EINSTEIN'S ROLE IN THE CREATION OF RELATIVISTIC COSMOLOGY

CHRISTOPHER SMEENK

1. Introduction

Einstein's paper, "Cosmological Considerations in the General Theory of Relativity" (Einstein 1917b), is rightly regarded as the first step in modern theoretical cosmology. Perhaps the most striking novelty introduced by Einstein was the very idea of a cosmological model, an exact solution to his new gravitational field equations that gives a global description of the universe in its entirety. Einstein's paper inspired a small group of theorists to study cosmological models using his new gravitational theory, and the ideas developed during these early days have been a crucial part of cosmology ever since. We will see below that understanding the physical properties of these models and their possible connections to astronomical observations was the central problem facing relativistic cosmology in the 20s. By the early 30s, there was widespread consensus that a class of models describing the expanding universe was in at least rough agreement with astronomical observations. But this achievement was certainly not what Einstein had in mind in introducing the first cosmological model. Einstein's seminal paper was not simply a straightforward application of his new theory to an area where one would expect the greatest differences from Newtonian theory. Instead, Einstein's foray into cosmology was a final attempt to guarantee that a version of "Mach's principle" holds. The Machian idea that inertia is due only to matter shaped Einstein's work on a new theory of gravity, but he soon realized that this might not hold in his "final" theory of November 1915. The 1917 paper should thus be read as part of Einstein's ongoing struggle to clarify the conceptual foundations of his new theory and the role of Mach's principle, rather than treating it only as the first step in relativistic cosmology.

Einstein's work in cosmology illustrates the payoff of focusing on foundational questions such as the status of Mach's principle. But it also illustrates the risks. In the course of an exchange with the Dutch astronomer Willem De Sitter, Einstein came to insist that on the largest scales the universe should not evolve over time—in other words, that it is static. Although he originally treated this as only a simplifying assumption, Einstein later brandished the requirement that any reasonable solution must be static to rule out an anti-Machian cosmological model discovered by De Sitter. Thus Einstein's concern with Mach's principle led him to introduce the first cosmological model, but he was also blind to the more dramatic result that his new gravitational theory naturally leads to *dynamical* models. Even when expanding universe models had been described by Alexander Friedmann and Georges Lemaître, Einstein rejected them as physically unreasonable. Einstein's work in cosmology was also not informed by a thorough understanding of contemporary empirical work. The shift in theoretical cosmology brought about by Einstein's work occurred at the same time as a shift in the observational astronomer's understanding of the nature of spiral nebulae and large scale structure of the cosmos. Others with greater knowledge of contemporary astrophysics, including Arthur Eddington, De Sitter, Lemaître, and Richard Tolman, made many of the more productive contributions to relativistic cosmology.

This paper describes the early days of relativistic cosmology, focusing on Einstein's contributions to the field. Section 2 gives an overview of the "Great Debate" in observational astronomy during the 10s and 20s. Section 3 describes Einstein's "rough and winding road" to the first cosmological model, with an emphasis on the difficulties with Newtonian cosmology and the importance of Mach's principle. Conversations and correspondence with De Sitter forced Einstein to consider the status of Mach's principle in his new theory, leading to the introduction of his cosmological model. Although Einstein hoped that this would insure the validity of Mach's principle, De Sitter quickly produced an apparent counterexample, a new model in which Mach's principle did not hold. Einstein's attempts to rule out this counterexample drew him into a tangle of subtle problems regarding the nature of singularities that also ensnared his correspondents Hermann Weyl and Felix Klein. Throughout the 20s, theoretical cosmology focused for the most part on understanding the features of the models due to Einstein and De Sitter and attempting to find some grounds on which to prefer one solution to the other. Section 4 takes up these debates, which came to an end with the discovery that several other viable solutions had been overlooked. This oversight can be blamed on Einstein's influential assumption that the universe does not vary with time, which was elevated from a simplifying assumption to a constraint on any physically reasonable cosmological model. This assumption kept Einstein, and many of the other leading lights of theoretical cosmology, from discovering expanding universe models. Section 5 describes the expanding models, and their relation to observational evidence in favor of expansion. In the concluding section I briefly discuss the fate of the consensus formed in the early days of relativistic cosmology in the ensuing decades, focusing on debates regarding the status of the so-called "cosmological principle."

2. The Great Debate

The first three decades of the twentieth century were a period of active debate within the astronomical community regarding the overall structure of the cosmos.[1] At the turn of the century, two leading observational cosmologists, the influential Dutch astronomer Jacobus Kapteyn and the German Hugo von Seeliger, had embarked on the ambitious project of determining the architecture of the Milky Way. Astronomers working in this line of research used sophisticated statistical techniques to overcome the fundamental problem facing the project, namely that of determining distances to the stars, in order to convert observational data into a three-dimensional map of the stellar system. The program culminated in the "Kapteyn Universe," according to which the galaxy is a roughly ellipsoidal distribution of stars, with the sun near the center.[2] Starting in 1918, Kapteyn and others had to contend with the alternative "Big Galaxy" view introduced by the American astronomer Harlow Shapley. Based on studies of large, dense groups of stars called globular clusters, Shapley argued that the Milky Way was a factor of ten larger than in Kapteyn's model, with the sun placed some distance from the center. Partially as a result of Shapley's challenge, the focus of observational cosmology shifted from determining stellar distributions in the Kapteyn tradition to using other objects (globular clusters and nebulae) as indicators of large scale structure. This era has been called the "second astronomical revolution" to reflect the impact of the new ideas and techniques introduced during this period. I will set the stage for Einstein's work by focusing on two aspects of this revolution, the debates regarding the spiral nebulae and related debates in cosmogony.

By 1930 astronomers had abandoned much of the turn-of-the-century conventional wisdom regarding "spiral nebulae." Around 1900 many astronomers held that the Milky Way galaxy was a unique system encompassing all observed objects, including the enigmatic spiral nebulae. The English astronomer Agnes Clerke, for example, confidently asserted that "no competent thinker" could accept that any spiral nebula is comparable to the Milky Way (Clerke, 1890, 368). But by 1930, few competent thinkers still held Clerke's opinion. A majority of astronomers instead accepted the "island universe" theory, according to which the spiral nebulae are similar in nature to the Milky Way.[3]

The nature of the spirals had been the subject of speculation and debate for over 150 years. The speculative proposal of Kant, Lambert, and Wright that the nebulae are "island universes," composed of stars and similar to the Milky Way in structure, first garnered support from the observations of William Herschel and Lord Rosse in the early to mid-19th century. Yet by the turn of the century, there were three major objections to this idea.[4] In 1885 a nova flared in the Andromeda Nebula, and at its brightest this star reached a luminosity of roughly one tenth that of the entire nebula. Faced with the implausible idea that a single star could outshine millions of others, many astronomers concluded instead that the Andromeda Nebula was not a group of stars like the Milky Way. Second, the technique of spectroscopy revealed bright emission lines in the spectra of some nebulae, characteristic of luminous gas rather than starlight. This suggested that the nebulae are regions of hot gas within the Milky Way, rather than groups of unresolved stars at a great distance. Although some nebulae have starlike spectra, this could be explained as the consequence of reflected starlight. Finally, the positions of the nebulae seemed to be related to the Galaxy itself: the nebulae shunned the plane of the Galaxy and clustered near the poles. Such a correlation would be unsurprising if the spirals were within the Galaxy, but it appeared to be inexplicable based on the island universe theory.

The fortunes of the island universe theory improved dramatically by the mid-teens. In his influential *Stellar Movements and the Structure of the Universe* (1914), the British astrophysicist Arthur S. Eddington argued that, despite the paucity of direct evidence in its favor, the island universe theory should be accepted as a useful "working hypothesis," whereas the idea that the spirals lie within the galaxy should be rejected because it led to an unproductive dead end. As Eddington emphasized, the classification of distinct types of nebulae undercut the second criticism above: the spiral nebulae might lie well beyond the galaxy even though other nebulae are part of the Milky Way. However, the other criticisms remained unanswered. Island universe advocates had to admit that novae could somehow produce tremendous energy, yet at least the novae discovered in spirals appeared to be significantly dimmer than those within the Milky Way.[5] The clustering of spirals near the Galactic poles remained a mystery throughout the 20s, but it was eventually explained as an observational selection effect due to the absorption of light by interstellar dust (Trumpler, 1930).[6]

Critics of the island universe theory marshalled two new and apparently quite damaging results in the late 10s. Two of the most prominent critics, Harlow Shapley and Adriaan van Maanen, attempted to put the final nails in the coffin of the island universe theory. As Shapley put it in a letter to van Maanen, "Between us we have put a crimp in the island universe, it seems—you in bringing the spirals in and I by pushing the Galaxy out" (quoted in Smith, 1982, 105). Starting in 1914, van Maanen measured what he took to be the rotational motion of objects within several spiral nebulae. These measurements only made sense if the nebulae were relatively small, nearby objects; if the nebulae were as large as the Milky Way and far away, as the island universe theory required, the motions would dramatically exceed reasonable physical limits. Shapley, for his part, argued that the Milky Way was much larger than had previously been assumed (by a factor of ten), leaving plenty of room for the spirals within the galaxy. Shapley reached the "Big Galaxy" view based on a novel astronomical yardstick: he used Cepheid variable stars as "standard candles." Henrietta Leavitt had established that the period of the variation in the brightness of these stars bears a fixed relationship to their intrinsic brightness (absolute magnitude). After calibrating the scale using nearby Cepheids, Shapley could then directly calculate the distances of remote Cepheids by determining their period.[7]

The decade closed with the so-called "Great Debate" between Shapley and a defender of the island universe theory, Heber Curtis. The event did not live up to its billing, since Shapley chose to avoid tackling the island universe theory directly.[8]

In fact, Shapley's "Big Galaxy" view was not directly incompatible with the island universe theory, as Shapley's own shifting views on the matter in the years leading up to the debate indicate.[9] Shapley did not present a detailed attack on the island universe theory in the published exchange (Shapley, 1921), although he briefly mentioned the clustering of the spirals around the galactic poles and the incompatibility of the island universe theory with van Maanen's measurements. The appeal to van Maanen's measurements of internal motion in the spiral nebulae had little chance of persuading Curtis, who responded that van Maanen's measurements would only be convincing once they had been successfully reproduced (Curtis, 1921, 214). Controversy regarding van Maanen's measurements continued throughout the twenties, until they were eventually rejected as artifacts of subtle systematic errors.

For his part in the debate, Curtis focused on the crucial issue of how to reliably measure distance to astronomical objects. Shapley had pioneered the use of Cepheid variables to establish distances to globular clusters, but Curtis questioned the internal consistency of this method as well as its relative reliability compared to other distance measures. Ironically, this controversial method would bring about the resolution of the debate within a few short years of Curtis and Shapley's exchange. Edwin Hubble observed a Cepheid variable in the Andromeda Nebula in 1923 with the powerful 100-inch telescope on Mount Wilson. Based on observations of variables in a number of galaxies, Hubble concluded that the spiral nebulae were much farther away than opponents of the island universe theory allowed.[10] This groundbreaking observational work convinced almost all astronomers of the validity of the island universe theory by the mid-20s. Hubble's work provided what had been lacking in earlier stages of the debate: a way of measuring distance to the spirals that island universe advocates and opponents both considered reliable.[11]

The great debate concerning the nature of the spiral nebulae was entangled with controversies in cosmogony.[12] This field focused on describing the origins and evolution of various astronomical structures, ranging in scale from the Earth-moon system to the solar system and beyond. After the turn of the century the Laplacian nebular hypothesis was under attack. Laplace had suggested that structures such as the solar system resulted from the condensation of gaseous nebulae, an account augmented by Herschel's claim that the spiral nebulae are proto-solar systems at an earlier stage of evolution. Theoretical development of the Laplacian idea drew on increasingly sophisticated mathematical treatments of a problem that is simple to state, if not to solve: what are the stable configurations of a rotating fluid according to Newtonian gravitational theory, and how do different configurations evolve over time?[13] The Americans Thomas Chamberlin and Francis Moulton coupled forceful criticisms of the nebular hypothesis with a new theory of the formation of the solar system proposed in 1905. Their account had two striking features: first, the planets were formed by the accretion of smaller hard fragments (planetesimals) rather than by condensation of a gaseous cloud, and second, an encounter with another star triggered formation of the planets. The Chamberlin-Moulton theory triggered refinements of the nebular hypothesis. James Jeans incorporated the second idea within a modified nebular hypothesis, according to which an encounter with a nearby star triggered the condensation of the hot gaseous cloud into the solar system.

The debate between the nebular hypothesis and Chamberlin-Moulton's alternative continued throughout the teens and twenties. Both theories appealed to a wide variety of astronomical phenomena to justify a number of speculative assumptions. In a 1913 review of two books on cosmogony, Karl Schwarzschild complained that the field was "heterogeneous" and "impenetrable" due to "prolixity, pretension, confusion, and the general lack of mathematical control" (Schwarzschild, 1913, 294). Many of the cosmogonists followed Herschel in appealing to observations of spiral nebulae, taken to be proto-solar systems at an earlier stage of evolution, to vindicate aspects of their accounts.[14] Schwarzschild singled out this idea as particularly dubious, and the developments described above vindicated his skepticism. But more significantly, the new understanding of the spiral nebulae dramatically increased the scale of the known universe. The concerns of cosmogony appear more parochial, although still important, once these differences of scale are recognized. Rather than appealing to results regarding the evolution of rotating fluids or the interactions of planetesimals, progress in the study of the structure and evolution of the universe at these largest scales came from an unexpected direction: a new theory of gravitation.

3. Relativistic Cosmology

The debates just described form the background for the reception and further development of relativistic cosmology. However, Einstein's foray into cosmology was not motivated by the problems described above, and he apparently had not kept abreast of new results regarding spiral nebulae and the ensuing controversies. Instead his groundbreaking 1917 paper began by highlighting an apparently embarrassing dilemma for Newtonian cosmology. Einstein emphasized this alleged flaw to cast his own newly minted gravitational theory in a better light, but the motivations for his first foray into cosmology lay elsewhere. Einstein's "rough and winding road" to cosmology was a continuation of the path he followed to the discovery of general relativity. Einstein created the field of relativistic cosmology as a final effort to guarantee that his new gravitational theory satisfied Machian ideas.[15]

3.1. **Paradoxes of Newtonian Cosmology.** Einstein (1917b) opened with the following dilemma. If matter is uniformly distributed throughout an infinite universe, Newtonian gravitational theory does not consistently apply, for reasons that we will consider shortly. Newtonian cosmology seems to allow only an alternative picture: an "island" of stars clumped together in an otherwise empty universe.

