An estimate of the size of the universe from a topological point of view

D. D. Sokolov and V. F. Shvartsman

(Submitted July 28, 1973)

An analysis of the observational data does not contradict the hypothesis that the space of our universe consists of a topologically closed structure of small size (regardless of the sign of the curvature) which is "glued together." Lower limits are obtained on the present size of the universe: The maximum and minimum values of the universe's "gluing parameters" are not smaller than 400 and 10 Mpc, respectively. Thus, it is possible that the maximum diameter of the universe is ten times smaller than the values commonly assumed for the distances to remote quasars. If the actual values of the "gluing parameters" are indeed of the magnitudes indicated above, then some of the sources with red shifts $z > 0.0025$ and all of the sources with $z > 0.1$ are "ghosts" of long extinct objects, and also of objects which exist at the present time close to us. In this connection a single object may correspond to several dozen ghosts on the celestial sphere.

1. INTRODUCTION

It is well known that, according to Einstein's equations, the presence of matter and its physical characteristics determine only the curvature of space-time, i.e., its local structure. F. Klein proved$^{[1]}$ in 1918 that one and the same local structure may correspond to globally different models of the universe. Later it was found that an infinite number of topologically different models of three-dimensional spaces with constant positive and negative curvature exist (see the reference cited in Appendix II). Up to now no mathematical classification of all these spaces has been given; an observational choice of the topological model which might describe the universe has not been made; it is not even clear whether one will be able to make such a choice in the future (see Secs. 4 and 5).

It is shown in the present article that an analysis of the observational data does not contradict the hypothesis that the space of our universe consists of a topologically closed structure of small size (regardless of the sign of the curvature) which is "glued together."$^{[1]}$ Let us recall the following: it is usually assumed that the size of the universe must certainly exceed the distance to the light horizon, $c t_0 = 4000$ Mpc, where $c$ is the velocity of light and $t_0$ is the age of the universe (the time from the initial singularity). The dimensions of a universe which is "glued together" may turn out to be substantially smaller. The lower bounds which we have obtained on the maximum and minimum values of the "gluing parameters" are as follows: $L \geq 400$ Mpc and $I \geq 10$ Mpc. Of course, the limitation which we found on the minimum value of the "gluing parameter" also pertains to an open universe which is "glued together."

In the present article we shall, as usual, call a spatial cross section, constructed by using a reference frame which is at rest with respect to the remnant microwave radiation, simply space. Below we shall consider only locally homogeneous and locally isotropic (Friedmann) models of the universe, because the remnant microwave radiation observed at this time is isotropic to a high degree of accuracy. In fact the class of spaces which are compatible with observations is even narrower since according to$^{[2]}$ our universe can only be described by models with an orientable space.

We emphasize that the dynamics of the expansion of locally homogeneous and locally isotropic universes which are glued together is the same as for the usual Friedmann models.

2. EXAMPLE OF A UNIVERSE WHICH IS GLUED TOGETHER

The reader can visualize the most important properties of universes which are glued together by considering as an example the space of an ordinary cylinder, i.e., the manifold which is obtained by gluing together the edges of an infinite sheet of a two-dimensional plane with Euclidean metric. Two points (the observer and the source) can be connected by an infinite number of geodesics that differ from one another by their lengths. We shall call the image of the object, corresponding to one of the shortest geodesics (in general there may be two of the shortest geodesics), the "original" object, and we shall call the remaining images its "ghosts." The ghosts can be labeled by the numbers $1, 2, 3, \ldots$ that increase monotonically with increasing distance to the ghost; we assign $n = 0$ to the original object. In the case of a cylinder, $n$ is equal to the number of complete half-turns of the geodesic around the cylinder, and the distance to the $n$-th ghost cannot be less than one-half of the circumference times the number of half-turns $n$; the angle between the original object and the ghost monotonically increases with increasing value of the number $n$.

It is easy to show that a similar picture holds in the case of a model universe whose space is obtained by identifying opposite faces of a three-dimensional cube. Of course, in connection with identifying the faces of more complicated regions than a cube, and also of regions consisting of pieces of space with nonzero curvature, a substantially more complicated picture of the distribution of the ghosts arises (i.e., more complicated than the one described above).

