THE ASTROPHYSICAL JOURNAL, 292:371–394, 1985 May 15

THE EVOLUTION OF LARGE-SCALE STRUCTURE IN A UNIVERSE DOMINATED BY COLD DARK MATTER

MARC DAVIS,[1,2] GEORGE EFSTATHIOU,[1,3] CARLOS S. FRENK,[1,4] AND SIMON D. M. WHITE[1,5]
Received 1984 August 20; accepted 1984 November 30

ABSTRACT

We present the results of numerical simulations of nonlinear gravitational clustering in universes dominated by weakly interacting, "cold" dark matter (e.g., axions or photinos). These studies employ a high resolution N-body code with periodic boundary conditions and 32,768 particles; they can accurately represent the theoretical initial conditions over a factor of 16 in length scale. We have followed the evolution of ensembles of models with $\Omega = 1$ and $\Omega < 1$ from the initial conditions predicted for a "constant curvature" primordial fluctuation spectrum. We also ran one model of a flat universe with a positive cosmological constant. Large filamentary structures, superclusters of clumps, and large low-density regions appear at certain times in all our simulations; however, we do not find large regions as extreme as the apparent void in Boötes. The evolution of the two-point correlation function, $\xi(r)$, is not self-similar; its effective power-law index becomes more negative with time. Models with $\Omega = 1$ are inconsistent with observation if galaxies are assumed to be unbiased tracers of the underlying mass distribution. The peculiar velocities of galaxies are predicted to be much too large. In addition, at times when the shape of $\xi(r)$ matches that observed, the amplitude of clustering is inferred to be too small for any acceptable value of the Hubble constant. Better agreement is obtained for $\Omega = 0.2$, but in both cases the rms relative peculiar velocity of particle pairs decreases markedly with pair separation, whereas the corresponding quantity for galaxies is observed to increase slowly. In all models the three-point correlation function ζ is found to fit the observed form, $\zeta \propto Q\xi^2$, but with Q depending weakly on scale. On small scales Q substantially exceeds its observed value. Consistent with this, the mass distribution of clusters is very broad, showing the presence of clumps with a very wide range in mass at any given time. The model with a positive cosmological constant closely resembles an open model with the same value of Ω. If galaxies are a random sampling of the mass distribution, none of our models is fully consistent with observation. An alternative hypothesis is that galaxies formed only at high peaks of the initial density field. The clustering properties of such "galaxies" are biased; they appear preferentially in high-density regions and so are more correlated than the overall mass distribution. Their two- and three-point correlation functions and their relative peculiar velocity distribution may be consistent with observation even in a universe with $\Omega = 1$. If this is an appropriate model for galaxy formation, it may be possible to reconcile a flat universe with most aspects of the observed galaxy distribution.

Subject headings: galaxies: clustering — galaxies: formation — numerical methods

I. INTRODUCTION

The fundamental problem of understanding the structure of our universe can be approached by phenomenological study of its present configuration or by analytical study of its origins. In recent years considerable progress has been made on both these fronts, but it has become clear that a major difficulty lies in bridging the gap between them. Large-scale surveys of galaxy redshifts have revealed many aspects of the three-dimensional morphology of the galaxy distribution. (See the recent reviews Oort 1983; Davis 1984.) At the same time there has been widespread enthusiasm in the cosmological community over new ideas injected both by Grand Unified Theories and by the possibility of an early inflationary epoch. The present ratio of photons to baryons may be explained by non-equilibrium particle interactions at grand unification energies (Weinberg 1979), while inflation can account for the homogeneity and flatness of the universe and for the low abundance of magnetic monopoles (Guth 1981). Thus some of the most fundamental parameters of our universe may be explained in a natural fashion. Inflationary models make a specific prediction for the form of the density perturbation spectrum from which observed structure must grow (Guth and Pi 1982; Hawking 1982; Starobinskii 1982; Bardeen, Steinhardt, and Turner 1983). Perturbations in curvature arise from quantum fluctuations during inflation. The time invariance of de Sitter space ensures that their amplitude as they are swept across the event horizon during the inflationary epoch, and consequently as they return across it in the present Friedmann universe, is independent of scale. Their distribution thus has the Harrison-Zel'dovich "constant curvature" form. Perturbations generated in this way are adiabatic, with no fluctuation in photon-to-baryon ratio. So far inflationary models have been able to predict the correct fluctuation amplitude only for very specific and finely tuned theories for the underlying particle physics. The prediction of an adiabatic scale-invariant spectrum is not, however, restricted to such models (see, e.g., Harrison 1970; Zel'dovich 1972; Kibble 1976; Turner and Schramm 1978; Press 1980).