Einstein argued that even this possibility can be ruled out, since the island would be unstable. Einstein presented this as an inescapable dilemma for Newtonian cosmology, and his discussion of it served the dual purpose of exposing the inconsistencies of a preceding theory and preparing readers of his paper for a modification of his own gravitational theory.[16]

The problem Einstein noted regarding Newtonian cosmology can be stated quite simply. Einstein discovered the problem independently, but it has a long history stretching back to shortly after the publication of Newton's *Principia*, and Hugo von Seeliger had recently given a systematic exposition of the problem (Seeliger, 1895).[17] Einstein gave a characteristically straightforward formulation of the problem in a popular work (Einstein, 1917a, 71–72): the number of lines of gravitational force passing through a sphere around a point mass O enclosing a uniform matter distribution is given by $\rho_0 V$, where ρ_0 is the mass density and V the volume of the sphere. The force per unit area is then proportional to $\rho_0 R$, where R is the radius of the sphere. As $R \rightarrow \infty$, the force also diverges—a result Einstein called "impossible." Although Einstein did not elaborate further, the failure of convergence can be elucidated more clearly by considering cases such as that illustrated in Fig. 1 (following Norton, 1999). The gravitational force on O at the origin is given by the sum of forces due to hemispherical shells of thickness d surrounding O. Since the increase of mass ($\propto r^2$) cancels the decrease in gravitational force with distance ($\propto r^{-2}$), the force due to each shell is the same—it is given by $F = G\rho\pi d$, where G is Newton's gravitational constant and ρ is the mass density. The total gravitational force is then given by an infinite sum of equal terms with opposite signs, and this sum fails to converge.

Einstein carried the argument one step further by applying statistical physics. The divergence described above can be avoided if the mass density falls off outside a central "island" of matter. To see this, consider the gravtational potential ϕ, a solution to Poisson's field equation for gravitation:

$$\nabla^2 \phi = 4\pi G \rho, \qquad (1)$$

where $\nabla^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, ϕ is the gravitational potential, and ρ is the mass density. If the mass density function is well-behaved, then the potential is given by:

$$\phi(r) = \int G \frac{\rho(r)}{r} dV. \qquad (2)$$

This integral converges if $\rho(r)$ falls off faster than $1/r^2$ as the distance r from the central concentration increases. However, Einstein argued that this alternative is not viable since an "island of matter" is not stable. The stars composing the island can be treated like the molecules of an ideal gas. The island would "evaporate" as individual stars acquired enough kinetic energy to escape the gravitational attraction of the other stars, just as water molecules evaporate into the air. If the potential ϕ converged to a small, finite value at infinity, at thermal equilibrium

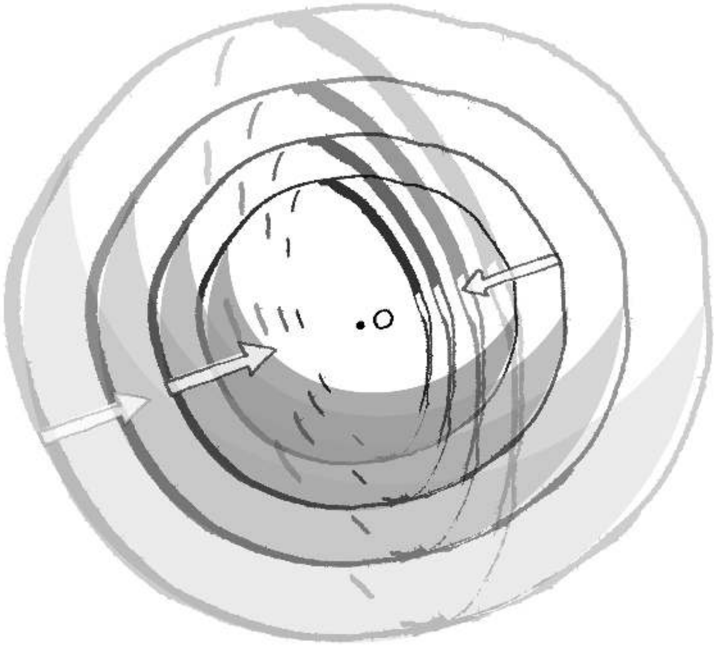

FIGURE 1. The total gravititional force on the point mass O is given by the sum of forces due to each hemispherical shell. Each half-shell contributes a force of equal magnitude, $F = G\rho\pi d$, and the sum does not converge.

the ratio between mass at the center and mass at infinity must take a finite value. Einstein concluded that requiring the mass distribution to vanish at infinity implied that it also vanished at the center, ruling out an "island" of concentrated matter.

Seeliger and his contemporaries had considered several ways to avoid the dilemma, from denying the universality of gravitation to ruling out an infinite, uniform matter distribution by *fiat*. Einstein and Seeliger both advocated a modification of Newton's inverse square law, effected in Einstein's case by simply adding a term to eqn. (1):[18]

$$\nabla^2\phi - \Lambda\phi = 4\pi G\rho. \tag{3}$$

The modified field equations admit a uniform field as a solution for a constant mass density ρ_0,

$$\phi = -\frac{4\pi}{\Lambda}G\rho_0. \tag{4}$$

However, Einstein and Seeliger did not have the same attitude towards a solution of the dilemma along these lines. Seeliger took the proposal seriously enough to suggest that astronomical measurements could constrain his modification of Newton's law. Einstein, on the other hand, commented that his solution "does not in itself claim to be taken seriously; it merely serves as a foil for what is to follow" (Einstein, 1917b, 543). Einstein was not primarily concerned with resolving the paradoxes of Newtonian cosmological theory, but the addition of a constant to Poisson's equation paves the way for the introduction of the infamous cosmological constant a few pages later.

Einstein's presentation of the dilemma facing Newtonian cosmology is at first compelling, but on closer examination the argument falls apart due to two significant oversights. First, a way between the horns of the dilemma had already been found. A decade before Einstein's paper, the Swedish astronomer Carl Charlier had explicitly constructed a "hierarchic cosmos," inspired by the speculations of Fournier D'Albe, that neatly avoided both problems discussed above (Charlier, 1908). The crucial trick was to produce a mass distribution which both avoided the divergences and lacked a preferred center. Charlier's model had a fractal structure: stars are grouped into spherical galaxies, galaxies into spherical metagalaxies, and so on (see Fig. 2). Charlier derived constraints on how densely the systems at one level could be "packed" into a system of the next level up without producing divergences. He showed that it was possible to build up a uniform cosmos in this way with an infinity of total mass, an average mass density of zero, and a convergent gravitational potential.[19] Newtonian gravitational theory could be consistently applied to such a hierarchic universe without leading to either of Einstein's problems.

Second, and more significantly, several later cosmologists have argued that the apparent inconsistency only reveals a shortcoming of a particular formulation of Newtonian cosmology. Some took Einstein's argument to establish only the incompatiblity of Newtonian cosmology with a *static* distribution of matter (see, e.g., Heckmann, 1942, 14). The "Neo-Newtonian" cosmological models developed in the 30s described a matter distribution that changed over time, and thus they seemingly avoided the inconsistency.[20] Geometric formulations of Newtonian theory, first introduced in the 20s, support the stronger claim that the apparent inconsistency can be avoided even for a uniform, static matter distribution.[21] There is a class of solutions to eqn. (1) for a uniform matter distribution with the following form,

$$\phi(\mathbf{r}) = \frac{2}{3}\pi G\rho|\mathbf{r} - \mathbf{r}_0|^2, \tag{5}$$

with distinct solutions corresponding to different choices of an arbitrary point $\mathbf{r}_0$.[22] The force $\mathbf{F}$ acting on a test particle with unit mass is then given by the gradient of the potential:

$$\mathbf{F} = -\nabla\phi = -\frac{4}{3}\pi G\rho(\mathbf{r} - \mathbf{r}_0), \tag{6}$$

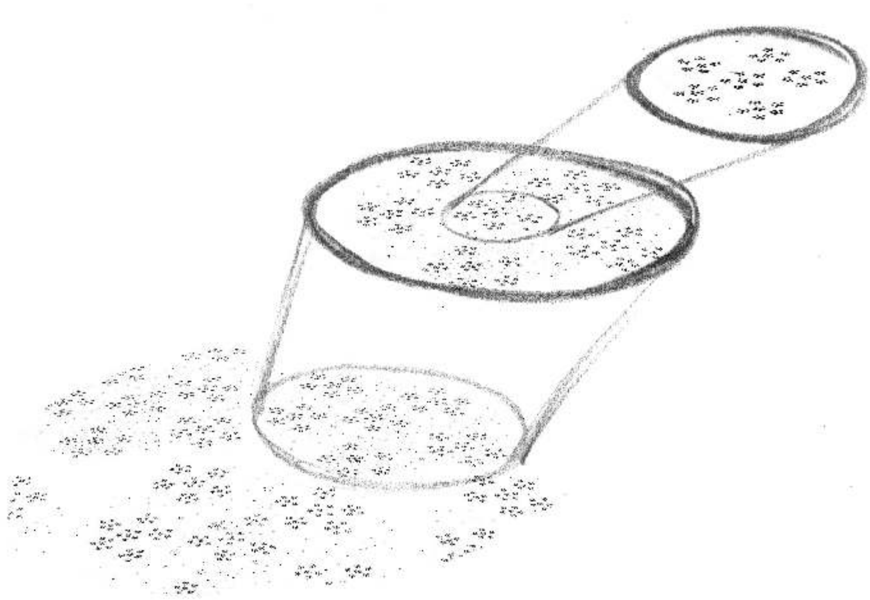

FIGURE 2. Charlier's hierarchical universe has a fractal-like structure, and it avoids divergences without introducing a preferred central region.

which represents a spherically symmetric field of force directed at the "center point" $\mathbf{r}_0$, where the force vanishes. The original problem resurfaces, in that a completely *uniform* matter distribution mysteriously leads to a *nonuniform* force field with an arbitrary preferred central point at $\mathbf{r}_0$. However, these peculiar properties of the solutions have no directly observable consequences. The geometric formulation modifies Newtonian theory to incorporate what Norton (1995) aptly called the "relativity of acceleration": the choice of how to decompose a given free-fall trajectory into inertial motion and gravitational deflection is conventional. The indeterminacy noted by Einstein is just a consequence of this feature, since there appears to be a physical difference between different choices of the decomposition. In the geometrical formulation the choice of a particular decomposition has as little significance as the choice of a rest frame within special relativity.

Einstein's response to the unravelling of his dilemma for Newtonian cosmology illustrates its relatively minor role in his thinking. Franz Selety (1922) spelled out how Charlier's hierarchic model avoided the dilemma in painstaking detail. In his brief response, Einstein (1922) conceded the point before reiterating the line of reasoning leading to his own cosmological model, giving Mach's principle the leading role.[23]

3.2. **Mach's Principle.** Einstein turned to cosmology after conversations with De Sitter in the fall of 1916 convinced him that an idea that had played a crucial role in his discovery of general relativity was at stake. "Mach's principle" (as Einstein later dubbed it) is difficult to formulate precisely, but Einstein hoped to capture Mach's idea that inertia derives from the distribution of matter rather than space itself. Einstein also called this the "relativity of inertia," which he put as follows: "In a consistent theory of relativity, there can be no inertia *relative* to 'space,' but only an inertia of masses *relative to one another*" (Einstein, 1917b, 145, original emphasis).[24] The discussions and correspondence with De Sitter forced Einstein to consider the status of this principle more carefully, and he struggled to give it a sharper formulation and to insure that it held in his new gravitational theory.[25] Einstein's famous cosmological model, and the infamous cosmological constant, were both by-products of this effort.

One threat to Mach's principle came from the need to stipulate boundary conditions in order to solve the gravitational field equations.[26] For example, Schwarzschild derived his expression for the gravitational field of a single body such as the sun by requiring (among other things) that the solution resembles Minkowski space-time, the flat space-time of special relativity, at infinity. This boundary condition captures the natural requirement that at great distances from the source the gravitational field should approach the "empty space" solution to the field equations. There is nothing unusual about imposing boundary conditions in order to find a solution to a set of field equations. But for Einstein introducing boundary conditions meant allowing a vestige of absolute motion to creep back into the theory. The inertia of a particle far from the sun in Schwarzschild's solution would be fixed by the space-time structure imposed at infinity as a boundary condition rather than just its relation with other matter. For Einstein this was contrary to the spirit, if not the letter, of his theory. The structure at infinity could be used to distinguish states of motion absolutely, without reference to other matter. On Einstein's view, there should be no such structures attributed to space itself in his new theory. Vacuum solutions clearly violated this requirement, because they possess structure that cannot be due to matter. But boundary conditions also violated this requirement by introducing spatial structure "at infinity" that is independent of the matter distribution.[27]

Einstein proposed two ways to banish this last remnant of "absolute space" in letters to De Sitter from 1916 and 1917. His first proposal was to stipulate that at infinity the metric field takes on "degenerate values" (all components equal to 0 or ∞) rather than approaching a flat, Minkowski metric. Einstein (1917b) argued that the degenerate boundary values follow from requiring that the inertia of a body drops to zero as it approaches infinity. This was intended to capture the idea that there would be no "inertia" relative to space itself.[28] A fundamental problem with this proposal is the difficulty with specifying the nature of a solution "at infinity" in a manner that does not depend on the choice of coordinates. A more obvious problem led Einstein to abandon this line of thought fairly quickly. A flat Minkowski metric appeared to be compatible with stellar motions on the largest scale yet observed. Einstein's proposal required postulating "distant masses" that would somehow reconcile degenerate values of the metric field "at infinity" with these observations. De Sitter did not share Einstein's Machian intuitions, and the lack of a Machian explanation of inertia did not bother him as much as this *ad hoc* introduction of such unobserved distant masses (CPAE 8, Doc. 272). Einstein soon dropped this first proposal in favor of a more radical way of handling the problem of boundary conditions.