3. THE CASE OF A THREE-DIMENSIONAL TORUS

The manifold which is obtained by formal identification of the opposite faces of a parallelepiped is called a three-dimensional torus. It will be shown below that present observations do not contradict the model of a topologically closed universe corresponding to a three-dimensional torus, which is glued together out of a cube whose diagonal $D = 800$ Mpc (for the present we assume the space to be flat). In this connection all sources with
the closer clusters being less rich. However, the estimate by a factor of two \(D/2\) is that both catalogs cover only approximately one-half the center of the cube to one of its vertices. The quantity of objects which we can use to scale (mark off) the universe, i.e., objects of such characteristics that it will be possible to identify the ghost with the original (we recall: the ghost will have a different \(z\), it is projected on a different segment of the celestial sky, and we view it at a different angle).

It is obvious that quasars, galaxies with active nuclei, remote radio sources, and also chains, rings, and other galaxy configurations of noteworthy form are not suitable for our purposes because the period of stationary activity of such objects (the period of existence of such configurations) is surely less than \(10^9\) years—a time during which a ray of light is able to go around the "small universe" from two sides \((t' = D/2c = 1.3 \times 10^9\) years). Individual galaxies may retain their external form unchanged over somewhat longer periods of time; however, there will be of the order of \(10^7\) pieces in a cube whose diagonal is 800 Mpc, and it is impossible in practice to carry out the identification.

Apparently the objects which are most convenient for identifying—these are rich, regular galactic clusters. Such clusters are actual formations, and not a consequence of topological (optical) focusing of the galactic background: They consist almost exclusively of elliptical and spherical galaxies;\(^\text{13}\) the fraction of giant galaxies in these clusters exceeds the analogous fraction for the galactic background by a factor of 30.\(^\text{12}\) Regular clusters have a spherical shape; therefore, their shape does not depend on the angle at which they are observed. Their individual characteristics are as follows: The total number of galaxies which they contain, the composition of the galactic population, the radius of the cluster, the extent to which galaxies are concentrated near the center of the cluster, etc. The radius of rich, regular clusters is, on average, about 5 Mpc, and the velocity of relative motion of the galaxies is on the order of 1000 km/sec; therefore, even if these clusters are unstable\(^\text{3}\) their evolution time \(t^* = R/v \approx 5 \times 10^9\) years, which is longer than the time required for light to cross the universe from both sides in our model. The latter fact indicates that at the present time not only the original regular clusters, but also their ghosts, can be observed.

There are at present two large catalogs of galactic clusters at the disposal of astronomers: Abell's catalog with 2712 rich clusters\(^\text{10}\) and the catalog by Zwicky and co-workers, with 9730 clusters.\(^\text{11}\) Both catalogs include only clusters with red shifts \(z \approx 0.2\). Thus, the data contained in contemporary catalogs certainly do not contradict the choice that the size of the universe (the length \(D\) of the cube's diagonal) is equal to 1600 Mpc (this corresponds to \(z = 0.2\) for the distance from the center of the cube to one of its vertices). The quantity \(D\) cannot be appreciably smaller because the analysis of the catalogs shows that regular galactic clusters, which are originals, exist at distances right up to \(r \approx 800\) Mpc. This follows from their richness, the closer clusters being less rich. However, the estimate of the parameter \(D\) can nevertheless be reduced by a factor of two \((D/2 \approx 400\) Mpc, \(z \approx 0.1\)). The point is that both catalogs cover only approximately one-half of the celestial sphere (regions with \(z < -27^\circ\) are not included in Abell's catalog, and regions with \(z < -3^\circ\) are not included in Zwicky's catalog, and the band obscured by the Milky Way is omitted from both); therefore, in a number of cases we are perhaps observing only the ghosts of clusters, but not the originals. It is also necessary to take into consideration the difficulty of identifying clusters that are projected on different segments of the sky, and also the difficulty of determining the boundaries of a cluster (for example, the diameter of the nearest \((r = 90\) Mpc) rich, regular galactic cluster (Coma) was estimated by Zwicky\(^\text{17}\) to be 12', was estimated by Abell\(^\text{6}\) to be \(\sim 5'\), and was estimated to be \(\sim 3.5'\) by Omer et al.;\(^\text{4}\) according to\(^\text{6}\) the number of terms of magnitude 18.5\(^\text{5}\) in this cluster amounts to 800, but according to\(^\text{6}\) the number is greater than 3000).

