figure 28. Typical galactic velocity distribution.
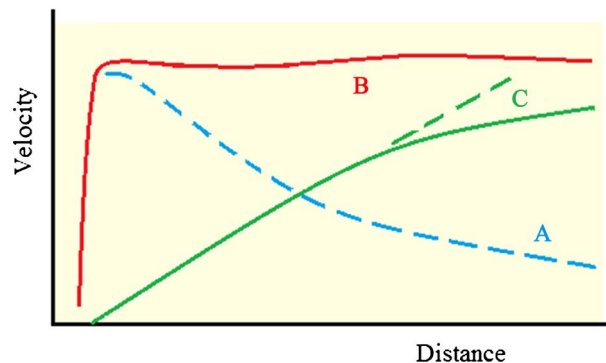
Figure 29. Sum of gravitational and spacetime rotations.

$$\mathsf{T}^{\mu\nu} = \left(\rho_{vac}c^2 + p_{vac}\right)\frac{u^\mu u^\nu}{c^2} + p_{vac}\mathsf{g}^{\mu\nu} + \rho_m c^2 \frac{v^\mu v^\nu}{c^2} \tag{14-2}$$

The arguments of all the metric functions, (r,ψ), have been suppressed for brevity. The angle ψ is defined as $\psi = \frac{\pi}{2} - \theta$ where θ is the usual spherical coordinates polar angle. (With this definition, the plane of the galaxy is defined by $\psi = 0$ which simplifies the specification of the boundary conditions). The vacuum quantities are denoted by the subscript "*vac*" and the matter by the subscript "*m*". Rather than attempt the general problem in which all the quantities are considered unknown, we will assume that the matter distribution is known leaving the unknowns to include the metric functions and the vacuum quantities.

For numbers, we will use the Milky Way as our example. The radius is $10^5/2 \, \text{ly} = 4.7 \times 10^{20} \, \text{m}$. The galaxy experiences differential rotation so there is no single angular velocity but a reasonable value for the period of rotation of the outer regions is about $3 \times 10^8 \, \text{yrs} = 9.5 \times 10^{15} \, \text{s}$ which yields an angular velocity of $\omega = 6.6 \times 10^{-16} \, \text{rad} \cdot \text{s}^{-1}$. The linear velocity at the outer edge of the galaxy is then $3.1 \times 10^5 \, \text{m} \cdot \text{s}^{-1}$ which is much less that c. Because of this, we can assume that coordinate time and proper time are the same and can approximate the 4-velocities as

$$u^\mu = \left(u^0, u^1, 0, 0\right) = \left(c, \dot\varphi_{vac}(r,\psi), 0, 0\right)$$
$$v^\mu = \left(v^0, v^1, 0, 0\right) = \left(c, \dot\varphi_m(r,\psi), 0, 0\right) \tag{14-3}$$

By observation, the particle velocity in the plane of the galaxy, $r\dot\varphi_m(r,0)$, is roughly constant away from the center, which is the whole point of this discussion, so $\dot\varphi_m(r,0) \propto r^{-1}$.

There are really two issues to be addressed. The first is to explain the rotation and the second is to account for the stability of the particle distribution given that rotation. Taking the rotation problem first, any small volume of vacuum energy will respond to the curvature of spacetime the same way as does a material particle. The geodetic equations for such a volume are,

$$\frac{du^0}{dt} = \Gamma^0_{00}u^0u^0 + 2\Gamma^0_{01}u^0u^1 + \Gamma^0_{11}u^1u^1 = 0$$

$$\frac{du^1}{dt} = \Gamma^1_{00}u^0u^0 + 2\Gamma^1_{01}u^0u^1 + \Gamma^1_{11}u^1u^1 = 0$$

$$\frac{du^2}{dt} = \Gamma^2_{00}u^0u^0 + 2\Gamma^0_{01}u^0u^1 + \Gamma^2_{11}u^1u^1 = 0$$

$$\frac{du^3}{dt} = \Gamma^3_{00}u^0u^0 + 2\Gamma^3_{01}u^0u^1 + \Gamma^3_{11}u^1u^1 = 0$$

$$\tag{14-4}$$

All the connection coefficients vanish in the first 2 of these equations so these equations just state that the velocity components are constant, which they must be given that we assumed a stationary metric. The LHS of the last two equations vanish because the velocity components are zero but the connection coefficients do not vanish so we have two equations that must be satisfied,

$$c^2 \left(u^0\right)^2 A^{(1,0)}\left[r,\psi\right] - \left(cu^1\left[r,\psi\right] - \omega\left[r,\psi\right]u^0\right)\left(cu^1\left[r,\psi\right]B^{(1,0)}\left[r,\psi\right]\right.$$
$$\left. -u^0\left(\omega\left[r,\psi\right]B^{(1,0)}\left[r,\psi\right] + 2B\left[r,\psi\right]\omega^{(1,0)}\left[r,\psi\right]\right)\right) = 0 \qquad (14\text{-}5a)$$

$$c^2 \left(u^0\right)^2 A^{(0,1)}\left[r,\psi\right] - \left(cu^1\left[r,\psi\right] - \omega\left[r,\psi\right]u^0\right)\left(cu^1\left[r,\psi\right]B^{(0,1)}\left[r,\psi\right]\right.$$
$$\left. -u^0\left(\omega\left[r,\psi\right]B^{(0,1)}\left[r,\psi\right] + 2B\left[r,\psi\right]\omega^{(0,1)}\left[r,\psi\right]\right)\right) = 0 \qquad (14\text{-}5b)$$

Since the angular velocities are very small, we expect these equations to be satisfied in the limit that ω and u^1 vanish. The consequence of that is that $A^{(1,0)}\left[r,\psi\right] = A^{(0,1)}\left[r,\psi\right] = 0$. After replacing u^0 and u^1, we now have

$$\left(\dot{\varphi}_{vac}\left[r,\psi\right] - \omega\left[r,\psi\right]\right)\left(\left(\dot{\varphi}_{vac}\left[r,\psi\right] - \omega\left[r,\psi\right]\right)cB^{(1,0)}\left[r,\psi\right]\right.$$
$$\left. + 2B\left[r,\psi\right]\omega^{(1,0)}\left[r,\psi\right]\right) = 0 \qquad (14\text{-}6a)$$

$$\left(\dot{\varphi}_{vac}\left[r,\psi\right] - \omega\left[r,\psi\right]\right)\left(\left(\dot{\varphi}_{vac}\left[r,\psi\right] - \omega\left[r,\psi\right]\right)cB^{(0,1)}\left[r,\psi\right]\right.$$
$$\left. + 2B\left[r,\psi\right]\omega^{(0,1)}\left[r,\psi\right]\right) = 0 \qquad (14\text{-}6b)$$

which are both satisfied if

$$\dot{\varphi}_{vac}\left[r,\psi\right] = \omega\left[r,\psi\right]. \qquad (14\text{-}7)$$

We find then that the vacuum energy is rotating as a result of inertial frame dragging. Actually, it would be more accurate to say the curvature is rotating but since all physical processes are a consequence of the curvature, it amounts to the same thing.

The geodetic equations for the particles will be exactly the same so the result will be,

$$\dot{\varphi}_m\left[r,\psi\right] = \omega\left[r,\psi\right]. \qquad (14\text{-}8)$$

and putting these results together, we find that *the curvature is differentially rotating and that the particles* (*stars or galaxies in the case of clusters*) *are at rest in that rotating curvature*. The original motivation for dark matter was to supply the mass thought to be needed to prevent the orbiting stars and galaxies from flying away from their hosts. From this new point of view, there is no issue of them flying away because the stars and galaxies are at rest.

We next calculate the norm of the 4-velocity which with (14-7) included, becomes $u^\mu u_\mu = -c^2 A\left[r,\psi\right] = -c^2$ so we find that $A\left[r,\psi\right] = 1$. The metric at this point is now

$$ds^2 = -\left(1 - \frac{B\omega^2}{c^2}\right)(cdt)^2 - 2\frac{B\omega(r,\psi)}{c}d\psi(cdt) + Bd\phi^2 + Cdr^2 + Dd\psi^2 \quad (14\text{-}9)$$

and the energy-momentum tensor is given by (14-2) with $u^0 = v^0 = c$, $u^1(r,\psi) = \omega(r,\psi)$ everywhere and $v^1(r,\psi) = \omega(r,\psi)$ everywhere that $\rho_m c^2(r,\psi) <> 0$. For large r, the vacuum energy density has its asymptotic value and $\omega(\infty,\psi) \to 0$.

After making those replacements, we find that the resulting Einstein equations

are dependent only on the *sum* of the vacuum and particle energy densities, $\rho_{\text{total}} c^2(r,\psi) = \rho_{vac} c^2(r,\psi) + \rho_m c^2(r,\psi)$. Asymptotically, this sum will be the vacuum energy we determined earlier which has a value of $O(10^{-10}) \, \text{J} \cdot \text{m}^{-3}$. The equivalent mass density of the galaxy, on the other hand, is of $O(10^{-3}) \, \text{J} \cdot \text{m}^{-3}$ and since the latter is much larger than the former, it is reasonable to set the boundary contribution in the interior of the galaxy to be

$\rho_{\text{total}} c^2(r,\psi) = \rho_m c^2(r,\psi)$ everywhere that matter exists. Away from the dense regions, the total energy density will be given by just the vacuum but this won't be the asymptotic vacuum because the equations will prevent the total energy density from dropping immediately to its asymptotic value. The fact that there is a halo of stars outside the galaxy proper will also contribute to the total energy density and help to prevent a rapid drop with increasing distance.

At this point, we would normally solve the equations with the necessary boundary conditions to determine, among other things, the vacuum energy density profile. Unfortunately, we have not been able to accomplish this task with the tools we have at hand. We are up against the same problem we ran into in Sec. 8, namely that although Mathematica does have the finite-element functionality needed to solve non-linear PDE boundary value problems, it can only do so for certain quasi-linear class of equations and these equations do not fall into that category. In this case, we are also limited by the huge amount of computer memory needed for the finite-element mesh. This being the case, in order to proceed, we were forced into the use of a more limited analysis to establish the stability.

To achieve this, we will examine the problem from the point of view of Newtonian forces in which the galaxy is assumed to be surrounded by a torus of vacuum energy as illustrated in **Figure 30** with the whole thing rotating with the curvature.