Einstein's ingenious suggestion was to do away with the problem of boundary conditions by getting rid of boundaries entirely. He introduced a cosmological model built up from spatial sections that describe the universe at a given cosmic time. These spatial sections are three-dimensional analogs of a two-dimensional sphere such as the Earth's surface. Like the Earth's surface, each spatial section is finite yet unbounded, leaving no place to assign boundary conditions. This cosmological model satisfied two strong idealizations. First, the model describes a simple uniform mass distribution without the lumps and bumps observed by astronomers.[29] Second, Einstein argued that the model should be static, which means that the properties of the model do not vary with time; each spatial section "looks like" any other.[30] This second requirement was justified with an appeal to astronomical observations. Since the observed relative velocities of the stars are much smaller than the speed of light, Einstein argued that there is a coordinate system such that all the stars are permanently at rest (at least approximately). This is a weak justification for such a crucial assumption. As De Sitter promptly objected in correspondence, "we cannot and must *not* conclude from the fact that we do not see any large changes on this photograph [our observational "snapshot" of stellar distributions] that everything will always remain as at that instant when the picture was taken" (CPAE 8, Doc. 321, original emphasis).[31] With the benefit of hindsight it is easy to fault Einstein for so quickly ruling out solutions that change over time on such weak grounds. Yet many of his contemporaries followed his lead in limiting consideration to static models. Even De Sitter, despite recognizing the weakness of Einstein's argument, nonetheless failed to study dynamical models until 1930. In any case, Einstein was not primarily interested in a detailed comparison between his cosmological model and observations. He frankly explained his motives in introducing the model in a letter to De Sitter: he called it "nothing but a spacious castle in the air," built to see whether his Machian ideas could be consistently implemented (CPAE 8, Doc. 311).

The castle in the air came at a price. Einstein was forced to modify the field equations of general relativity in order to admit this cosmological model as a solution. All that was required was the straightforward addition of a cosmological constant term, Λ, a modification Einstein had foreshadowed with the similar change to Poisson's equation discussed above:[32]

$$R_{\mu\nu} - \Lambda g_{\mu\nu} = -\kappa \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right). \tag{7}$$

This equation specifies the relationship between the space-time geometry (the left hand side) and the distribution of matter and energy (the right hand side).[33] Einstein showed that if these modified field equations are to hold in his model, the cosmological constant Λ is fixed by the mean density of matter ρ, namely $\Lambda = \kappa\rho/2$. The value of Λ is also related to the radius of the spatial sections R by the equation $\Lambda = 1/R^2$. Combining the two equations made it possible to estimate the "size of the universe" based on the mean density of matter.

The addition of the cosmological constant is a fairly natural generalization of the original field equations, and it satisifies the formal and physical constraints that Einstein had used in deriving the original equations.[34] Physically, a positive Λ term represents a "repulsive force" that counteracts the attraction of gravitation.[35] Observations available to Einstein could have been used to show that Λ must be *incredibly* close to zero, since even a slight deviation from zero affects the general relativistic predictions for motions within the solar system. (De Sitter argued that Λ was certainly less than 10^{-45}cm^{-2} and probably less than 10^{-50}cm^{-2} (CPAE 8, Doc. 327), and later calculations such as those reported in Tolman (1934) gave even tighter constraints.) However, even a very small nonzero value of Λ changes the space of solutions to the field equations. Einstein's static universe is a solution of the modified field equations, and a non-zero Λ term also rules out solutions of the original field equations such as Minkowski space-time, a welcome result for Einstein.[36]

Einstein did not have long to rest easy with the modified field equations and his ingenious way of saving Mach's principle. The introduction of Λ was not actually necessary for Einstein's way of eliminating boundary conditions to work; it was only necessary given the additional assumption that the model must be static. Einstein (1917b) concluded with a remark suggesting that it might be possible to construct time-varying models with closed spatial sections without the Λ term.[37] Such models would do away with the need to specify boundary conditions without Λ, undermining the rationale for its introduction. But Einstein must have been surprised to discover that the introduction of Λ was also not sufficient to guarantee that Mach's principle holds in general relativity. After learning of Einstein's solution, De Sitter promptly produced a second solution to the modified field equations. Since it was a vacuum solution (with non-zero Λ) it apparently violated Mach's principle. As we will see in the next section, until 1930 much of the literature in relativistic cosmology focused on understanding the properties of these two models.

4. Cosmological Models: *A* versus *B*

Einstein and De Sitter's published papers brought their debate to the wider scientific community. Eddington had invited De Sitter to provide a précis of Einstein's new gravitational theory in the *Monthly Notices*, and De Sitter responded with a series of three clear articles that introduced Einstein's theory—along with De

Sitter's own cosmological model—to the English-speaking world. By the mid 20s a handful of mathematical physicists had begun to study the properties of the Einstein and De Sitter solutions. Part of this literature focused on the question of whether the De Sitter solution could be dismissed on the grounds that it included a "singularity," with a debate regarding both the content of that charge and whether it applied. This topic was part of the broader debate regarding what properties a cosmological model should have to qualify as physically reasonable. The major conceptual innovation introduced by Einstein was the very possibility of a mathematical description of the universe as a whole, but it was not immediately clear what observational and physical content these abstract models possessed. With the benefit of hindsight, this early exploration of cosmological models was limited by illicit assumptions regarding what counts as physically reasonable. In particular, it was over a decade before the community overcame Einstein's insistence on static models. Several theorists followed De Sitter's lead in seeking out a stronger connection between cosmological models and astronomical observations. With his strong empiricist bent, De Sitter actively explored the connection between his model and observations, suggesting in particular an explanation of the redshift in the spectral lines of the spiral nebulae. This point of contact became far more important on the heels of Hubble's observational results published in 1929.

In the course of these explorations of cosmological models, this first generation of relativists encountered a number of the most striking novelties introduced by Einstein's theory of gravitation. To determine a particular solution's physical properties, and to find out whether it harbored any singularities, relativists had to differentiate artificial features due to the coordinates being used from genuine features of the space-time. This was no simple matter, given that they lacked the mathematical tools developed later for isolating the intrinsic features of a space-time. In addition, as Eisenstaedt (1989) has emphasized, in cosmology theorists faced a variety of situations far removed from familiar problems studied in Newtonian theory. (In fact, Newtonian cosmology and Newtonian analogs of the expanding universe models were studied systematically several years later, based on insights from general relativity.)

4.1. **The Many Faces of Model B.** De Sitter derived his solution as a variation on Einstein's theme: rather than taking only the spatial sections to be closed spheres, de Sitter treated the entire space-time manifold as a closed space.[38] In his letter to Einstein introducing the solution (CPAE 8, Doc. 313) as well as his publications (de Sitter, 1917a,c), De Sitter laid out the two solutions side by side—labeling Einstein's model "A" and his own model "B" in the published papers—to emphasize their similarities. However, B differed from A in several crucial respects. Perhaps most importantly, De Sitter's solution proved to be much harder to grasp, leaving most theorists to rely too heavily on particular coordinate representations; Eddington (1923) characterized Einstein's solution as "commonplace" compared to the complex geometry of De Sitter's solution.[39] De Sitter's solution was a vacuum solution of the modified field equations, with $\Lambda = 3/R^2$ and $\rho = 0$. Unlike Einstein's solution, the cosmological constant is *not* related to the matter density. For De Sitter this counted as an advantage of his solution, since he took Einstein's model to include "world matter" (responsible for producing the Λ term) in addition to ordinary matter. However, Einstein insisted that the uniform matter distribution in his own model was merely an approximation to ordinary matter. A more realistic model would be inhomogeneous but still static, and such a model would be to the cylinder universe as "the surface of a potato to the surface of a sphere" (CPAE 8, Doc. 356). De Sitter's attitude towards his own solution was at first fairly skeptical, and he professed a preference for model C, his label for flat Minkowski space-time, over both A and B, since model C was a solution of the original field equations without Λ. This initial skepticism eventually faded somewhat, and De Sitter took his model seriously enough to study its observational consequences, as we will see below.

The De Sitter solution appeared in a variety of guises over the following decade, as Einstein, De Sitter, and others used different coordinates to study its properties.[40] In retrospect it is easy to fault these discussions for failing to disentangle properties of the De Sitter solution from properties that reflect a particular coordinate representation of it. The participants in the debates were aware of the need to focus on invariant quantities that do not depend upon the choice of coordinates, but in practice this fundamental principle was often not heeded. The issue is a more subtle analogue of a familiar problem with maps of the Earth's surface. Representing the Earth's surface on a two-dimensional map inevitably produces distortions. Knowledge of the underlying geometry of the Earth's surface makes it easy to avoid confusing these distortions due to the map with actual features of the surface. For example, a Mercator projection breaks down at two points (such as the north and south poles), but it is easy to see that this is due to the projection and not to some special feature of the surface at those points. Einstein, De Sitter, and other cosmologists did not have a similar grasp of the geometry of the De Sitter solution to fall back on in assessing odd features of the solution revealed in different coordinate systems.

Einstein and De Sitter's discussion of model B eventually focused on a particular, and quite misleading, choice of coordinates. After originally presenting the solution geometrically, De Sitter wrote his solution in a static form to facilitate comparison with Einstein's model (CPAE 8, Doc. 355). The transition to this set of coordinates appears to have been driven by Einstein's preference for a static model, and in this guise the De Sitter solution appears to represent, like Einstein's model, a static space-time with uniform spatial sections. But this impression is misleading. Suppose that we consider freely falling particles that are initially at rest with respect to the background matter distribution. In Einstein's model the distance between such particles measured on successive spatial sections remains constant. This is just to say that the trajectories followed by the freely falling particles (timelike geodesics) coincide with the curves defined by fixed values of the spatial coordinates. This is not the case in De Sitter space-time. In De Sitter space-time, the distance between the particles changes over time—they either scatter or move closer together rather than staying at a fixed distance. This is a consequence of the fact that the curves defined by fixed values of the spatial coordinate, using static coordinates, are not timelike geodesics.[41]

Unlike Einstein, De Sitter and later Lemaître (1925) took the non-static character of the De Sitter solution to be a positive feature, since this might explain the receding motion of the spiral nebulae (an idea we will return to below). Lemaître's paper introduced a set of coordinates independently discovered by Robertson (1928). The De Sitter solution appears strikingly different in these two coordinate systems. The contrast between the static coordinates and the new coordinates is illustrated in Fig. 3. The global geometry of the De Sitter solution can be represented as the surface of a hyper-hyperboloid embedded in a five-dimensional Minkowski space (see Fig. 9 of Janssen's contribution).[42] The static coordinates used by De Sitter and Einstein only covered two wedge-shaped portions of the surface (see Fig. 3a). The spatial sections are bounded and overlap at two points in the diagram (corresponding to a 2-dimensional surface in the full solution), called the "mass horizon." The resulting degeneracy of the time coordinate reflects on the choice of coordinates rather than on the underlying geometry of the solution—the hyperboloid is entirely regular at these points. The solution takes on a very different guise in the Lemaître-Robertson coordinates. These coordinates cover only the upper half of the hyperboloid (see Fig. 3b). The diagram illustrates a striking contrast between the properties of spatial sections for the two cases: in the static coordinates, the spatial sections have finite volume, whereas in Lemaître and Robertson's coordinates the spatial sections have *infinite* volume. Historically the Lemaître-Robertson coordinate system served as much more than a reminder of the need to focus on invariant quantities in assessing the properties of a solution. For both Lemaître and Robertson, careful study of the De Sitter solution played a part in the discovery of more general non-static cosmological models.

4.2. **Singularities.** Starting with his first response to the new model, Einstein gave a series of arguments intended to rule out the De Sitter solution on physical grounds. Clearly Einstein hoped to do away with this counter-example to Mach's principle by revealing some fatal flaw unrelated to the fact that it was a vacuum solution. The most important and interesting charge was that model B harbored a "singularity," roughly a point or region where the metric field is ill-behaved.[43] If the metric field "blows up" to an infinite value or drops to zero, quantities appearing in the field equations would no longer be well-defined. In this sense a singularity would signal the breakdown of the theory, a result Einstein regarded as clearly unacceptable.

In order for the existence of singularities to be used as a criterion for separating physically reasonable cosmological models from the unclean, the concept of

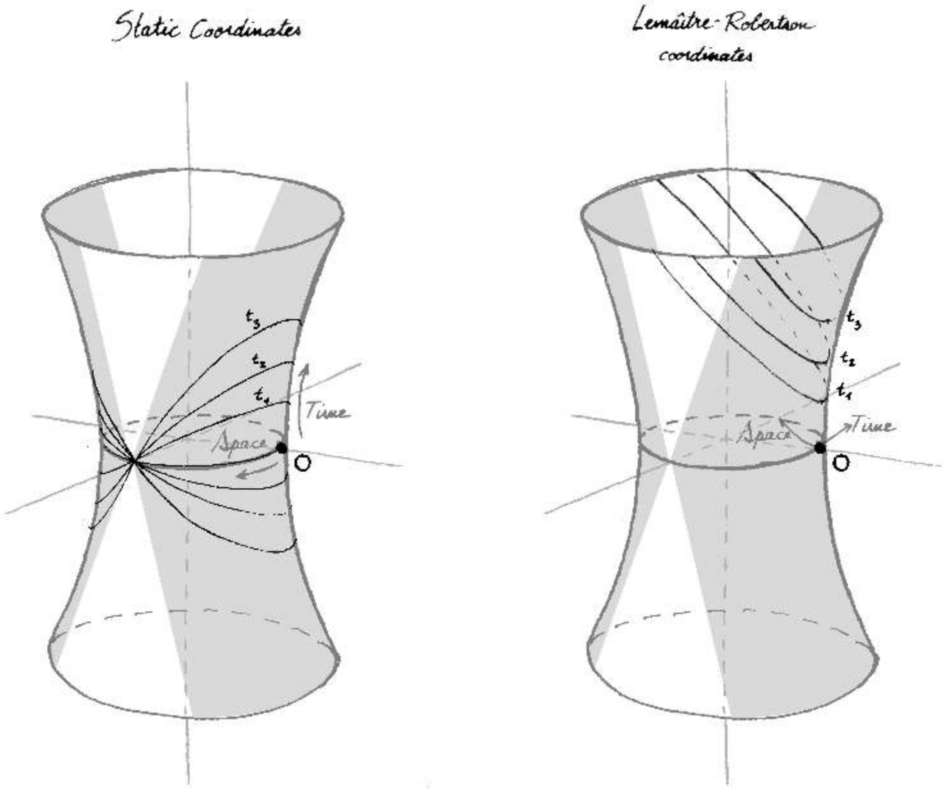

FIGURE 3. Coordinate representations of the De Sitter solution. (a) Static coordinates. These cover only the two wedge-shaped regions of the hyperboloid. The spatial slices have finite volume, and they overlap at the "mass horizon," the point p (and a point on the opposite side of the hyperbola). (b) Lemaître-Robertson coordinates. These cover only the upper half of the hyperboloid, and the spatial sections have infinite volume.

a "singularity" had to be defined more precisely. Giving such a definition has turned out to be a surprisingly subtle problem, and a widely accepted and useful definition of singularity was not in place until the mid 60s.[44] A difficult first step towards a satisfactory definition involves distinguishing coordinate effects from true singularities. For example, the apparent singularity at the origin of polar coordinates is clearly an artifact of the coordinates. Einstein took this first step in Einstein (1918c): for a space-time to qualify as non-singular, he required that for every point lying in regions within a "finite distance," there must be a coordinate system in which the metric is well-behaved.[45] More intuitively, for a space-time satisfying the definition, all the regions within a finite distance from some starting point could be covered by a patchwork of coordinate systems in which the metric is continuous. The metric for the De Sitter solution in static coordinates fails to satisfy this definition because it breaks down at the "mass horizon," the points in Fig. 3a (surfaces in the full solution) where the spatial sections intersect.[46] Einstein argued that the De Sitter solution should be rejected as unphysical due to this alleged singularity.