4. AN ESTIMATE OF THE GLUING PARAMETERS IN THE GENERAL CASE

In the general case the size of the closed universe which is glued together and the position of the observer inside it can be roughly characterized by two quantities: The minimum \((l)\) and maximum \((L)\) gluings parameter. Referring to Appendix II for the exact definitions of \(l\) and \(L\), we shall use their "astronomical" definitions below: At distances \(r < L\) from the observer, there isn't a single original; at distances \(r > l\) there isn't a single ghost.

The value \(D/2 > 400\) Mpc obtained above obviously pertains to the maximum gluing parameter. An estimate of the minimum gluing parameter can be obtained by carrying out a special analysis of the catalogs, the object being to search for ghosts among the clusters with \(z < 0.2\). Such an analysis is an extremely difficult problem (see the preceding section); hitherto, since such an analysis had not been made, it was impossible to exclude the possibility that a fraction of even the closest clusters known to us might be ghosts. The individual characteristics of the "quiet"\(^\text{6}\) giant galaxies nearest to us are much better known. A very preliminary analysis leads to the conclusion that there are apparently no ghosts among them up to distances \(\sim 10\) Mpc. Therefore, a rough lower limit on the minimum gluing parameter is \(l \geq 10\) Mpc.

It is possible that the specific analysis of the distribution of galaxies and of their clusters in the sky will permit us to determine the gluing parameters or even to raise their lower limits somewhat; the appropriate recommendations for observational astronomers are presented in the next section. It is important to emphasize, however, that at the present time there is no conceivable way to disprove the assertion that our universe is actually closed and small \((L < c t_r \approx 4000\) Mpc), but it appears "large" to us due to the existence of ghosts. Let us emphasize that the concept of gluing parameter and the size of the universe are not identical. See part 5 of Appendix II for the relationship between these two quantities.

5. RECOMMENDATIONS TO OBSERVATIONAL ASTRONOMERS

The object of the analysis of the observational data is to attempt to determine the global structure of the universe by searching the sky for originals and ghosts. The method of analysis—the establishment of an identity

\[\text{THE GENERAL CASE}\]

\[\text{AN ESTIMATE OF THE GLUING PARAMETERS}\]

\[\text{RECOMMENDATIONS TO OBSERVATIONAL ASTRONOMERS}\]
between the etalon sources: giant ‘quiet’ galaxies, rich regular galactic clusters, and also configurations of these objects. Although such an analysis is severely hindered by the fact that the number of topologically distinct structures with constant curvature is infinitely large (see Appendix II), it is difficult to overestimate the importance of any results at all in this direction.

The general characteristic of the appearance of configurations of ghosts in the closed universes which are glued together is the appearance of sections of space which can be matched with each other by means of a displacement of the configuration with respect to the universal covering surface, its reflection and rotation (see Appendix II). In various special cases the easily detectable characteristics might, for example, be the following: 1) The appearance of segments of the sky which can be matched with each other by means of a displacement of the configurations with respect to the celestial sphere, its reflection, rotation, and changes in the value of $z$; 2) the appearance of chains of sources having different values of $z$, the first source being the original and the subsequent ones are ghosts having monotonically increasing values of $z$; 3) the appearance in the sky of lines or points of symmetry in the angular distribution of the sources; 4) the appearance of a quasiperiodicity in the distribution of remote objects with respect to $z$; 5) and in the dependence between the number of far-distant radio sources and the flux density from them.