As shown in [27], to a good approximation the gravitational potential at a point on the galactic plane due to a circular torus is given by
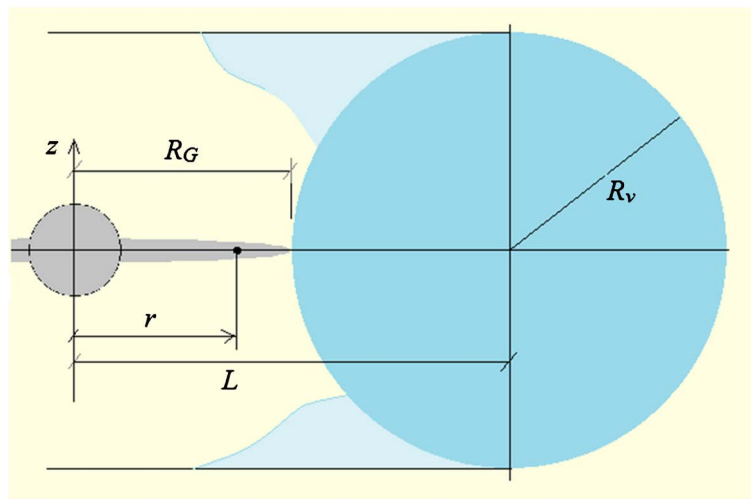
Figure 30. Coordinates. The dimensions shown correspond to a value of $R_v = R_G$.

$$\varphi_{\text{Torus}}(r) = -\frac{GM_v}{\pi L}\sqrt{\frac{L}{r}}m(r)K(m(r)) \tag{14-10}$$

where M_v is the total mass contained in the torus. The function $K(m)$ is the complete elliptical integral of the 1st kind,

$$K(m(r)) = \int_0^{\pi/2} \mathrm{d}\beta \Big/ \sqrt{1 - m(r)\sin^2(\beta)} \tag{14-11}$$

where

$$m(r) = \frac{4rL}{(r+L)^2}. \tag{14-12}$$

The force on a test particle at a distance r from the center due to the total equivalent mass of the torus is then

$$f_{\text{Torus}} = \frac{2GM_v}{\pi(r+L)}\left(\frac{\mathrm{d}K(m)}{\mathrm{d}m}\frac{\mathrm{d}m}{\mathrm{d}r} - \frac{K(m)}{r+L}\right). \tag{14-13}$$

Introducing two dimensionless parameters, $\xi = r/R_G$ and $\zeta = L/R_G$, these become

$$m(\xi,\zeta) = \frac{4\xi/\zeta}{(1+\xi/\zeta)^2}$$

$$R_v = R_G(\zeta - 1) \tag{14-14}$$

$$f_{\text{Torus}} = \frac{2GM_v}{\pi R_G^2}h_1(\xi,\zeta)$$

where

$$h_1(\xi,\zeta) = \frac{1}{(\xi+\zeta)^2}\left\{\frac{4(1-\xi/\zeta)}{(1+\xi/\zeta)^2}\frac{\mathrm{d}K(m)}{\mathrm{d}m} - K(m)\right\}. \tag{14-15}$$

The parameter ζ is not a measure of distance but instead defines the geometry. As ζ get larger, so does R_v and hence, so does the area of the torus. Finally, the equivalent mass of the torus is

$$M_v c^2 = \rho_{vac} c^2 2\pi^2 R_G^3 \zeta (\zeta - 1)^2 \tag{14-16}$$

Turning now to the galaxy, the disk accounts for most of the mass of the galaxy so for simplicity we will ignore the center bulge. From [28], the potential on the equatorial plane of a thin disk is given by

$$\varphi_{\text{Disk}}(r) = -2G\frac{M_G}{\pi R_G^2}\left\{(R_G + r)E(m_D) + (R_G - r)K(m_D)\right\}. \tag{14-17}$$

where

$$E(m_D(r)) = \int_0^{\pi/2} \mathrm{d}\beta\sqrt{1 - m_D(r)\sin^2(\beta)} \tag{14-18}$$

is the complete elliptical integral of the 2nd kind. In this case, $m_D(r) = 4rR_G\big/(r+R_G)^2 = m(\xi,1)$. Calculating the force on a test particle, we find

$$f_D = \frac{2GM_G}{\pi R_G^2} h_2(\xi) \tag{14-19}$$

with

$$h_2(\xi) = E(m_D) - K(m_D) + \frac{4(1-\xi)}{(1+\xi)^2}\left\{(1+\xi)\frac{dE(m_D)}{dm_D} + (1-\xi)\frac{dK(m_D)}{dm_D}\right\} \tag{14-20}$$

The function, $h_2(\xi)$ is negative reflecting the fact that the net force is towards the center of the galaxy. In order to have stability, the net force on a test particle at location r must vanish so

$$|f_D| = f_{\text{Torus}} \tag{14-21}$$

and from this, we obtain the following constraint on the vacuum energy density,

$$\rho_{vac}c^2 = \frac{|h_2(\xi)|}{2\pi^2\zeta(\zeta-1)^2 h_1(\xi,\zeta)}\left(\frac{M_G c^2}{R_G^3}\right) \text{J}\cdot\text{m}^{-3} \tag{14-22}$$

The result is the product of a dimensionless factor reflecting the geometry and a ratio that sets the magnitude of the energy density. The geometric factor is a positive number in the range of 1 - 10 and for the Milky Way, the magnitude factor is 5.2×10^{-6}. The results are shown in **Figure 31** for two values of ξ.

The short bars at the upper left and lower right indicate the energy densities of the disk and the asymptotic vacuum respectively. The horizontal axis does not represent a distance but instead defines the geometry. The red lines are the locus of values of the vacuum energy necessary to establish stability for the geometry indicated by the horizontal axis. For example, if the radius of the torus is $5R_G$ for $\xi = 0.5$, the necessary vacuum energy has a constant value of 1.6×10^{-5} J·m^{-3} everywhere in the interior of the torus. By assumption, the energy density is assumed to vanish outside the torus. We see is that the required value of the energy density is nearly constant for any value of the torus radius greater than 2 or so and that the curves for the two values of ξ are similar which lends support to the idea that a solution to the Einstein equations satisfying the required conditions exists. It would be a problem, for example, if the two curves were radically different.

To put these geometries in perspective, in order to account for mass of the Milky Way in terms of an accretion out of a uniform background of particles,
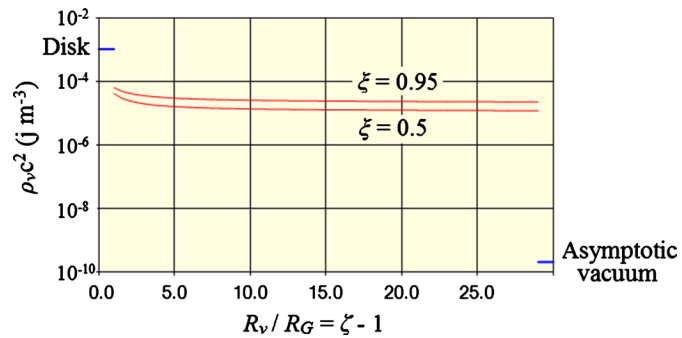
Figure 31. Solution of Equation (14-22) for two values of ξ.

starting with a density corresponding to a present-day density of 2 m^{-3} and leaving behind a residue equivalent to 1 m^{-3}, a spherical volume with a radius on the order of $R = 57R_G$ would have had to have been swept up and compared to this value, a torus radius of 2 - 5 R_G is quite small. We will show later, however, that accretion was not the primary mechanism by which galaxies were created. What is most important, however, is that the required vacuum energy density is *only about* 1% *of the equivalent energy density of the galactic matter.*

To help clarify this picture, in **Figure 32**, we show a hypothetical radial distribution of the total energy density for a torus radius of $2R_G$. This curve would be part of the solution of the Einstein equations were we able to solve them. The horizontal axis in this case is the actual distance from the center of the galaxy.

The blue line represents the equivalent torus mass energy density using in the calculation and the curved red line is a hypothetical to illustrate the idea of a smooth decay.

What we have shown is that the necessary stability can easily be obtained and thus a rotating curvature can readily account for the velocity distribution of spiral galaxies. This solution also explains why so-called dark matter always hovers just outside regions containing matter. Vacuum energy exists everywhere but its density is not uniform as we have explained because it is subject to accretion just as is ordinary matter.

Turning now to galaxy clusters where the idea of dark matter actually originated, to the extent that such a rotating cluster can be treated is a rotating disk, we can apply the same formalism. The only parameter in the model is the mass ratio and from **Table 3**, we find that for galaxy clusters, this ratio is of $O\left(10^{-9}\right)$ so we find that the required vacuum energy density is much smaller than in the spiral galaxy case and, in fact, is not significantly different from its asymptotic value. The fact that the required energy density is very small also allows abundant room for an adjustment of the geometric factor away from the thin disk model without the conclusion being affected.

We find then that the vacuum energy density can easily account for the observed rotation of galaxies and their contained stars and of galaxy clusters and their contained galaxies.

Dark matter is vacuum energy. Dark matter as a separate material *entity does*
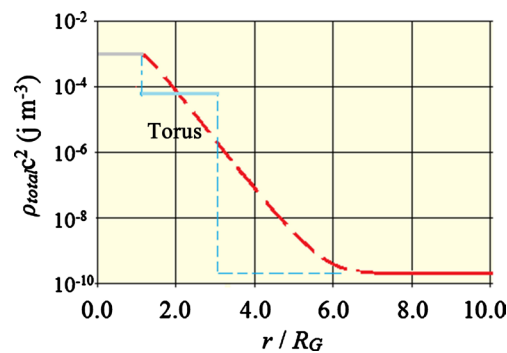
Figure 32. Radial dependence of vacuum energy density.

not exist.

15. CMB Spectrum

We have already discussed the origin of the CMB but didn't touch on its spectrum. In this section, we will show that the prominent features of the spectrum for angular sizes greater than 0.1˚ are a consequence of both the existence of superclusters, voids, and even larger structures on the one hand, and the energy uncertainty of the original Plank-sized regions at the end of the initial inflation on the other.

In Figure 33, we show the angular distribution of the CMB anisotropies from [29]. In the lower portion of the figure, we have enlarged a section of the distribution and added an 2˚ circle that gives a reference for the size of physical structures contributing to the spectrum.

For angular dimensions of 2˚ or less, the apparent features are consequences of physical structures. In the range between 2˚ and 45˚, the spectrum does not appear to be associated with any structure but is instead the consequence of the random, scale-invariant variance of the vacuum energy density which was set at the time of the initial inflation. We will refer to this as the Plank variance. The features with sizes of 45˚ and larger appear again to be related to actual structures.