Einstein called this surface the "mass horizon" since he expected that the singular behavior of the metric indicated the presence of matter. Anticipating results about to appear in Weyls book *Raum-Zeit-Materie*, which he was reading in proof, Einstein suggested that the De Sitter solution was after all *not* a vacuum solution (Einstein 1918c, 272). Weyl constructed a general solution by combining three different solutions, including one representing an incompressible fluid, and argued that the original De Sitter solution could be recovered in the limit as the fluid zone goes to zero volume.[47] Einstein welcomed this result, since it insured that the De Sitter solution could no longer be regarded as a counterexample to Mach's Principle. The De Sitter solution was no vacuum solution—the mass had just been "swept away into unobserved corners," as Eddington (1923) put it.

Einstein's claims regarding the singularity in the De Sitter solution provoked a number of critical responses. Klein's analysis brought out the central defect of Einstein's assessment quite clearly (at least for those familiar with projective geometry): in Klein's geometrical representation of the solution, the hyperboloid is an entirely regular surface without any kinks or discontinuities that would correspond to a genuine singularity. Thus the alleged singularity had to be an artifact of a poor choice of coordinates. This point was not clear to Klein's contemporaries. Eddington, for example, gave a geometrical description of the de Sitter solution similar to Klein's, followed with an inconclusive discussion of the mass horizon idea and singularities (Eddington, 1923, 164–166). Eddington suggested incorrectly that the De Sitter solution could not be fully covered by a single, entirely regular coordinate system. The young Hungarian physicist Cornelius Lanczos, drawing on Klein's work, showed that Eddington was mistaken by explicitly writing the De Sitter solution in a form that was globally singularity free (Lanczos, 1922).[48] Klein did convince Einstein of the error of his ways, and although Einstein admitted the mistake in letters to Klein and to De Sitter's colleague Paul Ehrenfest (CPAE 8, Doc. 664), he never publicly retracted his earlier criticism.

However, accepting that the De Sitter solution was free of singularities did not mean that Einstein accepted it as a physically reasonable model. Instead he rejected the De Sitter solution because it was non-static. The requirement that any physically reasonable model must be static seems to have emerged opportunistically as a way to rule out De Sitter's anti-Machian solution. At least, Einstein gives no indication of a deeper or more plausible motivation for this requirement. What started as a weakly motivated simplifying assumption in deriving his own model had thus turned into a substantive constraint, a constraint that Einstein did not critically examine for the next decade. If he had, he would have discovered that non-static models are much more natural in general relativity than the static model he preferred.

This early study of singularities shows the first generation of relativists grappling with difficult problems raised by Einstein's new theory, and the study of the De Sitter solution shaped other areas of research. Lemaître (1932) gave an insightful study of the alleged singularity in the Schwarzschild solution that drew on his earlier work regarding the De Sitter solution, as Eisenstaedt (1993) has emphasized, and the study of alternate coordinate representations of the De Sitter solutions was also connected with Lemaître's discovery of expanding models, as it was for H. P. Robertson. Furthermore, with the benefit of hindsight we can see in these early debates the first steps towards differentiating two distinct features of cosmological models. The odd behavior of the metric that Einstein took to signal the presence of a singularity instead shows that there is an "event horizon" in the De Sitter solution. De Sitter's initial response to Einstein, in which he emphasized that the alleged singularity could be ignored since it was physically inaccessible, contained a kernel of truth. The event horizon demarcates regions from which light signals can and cannot reach a given observer; regions beyond the event horizon are thus physically inaccessible to the observer. The event horizon is not an intrinsic feature of space-time, since it is defined relative to a given observer.[49] The study of horizons has been an important part of cosmology since its earliest days, and the presence of horizons in the standard cosmological models is one of the puzzles used to motivate a recent modification of the standard cosmological models known as inflationary cosmology.

4.3. **Redshift.** The previous sections focused on general criteria for what might count as a reasonable cosmological model. De Sitter, with his strong empiricist bent, added another way of comparing models A and B in one of his early discussions (de Sitter, 1917b): he predicted a redshift in the spectral lines of light emitted from distant objects, with the amount of redshift a function of the distance. In addition to calculating the effect, De Sitter took an important step in suggesting that the movements of the spiral nebulae rather than the stars (which Einstein had focused on) should be used as gauges of cosmic structure on the largest scales. The nature of the redshift effect and the precise functional dependence of redshift on distance for the De Sitter model were both matters of substantial controversy for the following decade and a half. But by 1930, Hubble's observational work had both convinced many astronomers that a linear relationship between redshift and distance obtained and promoted the idea that De Sitter's model explained the effect.

De Sitter recognized that the redshift effect would be a useful way to compare observations with the new cosmological models. Spectral lines have precise wavelengths that can be measured in the laboratory, making it possible to detect a systematic displacement towards either the red or blue end of the spectrum in light received from distant objects. One cause of such displacement is the Doppler effect, due to the relative motion of the emitter and receiver of a light signal. Light received from an object moving towards the receiver is shifted towards the blue end of the spectrum (by an amount that depends on the relative velocity of the emitter), whereas light from an object moving away from the receiver is shifted towards the red end of the spectrum. (The changing frequency of the siren of an ambulance approaching and receding illustrates this effect for sound waves.) Cosmology textbooks typically include a stern warning not to confuse Doppler shifts with cosmological redshift. The Doppler shift is caused by the motion of the emitter, whereas the cosmological redshift results from the stretching of the wavelength of light with the expansion of space (see Fig. 4). The nature of the cosmological redshift was, however, not so clear to De Sitter and his contemporaries.

The disputes regarding cosmological redshifts can be traced back to two basic problems.[50] The most important is that calculations of the redshift effect require choosing the worldlines representing the observer and the observed object, or more generally a set of worldlines for several objects. Without this specification a cosmological model is fundamentally incomplete. From a modern point of view choosing the appropriate wordlines is often completely straightforward; one might choose, for example, worldlines for freely falling test particles. These particles follow geodesics, lines of extremal length, through the space-time.[51] Weyl recognized the need to specify the class of curves traced out by objects in a cosmological model, and proposed using a set of geodesics that converge at a point in the past to represent the worldlines of galaxies, based on a fairly obscure causal argument (see Goenner, 2001). In the early debates these choices were not always clearly stated, and the confusion was exacerbated by the different coordinate representations of the models being considered. For example, in the static coordinates frequently used by De Sitter, the worldlines for particles at rest with respect to the coordinates are not geodesics. Particles following worldlines that appeared to be "at rest" with respect to static coordinates were in fact undergoing acceleration as they departed from geodesic motion. Second, within general relativity the cosmological redshift (or blueshfit) can be calculated for a specific cosmological model given a choice of curves traversed by the observer and emitter. While the result of this calculation is independent of the choice of coordinates, as it should be, identifying the causes of the calculated redshift does depend upon this choice. There are six distinct causes of redshift in a cosmological model.[52] In different coordinate frames, the calculated redshift effect will be traced back to different causes. Theorists at the time often did not take all of these different causes into account, failed to draw the appropriate distinctions among them, and did not appreciate how their arguments and calculations depended on the choice of coordinates. Calculations of the redshift effect in the same model often led to conflicting results due to these two problems. De Sitter initially argued that the redshift should increase quadratically with distance in his model B, whereas Weyl and Robertson (1928) later derived a linear redshift-distance relation.

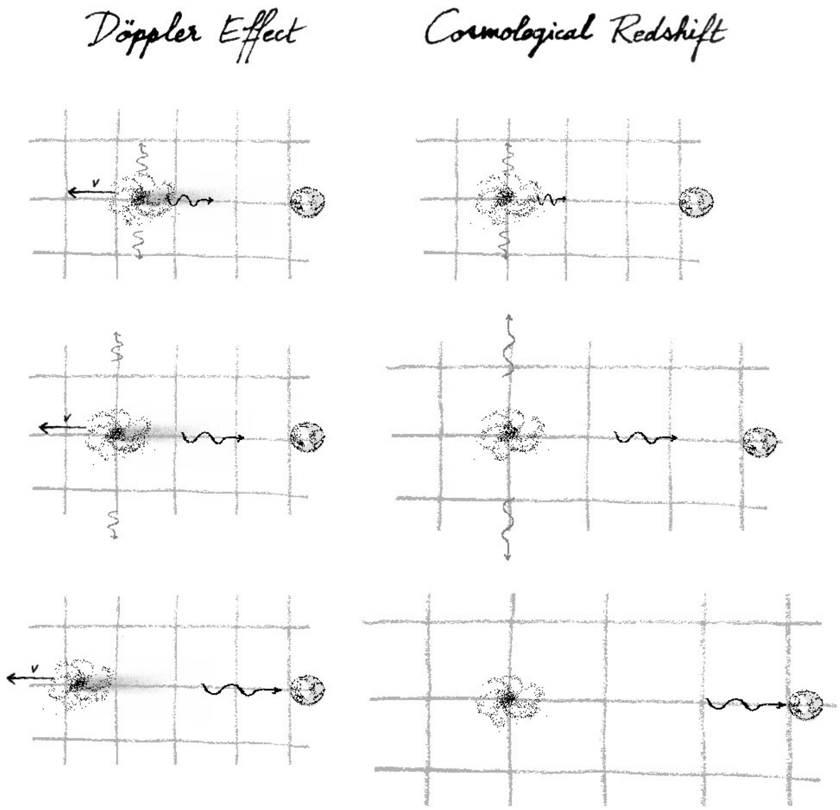

Figure 4. Contrast between Doppler effect and cosmological redshift. The Doppler effect is due to the motion of the emitter, which effectively "stretches" the wavelength of the emitted light. Cosmological expansion also leads to redshift, because light emitted from distant objects is "stretched" by the expanding space.

These disputes regarding how to calculate and understand the redshift occurred even among those who accepted general relativity and its cosmological models. There were also many astronomers who threw out cosmological redshifts with general relativity. The relativistic account of cosmological redshifts, such as it was, met with widespread skepticism among astronomers, leading to a variety of alternative explanations (such as Fritz Zwicky's "tired light" hypothesis). None of these alternative explanations has been successful.

On the observational side, the main difficulty was finding a way to reliably measure distances to the spiral nebulae. Vesto Slipher had begun measuring redshifts in the teens, and Eddington (1923) reported that of Slipher's 40 nebulae, 36 exhibited redshifts, on average quite large. Slipher made his observations with a small 24-inch refractor at the Lowell Observatory, and in 1928 Hubble and his collaborator M. L. Humason turned the 100-inch Mount Wilson telescope to the task of measuring redshifts. But in addition to the telescope power, Hubble was able to add another crucial component—reliable distances to the nebulae based on observations of Cepheid variables. By the time he published his results Hubble had redshift and distance measurements (of varying quality) for 46 nebulae, and the data fit a linear redshift-distance relation (Hubble, 1929).[53] Hubble had made one of the outstanding discoveries of twentieth century astronomy.

5. The Expanding Universe

After a decade of study and debate focused mainly on the Einstein and De Sitter models, several cosmologists belatedly re-examined their options. De Sitter's presentation to the Royal Astronomical Society in January of 1930 aptly summed up the sorry state of affairs (R.A.S., 1930). Rather than making a decisive case for one model or another, De Sitter concluded that *neither* model was completely satisfactory.

On the one hand, Einstein's static model failed to account for the observed redshifts of the spiral nebulae. Eddington had recently uncovered another serious defect of the model. Einstein's model treated the distribution of matter (represented by the cosmological constant) as perfectly uniform, and Einstein had assumed that it would be possible to construct a more realistic model that treated "normal" matter as lumps or condensations within this uniform background.[54] However, Eddington showed that Einstein's solution was unstable: a slight departure from uniformity would trigger runaway expansion or contraction (Eddington, 1930). Departures of the mass density from the value in Einstein's model ($\Lambda = \kappa\rho/2$) are enhanced via dynamical evolution, because the mass density changes while the cosmological constant remains fixed under expansion or contraction. A local concentration of matter triggers contraction; the effect of the contraction is to increase the matter density, which then differs even more from the (fixed) value of the cosmological constant, leading to further contraction. (Similarly for the case of a local deficit in density, triggering expansion.) As Eddington emphasized, the result follows directly from an equation Lemaître had derived from eqn. (7), and it is surprising that the instability escaped notice for so long.[55] Eddington did not immediately discard Einstein's solution due to the stability problem; instead, he followed Lemaître in suggesting that Einstein's model might describe the initial state of the universe, with some event triggering a transition to the De Sitter model. But once this instability was recognized, it was no longer possible to treat the model as Einstein had, as an approximation to the real, lumpy universe.

On the other hand, De Sitter's model predicted a redshift effect apparently compatible with observations. But it also failed to give a realistic approximation to the observed universe. De Sitter calculated the mean density of matter required in his model based on observed redshifts, and concluded that it was too high to be approximated as a vacuum.[56] As Eddington (1933, 46) put the question, "Shall we put a little motion into Einstein's world of inert matter, or shall we put a little matter into de Sitter's Primum Mobile?"

One answer to the dilemma was to drop one of the assumptions that led to such a short list of viable models. By the time of De Sitter's presentation, both Eddington and De Sitter had finally begun to question in print the assumption that any physically reasonable model must be static.[57] (Although De Sitter immediately criticized Einstein's argument that the universe must be static in 1917, he did not advocate a systematic study of dynamical solutions in print until 1930.) The American cosmologist Richard Tolman had reached a similar conclusion in the previous year, after surveying the problems facing all of the static models (Tolman, 1929b).[58] All three suggested that dynamical cosmological models were a suitable topic for further research—only to discover that these models had already been studied during the twenties. After reading the proceedings of the January 1930 Royal Astronomical Society meeting, Georges Lemaître wrote to remind Eddington of a study of dynamical solutions he had published three years earlier (Lemaître, 1927).[59] Eddington promptly advertised Lemaître's "brilliant work" in a letter to *Nature* and published a translation of Lemaître's paper in the *Monthly Notices* (Lemaître, 1931), where it would reach a wider audience than the Belgian journal in which it had originally appeared.[60] Lemaître had investigated evolving models without knowledge of an even earlier pair of papers by Friedmann (1922, 1924). Although Friedmann's papers appeared in the prominent German journal *Zietschrift für Physik*, they also garnered little attention. Einstein's (1922) initial claim that Friedmann's results rested upon a mathematical mistake can be partially blamed for this lack of interest. To paraphrase de Sitter (1931, 584), most cosmologists discovered expanding universe models in 1930, several years after they had first appeared in print.