We recall that according to Sec. 3 one should choose only rich, regular galactic clusters and also giant ‘quiet’ galaxies as the “standard sources.” Quasars, Seyfert galaxies, etc. can be used only by assuming that they possess a recurrent activity and by considering that the observed characteristics of these objects certainly do not change substantially during a time interval smaller than $3 \times 10^7$ years, which is the minimum time required for light to circumnavigate the universe in our model. Quasars and active galaxies may be used to determine the global structure of the universe in investigations of a statistical nature, but it is impossible to use them as standard sources, which are subject to identification with ghosts (see Sec. 3).

We wish to emphasize that gluing together the boundaries of spaces with constant curvature does not change the relations between the observable characteristics of objects (provided that the scale of the glue is larger than the dimensions of the object). For example, the relation between the angular dimensions of a cluster, its brightness, and its red shift is identical for originals and for ghosts; that is, even in a universe which is glued together this relation can be used to determine the curvature of space. In the event that our universe is indeed glued together, then there exists a remarkable opportunity to investigate the evolution of galaxies (including our own) and their clusters by means of the systematic observation of the sources’ ghosts.

APPENDIX I

THE ESTABLISHMENT OF A CAUSAL RELATIONSHIP IN UNIVERSES THAT ARE GLUED TOGETHER

Attempts have repeatedly been made to explain the observed isotropy of the remnant microwave radiation not on the basis of the initial conditions which existed at the time of the initial singularity, but on the basis of physical processes at more recent stages of the universe’s evolution. As is well known, the second explanation is impossible within the framework of the conventional Friedmann model. In fact, in the Friedmann model all points of space can be involved in a causal relationship (that is, they are able to exchange light signals) only in stages when the expansion of the universe alternates with its contraction (such a situation occurs if the curvature of the space is positive). However, right now the universe is expanding.

Regardless of what the sign of the curvature is, a universe which is glued together may have been encompassed by a causal relationship at a substantially earlier time (remark by Paal). Namely, the causal relationship is established beginning at the instant $t_1$ when

$$c t_1 = D(t_1)/2$$

Here $t_1$ is the time from the initial singularity, and $c t_1$ is the light radius of the universe. $D(t_1)$ is the maximum size of the region containing originals at the time $t_1$. Given the law $D(t)$ governing the expansion of space and assuming that at the present time $c t_0 = 4000$ Mpc and $D(t_0) = 2L(t_0) \approx 800$ Mpc, one can easily determine the time $t_0$ and the corresponding value $z_0$ of the redshift. At each moment $t$ the function $D(t)$ is determined by two parameters, the ratio of the average density $\rho_D(t)$ of matter in the universe to the critical density of matter $\rho_c(t)$, and the ratio of the radiation density $\rho_R(t)$ to the matter density $\rho_D(t)$.

Simple calculations show that for $2c t_0/D(t_0) = 10$ and $\rho_D(t_0) = 10^{-26}$ to $10^{-31}$ $g/cm^3$, $\rho_R(t_0) = 10^{-28}$ to $10^{-33}$ $g/cm^3$, the time when the causal relationship was established corresponds to a value of $z_1$ between 10 and 100. Thus, since $D(t_0)/2 \leq 400$ Mpc, in a glued-together universe the causal relationship is established after the separation of the remnant radiation from the matter (which occurs at $z^* = 1400$). Therefore the exclusively isotropic background radiation which is observed at the present time is not a consequence of the possible smallness of our universe. On the other hand, it should be emphasized that the establishment of the causal relationship for $z > 10$ facilitates equalization of the densities in the universe. In this connection it is curious to note that the far-distant radio sources apparently produce a uniform background, but they are not grouped in clusters like galaxies. One can attempt to explain the isotropy of the remnant background radiation with the aid of the hypothesis about a glued-together universe only by assuming that secondary heating of the plasma occurred in the epoch $z \leq 10$, and also the density of the universe was large enough to guarantee an effective interaction between the remnant photons and the plasma electrons.

It is also of interest to investigate the result of the general solution by Belinskii, Lifshitz, and Khalatnikov in glued-together universes on the Friedmann solution.

APPENDIX II

METRIC AND TOPOLOGICAL PROPERTIES OF UNIVERSES THAT ARE GLUED TOGETHER

The goal of the present Appendix is to give some kind of idea about the diversity of glued-together uni-
verses and to give an exact formulation of the results of the work.