If any inflationary model is valid, the present value of the cosmological density parameter Ω is expected to be very close to unity. Constraints from Big Bang nucleosynthesis suggest, however, that the density of baryonic material is an order of

[1] Institute for Theoretical Physics, University of California, Santa Barbara.
[2] Astronomy and Physics Departments, University of California, Berkeley.
[3] Institute of Astronomy, University of Cambridge.
[4] Astronomy Centre, University of Sussex.
[5] Steward Observatory, University of Arizona.

magnitude lower (Yang *et al.* 1984). Furthermore, in a purely baryonic universe an adiabatic constant-curvature spectrum leads to small-scale anisotropies in the microwave background in excess of the reported upper limits (Wilson and Silk 1981; Uson and Wilkinson 1984). These apparent inconsistencies are avoided if the dark matter that appears to dominate the dynamics of galaxy clusters and of the universe as a whole is made up of weakly interacting, nonbaryonic elementary particles. Candidate particles can be grouped into three categories on the basis of their effect on the fluctuation spectrum (Bond *et al.* 1983). If the dark matter is composed of abundant light particles which remain relativistic until shortly before recombination, then it may be termed "hot." The best candidate for hot dark matter is a neutrino with a mass in the range $10 \text{ eV} < m_\nu < 100$ eV. The phase mixing of freely streaming relativistic particles erases density fluctuations of scale smaller than the horizon, and for massive neutrinos detailed calculations show that fluctuations are damped out if their present wavelength is less than the critical damping scale:

$$\lambda_c = 41(m_\nu/30 \text{ eV})^{-1} \text{ Mpc} \quad (1)$$

(Peebles 1982; Bond and Szalay 1983). A second possibility is for the dark matter particles to interact more weakly than neutrinos, to be less abundant, and to have a mass of order 1 keV. Such particles are termed "warm" dark matter, because they have lower thermal velocities than massive neutrinos and are able to erase fluctuations only on wavelengths smaller than about 1 Mpc, corresponding roughly to the mass of a bright galaxy. There are, at present, few candidate particles which fit this description. Gravitinos and photinos have been suggested (Pagels and Primack 1982; Bond, Szalay, and Turner 1982), but currently it seems more likely that supersymmetric particles have a mass in excess of 1 GeV. Any particles which became nonrelativistic very early, and so were able to diffuse a negligible distance, are termed "cold" dark matter (CDM). There are many candidates for CDM, including supersymmetric particles (with masses above 1 Gev) and axions with a mass of 10^{-5} eV (Preskill, Wise, and Wilczek 1983).

A priori the most plausible elementary-particle candidate for the dark matter would seem to be the neutrino. Extensive studies of a neutrino-dominated universe have been carried out using analytic and N-body techniques, while related studies have focused on the hydrodynamics of pancake formation and cooling (Klypin and Shandarin 1983; Frenk, White, and Davis 1983; Centrella and Melott 1983; White, Frenk, and Davis 1983; Shapiro, Struck-Marcell, and Melott 1983; Bond *et al.* 1983). The transient filamentary structure that forms in these models is reminiscent of the long supercluster chains apparent in the galaxy distribution. However, it is not possible to reconcile the relatively small length-scale of galaxy clustering with the large length-scale of equation (1), unless galaxies formed at unacceptably recent epochs and in a manner very different from that predicted in the standard pancake picture. For these and other reasons, it is very unlikely that stable massive neutrinos dominate the universe (White, Frenk, and Davis 1983; White, Davis, and Frenk 1984).

Detailed calculations of the linear evolution of the CDM density-perturbation spectrum have been given by several authors (Peebles 1982, 1984; Blumenthal and Primack 1983; Bond and Efstathiou 1984). During the radiation-dominated epoch, fluctuations grow significantly only while their wavelength exceeds the horizon scale. However, once the universe is matter-dominated, all linear modes grow at the same rate. As a result the fluctuation spectrum develops a bend at a scale corresponding to the horizon size at the transition between the two epochs. For a constant-curvature initial spectrum, the power density at late times at comoving wavenumber k tilts gradually from the Harrison-Zel'dovich form, $|\delta_k|^2 \propto k$, to $|\delta_k|^2 \propto k^{-3}$ over more than two decades in scale. It is well approximated by the following expression:

$$|\delta_k|^2 = Ak/(1 + \alpha k + \beta k^{3/2} + \gamma k^2)^2 , \quad (2)$$

where A is a normalization constant. If k is expressed in terms of present length units, the parameters in equation (2) take the values $\alpha = 1.7l$, $\beta = 9.0l^{3/2}$, $\gamma = 1.0l^2$, where $l = (\Omega h^2 \theta^{-2})^{-1}$ Mpc, $h = H_0/100$ km s^{-1} Mpc^{-1}, Ω is the present cosmological density parameter, and θ is the microwave background temperature in units of 2.7 K. This fit is based on the numerical calculations of Blumenthal and Primack (1983) and of Bond and Efstathiou (1984), which have $\Omega = 1$ and $\theta = 1$ and include three massless neutrino flavors. If the baryon density is small compared to that of the CDM, Silk damping has a negligible effect on the CDM fluctuation spectrum and the scaling with l becomes exact; otherwise it is approximate. The scaling of l with $\Omega h^2 \theta^{-2}$ simply reflects that of the horizon diameter when matter and radiation had equal energy density; it is independent of the nature of the CDM. As a result, it may be possible to constrain this combination of cosmological parameters by requiring the nonlinear structure that forms at late times to agree with observation. Alternatively, it may be possible to reject the CDM model, if we find no match to the observations for acceptable values of the cosmological parameters. Since the overall amplitude of fluctuations cannot, at present, be calculated *a priori*, we are free to adjust it to give the correct level of clustering today. This is the only free parameter of the model.

In this paper we study the nonlinear evolution of clustering from the linear initial conditions characterized by equation (2). Analytic treatments of this problem can be carried through only for simple, highly symmetric models of clusters or voids. Peebles (1984) and Blumenthal *et al.* (1984) have investigated the cosmogony of CDM universes by applying scaling arguments based on such models to the predictions of linear theory. While these studies give considerable insight into galaxy formation, clustering patterns in the real universe show no special symmetry, and direct simulations are required to study their morphology and their statistical properties. We have studied the full nonlinear problem by running ensembles of N-body simulations of Einstein–de Sitter, and of open Friedmann models, together with one simulation with zero curvature but with a nonzero cosmological constant, $\Lambda > 0$. Our models have an initial resolution limit corresponding to a present scale in excess of 2 Mpc and so are insensitive to details of the initial spectrum below this wavelength. As a result, the nonlinear clustering that they develop differs very little from that expected in a universe dominated by warm dark matter.

Results from a single simulation of an "axion/gravitino/photino–dominated universe" have been published previously by Melott *et al.* (1983). Unfortunately, these authors chose an initial power spectrum (power-law segments with $n = 1$ and $n = -3$ separated by a sharp break) which is an extremely poor approximation to that expected in a universe dominated by cold dark matter (eq. [2]). In addition they used a particle-mesh code on a 32^3 grid, and so had too little resolution to be able to follow the hierarchical growth of structure from small scales (see Efstathiou *et al.* 1985, hereafter EDFW). As a result, the relevance of their work to CDM universes is unclear.

In § II we describe our numerical methods and the way in which we generate initial conditions, and we catalog the simulations we have performed. The evolution of the fundamental statistical properties of the models (power spectra, two- and three-point correlation functions, velocity correlations, multiplicity functions, cluster properties, and cell-count distributions) is described in § III, where we also discuss their compatibility with observation. In § IV we present graphical comparisons of our open models with the observed galaxy distribution in a large redshift survey. We conclude from the data in these two sections that an $\Omega = 1$ model is quite unacceptable if galaxies trace the mass distribution and that models with $\Omega \approx 0.2$, while better, still do not provide a fully acceptable match to observation. Finally, in § V we investigate the consequences of relaxing the assumption that the galaxy distribution follows the mass. Instead we identify "galaxies" with high peaks in the initial density distribution; such points form a biased subset of the overall mass distribution and have enhanced correlation properties. This procedure models a situation in which galaxy formation was suppressed except in sufficiently dense regions. It leads to models which can agree with observation quite well even for $\Omega = 1$. Conclusions and general remarks are given in § VI.