These models can also be derived elegantly by appealing to symmetries, as Robertson and A. G. Walker showed in 1935. As a result they are often called the Friedmann-Lemaître-Robertson-Walker (FLRW) models. Robertson (1929) arrived at expanding models in a mathematical study of all possible solutions to Einstein's field equations with uniform spatial sections. Six years later he considered the relationship between these relativistic models and "kinematical relativity" advocated by E. A. Milne (1935). Milne aimed to derive cosmological models by extending the kinematical principles of *special* relativity. The project was motivated by an operationalist methodology and a conviction that Einstein's *general* relativity was unsound.[61] Robertson (1935, 1936a,b) and Walker (1935, 1936) independently showed that Milne's distinctive approach led, ironically, to the same set of basic models already studied in relativistic cosmology.[62] Walker showed that the FLRW metric follows from Milne's basic principles: the "cosmological principle," which requires observational equivalence among fundamental observers, along with a further symmetry principle stating that the model is spherically symmetric around each fundamental observer.[63] Robertson and Walker applied group theory to derive the consequences of such strong symmetry principles for the space-time metric. Their work showed how the FLRW metric follows from simple geometrical assumptions.

The discovery of the FLRW models made it clear that the Einstein and De Sitter solutions are special cases of a more general class of solutions. Friedmann retained two of Einstein's starting assumptions: that the space-time can be decomposed into three dimensional spatial sections corresponding to different cosmic times t, and that these sections are uniform, without preferred locations or directions in space.[64] He then showed that these two assumptions could be satisfied in dynamical models, in which the curvature of the spatial sections changes with time. The change in curvature corresponds to expansion or contraction in the sense that the distance measured on any spatial section between freely falling test bodies changes with time; the so-called "scale factor" $R(t)$ measures the rate of change of this distance (see Fig. 5).[65] Using the simplifying assumptions stated above, Friedmann showed that Einstein's field equations—in general a complicated set of 10 non-linear, coupled partial differential equations—reduced to a simple pair of differential equations for $R(t)$.[66]

The set of solutions to these equations describe simple dynamical models that have since become the standard models of relativistic cosmology. There are three classes of solutions to this pair of equations, typically classified according to the curvature of the spatial sections and the corresponding geometry. Friedmann (1922) discovered what are now called the spherical models, whose spatial sections have positive constant curvature and finite volume (like a sphere), and Friedmann (1924) introduced hyperbolic models, with infinite spatial sections of constant negative curvature. Robertson (1929) discovered the intermediate flat case between these two, namely infinite spatial sections with zero curvature. For all three cases, the two dynamical equations relate the scale factor $R(t)$ to the properties of the matter filling the space-time. From the equations for $R(t)$ one can define a critical density that divides the three cases described above. In addition, normal matter[67] leads the expansion of the universe to decelerate, i.e., $\ddot{R}(t) < 0$. The flat solution corresponds to a universe with exactly the critical density, and the initial velocity of the expansion is delicately balanced with this deceleration due to gravitational attraction. Models with less than the critical density have hyperbolic spatial sections and expand forever, whereas the spherical models have sufficient matter density to stop and reverse the initial expansion. The equations for $R(t)$ also illustrate that the simplest cosmological models are inherently dynamical. Einstein's preference for a static model was satisfied by choosing precisely the value of Λ needed to counteract the attraction of gravitation and yield $\ddot{R}(t) = 0$.

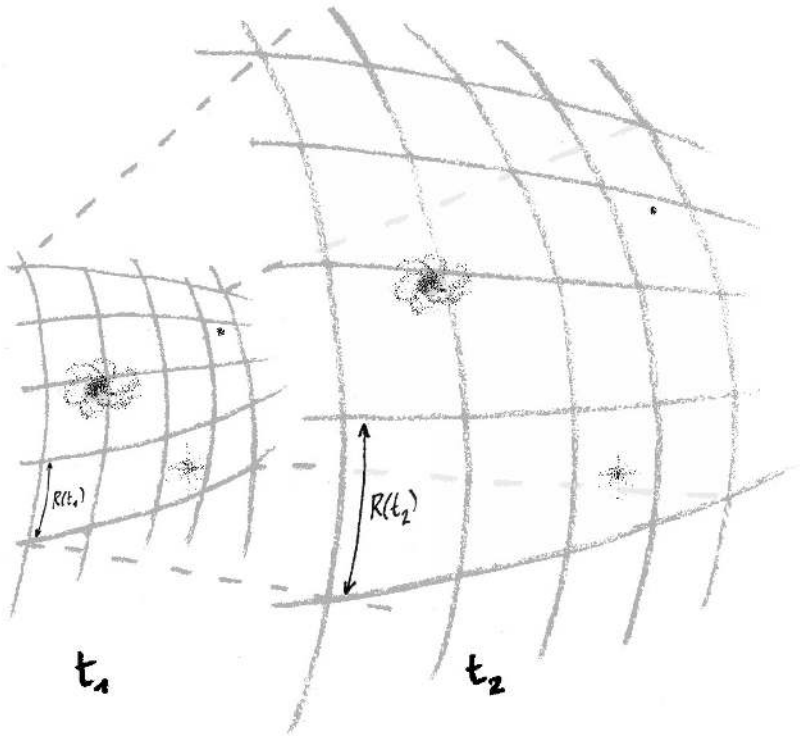

Figure 5. The scale factor $R(t)$ measures the rate of change of distance between freely falling particles as a function of time. Although the cosmological expansion has dynamical effects on local systems, these effects are *incredibly* small—far too small to "stretch" systems along with expansion.

Einstein did not overlook Friedmann and Lemaître's work before 1930, but he rejected expanding models as unphysical due to his persistent preference for a static model. Even after he retracted his initial claim (Einstein, 1922) that Friedmann had made an error, he clearly did not think that the expanding models were of any use in describing the observed universe. Stachel (1986) discovered that Einstein's draft retraction closed by stating that no physical significance could be attributed to Friedmann's models.[68] Fortunately this strong statement did not appear in print, but several years later Einstein expressed much the same view in conversation with Lemaître. According to Lemaître's account of their 1927 meeting, Einstein acknowledged that he had to withdraw his criticism of Friedmann's work but still found expanding models "totally abominable" from a physical point of view.[69] Friedmann seems to have shared, to a degree, Einstein's reluctance to attribute physical significance to the expanding models; his papers explored the properties of the hyperbolic and spherical solutions without linking them to the observed universe. Friedmann approached cosmology like a mathematician exploring the space of solutions to a differential equation. By way of contrast, Lemaître (1927) clearly regarded the spherical model (derived independently of Friedmann) as possibly giving a description of the universe's evolution. Unlike Friedmann, Lemaître explicitly derived the redshift effect and argued that his model naturally accounted for observed redshifts, and he also speculated about the physical origins of the expansion.

In the early 30s theorists began to develop a richer account of the evolution of the universe based on the expanding models. Hubble's results qualitatively agreed with the redshift effect calculated in these models, but the utility of the simple dynamical models depends upon whether the universe is approximately uniform. The status of this assumption was the focus of lively debate, partly due to Milne's advocacy of an alternative theory based on what he called the "cosmological principle." Relativistic cosmologists regarded the idealized uniformity of the FLRW as a simplifying assumption compatible with observations but not justified by them (see, e.g., Tolman, 1934, 332), rather than a methodological principle. The unrelenting uniformity built into the FLRW models conflicts with the clear non-uniformities of the stars, star clusters, and galaxies of the local universe, but the models might still serve as a useful approximation if the non-uniformities are negligible at larger scales. In 1926 Hubble initiated an observational program to measure the uniformity of the large scale distribution of nebulae, but the results of this program were challenged by Shapley.[70] Shapley and Hubble disagreed about the significance of observed small-scale clumping. In any case, the observational debates did not force theorists to abandon the simple expanding models.

There were, however, two widely acknowledged limitations of these simple models. First, extrapolating the models backwards in time leads to a genuine singularity (later derisively called the "big bang" by Fred Hoyle) that cannot be blamed on a poor choice of coordinates. Eddington and Lemaître avoided this initial singularity by suggesting that the universe began in an Einstein static state and then decayed into an expanding model. Einstein's cosmological constant could still serve the useful purpose of avoiding this initial singularity, and throughout the 30s Lemaître and others studied a wide variety of evolving models with non-zero Λ.[71] Lemaître later proposed a speculative theory of the initial state; on this account, the universe began as a "primeval atom" and has a finite age. A more common response, advocated by Einstein as well as Tolman, blamed the existence of the singularity on the strong idealizations built into the simple models. Perhaps a more detailed model describing the lumps and bumps of the real universe would not have evolved from an initial singularity, or so the suggestion went. (The singularity theorems of the 1960s showed that the singularity could not be blamed on the symmetries of the standard models.) The second limitation is that the uniform models provide no insight into the formation of non-uniformities, such as galaxies and stars. The cosmological models did not directly answer the cosmogonic problem that had concerned Jeans and others before the introduction of general relativity. Starting in the early 30s, Lemaître, McCrea, and others studied the growth of small clumps of matter in a background expanding model with the hope of shedding some light on the formation of galaxies.

Einstein missed the chance to discover the expanding models, and Friedmann and Lemaître's papers did not convince him of their physical importance. Throughout the 20s he apparently did not waver in his insistence on static models. He did reconsider the status of the cosmological constant but did not abandon it. He acknowledged in Einstein (1919a) that his earlier introduction of the cosmological constant was "gravely detrimental to the formal beauty of the theory," but immediately went on to argue that it could be treated more satisfactorily as a constant of integration.[72] Hubble's observations convinced him to abandon his preference for static models as well as the cosmological constant. During a trip to Pasadena from December, 1930 to March, 1931, Einstein learned about the latest observations at Mount Wilson first-hand from Hubble and his colleagues. Tolman persuaded Einstein that dynamical models were preferable to static models for describing the observed universe.[73] Shortly after returning to Berlin, Einstein published a paper that discussed the expanding models favorably and noted the shortcomings of his own static model (Einstein, 1931). During another visit to CalTech the following year, Einstein wrote a brief paper with De Sitter describing the properties of a simple expanding model with zero spatial curvature (Einstein and de Sitter, 1932). Both of these papers emphasized that Hubble's results completely undermined the original rationale for introducing the cosmological constant, which was to insure the possibility of a static universe with a finite mean matter density. Einstein later reportedly regarded its introduction as his "biggest blunder."[74]

After this decade and a half of sometimes intense work on cosmology, Einstein returned to the subject only occasionally in his later years. His most significant later contribution was a discussion of the impact of cosmological expansion on the gravitational field surrounding a star. The Schwarzschild solution used to describe such a field asymptotically approaches Minkowski spacetime as the distance from the star goes to infinity. But this cannot be correct for a star within an expanding universe. What are the consequences of treating the star as part of an expanding model? Einstein and Straus (1945) showed that the Schwarzschild solution could be embedded in an expanding FLRW model, and that despite the time dependence of the background cosmological model the field near the star remains static. This was an important first step in understanding the impact of global cosmological expansion on local physics. Einstein's other research in general relativity periodically touched on cosmology, but it was no longer a major focus of his work.

6. Conclusion

Within fifteen years after the introduction of Einstein's theory, theorists had surveyed the properties of a number of simple, idealized solutions of the new gravitational field equations. Einstein's groundbreaking work initiated the study of cosmological models, although several contemporaries were more clear sighted in understanding the properties of these models, and Einstein's preference for a static model kept him from discovering or initially accepting the expanding models. The exploration of these models touched on a number of the novel features of general relativity, including event horizons and singularities, that have been a focus of further research. The work described above has been aptly called "geometrical cosmology," given its very mathematical style. However, Hubble's discovery of the redshift - distance relation provided a link with astronomical observations that encouraged many astronomers and theorists to take these models seriously. The new consensus regarding the expanding models was codified in systematic reviews (such as Robertson, 1933) and textbooks (Tolman, 1934).[75]

However, this consensus faced a number of objections, ranging from empirical anomalies to questions regarding the legitimacy of taking cosmology to be the study of idealized solutions of Einstein's gravitational field equations. As critics of relativistic cosmology rightly pointed out, the links between the expanding models and observations were remarkably tenuous. The expanding models provided a natural explanation of the observed redshifts of the spiral nebulae. However, the use of these observations to constrain parameters of the models led to a striking problem: for expanding models with a finite age, the "age of the universe" often turned out to be far less than the age estimated for various astronomical objects! The eventual resolution of this problem two decades later came not from a change in the cosmological models but a recalibration of the distance scale based on Cepheids. Prior to this recalibration the opponents of the expanding universe models often cited the age problem as a serious anomaly (among others) for relativistic cosmology.

Debates in the next two decades extended beyond such empirical problems. Critics of relativistic cosmology introduced rival cosmological theories based on allegedly sounder scientific methodology than that followed by Einstein and the other relativists. Although I do not have space to explore these debates in detail, a brief discussion of the status of the cosmological principle will convey some of the issues at stake. Within the standard approach to relativistic cosmology, the cosmological principle was typically treated as a simplifying assumption used to arrive at mathematically tractable models. Although the assumption of uniformity could be precisely characterized mathematically, there was admittedly no physical motivation for demanding such a high degree of symmetry. As a result, it was unclear what, if anything, the failure of the observed universe to satisfy the uniformity assumption indicated regarding the status of the models, and whether the models could be refined to give more realistic descriptions of the observed matter distribution. By way of contrast, the main alternative to relativistic cosmology throughout the 30s, Milne's kinematic relativity, elevated a version of the cosmological principle to a fundamental axiom. Milne's unabashed rationalism in treating the cosmological principle as an *a priori* axiom led to a series of heated debates regarding method in cosmology, and Milne's distinctive approach influenced work in relativistic cosmology (as noted above) even though kinematic relativity never won widespread acceptance. Admiration of Milne's approach also played a role in the discovery of the second major alternative to relativistic cosmology, the steady state theory introduced in 1948 by Herman Bondi, Thomas Gold, and Fred Hoyle (Bondi and Gold, 1948; Hoyle, 1948).