1. Let us recall beforehand the meaning of the terms “topologically” and “metrically equivalent spaces.” Two spaces are called topologically equivalent if they mutually can be uniquely and continuously transformed one into the other (with an arbitrary change of the metric). Two spaces are called globally isometric (metrically equivalent) if they mutually can be uniquely transformed one into the other in such a manner that the distance between two arbitrary points of one space and the distance between their images in the other space are corresponding equal. Two spaces are called locally isometric if they can be transformed into one another in such a manner that the metric tensors at corresponding points will be identical. It is obvious that one and the same topological type may correspond to infinitely many locally isometric, but globally nonisometric, spaces (for example, cylinders of different radii) and, consequently, may correspond to substantially different observational pictures.

2. Let us outline more exactly the class of spaces which we are considering. In the present article we confine our attention to three-dimensional, complete⁸ spaces of constant curvature (it is customary to refer to these spaces as Clifford–Klein spaces[1])⁸. As was shown by Hopf[¹¹] such spaces can be characterized by the concepts of a fundamental region, i.e., an “elementary cell,” and a universal covering surface, i.e., a surface on which the fundamental region can be expanded by means of repeated actions of the fundamental group. In the case of zero, positive, and negative curvatures the universal covering spaces correspond, respectively, to three-dimensional Euclidean space, the three-dimensional sphere, and three-dimensional Lobachevski space. In this connection the fundamental groups are those discrete subgroups of the group of motions admissible by the universal covering space which do not contain fixed points. Points of the universal covering space, which can be reduced into one another by the action of the fundamental group, are called equivalent. Any arbitrary region on the universal covering space, which contains one and only one equivalent point, is called a fundamental region.

The Clifford–Klein manifold is obtained upon identifying equivalent points of the universal covering space. One can think of it as a fundamental region for which certain points on the boundary have been identified. Therefore, for purposes of clear visualization in the main text of this article we refer to such manifolds as spaces which have been glued together (in this connection, of course, it is understood that the universal covering spaces enter into their number). We also recall that spaces which have been glued-together are multiply-connected.

It is necessary to emphasize that in the three-dimensional case there exists infinitely many topologically and metrically different spaces[16,18] for each nonvanishing fixed value of the curvature; for the case of zero curvature there are 18 topologically distinct types of spaces but an infinite number of globally nonisometric structures can be defined on them. A number of examples are analyzed in¹¹[¹⁹].

3. Now let us consider the definition of the concept of a ghost. A class of isolated equivalent points, corresponding to the observer, exists on the universal covering space; we shall call one of these points the original observer (B), and the other points will be referred to as its ghosts (of the first kind).

The picture of the universal covering space is given to us in direct observation: We perceive the light rays which travel along different paths from a single object as emitted from different sources which are located at equivalent points on the universal covering surface. Let us call the original object which is closest to B the source corresponding to this object, and the sources located at equivalent points will be called its ghosts (of the first kind). Sometime a single object may correspond to several points on the universal covering surface which are located at equal, minimum distances from B; than any of these points can be regarded as the original. It is very important to note that the distinction between a ghost and the original is not absolute, but in general depends on the observer’s position.

In the case of a space of constant positive curvature, even on the universal covering surface, which is a sphere, two points can be joined by two geometrically distinct segments of the geodesic, the two segments together forming a great circle. We also perceive the images corresponding to these distinct segments as ghosts, but they are obviously not the same as ghosts of the first kind. We shall call them ghosts of the second kind. Ghosts of the second kind can also arise in connection with multiple circuits of the light ray around the great circle. We emphasize that in the stage of the universe’s monotonic expansion (which has apparently been taking place from the initial singularity up to now) ghosts of the second kind have not been able to appear since light has not been able to cross one-half of the great circle during this time interval[²⁰]. Therefore, we did not refer to ghosts of the second kind in the main text of this article. In particular, the articles by Solheim[²¹] and Piper[²²] are devoted to unsuccessful searches for ghosts of the second kind.