In Figure 34 from [30], we see that the power spectrum consists of a flat region for angles between 6˚ and 45˚ of arc, a large peak centered at about 1˚ of arc and then a series of lower peaks extending to smaller angles. There is also a hint of a low peak beginning at 45˚ and extending to larger angles but the error bars are large.
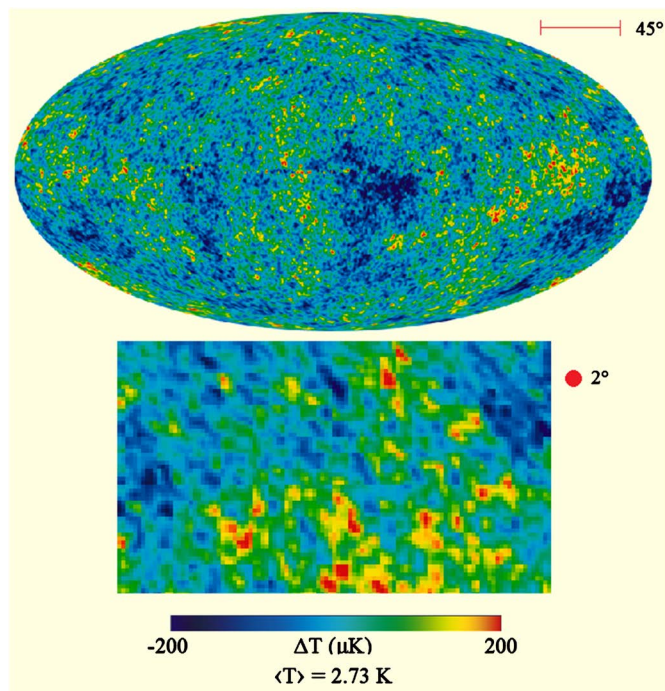
Figure 33. CMB anisotropy.
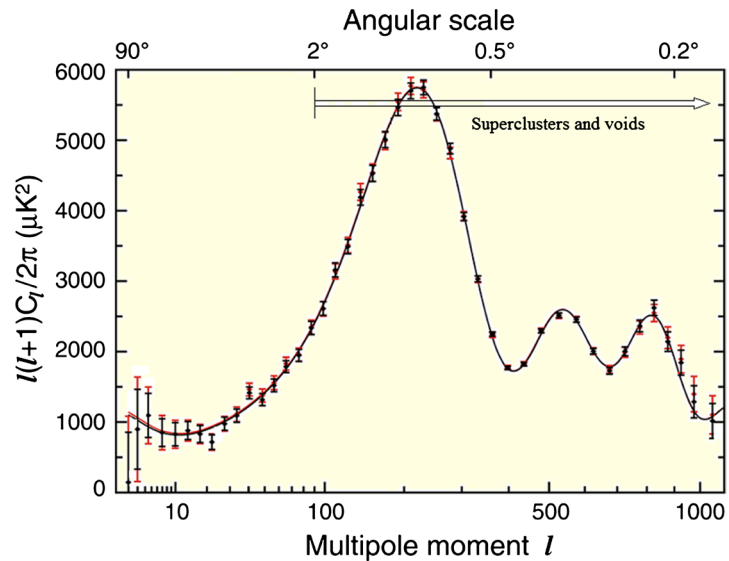
Figure 34. The power spectrum of the CMB anisotropy. Note that neither the upper nor lower scale is actually logarithmic. Angles are related to the moment by $l = \pi/\theta_{\mathrm{rad}} = 180/\theta_{\mathrm{deg}}$.

The magnitude of the spectrum sets the relative temperature variance to be $\delta T/T = O\left(10^{-5}\right)$ all across the spectrum. In fact, because the spectrum is a proportional to the square of the temperature variance, the difference between the variance at the peak and that of the large angle portion of the spectrum is less than a factor of 2.5.

The peaks are strongly suggestive of physical structures so in order to understand these peaks it will be necessary to establish the connection between the size of such structures and the angular size of the resultant anisotropies. Recombination took place everywhere and the CMB radiation fills all space so it might not be immediately obvious what the interpretation of the angular distribution of the CMB might be. The answer comes from simple geometry and is much simpler than is sometimes suggested in the literature. A discussion of the rather over-complicated FRW viewpoint is given in [31]. The fact is that we are observing light today that was emitted at time t_{rec} by a spherical shell of spacetime centered at our location. If we could travel back in time to t_{rec}, the universe would get progressively smaller but the angular position of all sources would remain unchanged. To fix the angular size of any particular structure, then, we only need to know *at* $t = t_{rec}$, our distance to the shell of sources and their size. For the first, we use the results shown in **Figure 7** which gives us the radial coordinate of a source whose light we are receiving at the present time. We see that at t_{rec} the coordinate distance was about 0.6 and since the universe is at rest, this value doesn't change with time. The proper distance from our vantage point at $t = t_{rec}$ to the source would then have been

$$S\left(t_{rec}\right) = 0.6a\left(t_{rec}\right). \tag{15-1}$$

Simple geometry then tells us that, for a structure of size $D\left(t_{rec}\right)$, the subtended angle would be

$$\theta = \frac{D(t_{rec})}{S(t_{rec})}(360/2\pi). \tag{15-2}$$

But the size of the structure varies with the scaling in a known way so we have

$$\theta = \frac{D(t_0)}{0.6a(t_{rec})}\frac{a(t_{rec})}{a(t_0)}(360/2\pi) = 95.5\frac{D(t_0)}{a(t_0)}\deg. \tag{15-3}$$

As we travel back to the present, the sources get further and further away because of the expansion while their light travels towards us along paths of constant angle until eventually, we and the light arrive at our present location at the same moment.

We will now consider the present-day size of actual structures. Table 3 lists typical dimensions and in Table 4, the corresponding angular sizes are given. Groups and clusters are roughly spherical in shape so their angular size will be representative of their influence on the CMB spectrum. Superclusters, on the other hand, are not spherical so the effective angular size of any particular structure will depend on its orientation relative to the line of sight to the earth. On the other hand, there are a lot of superclusters so the orientations should tend to average out.

From the table, we see that galaxies and even clusters are far too small to have any impact on the spectrum within the displayed range of angles. Superclusters and voids, on the other hand, are large enough to account for the peaks and in fact, these are the only known structures that *are* large enough. We also see from the expanded portion of Figure 33 that the individual, well-defined structures are comparable in size to the largest superclusters which reinforces the same idea.

Of course, not even stars existed at the point in time that the spectrum was fixed so the structures we are speaking of are not their present-day manifestations. Instead, what we are detecting are *precursor imprints in the vacuum that later developed into the present-day structures*. In the next section, we will develop this idea further.

We will soon show that superclusters and voids do indeed provide a convincing explanation for the peaks in the spectrum but we should mention that there exists a commonly believed alternative which supposes that the peaks are the result of acoustic oscillations of the densities of photons and protons. In order for this to have happened, however, regions of space as large as superclusters would have had to repeatedly pass signals back and forth. A review of Table 3, on the other hand, shows that even the smallest supercluster was 5 times larger than any possible signal distance at that time so the largest angular-sized anisotropy that such a mechanism could account for would be no larger than a cluster and probably considerably smaller. *The conclusion is that acoustic oscillations on the scale required to explain the first peaks were not possible.*

The 2nd and 3rd peaks have roughly a harmonic distribution relative to the first peak which suggests that they are reflection of multipole distributions of tem-

perature variances within the superclusters and voids since even the 3rd peak represents a size still much larger than the largest cluster. These peaks provide evidence that the temperature is nearly uniform over the expanse of the superclusters since if it wasn't, these secondary peaks would be much larger. We can also see this in the expanded portion of Figure 33 where a significant fraction of the 1° - 2° sized structures appear to have a single temperature.

The same dimensional arguments apply to the voids with the only difference being that they are cooler than the average rather than warmer. They contribute in the same way to the anisotropy, however, because the spectrum is proportional to the square of the temperature variance.

For our next argument, we need an estimate of the number of superclusters contributing to the CMB and this we can obtain in two ways. First, we can compute the average density of superclusters/voids based on an estimate of their total number. Then with that density, we can compute the number of superclusters/voids in the CMB shell by multiplying the area of the shell, the density, and the thickness of the shell. For the latter, we can assume that the size of a supercluster is a reasonable value. Allowing for the fact that these exist in the 20% of the total volume of the universe that contains most of the matter, we have

$$N_{\text{CMB}} = \frac{4\pi S\left(t_{rec}\right)^2 D_{sc}\left(t_{rec}\right)}{0.2\left(4\pi/3\right)a\left(t_{rec}\right)^3} \times 10^7 = 5.4 \times 10^5. \qquad (15\text{-}4)$$

where we have used a common estimate that there are around 10^7 superclusters/voids. A second method simply divides the area of the shell by the area of a supercluster. This gives

$$N_{\text{CMB}} = \frac{4\pi S\left(t_{rec}\right)^2}{\pi R_{sc}\left(t_{rec}\right)^2} = 5.8 \times 10^4. \qquad (15\text{-}5)$$

The latter method seems less likely to be in error so the difference between these values suggests that the total number of superclusters/voids is closer to 10^6 than to 10^7.

Using this number, we ran a number of simulations to determine how the CMB would appear if the temperatures of the superclusters were random. The results are shown in Figure 35. Each rectangle contains 10^4 superclusters. In the

Table 4. Angular sizes of various structures.