Bondi and Gold (1948) proposed the "perfect cosmological principle" as a response to an apparent threat to the extrapolation of local physical laws to cosmological scales. Relativistic cosmology certainly involves an enormous extrapolation of Einstein's theory, from its empirical testing grounds of roughly solar-system scale to cosmological length scales such as the Hubble radius (the length scale of the observable universe), roughly 14 orders of magnitude larger. The steady state theorists were not primarily concerned with the threat of new physics arising at greater length scales; instead, they followed Milne in insisting that cosmology faced distinctive methodological problems due to the uniqueness of the universe. In particular, they argued that the distinction between "laws" and "initial conditions" familiar in other areas of physics does not carry over to cosmology. Since this distinction could not be drawn clearly, perhaps there would be some form of "interaction" between local physical laws and the features usually described as initial conditions, such as the global distribution of matter?[76] The perfect cosmological principle was introduced to rule out the possibility that the local physical laws evolve along with the changing universe by simply stipulating that the universe does not change. Bondi and Gold argued for this principle on methodological grounds: if the principle fails, they argued, any extrapolation of local laws to cosmological scales is unjustified, since it might neglect possible interactions. The principle leads to a "steady state" universe with unchanging global features, and the reconciliation of this picture with the observed expansion forced the steady state theorists to relinquish conservation of matter. The steady state theory provoked sharp debates throughout the 50s, focused in part on whether cosmology should employ such a distinctive methodology. The perfect cosmological principle provided a tight constraint on cosmological theorizing, as the steady state theorists had hoped, but by the early 60s the constraint proved too tight—and several advocates of the theory abandoned it in light of new observational evidence.

The status of the cosmological principle changed again in 1965 with the discovery of very low temperature radiation by Arno Penzias and Robert Wilson. The Princeton physicist Robert Dicke immediately interpreted this radiation as the remnant of a hot big bang in the early universe, and it provided a final piece of evidence against the steady state theory.[77] But more importantly, the discovery of the background radiation provided an empirical touchstone that encouraged a dramatic increase in research effort devoted to cosmology. This effort included the development of detailed accounts of the nuclear reactions taking place in the first three minutes of the universe's history (an idea first introduced by Gamow, Alpher, and Herman a generation earlier, but largely ignored), and theorists also explored competing accounts of the formation of galaxies and other large-scale structure. The uniform temperature of the background radiation indicated that the simple mathematical assumption of uniformity built into the evolving models was far more accurate than cosmologists had expected, and it also provided evidence that these models applied to the universe at very early times.

Cosmologists had reason to be puzzled rather than exhilirated by this result, since they lacked an explanation of why the universe should be in such a highly symmetric state soon after the big bang. The attempt to resolve this final puzzle has pushed cosmological theories to incredibly early times ($t \approx 10^{-35}$ seconds after the big bang!), where exotic features of particle physics theories become relevant. "Particle cosmology" has been an active area of research for the last twenty years, partially because the "poor man's accelerator" (to use the Soviet cosmologist Zel'dovich's apt description of the early universe) provides perhaps the only way of testing aspects of particle physics well beyond the reach of Earth-bound accelerators. The field has been dominated by inflationary cosmology, introduced in Guth (1981), which explains the puzzling uniformity of the early universe as the consequence of a brief period of exponential ("inflationary") expansion in the early universe. In light of Einstein's early research in cosmology, the theory of inflationary cosmology is doubly ironic, since it re-introduces Einstein's infamous cosmological constant Λ (admittedly for very different reasons) and it makes use of the De Sitter solution, the model Einstein argued so strenuously against.

Acknowledgments

I have benefited from comments and discussions regarding earlier drafts from Tyler Burge, Sarah Gallagher, David Malament, Charles Midwinter, Sheldon Smith, and especially Michel Janssen. I am grateful to Laurent Taudin, who drew the diagrams for this paper. I also gratefully acknowledge the debt to my two mentors in history and philosophy of physics, John Earman and John Norton, from whom I learned a great deal about the topics discussed above.

Notes

[1]This brief overview draws on several comprehensive historical accounts, primarily Smith (1982); North (1965); Paul (1993); see also North (1995). Macpherson (1929) and Clerke (1890) offer well-informed contemporary overviews of the field that bracket this time period.

[2]Jeans coined the name following Kapteyn's publications in the early 20s, and Seeliger's research supported a similar model (see Paul, 1993, 141–150). Estimates of the scale of the system were subject to a great deal of uncertainty. Eddington (1914, 31) argued that the dimensions are roughly 1100 light-years (along the shorter axis) by 3200 light-years (along the galactic plane),

where the distance is measured from the center to the region where the density of stars drops to below one fifth of the maximum density. Six years later Kapteyn and van Rhijn (1920) estimated the size of the galaxy to be roughly 6,100 by 31,000 light-years.

[3]Advocates and opponents of the island universe theory advanced a variety of ideas regarding the size, distance and nature of the spiral nebulae and other astronomical objects; in this sense there was not a single "island universe theory" throughout this period. For example, one disputed topic was whether globular clusters also lie outside the Milky Way and have some connection with the spirals. For a more fine-grained analysis of these debates that highlights such differences, see Smith (1982).

[4]For further discussion of these objections and references to the original literature, see Smith (1982, 1–54), North (1965, 4–15).

[5]The observational situation was muddied by the failure to distinguish novae from much more luminous supernovae. At the time the small sample of novae used in the comparison included two supernovae, leaving Shapley and other island universe opponents ample reason to doubt that novae in the spirals were dimmer than those within the Galaxy (North, 1965, 2–13).

[6]There is much less dust along the line of sight to objects near the Galactic poles than there is to objects near the Galactic plane. Even if the spirals were uniformly spread throughout the sky, they would appear to be clustered around the Galactic poles due to this differential absorption.

[7]The observed brightness (apparent magnitude) of the star decreases with distance, so if the intrinsic brightness (absolute magnitude) is known the distance can be calculated.

[8]The published exchange appeared a year later (Curtis 1921, Shapley 1921). Hoskin (1976) argues that Shapley tailored his performance to secure a position as the director of the Harvard College Observatory, which had recently opened up with the death of E. C. Pickering. See also Smith (1982, 77–90) for a detailed discussion of the Great Debate and its context.

[9]In 1917 Shapley accepted the island universe theory, but became increasingly critical in 1918 and 1919 (Smith, 1982, 85–86).

[10]Hubble measured the distance of the spiral nebula Messier 33 to be roughly $930,000$ lightyears, compared to van Maanen's estimate of $9,800$ lightyears.

[11]Three decades after Hubble's pioneering work, Walter Baade showed that earlier distance measurements had failed to take into account the existence of two *distinct* populations of variable stars, each with a characteristic period-luminosity relation (Baade, 1952). Taking this into account dramatically changed the distance scale, literally doubling the earlier distant estimates.

[12]Here I draw on Brush's (1996a,b) rich historical account of the nebular hypothesis and the competing ideas introduced in the early twentieth century.

[13]The subject of the 1917 Adams Prize at Cambridge was "the course of configurations possible for a rotating and gravitating fluid mass, including the discussion of the stabilities of the various forms" (Jeans, 1919, v); Jeans' innovative 1919 book extended his prize-winning essay. See Smith (1982) for an account of Jeans's influence and his interactions with van Maanen, Curtis, and others.

[14]For example, from roughly 1916 to 1924 Jeans thought that van Maanen's measurements of radial velocities in the spiral nebulae fit naturally into his account of the dynamical evolution of spiral nebulae (Smith, 1982), an unfortunate association given the later problems with van Maanen's measurements.

[15]See Torretti (2000) for an insightful discussion of the nature of cosmological models and what he memorably calls Einstein's "unexpected lunge for totality" in introducing the first cosmological model. I am indebted to Torretti's article and previous work, especially in this section and in Section 5 below.

[16]In this subsection, I draw on Norton's (1999) comprehensive discussion of the difficulties with Newtonian cosmology and their history; cf. Jaki (1990); North (1965).

[17]Einstein was unaware of Seeliger's papers when he wrote his cosmology paper (Einstein 1917b), but after they came to his attention he duly cited them (see, in particular, Einstein 1919b, 433, footnote 1, and CPAE 8, Doc. 400). In 1692 Richard Bentley raised the problem in correspondence with Newton while preparing his inaugural Boyle lectures for publication (Turnbull, 1959-1977, Vol. III, 233–56). He pressed Newton to consider the following question. For a given particle of matter in a uniform, infinite mass distribution, the gravitational force in any chosen direction is infinite. But would the pull exerted by the infinity of mass lying in the opposite direction cancel this first infinite force, so that the forces in different direction would be in equilibrium? Bentley balked at comparing infinities, but Newton answered with a definite yes, apparently without recognizing the deeper problem. The force on the particle is found by integrating over the contributions due to all other masses, but this integral fails to converge for a uniform, infinite distribution of mass.

[18]Seeliger proposed adding an exponential decay term to Newton's gravitational force law. Einstein's modification is not equivalent to this, although Carl Neumann had proposed a modification equivalent to Einstein's, with different motivations, in 1896 (see Norton, 1999, 293–298).

[19]See Norton (1999, 302–313), for a clear description of the properties of these hierarchic models and an explanation of how these models avoid the divergences.

[20]Treating time-varying matter distributions is clearly not sufficient to avoid the difficulty, since even in an evolving model the universe may pass through a uniform state. Milne (1934); McCrea and Milne (1934) do not address the question directly, and Layzer (1954) argued that the models do not apply consistently to an infinite universe, even though they can be consistently applied to a finite mass distribution. McCrea (1955); Heckmann and Schücking (1955, 1956) aptly defended the Neo-Newtonian models along lines similar to those discussed in the text, but without appealing to the geometric formulation (cf. Layzer, 1956; Heckmann and Schucking, 1956).

[21]See Malament (1995); Norton (1995, 2003) for thorough treatments of the geometric formulation and its implications for Newtonian cosmology.

[22]Additional solutions to eqn. (1) can be obtained by adding a harmonic function (a function ψ such that $\nabla^2\psi = 0$) to one of these solutions. These additional solutions can be eliminated by requiring that ϕ is isotropic around $\mathbf{r}_0$.

[23]One other feature of Einstein's note is worth highlighting: he comments that the latest observations refute the "hypothesis of the equivalence of the spiral nebulae and the Milky Way," although it is not clear which observational results he had in mind.

[24]Einstein later added Mach's principle—"The metric field is determined *without residue* by the masses of bodies"—to the list of three foundational principles of general relativity (Einstein, 1918e, 242, original emphasis). There are problems with both formulations; in particular, it is not clear how to characterize the matter distribution without employing the metric. In what follows, we will focus almost exclusively on the connection between Mach's principle and the status of boundary conditions and vacuum solutions. For an entry into the extensive literature on Mach's principle, its historical role in Einstein's path to general relativity, and its relation to other foundational principles, such as the requirement of general covariance and the equivalence principle, see Janssen's contribution to this volume, Torretti (2000), and Hoefer (1994).

[25]The editorial note "The Einstein-De Sitter-Weyl-Klein Debate" (CPAE 8, 351–357) gives a thorough discussion of the correspondence between Einstein and De Sitter, and the later participants Weyl and Klein. See also Section 5 of Janssen's contribution to this volume.

[26]Einstein (1915h) introduced the boundary conditions of the Schwarzschild solution without commenting on the potential conflict with his Machian ideas. However, there are hints that Einstein had already begun to worry about these issues (see Hoefer, 1994, 298–303), in both his

comments regarding the status of Minkowski space-time in the context of general relativity and a letter to Lorentz from the winter of 1915 (CPAE 8, Doc. 47).

[27]As Torretti (2000) has emphasized, Birkhoff's (1923) theorem shows that the Schwarzschild solution can be derived from the assumption of spherical symmetry *without* further stipulating boundary conditions at infinity. Thus in the case Einstein had in mind, the boundary conditions are not logically independent from the symmetry requirements.

[28]Einstein and De Sitter also studied the transformation behavior of different boundary conditions as a way of testing their consistency with Mach's principle. Einstein apparently thought that degenerate values would be generally covariant, but De Sitter showed that this claim was incorrect. de Sitter pointed out that even the "degenerate values" were not invariant under certain transformations (see CPAE 8, Doc. 272 and De Sitter (1916c, 1917a)).

[29]More precisely, Einstein takes the matter distribution and the corresponding metric field to be homogeneous and isotropic. Roughly, homogeneity is satisfied if there are no "preferred locations" on a spatial section, and isotropy holds that at a given location there are no "preferred directions" in space. The modern treatment of these symmetries traces back to the work of Robertson and Walker described briefly below.

[30]A few years later Weyl introduced a contrast between "static" and "stationary" solutions that has since become standard: in a stationary solution, the spatial components of the metric are not functions of time but the "mixed" space-time components may be functions of time, whereas for a static solution only the time-time component of the metric is allowed to vary with time. More intuitively, spatial distances remain fixed under time translations in a stationary solution, and this is compatible with a perfectly rigid rotation; there is no such rotation in static solutions.

[31]In a later letter (Doc. 355) he considered the constraints that low stellar velocities would put on a cosmological solution, and argued that they do not lead to a static solution as Einstein had thought.

[32]This change is not "perfectly analogous" to the addition of a constant in Poisson's equation, as Einstein claimed. The Newtonian limit of the modified field equations yields $\nabla^2\phi + \Lambda\phi = 4\pi G\rho$, rather than eqn. (3) (Trautman, 1965, 230).

[33]The Ricci tensor $R_{\mu\nu}$ is defined in terms of the metric field $g_{\mu\nu}$ and its first and second derivatives, and the metric field characterizes the geometrical properties of space-time. The stress-energy tensor $T_{\mu\nu}$ represents the distribution of matter, energy, and stress throughout the space-time.

[34]The left hand side of eqn. (7) is the most general symmetric tensor that can be constructed from the metric $g_{\mu\nu}$ and its first and second derivatives, such that the conservation law for $T_{\mu\nu}$ holds. See Renn (2007a, Vol. 2, 493–500) for further discussion of the physical requirements that guided Einstein's derivation of the original field equations.

[35]To be more precise, Λ enters into equations governing, for example, the behavior of nearby test particles with the opposite sign as matter or energy density. The presence of normal matter or energy leads paths of nearby test particles to converge, and this captures the idea that gravitation is a "force of attraction." A positive Λ term leads to divergence of these paths (if it is strong enough to counteract gravitational attraction), and in this sense it is a "repulsive force."

[36]Minkowski space-time apparently constitutes a violation of Mach's principle as formulated above, since it is a vacuum solution. Einstein did not explicitly mention that Minkowski space-time is not a solution with $\Lambda \neq 0$, although he undoubtedly appreciated this point (as did Pauli, 1921).