II. THE SIMULATIONS

All the simulations discussed here were done on VAX computers using a particle-particle/particle-mesh code (P^3M; Efstathiou and Eastwood 1981; Hockney and Eastwood 1981). This code was designed to have high force resolution and uses a comoving formulation of the equations of motion to follow evolution within the fundamental cube of a triply periodic universe; a detailed examination of its performance and a comparison with simpler particle-mesh codes are given by EDFW. In our models the long-range force field was calculated by Fourier methods on a grid of 64^3 cells. Inclusion of a short-range force correction resulted in point-mass interactions between particles separated by more than a softening length, $\eta = L/213$, where L is the comoving length of the side of our computational volume. At separations smaller than η our force law corresponded to the interaction of two interpenetrating spheres with radii $\eta/2$ and linear density profiles. We applied the short-range correction only at separations below 9η; the discontinuity in the total force at this separation was less than 2%, and the force field was isotropic on all scales to much better accuracy than this. All our simulations integrated the motion of 32,768 particles using a leapfrog scheme with a, the expansion parameter, as the time variable and a timestep of $\Delta a = 0.02$. Their accuracy can be judged using the Layzer-Irvine cosmic energy equation; typically the integration constant in the integral form of this equation had changed by less than 0.1% of the current potential energy at any stage of a run (see EDFW). The longest simulations were integrated until $a = 8.4$ and took about 70 hours of CPU time on a VAX 11/750.

We generated initial conditions using the Zel'dovich (1970) approximation, following the algorithm described by EDFW. Starting with the particles on a cubic lattice, we assigned a displacement and peculiar velocity to each in proportion to the gradient of the potential derived from a random realization of the linear fluctuation distribution. The latter corresponded to a superposition of 64^3 waves with random phases and with amplitudes distributed normally with variance given by equation (2). In all cases velocities were set so that only the growing mode was present. Since we used 32^3 particles, waves with a projected wavelength along any coordinate axis less than $L/16$ were undersampled; initial perturbations cannot be generated at such high frequencies. We therefore chose the amplitude of the initial perturbations so that the theoretically predicted power would be equal to the white-noise level at a wavelength of $L/16$. Tests show that this choice leads to the best possible random-phase representation of the theoretical spectrum and that the subsequent evolution of all structures made up of more than a few particles is unaffected by residual artificialities of the method (see EDFW § V, and Figs. 2 and 5 below). For our Einstein–de Sitter models we used equation (2) with the parameters as given; for our open models we used parameters obtained by fitting this equation to Blumenthal and Primack's (1983) results for $\Omega = 0.2$. The resulting spectra agree quite well with the expected scaling; small residual differences are reflected in a reduction of power in the open models by about 10% at long wavelengths (see Fig. 2*b*).

Because the linear CDM fluctuation spectrum has considerable power on large scales, statistically equivalent initial conditions can lead to substantially different levels of clustering even after expansion by quite a modest factor. Thus several simulations from different realizations of the same initial spectrum must be compared to assess the statistical uncertainties in the results. Furthermore, although the spectrum behaves asymptotically as $|\delta_k|^2 \propto k^n$, with $n = 1$ on large scale and $n = -3$ on small scale, the transition between these two regimes is quite slow; the length scales where $n = 0.5$ and $n = -2.5$ differ by a factor of 100 (cf. eq. [2]). Since we can reproduce this spectrum only over a factor of 16 in scale, simulations in different-size "boxes" are useful to extend the dynamic range of our results and to study the effect of their limited resolution.