Object	$\theta\left(\deg\right)$
Milky Way	0.0001
Groups	0.007 - 0.013
Clusters	0.013 - 0.065
Superclusters	0.2 - 2.0
voids	0.6 - 1.6
Extreme structures	>45
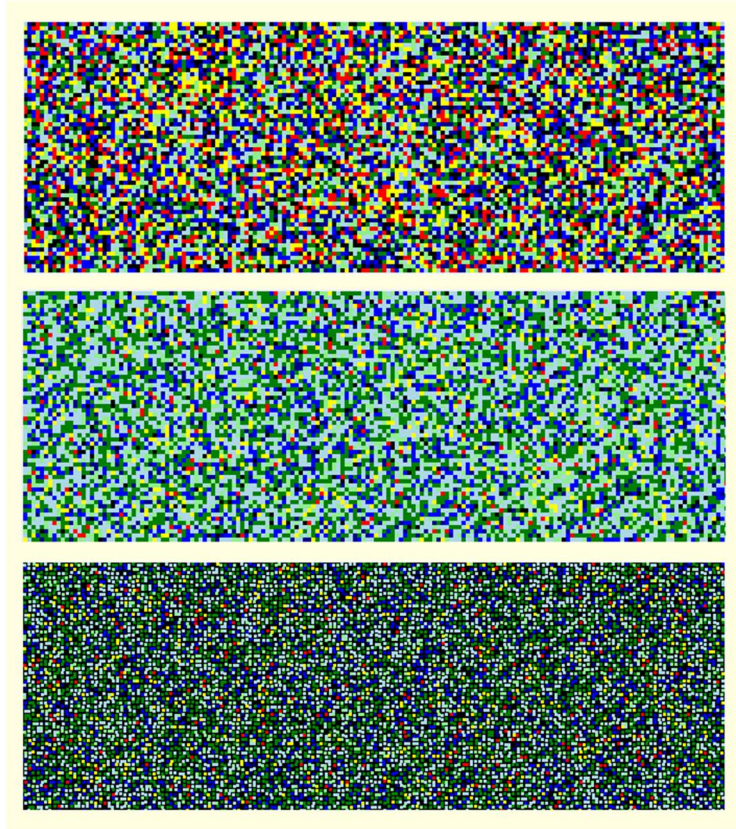
Figure 35. Random distributions of temperatures.

first, the temperatures were selected at random with no spacing between the superclusters. In the second, the temperatures were heavily biased towards the blue and green again with no spacing and in the third, a random spacing between superclusters was introduced equal to $1/4^{th}$ of the size of the supercluster with the resulting voids filled with black. Of course, each rectangle is a particular sample but because the variance is on the order of 4.1×10^{-3}, successive samples will appear much the same.

What we find is that none of these looks much like **Figure 33**. The second rectangle seems to give a reasonable representation of the proportions of temperatures but the distribution is clearly wrong. None of these shows any tendency towards the very large-scale clustering we see in the CMB. The conclusion is that the clustering of superclusters with a common temperature *is not random* which implies that there must exist structure on scales much larger than the size of a supercluster.

We will now turn to the details of the statistical analysis that leads to a description of the CMB spectrum. We begin by working out the spectrum of an ensemble of sources of some fixed size. Our starting point is the Fourier transform representation of the temperature spectrum of some source,

$$\delta T\left(\vec{x}\right)/T\left(\vec{x}\right) = \left(2\pi\right)^{-3/2} \int \mathrm{d}^3 k \, \mathrm{e}^{i\vec{k}\cdot\vec{x}} g\left(\vec{k}\right). \qquad (15\text{-}6)$$

The 2-point expectation value is

$$\left\langle \frac{\delta T(\vec{x}')}{T} \frac{\delta T(\vec{x}'')}{T} \right\rangle = (2\pi)^{-3} \int \mathrm{d}^3 k' \int \mathrm{d}^3 k'' \mathrm{e}^{i\vec{k}'\cdot\vec{x}'} \mathrm{e}^{-i\vec{k}''\cdot\vec{x}''} \left\langle g(\vec{k}') g^*(\vec{k}'') \right\rangle. \quad (15\text{-}7)$$

If we assume that statistically, the universe is the same everywhere, the expectation value will depend only on $|\vec{x}' - \vec{x}''| = |\vec{x}|$. This implies that $\left\langle g(\vec{k}') g^*(\vec{k}'') \right\rangle \propto \delta(\vec{k}' - \vec{k}'')$ so we have

$$\left\langle \frac{\delta T(\vec{x}')}{T} \frac{\delta T(\vec{x}'')}{T} \right\rangle = (2\pi)^{-3} \int \mathrm{d}^3 k \, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \left\langle \left| g(\vec{k}) \right|^2 \right\rangle. \quad (15\text{-}8)$$

We now want to express the idea that the variations are everywhere uncorrelated in which case, the expectation value will have its maximum value at $|\vec{x}| = 0$ and will decay smoothly away from that point at a rate determined by some length scale of the source. A convenient way of expressing this idea is with a Gaussian distribution,

$$\left\langle \frac{\delta T(\vec{x}')}{T} \frac{\delta T(\vec{x}'')}{T} \right\rangle = \left| \frac{\delta T(0)}{T} \right|^2 \mathrm{e}^{-(x/R)^2}. \quad (15\text{-}9)$$

(For now, we will drop the magnitude factor and add it back in later). Solving for the spectral density, we find

$$\left\langle \left| g(\vec{k}) \right|^2 \right\rangle = \pi^{3/2} R^3 \mathrm{e}^{-\frac{(kR)^2}{4}}. \quad (15\text{-}10)$$

These results apply to the universe as a whole. We next need to compute the 2-point expectation on just the spherical shell that emitted the radiation that we detect as the CMB. Denoting the proper distance vector to the shell by $\vec{s} = S(t_{rec})\hat{s}$, (15-8) becomes

$$\left\langle \frac{\delta T(\vec{s}')}{T} \frac{\delta T(\vec{s}'')}{T} \right\rangle = (2\pi)^{-3} \int \mathrm{d}^3 k \, \mathrm{e}^{i\vec{k}\cdot(\vec{s}'-\vec{s}'')} \left\langle \left| g(\vec{k}) \right|^2 \right\rangle. \quad (15\text{-}11)$$

We now introduce the identities

$$\mathrm{e}^{i\vec{k}\cdot\vec{s}} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(ks) P_l(\hat{k}\cdot\hat{s}) \quad (15\text{-}12\mathrm{a})$$

$$P_l(\hat{k}\cdot\hat{s}) = \frac{4\pi}{2l+1} \sum_m Y_{lm}^*(\hat{k}) Y_{lm}(\hat{s}) \quad (15\text{-}12\mathrm{b})$$

$$\int \mathrm{d}\Omega_k Y_{lm}^*(\hat{k}) Y_{l'm'}(\hat{k}) = \delta_{ll'} \delta_{mm'} \quad (15\text{-}12\mathrm{c})$$

where $j_l(ks)$ is the spherical Bessel function of order l. After substituting and performing the angular integrations, etc, we eventually find

$$\left\langle \frac{\delta T(\vec{s}')}{T} \frac{\delta T(\vec{s}'')}{T} \right\rangle = \sum_l \frac{2l+1}{4\pi} P_l(\hat{s}\cdot\hat{s}') C_l \quad (15\text{-}13)$$

where, after substituting (15-9),

$$C_l = 2\sqrt{\pi} R^3 \int \mathrm{d}k \, k^2 \mathrm{e}^{-\frac{(kR)^2}{4}} j_l^2(kS). \quad (15\text{-}14)$$

Making a change of variables to $w = kR$, with the variance restored, this be-

comes,

$$C_l = \left(\frac{\delta T(0)}{T}\right)^2 2\sqrt{\pi} \int \mathrm{d}w\, w^2 \mathrm{e}^{-\frac{w^2}{4}} j_l^2\left(wS/R\right). \qquad (15\text{-}15)$$

The next step is to set the limits of the integration. We know that $k = 2\pi/\lambda$ where λ is a characteristic dimension. The largest dimension is the size of the spherical shell so we have $k_{min} \sim 2\pi/2S\left(t_{rec}\right)$. At the other end of the scale, the smallest relevant size is the size of the structure $k_{max} = 2\pi/2R$. Because of the Gaussian in the formula, we could extend the upper limit to infinity but in this case where the integral must be computed numerically, it would be a disadvantage to do so. The values plotted are $\left(l(l+1)/2\pi\right)C_l$ and we also need to fix up the units. The calculation was based on the 2-point expectation value of the relative temperature variation $\delta T/T$ whereas from the units of the graph, it is apparent that the plotted values are the actual variations rather than the ratios. The actual temperature is 2.73 K so we need to multiply (15-15) by $(2.73 \times 10^6)^2$. After doing this, we end up with

$$K_l = \left(\frac{\delta T(0)}{T}\right)^2 \left(2.73 \times 10^6\right)^2 \overline{K}_l \qquad (15\text{-}16)$$

where

$$\overline{K}_l = \frac{l(l+1)}{\sqrt{\pi}} \int_{\pi(R/S)}^{\pi} \mathrm{d}w\, w^2 \mathrm{e}^{-\frac{w^2}{4}} j_l^2\left(wS/R\right). \qquad (15\text{-}17)$$

Note that this result depends only on the ratio, S/R.

We now want to apply this result to superclusters. Using values from **Table 3**, we find that S/R must fall somewhen within the range of 30 to 257. Our procedure was to try various values until we found the value for which the peak of the calculated spectrum best matched the position of the 1st peak of the actual spectrum. As the ratio is changed, both the position and, to some extent, the shape of the peak change. After a few trials, we found that a value of about $S/R = 120$ which falls near the middle of the range seemed to provide the best fit.

We next want to plot the predicted curve but we must first take into account the flat, large angle background. It will become apparent later that source of this background with a value of about 830 extends to all angles so the peak is actually sitting on top of this. The final displayed value is then

$$V_l = 830 + K_l \qquad (15\text{-}18)$$

Table 5 gives the numerical values and **Figure 36** shows the curve normalized to the peak value which, in this case, implies a temperature variation of $\delta T/T = 3.12 \times 10^{-5}$.

We see that the resulting curve matches the shape of the observed peak reasonably well. The calculated curve is slightly broader than the actual peak which is probably a consequence of assuming a spherical distribution for the superclusters. A more detailed model would replace (15-9) with a non-spherical distribution and include integrals over the orientations in the various expectations.

Table 5. Numerical values with single supercluster size.

Moment I	$\overline{K}_I$	V_I
40	0.054	1222
60	0.116	1671
80	0.195	2244
100	0.284	2889
150	0.509	4521
200	0.653	5565
210	0.666	5659
220	0.673	5710
250	0.663	5637
300	0.538	4731
350	0.292	2947
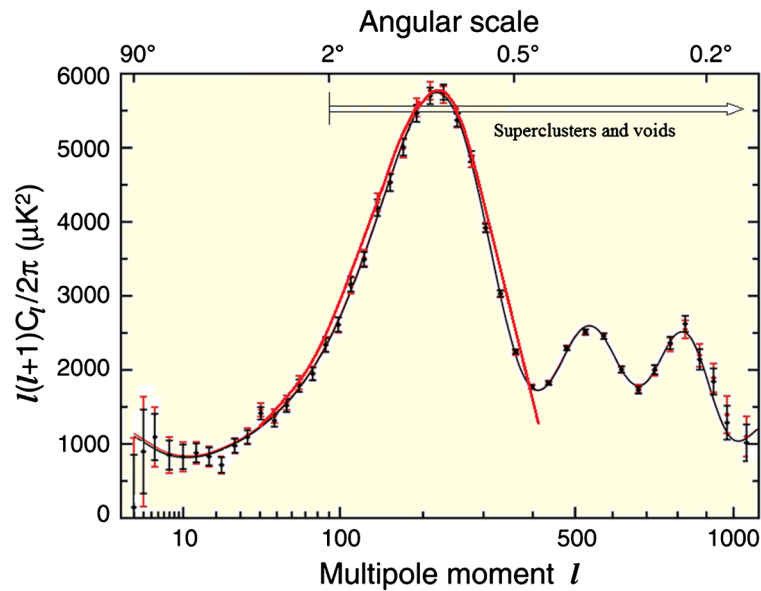
Figure 36. The predicted power spectrum after normalizing to the peak value.