[37]"It is to be emphasized, however, that our results give a positive curvature of space even if the the supplementary term [Λ] is not introduced; that term is necessary only for the purpose of making a quasi-static distribution of matter possible, as required by the fact of the small stellar

velocities" (Einstein, 1917b, 152). Torretti (2000, 178) suggested that "Einstein could hardly have published that remark" without knowing of a non-stationary solution with closed spatial sections. Torretti's point is plausible, but as far as I know there is no evidence that Einstein had obtained such a solution in 1917.

[38]Prior to the introduction of his own model, De Sitter had focused on the status of the time coordinate in Einstein's solution, and he characterized his solution as treating the temporal coordinate as "closed" in the same manner that Einstein closed the spatial dimensions. He gives credit to Ehrenfest for the idea of getting rid of temporal as well as spatial boundary conditions by postulating a "spherical" space-time (de Sitter, 1917c).

[39]I am setting aside debates regarding the topology of De Sitter's space-time; Klein, Einstein and others discussed whether antipodal points should be identified or treated as distinct points. (Antipodal points lie on opposite "sheets" of the hyperboloid and are connected by a line passing through the center of the hyperboloid.)

[40]See, in particular, Janssen's contribution to this volume, Earman and Eisenstaedt (1999), and Goenner (2001) for further discussions of the debates regarding De Sitter's solution.

[41]For discussion of geodesics in the De Sitter solution, see Schrödinger (1956). Geometrically, referring to Fig. 9 of Janssen's contribution to this volume, timelike geodesices are defined by the intersection of the hyperboloid and vertical planes through the origin. These are curves of constant spatial coordinates in pseudo-polar coordinates, which cover the entire hyperboloid. But the curves of constant spatial coordinates in static coordinates are not geodesics, as can be shown by considering the coordinate transformation from pseudo-polar to static coordinates.

[42]See Janssen's contribution to this volume regarding Klein's contribution and his exchange with Einstein, and for further discussion of the properties of the De Sitter solution. See also Schrödinger (1956) for a particularly clear treatment of the De Sitter solution.

[43]See "The Einstein-De Sitter-Weyl-Klein Debate" (CPAE 8, 352–354) for a discussion of Einstein's other criticisms and De Sitter's responses.

[44]See Earman (1999) for a history of the problems related to singularities, culminating with a discussion of the famous Penrose-Hawking singularity theorems.

[45]See Earman and Eisenstaedt (1999, 189–193) for further discussion of Einstein's definition and its relation to an alternative, more stringent definition proposed by Hilbert.

[46]Written in static coordinates, the g_{tt} component of the metric is given by $\cos^2(r/R)$ (where r is a coordinate and R is a constant), and this goes to zero at the "mass horizon" $r = \pi R/2$. See also Section 5 of Janssen's contribution to this volume.

[47]The construction only appeared in the second edition of *Raum-Zeit-Materie*. Weyl and Einstein exchanged several letters regarding Weyl's results, which Einstein foreshadowed in the closing remarks of Einstein (1918c). See Goenner (2001) and Earman and Eisenstaedt (1999), as well as the correspondence and the editorial headnote in CPAE 8, for discussions of Weyl's result and its role in Einstein's thinking.

[48]Lanczos criticized the mass horizon idea in a second paper (Lanczos, 1923); see also Stachel (1994); Eisenstaedt (1993).

[49]Rindler (1956) is the classic paper regarding horizons in cosmology; see Ellis and Rothman (1993) for a more recent and extremely clear discussion.

[50]I do not have space to review the disputes regarding redshift during the twenties. See North (1965), Goenner (2001) and Ellis (1989) for thorough assessments and references to the original papers.

[51]See Section 3 of Janssen's contribution for a brief explanation of the notion of a geodesic.

[52] Contributions to the overall redshift come from the following six sources (Ellis 1989, 374–375): the Doppler shift due to relative motion of the source, gravitational redshift due to inhomogeneity near the source and near the observer, Doppler shift due to the peculiar velocity of

the observer, the cosmological expansion or contraction, and the gravitational redshift produced by large-scale inhomogeneities.

[53]See Smith (1982), Chapter 5, regarding the background of Hubble's work and its reception.

[54]De Sitter had voiced some skepticism about this in correspondence with Einstein (CPAE 8, Doc. 355), and Gustav Mie also broached the issue (CPAE 8, Docs. 465, 488), although he argued incorrectly that any inhomogeneities in Einstein's solution would evolve back to a uniform matter distribution.

[55]The conclusion follows from Lemaître's equations for the evolution of the scale factor given a perfect fluid with zero pressure: $3\ddot{R}/R = \Lambda - (\kappa\rho/2)$, given that ρ varies with $R(t)$ as $\rho \propto R^{-3}(t)$ whereas Λ remains constant.

[56]Tolman (1929a) also criticizes the De Sitter model. Unlike De Sitter, Tolman argued that predicting a correct redshift-distance relation requires *ad hoc* assumptions regarding the world-lines of the spiral nebulae, but like De Sitter he expressed unease at treating galaxies as test bodies within a vacuum solution.

[57]At this time the De Sitter solution was regarded as a static solution. Following De Sitter's presentation, Eddington expressed puzzlement regarding why there should be only two solutions, adding that "I suppose the trouble is that people look for static solutions" (R.A.S., 1930, 39). Similarly, de Sitter (1930) concluded by advocating dynamical solutions in light of the difficulties facing both models; cf. a letter from De Sitter to Shapley quoted in Smith (1982, 187).

[58] Tolman remarked that "the investigation of non-static line elements would be interesting" (Tolman, 1929b) in the final line of a paper presenting the difficulties for static models; in a slightly later paper (Tolman, 1929a), Tolman argued that the De Sitter solution does not give a "simple and unmistakably evident explanation" of the distribution and motion of the spiral nebulae, as observed by Hubble.

[59]Lemaître had been Eddington's student in 1923–24, and Eddington later remarked that he had probably seen but not appreciated Lemaître's paper before 1930; see Kragh (1996, 31–33); Eisenstaedt (1993).

[60](Lemaître, 1931) is not, however, a complete translation of the original paper; Eddington dropped Lemaître's determination of what we now call the Hubble constant, based on the redshift-distance relation for 42 galaxies.

[61]Milne's work and the larger methodological debates it sparked have been the focus of a series of papers by George Gale and various collaborators—see, in particular, Gale and Urani (1993); Gale and Shanks (1996), Gale (2002) for a general overview, and Lepeltier (2006) for a critical response.

[62]Despite the significant overlap in mathematical content in these sets of papers, their methodological motivations were quite different. Robertson adopted Milne's operationalist standpoint only to rebut Milne's supposed alternative to general relativity, highlighting its similarities at the level of kinematics and the difficulties facing Milne's kinematical-statistical theory of gravitation. Milne's protegé Walker was, unsurprisingly, much more sympathetic to Milne's approach.

[63]In current terminology, the metric is required to be homogeneous and isotropic. Robertson and Walker both set aside Milne's further requirement that the (operationally defined) coordinates assigned by distinct fundamental observers are related by Lorentz transformations, as this singles out the $k = -1$ FLRW model.

[64]Torretti (1983, 202–210) describes Friedmann's work in greater detail and compares it to the work of Robertson and Lemaître; cf. Kragh (1996, 22–38).

[65]In Einstein's static model the distance between test particles remains constant; De Sitter's solution only appears to be static if one mistakenly focuses on test particles that do not move along geodesics.

[66]Friedmann studied solutions of the modified field equations including the cosmological constant Λ, although his results establish the existence of the solutions described below with $\Lambda = 0$.

[67]Roughly speaking, normal matter is defined as having "positive total stress-energy density." The presence of such matter produces gravitational attraction, other things being equal, in the sense of convergence of the trajectories of freely falling test particles. There are different ways of making the notion of "normal matter" more precise, leading to various energy conditions in general relativity. Matter which satisfies the so-called strong energy condition contributes to the dynamical equations such that it contributes to deceleration. By way of contrast, a non-zero cosmological constant term does not satisfy this energy condition and does not qualify as normal matter.

[68]The last line of the brief correction reads, "It follows that the field equations admit, as well as the static solution, dynamic (that is, varying with the time coordinate) centrally symmetric solutions for the spatial structure, [to which a physical significance can hardly be ascribed]." The phrase in square brackets is canceled in the manuscript (EA 1-206) and does not appear in the printed version (Einstein, 1923).

[69]Lemaître described the encounter to Eddington in a letter written in early 1930 accompanying the reprint of his 1927 article. Eisenstaedt (1993, 361) quotes a draft of this letter from Lemaître's papers.

[70]See Peebles (1980, 3–11) for an overview of the debate, and Tolman (1934, Sections 177–185) for a contemporary assessment of the evidence.

[71]See Earman (2001) for a discussion of these models and different attitudes to the cosmological constant.

[72]Schrödinger (1918) had pointed out another way of treating the cosmological constant: moving it from the left-hand side of eqn. (7), where it represents a contribution to space-time curvature, to the right-hand side, where it represents a contribution to the energy-matter distribution. Then it would correspond physically to a kind of cosmic pressure. In his response to Schrödinger, Einstein (1918d) wrote that he had considered this option too but had rejected it as artificial. Einstein (1919a) argues that the cosmological constant term, treated as a constant of integration, also solves a problem in matter theory by contributing a negative pressure term to keep the electrodynamic forces on a small charged particle in equilibrium. For further discussion of the proposal in Einstein (1919a), see Earman (2003).

[73]As Stachel (1986) has pointed out, Einstein's diary from the trip records that he initially had doubts regarding Tolman's argument, but that during this trip he was convinced that Tolman was in fact correct (EA 29-134, 27–28). Einstein had a positive impression of the astronomers at Mount Wilson; see, e.g., his letter to Besso written in March of 1931, quoted in Stachel (1986).

[74]The canonical source for this story is George Gamow's autobiography: "Einstein's original gravity equation was correct, and changing it was a mistake. Much later, when I was discussing cosmological problems with Einstein, he remarked that the introduction of the cosmological term was the biggest blunder he ever made in his life. But the "blunder," rejected by Einstein ... rears its ugly head again and again and again" (Gamow, 1970, 149–150).

[75]In 1931 the consensus favored a multi-stage evolutionary model rather than a simple big bang model; see Gale and Shanks (1996) regarding the discussion of this emerging consensus at a 1931 British Association for the Advancement of Science meeting.

[76]Mach's principle returned to the limelight as one of the steady state theorists' main examples of this type of interaction (see, for example, Sciama 1957). In the case of Mach's principle, the concern was that local inertial structure would change with the overall change in matter distribution in an evolving cosmos. For example, Sciama developed a theory according to Newton's gravitational constant G depended upon the density and expansion velocity of the global matter distribution, such that in an expanding model the current value of G would differ from that at

earlier or later times. Bondi and Gold were concerned that other local "laws" of physics may also "evolve" in this sense. Whether the ideas of interaction and evolving laws invoked here make sense is a contentious issue; see Balashov (2002) for a discussion that is sympathetic to the steady state theorists.

[77]Popular accounts of the controversy typically overplay the importance of the background radiation; as Kragh (1996) establishes, the background radiation was the last of several empirical anomalies facing the steady state theory, and all those who could be convinced to abandon the theory had already done so before 1965.

REFERENCES

Baade, Walter (1952). "Extragalactic Nebulae. Report to IAU Commission 28." *Transactions of the International Astronomical Union of Astronomy* 8, 397–399.

Balashov, Yuri (2002). "Laws of physics and the universe." In *Einstein Studies in Russia*, Yuri Balashov and Vladimir Vizgin, eds., vol. 10 of *Einstein Studies*. Basil: Birkhäuser, 107–148.

Birkhoff, G. (1923). *Relativity and Modern Physics*. Cambridge: Harvard University Press.

Bondi, Herman and Gold, Thomas (1948). "The steady-state theory of the expanding universe." *Monthly Notices of the Royal Astronomical Society* 108, 252–270.

Brush, Stephen G. (1996a). *Fruitful Encounters: The origin of the Solar System and of the Moon from Chamberlin to Apollo*. Cambridge: Cambridge University Press.

— (1996b). *Nebulous Earth: The origin of the Solar System and the core of the Earth from Laplace to Jeffreys*. Cambridge: Cambridge University Press.

Charlier, Carl V. L. (1908). "Wie eine unendliche Welt aufgebaut sein kann." *Arkiv för Matematik, Astronomi och Fysik* 4, 1–15.

Clerke, Agnes (1890). *The System of the Stars*. London: Longmans, Green and Co.

Curtis, Heber (1921). "The scale of the universe." *Bulletin of the National Research Council* 2, 194–217.

De Sitter, Willem (1917a). "Over de relativiteit der traagheid: Beschouwingen naar aanleiding van Einsteins laatste hypothese." *Koninklijke Akademie van Wetenschappen te Amsterdam. Wis- en Natuurkundige Afdeeling. Verslagen van de Gewone Vergaderingen.* 25, 1268-1276. Reprinted in translation as "On the relativity of inertia: remarks concerning Einstein's latest hypothesis." *Koninklijke Akademie van Wetenschappen te Amsterdam. Section of Sciences. Proceedings.* 19, 1217–1225.

— (1917b). "Over de kromming der ruimte." *Koninklijke Akademie van Wetenschappen te Amsterdam. Wis- en Natuurkundige Afdeeling. Verslagen van de Gewone Vergaderingen.* 26, 222-236. Reprinted in translation as "On the curvature of space." *Koninklijke Akademie van Wetenschappen. Section of Sciences. Proceedings.* 20, 229–243.

— (1917c). "On Einstein's theory of gravitation and its astronomical consequences, III." *Monthly Notices of the Royal Astronomical Society* 78, 3–28.

— (1930). "On the distances and radial velocities of extra-galactic nebulae, and the explanation of the latter by the relativity theory of inertia." *Proceedings of the National Academy of Science* 16, 474–488.

— (1931). "The expanding universe." *Scientia* 49, 1–10.

Earman, John (1999). "The Penrose-Hawking singularity theorems: History and implications." In *The Expanding Worlds of General Relativity*, Jürgen Renn, Tilman Sauer and Hubert Gönner, eds. No. 7 in Einstein Studies, Boston: Birkhäuser, 235–270.

— (2001). "Lambda: The constant that refuses to die." *Archive for History of Exact Sciences* 55, 189–220.

— (2003). "The Cosmological Constant, the Fate of the Universe, Unimodular Gravity, and All That." *Studies in History and Philosophy of Modern Physics* 55, 34, 559–577.

Earman, John and Eisenstaedt, Jean (1999). "Einstein and singularities." *Studies in the History and Philosophy of Modern Physics* 30, 185–235.

Eddington, Arthur S. (1914). *Stellar Movements and the Structure of the Universe*. London: Macmillan.