4. Let us denote the totality of all originals by H(B). It is obvious that H(B) is the fundamental region. The size of the region H(B) and the position of observer B inside it can be roughly characterized by two quantities: The maximum L(B) and the minimum L(0) of “gluing parameters” which are equal, respectively, to the maximum and minimum distances from B to the boundary of H(B). Generally speaking L ≥ l; for L = l the region H(B) is a sphere. It is obvious that there won’t be even a single ghost at distances r < l from the observer, and there will not be even a single original at distances r > L. The exact formulation of the results of the present article is: 1) L ≥ 400 Mpc; 2) l ≥ 10 Mpc.

5. It is reasonable to agree to call the maximum distance between two points of space the maximum size of the universe D (here the distance between the two points is measured along the minimum geodesic connecting them). It is obvious that L(B) ≤ D ≤ 2L(B). Furthermore, to each point B of space we associate the length d(B) of the segment of the minimum geodesic connecting the point B with itself. One can show that this quantity is exactly equal to double the minimum gluing parameter, i.e., 2l(B). One can call the minimum value of d(B) for all points B the “minimum size” of the universe. It is obvious that 0 ≤ d ≤ 2l(B).

It is interesting to note that among the spaces of constant negative curvature there are some for which
D = ∞ but d = 0. These spaces can have both finite and infinite volume.

The authors express their sincere gratitude to I. N. Bernshteln for numerous consultations and help in the work, and to Ya. B. Zel'dovich for his interest in the work and for advice. The authors are grateful to V. L. Ginzburg, A. B. Ivanov, I. D. Karachentsev, N. S. Kardashev, I. D. Novikov, L. M. Ozernol, Yu. N. Parilskii, E. G. Poznyak, and B. Ya. Shitvel'man for a discussion of the work and helpful comments, and we are also grateful to G. Paal for a stimulating discussion.

1) We shall call smooth spaces "glued together" if geometrical models of these spaces can be formally obtained by identifying segments of the boundaries of open manifolds. We refer to Appendix II for the precise formulation.

2) We assume that the observer is located at the center of the cube, and the Hubble constant $H = 75$ km/sec-Mpc.

3) For reasons that are not understood, the observed density of such clusters is one or two (!) orders of magnitude smaller than their virial density. [5]

4) See the preceding section for a discussion of the impossibility of using active objects as etalon sources.

5) As an example we mention the work of Burbridge and O'Dell, [10] in which it is asserted that the Fourier spectrum of the observed red shifts of quasi-stellar objects contains peaks corresponding to a wavelength of 0.07 in z. See the references cited there for similar articles by other authors.

6) In this connection, see Sandage's recent analysis of the observational data. [11]

7) We assume below that the cosmological constant $\Lambda = 0$.

8) Zel'dovich [13] has called attention to the possible existence of a dependence between the closed nature of the universe and dynamical perturbations of the density. The relation between topological and local physical laws is discussed by Dautcourt. [14]

9) A space is called complete if the limit of every fundamental sequence of its points belongs to this space.

10) The proof is based on the fact that the distances from the point B to any arbitrary point A on the boundary of $H(B)$ and to its nearest ghost $A'$ are identical, and it is also based on the fact that the shortest geodesic connecting point B with the boundary of $H(B)$ is perpendicular to the boundary of $H(B)$.


12) Ya. B. Zel'dovich and I. D. Novikov, ZhETF Pis. Red. 6, 772 (1967) [JETP Lett. 6, 236 (1967)].


14) Ya. B. Zel'dovich and I. D. Novikov, ZhETF Pis. Red. 6, 772 (1967) [JETP Lett. 6, 236 (1967)].

15) Ya. B. Zel'dovich and I. D. Novikov, ZhETF Pis. Red. 6, 772 (1967) [JETP Lett. 6, 236 (1967)].


21) Ya. B. Zel'dovich and I. D. Novikov, ZhETF Pis. Red. 6, 772 (1967) [JETP Lett. 6, 236 (1967)].

22) Ya. B. Zel'dovich and I. D. Novikov, ZhETF Pis. Red. 6, 772 (1967) [JETP Lett. 6, 236 (1967)].

Translated by H. H. Nickle