At this point, we recognize that since this result is the spectrum of an ensemble of structures with a single, fixed size, the agreement with the observed spectrum is perhaps fortuitous because superclusters and voids exist with a range of sizes. That being the case, we need to calculate the spectrum for a distribution of sizes. In Figure 37, we display the size distribution of the compilation of 35 superclusters and 36 voids listed in [32]. This list, of course, is not definitive both because of on-going observations that add new structures and also because of the difficulties involved with measuring the dimensions of structures which are only hazily defined but it will be sufficient for our purposes. We have included only those structures from the list that consist of collections of galaxies. A few tentative larger structures are also listed which we have not included. Earlier we established
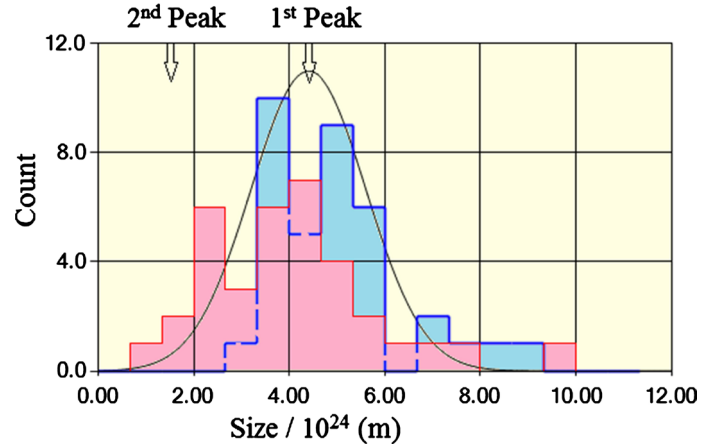
Figure 37. Count of observed superclusters (red) and voids (blue).

that there are around 10^6 superclusters of which about 10^4 contributed to the CMB so 35 is an extremely small sample. It is also worth noting that only those superclusters and voids with the correct redshift would have contributed to our observed CMB and that these particular observed superclusters would not be among those that did.

The indicated position of the first peak was taken from the spectrum and as can be seen, that value corresponds very closely with the center of the sample.

The voids appear to have somewhat narrower range of sizes than does that of the superclusters but that could easily be just a consequence of the small sample size. We also see that the 2nd peak does not correlate with the size of any structure which we already determined by reviewing Table 3.

Because the sources are independent, we obtain the ensemble expectation value by combining the intensities rather than the field values of the photons. Thus, we have

$$\left\langle \bar{K}_l \right\rangle = \int_0^\infty dR \, P(R) \, \bar{K}_l \left(S/R \right) \tag{15-19}$$

Because of the large number of superclusters and voids, no matter what their local distribution happens to be, by the Central Limit theorem, the ensemble distribution must be Gaussian. In the figure, we show the distribution

$$P(R) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left(R - R_p\right)^2}{2\sigma^2}} \tag{15-20}$$

from which we obtain an estimate of $\sigma/R_p \approx 0.19$. Substituting and making a change of variable to $y = R/R_p$, we have

$$\left\langle \bar{K}_l \right\rangle = \frac{R_p}{\sqrt{2\pi}\sigma} \int dy \exp\left[-\frac{1}{2}\left(\frac{R_p}{\sigma}\right)^2 (y-1)^2 \right] \bar{K}_l \left(\frac{S}{yR_p}\right) \tag{15-21}$$

Note that we are using a single value of the temperature variance which is equivalent to assuming that the temperatures of superclusters are independent of their size. Table 6 tabulates the numerical results and Figure 38 shows the

Table 6. Numerical values with a spectrum of sizes.

Moment I	$\langle \overline{K}_I \rangle$	V_I
40	0.054	1227
60	0.116	1684
80	0.195	2265
100	0.284	2920
150	0.505	4547
200	0.644	5570
210	0.656	5658
220	0.663	5710
250	0.649	5607
300	0.523	4680
350	0.269	2810
400	0.054	1227
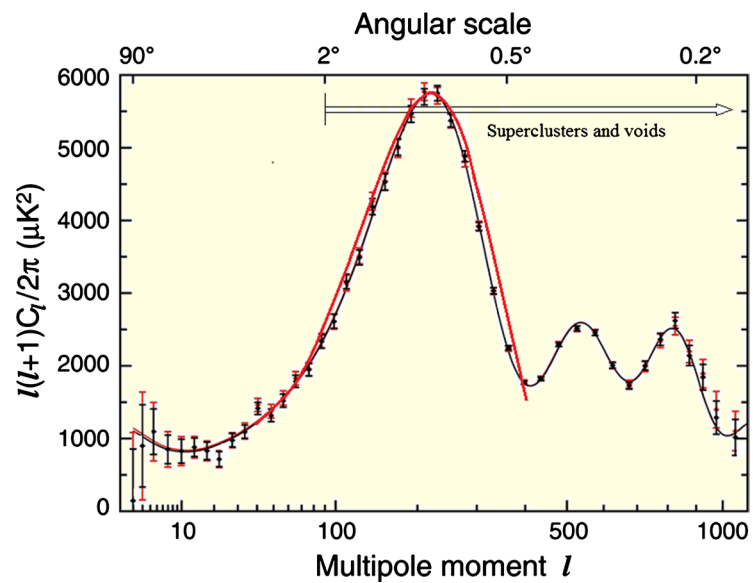
Figure 38. Ensemble average supercluster/void CMB spectrum.

resulting curve using a variance of $\delta T/T = 3.14 \times 10^{-5}$. Note that this value is essentially the same as the previous value.

Based on our experience with the single-size calculations in which the position of the calculated peak varied with S/R, we would have expected to find some broadening of the peak. Comparing the figures, however, we find that there is almost no broadening at all.

In this case, the position of the peak is fixed by the distribution of Figure 37 so the agreement with the spectrum is now elevated to a *prediction* rather than the result of curve fitting. The conclusion is that *the ensemble of superclusters and voids are responsible for the primary peak of the spectrum.*

It would make things easy if we could now apply these same equations to the 2nd and 3rd peaks with an adjust distribution function but that is not the case. The issue is that, because these secondary peaks result from comparisons between different regions in a single supercluster, they are not uncorrelated and so the assumptions leading to (15-16)-(15-17) are not valid. These must be extended to encompass a multipole expansion of the temperature distribution within a supercluster.

Finally, we come to the flat spectrum for all angles larger than about 6°. This flat region is explained in the FRW model by a process involving quantum fluctuations of exotic meson fields during the FRW inflation. But according to this new model, none of that actually happened.

To proceed, we need to derive the spectrum that results from a scale-invariant source in the absence of any structure. If we focus on just a single originally Plank-sized region, we would calculate a peak similar to the peak in the previous figure but at an extremely small angle because even at $t = t_{rec}$, the Plank regions were still very small. But the size of those regions is not the relevant dimension in this case because each of the multipole moments is dimensioned by its wavelength. The fundamental wavelength of the temperature distribution is given by the size of the Plank regions, $D_P(t_{rec})$, but because of the periodic nature of the spherical harmonics, the spectrum consists of a sum of surface integrals over regions whose size is fixed by the wavelength of the multipole moment rather than by D_P. Thus, instead of the distribution of (15-17) where $R = D_P/2$, R must reflect the fact that each moment is sampling regions containing multiple D_P-sized regions. The size of such regions is given by

$$R(l) = \frac{D_P}{2}\left(\frac{\theta_l}{\theta_P}\right) = \frac{D_P}{2}\left(\frac{l_P}{l}\right) \tag{15-22}$$

so, the length ratio parameter becomes

$$S/R = (2S/D_P)\left(\frac{l}{l_P}\right). \tag{15-23}$$

But

$$l_P = \frac{180}{\theta_P} = \frac{180}{(D_P/S)(180/\pi)} = \frac{S\pi}{D_P} \tag{15-24}$$

so finally, we have

$$S/R = \left(\frac{2l}{\pi}\right) \tag{15-25}$$

which in independent of the size of the Plank regions. We now invoke the essential fact of the scale-invariance which means that each multipole region acts like a random variable with an expectation value that is independent of the size of the region and so can be described by (15-17) with a single value for $\delta T/T$. Table 7 presents the calculated spectrum and Figure 39 which follows shows both the scale-invariant spectrum and the sum of that with the peak spectrum.

Table 7. Numerical values for the structure independent spectrum.

Moment l	$\overline{K}_l$	V_l
2	0.22	285
3	0.59	765
4	0.66	856
5	0.64	830
7	0.65	843
≥ 10	0.64	830
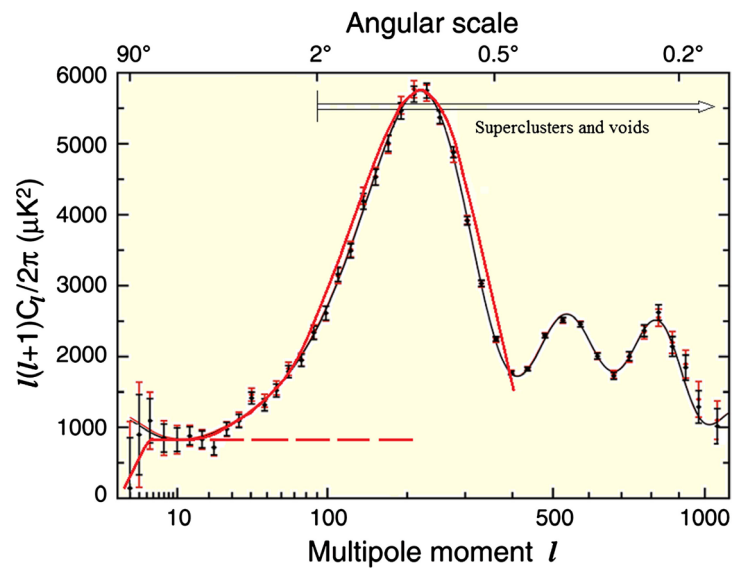
Figure 39. Large angle spectrum of the CMB.