— (1923). *The Mathematical Theory of Relativity*. Cambridge: Cambridge University Press.

— (1930). "On the instability of Einstein's spherical world." *Monthly Notices of the Royal Astronomical Society* 90, 668–678.

— (1933). *The Expanding Universe*. New York: MacMillan.

Einstein, Albert (1915h). "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie." *Sitzungsberichte der Preussische Akademie der Wissenschaften (Berlin)* , 831–839.

— (1916e). "Die Grundlage der allgemeinen Relativitätstheorie." *Annalen der Physik* 49, 769–822. Reprinted in translation in Lorentz et al. 1923.

— (1917a). *Über die spezielle und die allgemeine Relativitätstheorie*. Braunschweig: Vieweg.

— (1917b). "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie." *Preussische Akademie der Wissenschaften (Berlin). Sitzungsberichte* , 142–152. Reprinted in translation in Lorentz et al. 1923.

— (1918c). "Kritisches zu einer von Herrn de Sitter gegebenen Lösung der Gravitationgleichungen." *Sitzungsberichte der Preussischen Akademie der Wissenschaften* .

— (1918d). "Bemerkung zu Herrn Schrödinger's Notiz "Über ein Lösungssystem der allgemein kovarianten Gravitationsgleichungen."" *Physikalische Zeitschrift* 19, 165–166.

— (1918e). "Prinzipielles zur allgemeinen Relativitätstheorie." *Annalen der Physik* 55, 241–244.

— (1919a). "Spielen Gravitationsfelder im Aufber der materiellen Elementarteilchen eine wesentliche Rolle?" *Sitzungsberichte der Köngliche Preussiche Akademie der Wissenschaften (Berlin)* , 349–356. Translated in *The Principle of Relativity.*

— (1919b). "Bemerkung über periodische Schwankungen der Mondlänge, welche bisher nach der Newtonschen Mechanik nicht erklärbar schienen." *Sitzungsberichte der Köngliche Preussiche Akademie der Wissenschaften (Berlin)* , 433–436.

— (1922). "Bemerkung zu der Arbeit von A. Friedmann, 'Über die Krümmung des Raumes'." *Zeitschrift für Physik* 11, 326.

— (1923). "Notiz zu der Arbeit von A. Friedmann 'Über die Krummung des Raumes'." *Zeitschrift für Physik* 16, 228.

— (1931). "Zum kosmologischen Problem der allgemeinen Reitivitätstheorie." *Sitzungsberichte der Preussische Akademie der Wissenschaften (Berlin)* , 235–237.

Einstein, Albert and de Sitter, Willem (1932). "On the relation between the expansion and the mean density of the universe." *Proceedings of the National Academy of Science* 18, 213–214.

Einstein, Albert and Straus, Ernest G. (1945). "The influence of the expansion of space on the gravitation fields surrounding individual stars." *Reviews of Modern Physics* 18, 148–152.

Eisenstaedt, Jean (1989). "Cosmology: A space for thought on general relativity." In *Foundations of Big Bang Cosmology*, F. Walter Meyerstein, ed. Singapore: World Scientific, 271–295.

— (1993). "Lemaître and the Schwarzschild Solution." In *The Attraction of Gravitation*, John Earman, Michel Janssen and John D. Norton, eds., vol. 5 of *Einstein Studies*. Birkhäuser, 353–389.

Ellis, G. F. R. (1989). "The expanding universe: A history of cosmology from 1917 to 1960." In *Einstein and the History of General Relativity*, Don Howard and John Stachel, eds., vol. 1 of *Einstein Studies*. Boston: Birkhäuser, 367–431.

Ellis, G. F. R. and Rothman (1993). "Lost horizons." *American Journal of Physics* 61, 883–893.

Friedmann, Alexander (1922). "Über die Krümmung des Raumes." *Zeitschrift für Physik* 10, 377–386.

— (1924). "Über die Möglichkeit einer Welt mit Konstanter negativer Krümmung des Raumes." *Zeitschrift für Physik* 21, 326–332.

Gale, George (2002). "Cosmology: Methodological debates in the 1930s and 1940s." Published online at plato.stanford.edu/entries/cosmology-30s/.

Gale, George and Shanks, Niall (1996). "Methodology and the birth of modern cosmological inquiry." *Studies in the History and Philosophy of Modern Physics* 27, 279–296.

Gale, George and Urani, John (1993). "Philosophical midwifery and the birthpangs of modern cosmology." *American Journal of Physics* 61, 66–73.

Gamow, George (1970). *My World Line*. New York: Viking.

Goenner, Hubert (2001). "Weyl's contributions to cosmology." In *Hermann Weyl's Raum-Zeit-Materie and a General Introduction to his Scientific Work*, Erhard Scholz, ed. Boston: Birkhäuser, 105–137.

Guth, Alan (1981). "Inflationary universe: A possible solution for the horizon and flatness problems." *Physical Review D* 23, 347–56.

Heckmann, O. and Schücking, E. (1956). "Bemerkungen zur Newtonschen Kosmologie. II." *Zeitschrift für Astrophysik* 40, 81–92.

Heckmann, O. and Schucking, E. (1956). "Reply to Layzer." *The Observatory* 76, 74–75.

Heckmann, Otto (1942). *Theorien der Kosmologie.*. Berlin: Springer Verlag.

Heckmann, Otto and Schücking, E. (1955). "Bemerkungen zur Newtonschen Kosmologie. I. Mit 3 Textabbildungen in 8 Einzeldarstellungen." *Zeitschrift für Astrophysik* 38, 95–109.

Hoefer, Carl (1994). "Einstein's Struggles for a Machian Gravitational Theory." *Studies in the History and Philosophy of Science* 25, 287–335.

Hoskin, Michael (1976). "The 'Great Debate': What Really Happened." *Journal for the History of Astronomy* 7, 169–182.

Hoyle, Fred (1948). "A New Model for the Expanding Universe." *Monthly Notices of the Royal Astronomical Society* 108, 372–382.

Hubble, E. (1929). "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae." *Proceedings of the National Academy of Science* 15, 168–173.

Jaki, Stanley L. (1990). "The gravitational paradox of an infinite universe." In *Cosmos in Transition*, vol. 5 of *Pachart History of Astronomy Series*. Tucson: Pachart, 189–212.

Jeans, James H. (1919). *Problems of Cosmogony and Stellar Dynamics*. Cambridge: Cambridge University Press.

Kapteyn, J. C. and van Rhijn, P. J. (1920). "On the Distribution of the Stars in Space Especially in the High Galactic Latitudes." *Astrophysical Journal* 52, 23–38.

Klein, Felix (1919). "Über die Integralform der Erhaltungssätze und die Theorie der räumlich-geschlossenen Welt." *Nachtrichten von der Königlichen Gesellschaft der Weissenschaften zu Göttingen, Mathematisch-physikalische Klasse* , 393–423.

Kragh, Helge (1996). *Cosmology and Controversy*. Princeton: Princeton University Press.

Lanczos, Cornelius (1922). "Bemerkung zur de Sitterschen Welt." *Physikalische Zeitschrift* 23, 539–543.

— (1923). "Über die Rotverscheibung in der de Sitterschen Welt." *Zeitschrift für Physik* 17, 168–189.

Layzer, David (1954). "On the significance of Newtonian cosmology." *Astrophysical Journal* 59, 268–270.

— (1956). "Newtonian cosmology." *The Observatory* 76, 73–74.

Lemaître, Georges (1925). "Note on de Sitter's Universe." *Journal of Mathematical Physics* 4, 188–192.

— (1927). "Un universe homogène de masse constante et de rayon croissant, rendant compte de la vitesse radial des nébuleuses extra-galactiques." *Annales de la Société Scientifique de Bruxelles* , 49–56.

— (1931). "Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae." *Monthly Notices of the Royal Astronomical Society* 91, 483–490.

— (1932). "L'Universe en Expansion." *Publication du Laboratoire d'Astronomie et de Géodésie de l'Université de Louvian* 9, 171–205.

Lepeltier, Thomas (2006). "Edward Milne's influence on modern cosmology." *Annals of Science* 63, 471–481.

Macpherson, Hector (1929). *Modern Cosmologies: A Historical Sketch of Researches and Theories concerning the Structure of the Universe*. London: Oxford University Press.

Malament, David (1995). "Is newtonian cosmology really inconsistent?" *Philosophy of Science* 62, 489–510.

McCrea, William H. and Milne, Edward A. (1934). "Newtonian universes and the curvature of space." *Quarterly Journal of Mathematics* 5, 73–80.

McCrea, William H. (1955). "On the Significance of Newtonian Cosmology." *Astronomical Journal* 60, 271–274.

Milne, Edward A. (1934). "A newtonian expanding universe." *Quarterly Journal of Mathematics* 5, 64–72.

— (1935). *Relativity, Gravitation, and World-Structure*. Oxford: Clarendon Press.

North, John (1995). *The Norton History of Astronomy and Cosmology*. New York: W. W. Norton & Co.

North, John D. (1965). *The Measure of the Universe*. Oxford: Oxford University Press.

Norton, John (1995). "The Force of Newtonian Cosmology: Acceleration is Relative." *Philosophy of Science* 62, 511–522.

— (1999). "The cosmological woes of Newtonian gravitation theory." vol. 7 of *Einstein Studies*. Boston: Birkhäuser, 271–323.

— (2003). "The N-Stein Family." In *Revisiting the Foundations of Relativistic Physics*, Abhay Ashtekar, Robert S. Cohen, Don Howard, Jürgen Renn, Sahotra Sarkar and Abner Shimony, eds., vol. 234 of *Boston Studies in the Philosophy of Science*. Dordrecht: Kluwer, 55–67.

Paul, Erich (1993). *The Milky Way Galaxy and Statistical Cosmology, 1890-1924*. Cambridge: Cambridge University Press.

Pauli, Wolfgang (1921). "Encyklopädie der mathematischen Wissenschaften, mit Einschluss ihrer Anwendungen." vol. 5, chap. Relativitätstheorie. Leipzig: Teubner, 539–775.

Peebles, Phillip James Edward (1980). *Large-scale Structure of the Universe*. Princeton: Princeton University Press.

R.A.S. (1930). "Report of the Royal Astronomical Society meeting in January 1930." *The Observatory* 53, 33–44.

Renn, Jürgen, ed. (2007a). *The Genesis of General Relativity*. 4 Vols. New York, Berlin: Springer.

Rindler, Wolfgang (1956). "Visual horizons in world models." *Monthly Notices of the Royal Astronomical Society* 116, 662–677.

Robertson, H. P. (1928). "Relativistic cosmology." *Philosophical Magazine* 5, 835–848.

— (1929). "On the foundations of relativistic cosmology." *Proceedings of the National Academy of Science* 15, 822–829.

— (1933). "Relativistic cosmology." *Reviews of Modern Physics* 5, 62–90.

— (1935). "Kinematics and World Structure." *Astrophysical Journal* 82, 284–301.

— (1936a). "Kinematics and World Structure. II." *Astrophysical Journal* 83, 187–201.

— (1936b). "Kinematics and World Structure. III." *Astrophysical Journal* 83, 257–271.

Röhle, Stefan (2002). "Mathematische Probleme in der Einstein - de Sitter Kontroverse." Max-Planck-Institut für Wissenschaftsgeschichte Preprint 210, Berlin.

Schrödinger, Erwin (1956). *Expanding Universes*. Cambridge: Cambridge University Press.

Schwarzschild, Karl (1913). "Review of Leçons sur les hypothèses cosmogoniques by H. Poincaré and Researches on the Evolution of the Stellar Systems by T. J. J. See." *Astrophysical Journal* 37, 294–298.

Sciama, Dennis (1957). "Inertia." *Scientific American* 196, 99–109.

Seeliger, H. (1895). "Über das Newton'sche Gravitationsgesetz." *Astronomische Nachrichten* 137, 129–136.

Selety, Franz (1922). "Beitrage zum kosmologischen Problem." *Annalen der Physik* 72, 281–334.

Shapley, Harlow (1921). "The scale of the universe." *Bulletin of the National Research Council* 2, 171–193.

Smith, Robert W. (1982). *The Expanding Universe: Astronomy's 'Great Debate' 1900-1931*. Cambridge: Cambridge University Press.

Stachel, John (1986). "Eddington and Einstein." In *The Prism of Science*. No. 2 in The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Boston: D. Reidel, 225–250. Reprinted in *Einstein from 'B' to 'Z'*.

— (1994). "Lanczos's Early Contributions to Relativity and his Relationship with Einstein." In *Proceedings of the Cornelius Lanczos International Centenary Conference*, J. David Brown, Moody T. Chu, Donald C. Ellison and Robert J. Plemmons, eds. Philadelphia: SIAM, 201–221. Reprinted in *Einstein from 'B' to 'Z'*.

— (2002). *Einstein from 'B' to 'Z'*, vol. 9 of *Einstein Studies*. Boston: Birkhäuser.

Tolman, Richard C. (1929a). "On the Astronomical Implications of the de Sitter Line Element for the Universe." *Astrophysical Journal* 69, 245–274.

— (1929b). "On the possible line elements for the universe." *Proceedings of the National Academy of Science* 15, 297–304.

Tolman, Richard C. (1934). *Relativity, Thermodynamics, and Cosmology*. Oxford: Oxford University Press. Reprinted by Dover.

Torretti, Roberto (1983). *Relativity and Geometry*. Oxford: Pergamon Press. Reprinted by Dover, 1996.

— (2000). "Spacetime models of the world." *Studies in the History and Philosophy of Modern Physics* 31, 171–186.

Trautman, Andrzej (1965). "Foundations and Current Problems of General Relativity." In *Lectures on General Relativity*, A. Trautman, F. A. E. Pirani and H. Bondi, eds., vol. 1 of *Brandeis Summer Institute in Theoretical Physics 1964*. Englewood Cliffs (NJ): Prentice Hall, 1–248.

Trumpler, R. J. (1930). "Absorption of Light in the Galactic System." *Publications of the Astronomical Society of the Pacific* 42, 214–227.

Turnbull, H. W., ed. (1959-1977). *The Correspondence of Sir Isaac Newton*, vol. 7 volumes. Cambridge: Cambridge University Press. Edited by H. W. Turnbull (Vols. 1-3), J.W. Scott (Vol. 4), and A. Rupert Hall and Laura Tilling (Vols. 5-7).

Walker, A. G. (1935). "On the formal comparison of Milne's kinematical system with the systems of general relativity." *Monthly Notices of the Royal Astronomical Society* 95: 263–269.

— (1936). "On Milne's theory of world structure." *Proceedings of the London Mathematical Society, Series 2* 42: 90–127.