The normalization of the scale-invariant portion is $\delta T/T = 1.32 \times 10^{-5}$.

We see that the curve drops off for small l at a value $l \approx 4$ which corresponds to an angle of 45° which is indicated in **Figure 33**. Referring back to that figure, we see that 45° is also representative of the largest features in the CMB. While the presence of these is obvious from the figure, their contribution to the spectrum is much less obvious but they could account for the hint of a peak at the point where the flat spectrum drops off. The error bars are large, however, and because of their size, there are relatively few such structures compared to the superclusters so the statistical-based formalism just developed is perhaps not applicable for calculating their contribution to the spectrum. The flat spectrum between 6° and 45° indicates that there are no significant structures with sizes lying between those of the extreme structures and the superclusters.

There now remain the issues of explaining the temperature variances implied by the spectrum and the more significant problem of the accounting for the very existence of the superclusters and larger-sized structures at a time many orders of magnitude earlier than the time of star formation. This will be the subject of the next section.

Here, we will conclude with the problem of accounting for the temperature variance of the flat region. At the end of the initial inflation, the Plank variance was $\delta\rho_v c^2 / \rho_v c^2 = e^{-4.2} = 1.5 \times 10^{-2}$. This is clearly much larger that the observed variance so some process must have intervened between the inflation and recombination to reduce the Plank variance. Referring again to Table 3, we see that the vacuum energy completely dominated spacetime during that period so the reduction must have been a consequence of the vacuum itself. One possibility is a merging of spacetime at the time of neutron formation. Each originally Plank-sized region would have expanded along with everything else after the inflation and by the time of neutron formation, the size of each such region would have been

$$D_{\text{Plank}}(t_n) = 4.4 \times 10^{-17}\, \text{m} \tag{15-26}$$

which is clearly much smaller than a neutron. We might now suppose that during the formation of the neutrons (and antineutrons), in order to accommodate their size, the Plank regions merged into regions the size of a neutron. The number of Plank regions in each such region would then be

$$N = \left(\frac{1 \times 10^{-15}}{4.4 \times 10^{-17}} \right)^3 = 1.2 \times 10^4. \tag{15-27}$$

The energy variance of the new regions would subsequently be reduced by the square root of this number so we have

$$\frac{\delta\rho_v c^2}{\rho_v c^2} = \frac{1.5 \times 10^{-2}}{\sqrt{1.2 \times 10^4}} = 1.4 \times 10^{-4}. \tag{15-28}$$

Finally, energy variance is related to temperature variance by $\delta\rho c^2 / \rho c^2 = 4\,\delta T / T$ so we end up with a value of

$$\frac{\delta T}{T} = 3.4 \times 10^{-5}. \tag{15-27}$$

While this is larger than the actual value, it is close and a small change in the size of the merged region or the time that the inflation ended or both could account for the difference.

16. Tying Things Together

We will now summarize the picture of cosmology we have developed and then present data that ties these ideas together. The generally held view of the development of the universe is one of accretion initiate by small in homogeneities in an otherwise uniform distribution of particles at the end of nucleosynthesis. Once the process started, it is supposed that particles coalesced via gravitational interaction into larger and larger structures. We have already shown in Sec. 12 that such an idea won't work but the really insurmountable problem with this concept is that it cannot explain the existence of superclusters much less the much larger structures evidenced by the CMB. As we noted earlier, at the time of recombination, even the smallest superclusters were 5 times larger than the sig-

nal distance so their existence cannot be explained by any process involving accretion and the problem only gets worse as one goes back in time because, as one does so, structures gets larger and larger relative to the signal distance. Accretion won't work and we have also shown in the previous section that no random process can account for the structures either.

The conclusion we reached was that *the existence of all large structures was imprinted on spacetime during the initial inflation and it was this imprint that regulated the creation of neutrons and antineutrons at the time, t_n, in such a manner that the resulting distribution eventually developed into the structures we now see.*

From this perspective, all large structures were born with more or less their final sizes and masses with accretion playing only a subsidiary role. In fact, we will show that this process was responsible for all cosmic structures and not just the very large.

The three quantities that regulated the distribution of matter were the total vacuum energy, the fraction of that energy that was converted into neutrons and antineutrons and the fraction of that which determined the ratio by which the number of neutrons exceeded the number of antineutrons. On average across the entire universe, the total energy is given by (8-24), the creation fraction was on the order of 10^{-3}, and the asymmetry fraction was on the order of 10^{-8}. Another measure of the magnitude of the variances follows from the CMB spectrum. The observed temperature variance is $\delta T/T = (1.3\text{-}3.1)\times10^{-5}$ which differs by less than a factor of 2.5 all across the universe which is another argument for an origin in which length scales were not constrained by the speed of light. The variance in the total energy density necessary to explain the spectrum is of $O\left(10^{-7}\right)$.

We determined that the matter/antimatter asymmetry factor always had the same "sign" but that it too varied in magnitude from one place to another. In the regions with the greatest particle density, its value was around 2.4×10^{-8} whereas in the voids, the factor was on the order of 1.9×10^{-10}. We found then that the total number of neutrons and antineutrons initially created was much the same everywhere with a variance no larger than one part in 10^{-7} and that the differences between the high-density regions and the voids are almost entirely a result of differences in the asymmetry factor. From observations, we know that the high-density regions tend to be warmer and vice-versa so these factors appear to be correlated. Referring again to Table 3, we see that, on a logarithmic scale, the Milky Way is actually much closer in size to clusters and even superclusters than it is to the size of a "t_n" cell. This means that the dimensions that characterize these very small in magnitude imprint variances are vastly larger than the dimensions that characterized particle creation and nucleosynthesis.

If this was all there was to it, the universe would have ended up with a more or less uniform distribution of matter with no structure; a result that follows from the fact that small variances in matter density alone could not have initiated ac-

cretion. That being the case, it follows that the controlling factor must have been largely or wholly a matter of extremely small variances in the properties of the vacuum and the fact that these variances were smooth on length scales vastly in excess of Lorentz limitations implies that they must have originated during the initial Plank era inflation.

Having reached this conclusion, we will now consider observational data that supports these ideas. What we will show is that the distribution of cosmic structures places significant constraints on possible structure creation models. In Figure 40 and Figure 41, we show the count of cosmic structures as a function of their size and mass respectively. Combining these gives the size as a function of mass with the result shown in Figure 42.
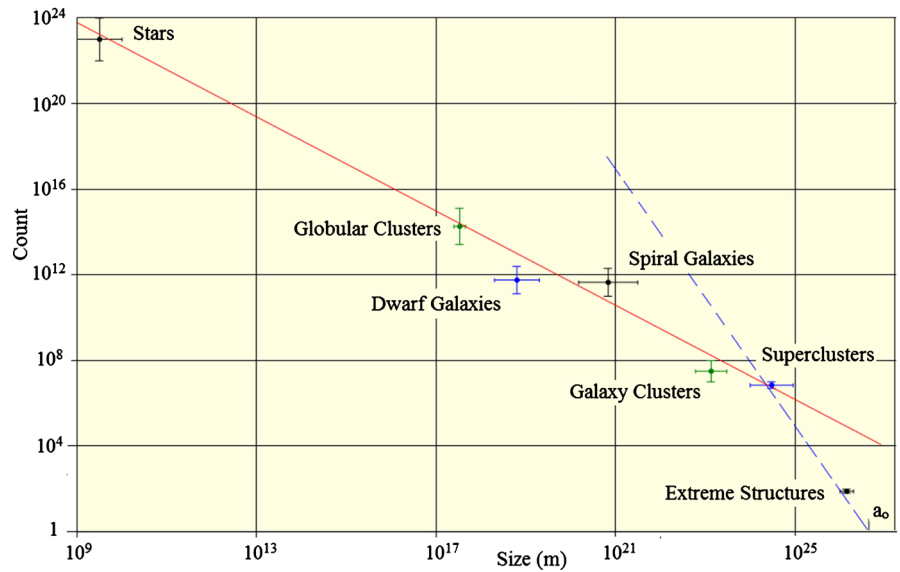
Figure 40. Count of structures vs size.
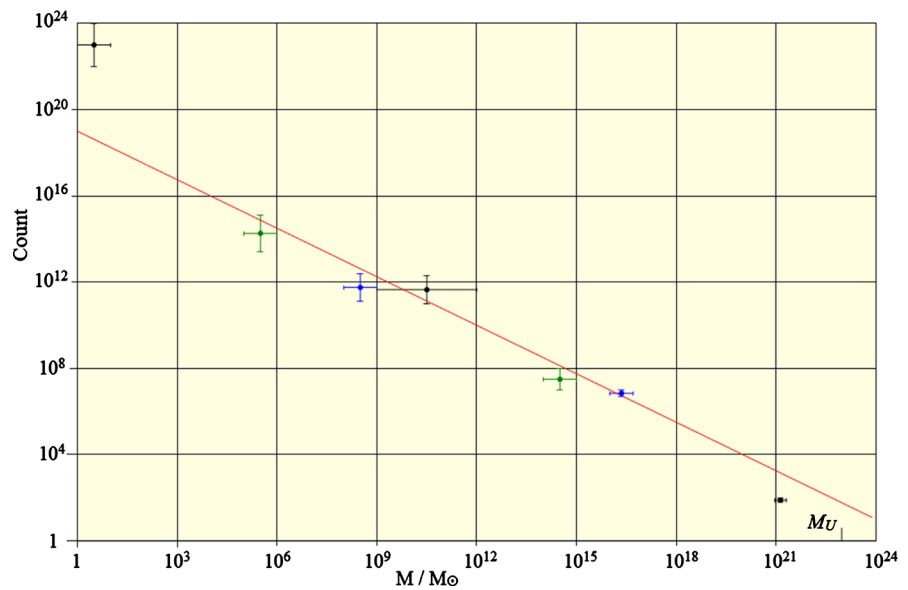
Figure 41. Count of structures vs mass.
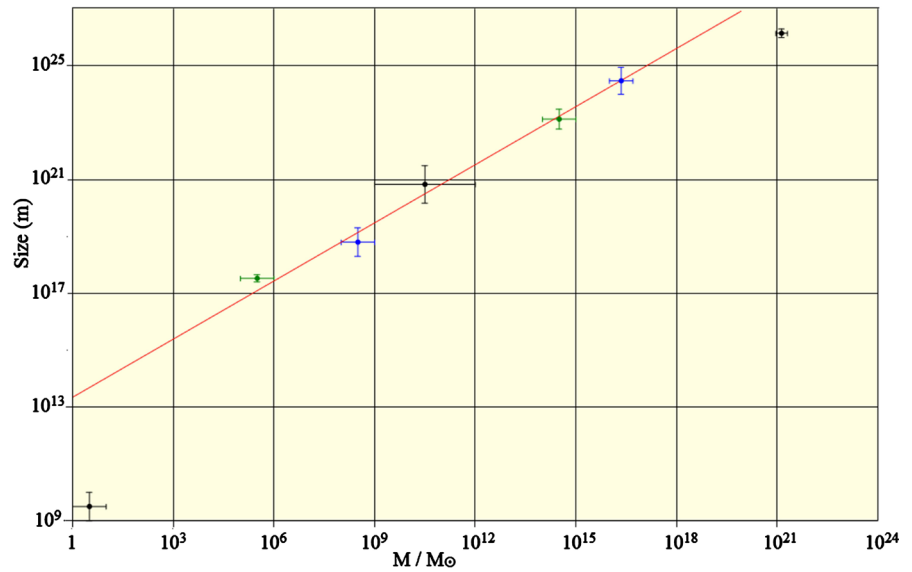
Figure 42. Size of structures vs mass.

To make these plots, we needed to know the sizes, masses, and counts of all types of structures. The sizes and masses of each type are reasonably well known. The counts are less reliable and in some case are just estimates based on local densities. For example, to estimate the total number of dwarf galaxies, we used the fact that there are roughly 50 associated with the Milky Way while some other large galaxies are thought to have counts as high as 10^5. Allowing for a range of values and multiplying by the number of large galaxies gave us an estimate of the total. The extreme structures are representative of the apparent 45° structures visible in the CMB anisotropy map.

Referring first to Figure 40, what is remarkable is that, with the exception of the extreme structures, all these structures with their vast range of sizes lie on a power law curve. The extreme structures lie below the curve but this is just a consequence of the finite size of the universe since the maximum count of any structure cannot exceed the number that would fill the universe.

Similarly, in Figure 41, with the exception of stars and again the extreme structures, the mass distribution also follows a power law curve. The extreme structures lie below the curve because of finite mass of the universe. Stars are the exception because they are obviously far more massive relative to their size that any of the other structures and certainly in their case, accretion was, and is, significant factor.

The following formulas for the curves give the corresponding power law coefficients. These are not model predictions but rather parameterized curves adjusted to match the data. We chose to use superclusters as the reference.

$$C(s) = 5.7 \times 10^6 \left(s_{Sc}/s \right)^{1.1}. \tag{16-1}$$

$$C(m) = 9.2 \times 10^6 \left(m_{Sc}/m \right)^{0.75}. \tag{16-2}$$

where s is the size and $m = M/M_\odot$. The subscripts "Sc" refer to supercluster

mean values. By combining these two equations, we obtain

$$s(m) = 1.9 \times 10^{24} (m/m_{Sc})^{0.68} m. \tag{16-3}$$

What we are going to argue now is that these results not only support the notion of a Plank era imprint being responsible for the distribution of structures but also that the imprint is correctly described as a fractal geometry. Concentrating now on Figure 40, there are two model curves shown. The dashed blue line which is given by

$$C_{\text{filled}}(s) = (a_0/s)^3. \tag{16-4}$$

gives the count of structures necessary to fill the entire volume of space as a function of their size. The extreme structures line on this curve by definition but what is more interesting is that superclusters also lie on this curve which implies that in an order of magnitude sense, they fill all space. The model line of (16-1), however, is where it gets interesting.

We now want to introduce the idea of fractal dimension. Equation (16-4) is a simple formula that gives the count of objects of a given size need to fill a 3-dimensional space. Similarly, the number needed to fill a 2-dimensional surface would be $(a/s)^2$. We can write this generally as

$$C = r^d. \tag{16-5}$$

where r is the magnification factor and d is the dimension of the space which in common usage would be an integer. The idea of a fractal geometry is one in which the same general formula holds but the dimension can have any value, not just an integer value.

But this is exactly the form of (16-1) and from this, we learn that the initial *imprint that defined all the structures we observefrom stars on up to superclusters was a fractal geometry with a (box) dimension of $d_F = 1.1$.*

There are a few consequences that follow immediately. First, not only are fractal geometries non-differentiable but it has been proven that all non-differentiable geometries are fractal (see e.g. [33]) so this model uniquely satisfies our earlier contention that the Plank era must be described in terms of a non-differentiable manifold.

The second point is that with a dimension only slightly larger than one, the basis of cosmic structures must be in the form of filaments. Thus, we find that two seemingly unrelated facts, namely the count distribution and the filament structure of space have a single origin. *It is unavoidable that a universe with the counts we observe must also have a filament structure and vice-versa.* Either way, it is fractal.

The third point follows from the fact that the scatter away from the line is not large. Remember that the fractal imprint cannot be responsible for more than the initial size and mass of the structures and that these structures would be subject to subsequent gravitational influences from that point on. What the small scatter tells us is that, with the exception of stars, the subsequent interac-

tions had little effect on their sizes and masses or, in other words, that accretion was not the overriding factor in their development. Another fact in support of these ideas is that the volume of background space (1 m^{-3}) necessary to form a single star solely by accretion is roughly volume of a globular cluster.

A fourth point hinting at a common origin of the structures is that they have distinct sizes with no overlap. Put another way, if the structures were purely the result of accretion, one would expect to find a continuum of sizes instead of, in some vague way, the multipole distribution that we observe.

A fifth point is that we again find an equivalence between vacuum energy and dark matter but with a far more detailed understanding of how the filament structure came to exist.

We will now return to the issue of causality. The expansion of the scaling occurs at every point independent of any influence from any other point. In order to form structures, on the other hand, coordination between different locations is necessary which in a normal situation would imply an exchange of signals. Given the results of (4-12), the speed of any such signals would be scaled by $a_I/t_I = 10^{26} \text{ m} \cdot \text{s}^{-1}$ which in practical terms is approaching infinity so the whole concept of normal exchange is probably wrong. What seems more likely is that, because of the uncertainties of time and dimension, different regions had, in some way, effectively a zero separation so a change in one location was a change over a region. This, however, is total speculation at this point and given our lack of even a framework to work with, we really can't say what process accounted for the formation of structures.

So, at the end, we are back where we started. We need a new understanding of the Plank era to make further progress because it was during that era that the "DNA" that defined the universe originated. What we have learned is that there must have been a Plank era during which an exponential inflation occurred. Not the least of the arguments for that inflation is the fact that without it, the present-day size of the universe would be measured in fractions of a meter. We have seen that during that era, the normal ideas of causality did not apply and that structure in the vacuum energy developed that exhibits a fractal geometry. We have thus defined a number of constraints that must be satisfied but do not yet have a model of how this all happened.

The fact of a Plank era, however, leaves us with another problem. It has been noted by many people that expressions such as $l_P = \left(\hbar G/c^3\right)^{1/2}$ is just a combination of constants so one is faced with the problem of explaining how these combinations of constants just happen to match up with the reality of the Plank era. It is beyond imagining that the agreement could be just a coincidence so we are led to the idea that l_P, t_P, etc. are, in fact, the fundamental entities and physical constants such as c are properties of the vacuum that derive from these entities, e.g. $c = l_P/t_P$. In other words, we are reading the Plank relations the wrong way around. This notion also hints at a solution of the causality problem because, according to our thinking, the Plank quantities were initially subject to

uncertainty so it follows that the value of c, for example, was also uncertain and it did not obtain its final (certain) value until after the end of the inflation when the uncertainties became negligible compared to the age of the universe.

We mentioned in the introduction that much effort has gone into the study of non-commutative geometry with the aim of formalizing the notions of coordinate uncertainty and non-differential manifolds. While this shows that people have been thinking about that Plank era problem for a considerable period of time, the results so far have been nil as far as any application to Plank era physics is concerned and do not even begin to approach the problem of explaining the existence of the very large, relative to Lorentz limits, and also very smooth structures that must have existed.

17. Alternate Theories

Over the years a number of extensions of the original theory of gravitation have been directed towards solving a range of shortcomings of the standard model. As pointed out in [34], these extensions can be grouped into those in which the left-hand side of the equation is modified, for example, by the addition of higher order powers of the Riemann tensor and those in which additional contributions to the right-hand side are included.

Left-hand extensions have, for example, have been applied by those seeking to achieve a unification of gravity with the other fields. None of these efforts, however, have achieved any success which, in our opinion isn't surprising because we believe that gravitation is fundamentally different from other fields and that a unified theory is just wishful thinking. There is certainly no observational evidence that any such unification exists.

Right-hand extensions, on the other hand, have led to the development of theories incorporating *ad hoc* entities such as dark energy (cosmological constant) and dark matter. These entities are considered to be unrelated with dark energy distributed uniformly and dark matter distributed in clusters. The models do not explain what those entities are but calculations incorporating those entities can be made to match observations by adjusting various parameters. There are a number of problems with such models, however. For example, it is considered a mystery that the magnitude of the cosmological constant is so small. The models also do not explain why dark matter is always in close association with ordinary matter. As just noted, all the results obtained are dependent on curve fitting and it is a serious defect of these models that they do not actually predict anything solely on the basis of the metric and Einstein's equations. By choosing appropriate parameters any sort of evolution can be obtained.

We will now compare it with the new model. Leaving aside the problem of the Plank era, what we have shown is that by formulating a model that incorporates time-varying curvature, a significant number of the outstanding problems are solved. For example, the acceleration of the scaling is a parameter independent *prediction* of the model which has nothing to do with a cosmological constant or

equivalently, dark energy. In fact, the concept of dark energy in the standard model sense simply does not exist. What does exist is *time-varying* vacuum energy whose present-day energy density is *predicted* to be close in magnitude to that so-called dark energy so the smallness of the magnitude is no mystery at all. We have also shown that so-called dark matter is, in fact, just another manifestation of the same vacuum energy and that is association with ordinary matter is easily explained. Finally, in contrast to the *ad hoc* models in which any sort of evolution is possible, the new model is totally constrained by Einstein's equations; there are *no* adjustable parameters, and only *one* evolution is possible. This model stands as an alternative to the extended models of gravitation that solves many of the outstanding problems while, at the same time, bringing us back to the original concept of gravitation.

So, is this model is the final answer? It certainly appears to be closer to the truth than any of the other models thus far proposed but this paper represents only the starting point of a new direction in cosmology. In particular, as shown in the above reference, gravitational wave astronomy has the potential for detecting small model deficiencies and with this in mind, in a subsequent paper, we plan to examine gravitational waves within the context of our new model.

18. Conclusions

In this paper, we present a new model of cosmology based on very few assumptions that completely avoids any type of exotic particle, field theory, or cosmological constant. A considerable number of predictions have been made that are in agreement with observations. Among the highlights, the new model.

1) proposes that the Big Bang began with a Plank era period of exponential inflation driven by uncertainty principle effects and time-varying spacetime curvature. It is shown that time variation of the curvature is a decisive factor driving the evolution of the universe and that the present-day structure of the universe had its origin in very small in amplitude but exceeding large, in dimension, variances that came into existence during the inflation.

2) presents an exact solution of Einstein's equations that predicts an acceleration of the present-day expansion of the universe. This model has *no* adjustable parameters. The solution reconciles the homogeneity and isotropy of spacelike hypersurfaces with time-varying curvature and produces a number of exact results including the prediction that the curvature is proportional to the sum of the vacuum energy density and pressure and that the curvature always has its maximum possible value. The model also makes a prediction of the luminosity distance that matches the data and points to a solution to the problem researchers are having in trying to determine the Hubble constant.

(3) shows that all physical quantities such as the scaling, the curvature of spacetime, and the motion of particles are dependent on only the *sum* of the vacuum energy, pressure, and particle mass energy equivalent at any point in spacetime and that this sum varies as with time as t^{-2} independent of the scaling.

4) proposes an origin of ordinary matter that is in no way connected with conventional field theory. A detailed model of nucleosynthesis is presented that accounts for both the CMB and the matter-antimatter asymmetry. Although it is a minor point, we also show that the so-called Lithium problem is actually nothing more than a procedural issue.

5) shows that the phenomena that dark matter was proposed to explain can be readily understood as consequences of the vacuum energy thereby establishing the fact that dark matter is vacuum energy.

6) proposes a new explanation for the CMB spectrum. We show that the large peaks are a consequence of superclusters and voids and that the large angle flat spectrum is a consequence of energy uncertainties embedded in spacetime at the termination of the initial inflation.

7) shows that the basis for all cosmic structure was a fractal geometry imprint that originated during the initial Plank era inflation.

Acknowledgements

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Wald, R.M. (1984) General Relativity. The University of Chicago Press, Chicago.

[2] Hobson, M.P., *et al.* (2006) General Relativity, an Introduction for Physicists. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511790904

[3] Arnowitt, R., Deser, S. and Misner, C.W. (2004) The Dynamics of General Relativity. https://arxiv.org/abs/gr-qc/0405109

[4] Riess, A.G., *et al.* (1998) Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astrophysical Journal*, **116**, 1009-1038. https://doi.org/10.1086/300499

[5] Nielsen, J.T., Guffanti, A. and Sarkar, S. (2016) Marginal Evidence for Cosmic Acceleration from Type 1a Supernovae. *Scientific Reports*, **6**, Article No. 35596. https://doi.org/10.1038/srep35596

[6] NASA. Cosmic Web, No Date. https://www.nasa.gov/images/content/228352main_cosmicweb_HI.jpg

[7] Meiksin, A.A. (2009) The Physics of the Intergalactic Medium. *Reviews of Modern Physics*, **81**, 1405-1469. https://doi.org/10.1103/RevModPhys.81.1405

[8] Simcoe, R.A. (2004) The Cosmic Web: Observations and Simulations of the Intergalactic Medium Reveal the Largest Structures in the Universe. *American Scientist*, **92**, 30-37. https://doi.org/10.1511/2004.1.30

[9] Weisstein, E.W. (2020) Random Walk—1 Dimensional. http://mathworld.wolfram.com/RandomWalk1-Dimensional.html

[10] Greene, G.L. and Geltenbort, P. (2016) The Neutron Enigma. *Scientific American*,

314, 36-41. https://doi.org/10.1038/scientificamerican0416-36

[11] Cyburt, R.H. (2004) Primordial Nucleosynthesis for the New Cosmology: Determining Uncertainties and Examining Concordance. *Physical Review D*, **70**, Article ID: 023505. https://doi.org/10.1103/PhysRevD.70.023505

[12] Rupak, G. (1999) Precision Calculation of np $\rightarrow$ dγ Cross Section for Big-Bang, Nucleosynthesis.

[13] Abramovich, S.N., Guzhovskij, B.Ya., Zherebtsov, V.A. and Zvenigorodskij, A.G. (1989) Nuclear Physics Constants for Thermonuclear Fusion—A Reference Handbook. Central Scientific Research Institute on Information and Techno-Economic Research of Atomic Science and Technology, State Committee on the Utilization of Atomic Energy of the USSR.

[14] Caughlan, G.R. and Fowler, W.A. (1988) Thermonuclear Reaction Rates V. *Atomic Data and Nuclear Data Tables*, **40**, 283-334.
http://www.nuclear.csnb.cn/data/CF88
https://doi.org/10.1016/0092-640X(88)90009-5

[15] Kopecky, J. (1997) Atlas of Neutron Capture Cross Sections, JUKO Research, International Nuclear Data Committee.

[16] Sadeghi, H. and Bayegan, S. (2005) Nd $\rightarrow$ ^{3}H Gamma with Effective Field Theory. *Nuclear Physics A*, **753**, 291-304. https://doi.org/10.1016/j.nuclphysa.2005.03.004

[17] Nematollahi, H., Bayegan, S., Mahboubi, N. and Moeini Arani, M. (2016) p + d $\rightarrow$ ^{3}He + Gamma Reaction with Pionless Effective Field Theory. *Physical Review C*, **94**, Article ID: 054004. https://doi.org/10.1103/PhysRevC.94.054004

[18] Bhatia, V.B. (2001) Textbook of Astronomy and Astrophysics with Elements of Cosmology. Alpha Science Intl. Ltd., Oxford, 203.

[19] Laborie, J.M., Varignon, C., Ledoux, X. and Arnal, N. (2007) Measurement of the D(n, 2n)p Reaction Cross Section up to 30 MeV. *Proceedings of the* 2007 *International Conference on Nuclear Data for Science and Technology*, Nice, 22-27 April 2007, 437-440.
https://nd2007.edpsciences.org/articles/ndata/pdf/2007/01/ndata07313.pdf

[20] Skibinski, R., Golak, J., Topolnicki, K., Witala, H., Epelbaum, E., Kamada, H., Krebs, H., Meissner, Ulf-G. and Nogga, A. (2016) Selected Two- and Three-Body Electroweak Processes with Improved Chiral Forces.
https://arxiv.org/abs/1604.03395v1

[21] Fukugita, M. and Kajino, T. (1990) Contribution of the He3(t, γ)Li6 Reaction to Li6 Production in Primordial Nucleosynthesis. *Physical Review D*, **42**, 4251.
https://doi.org/10.1103/PhysRevD.42.4251

[22] Malaney, R.A. and Fowler, W.A. (1989) On Nuclear Reactions and Be9 Production in Inhomogeneous Cosmologies. *Astrophysical Journal, Part* 2 *Letters*, **345**, L5-L8.
https://doi.org/10.1086/185538

[23] Barbagallo, N., *et al*. (2014) Measurement of ^{7}Be(n, alpha) and ^{7}Be(n, p) Cross Sections for the Cosmological Li Problem in EAR2@n_TOF. CERN INTC Meeting, June 25.
https://www.researchgate.net/publication/319674486_7Bena_and_7Benp_cross_section_measurement_for_the_cosmological_lithium_problem_at_the_n_TOF_facility_at_CERN

[24] Hou, *et al*. (2015) A Revised Thermonuclear Rate of ^{7}Be(n, a)^{4}He Relevant to Big-Bang Nucleosynthesis. *Physical Review C*, **91**, Article ID: 055802.
https://doi.org/10.1103/PhysRevC.91.055802

[25] Eilers, A.-C., Hogg, D.W., Rix, H.-W. and Ness, M.K. (2019) The Circular Velocity

Curve of the Milky Way from 5 to 25 kpc. *The Astrophysical Journal*, **871**, 120.
https://doi.org/10.3847/1538-4357/aaf648

[26] Hibbs, P. (2005) Galactic Rotation.
https://commons.wikimedia.org/wiki/File:GalacticRotation2.svg

[27] Bannikova, E.Yu., Vakulik, V.G. and Shulga, V.M. (2012) Gravitational Potential of a Homogeneous Circular Torus: New Approach. *MNRAS*, **411**, 557.
https://arxiv.org/abs/1009.4324

[28] Lass, H. and Blitzer, L. (1983) The Gravitational Potential Due to Uniform Disks and Rings. *Celestial Mechanics*, **30**, 225-228. https://doi.org/10.1007/BF01232189

[29] NASA, CMB Anisotropy Map.
http://map.gsfc.nasa.gov/media/121238/index.html

[30] Hinshaw, G., *et al.* (2013) Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results.
https://lambda.gsfc.nasa.gov/product/map/dr5/pub_papers/nineyear/cosmology/wmap_9yr_cosmology_results.pdf

[31] Riotto, A. (2002) Inflation and the Theory of Cosmological Perturbations. *Lectures Given at the*: *Summer School on Astroparticle Physics and Cosmology*, Trieste, 17 June-5 July 2002. https://arxiv.org/abs/hep-ph/0210162

[32] Wikipedia (2020) List of Largest Cosmic Structures.
https://en.wikipedia.org/wiki/List_of_largest_cosmic_structures

[33] Nottale, L. (2003) The Theory of Scale Relativity: Non-Differentiable Geometry and Fractal Space-Time. *AIP Conference Proceedings*, **718**, 68.
https://doi.org/10.1063/1.1787313
https://pdfs.semanticscholar.org/06b2/074a6619387974b9e54c5ecaec698c15e203.pdf

[34] Corda, C. (2009) Interferometric Detection of Gravitational Waves: The Definitive Test for General Relativity. *International Journal of Modern Physics D*, **18**, 2275-2282. https://doi.org/10.1142/S0218271809015904