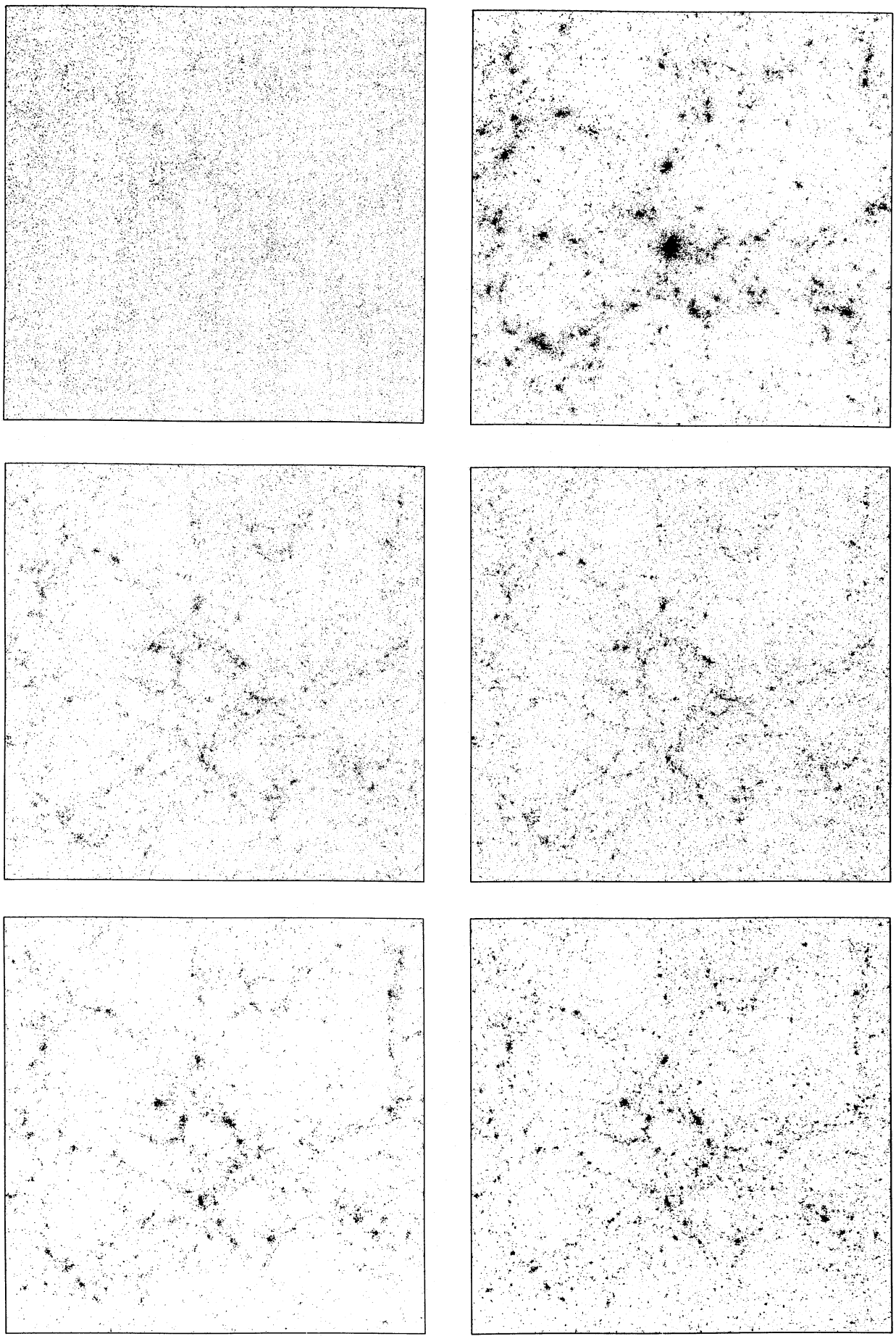
We ran three ensembles of simulations with the physical parameters given in Table 1. While h and θ affect the simulations only through their relation to the length scale L, the value of Ω influences the dynamical evolution and so must be specified separately. Henceforth we assume $\theta = 1$. Models EdS1–5 represent an $\Omega = 1$ universe with $L = 32.5h^{-2}$ Mpc in present units. Models O1–4 were evolved from similar initial conditions, but the mass of each particle was reduced so that $\Omega = 0.2$ after an expansion of 3.2. If this time is identified with the present, then $L = 162.5h^{-2}$ Mpc. Output times were chosen so that the linear growth factors in the two ensembles would match. Models EdS6–8 also had $\Omega = 1$, but the present size of the box was taken to be $L = 65h^{-2}$ Mpc. Again the output times were chosen so that one could compare these results to EdS1–5 at the same linear fluctuation amplitude. Finally, we ran a single model (L1) of a universe with zero curvature and a

TABLE 1

MODEL PARAMETERS

Model	L (l)	η (l)	M_T (m)	m_p ($10^{-4}m$)	Ω_i
EdS1–EdS5	32.5	0.15	9.4	2.9	1
EdS6–EdS8	65.0	0.30	75	23	1
L1	32.5	0.15	9.4	2.9	0.796
O1–O4	32.5	0.15	9.4	2.9	0.446

NOTE.—L is the side of the computational box, M_T the total mass it contains, η the softening length of the force law, and m_p the mass of an individual particle. The length and mass units are defined by $l = (\Omega h^2\theta^{-2})^{-1}$ Mpc and $m = 10^{15}(\Omega h^2)^{-2}\theta^6\ M_\odot$, where Ω, h, and θ denote the *present* value of the cosmological parameters. Ω_i is the value of Ω at the start of the simulations.



positive cosmological constant. Its initial condition matched those of EdS1 and O1, and the initial values of Ω and Λ were chosen so that when $\Omega = 0.2$ the linear fluctuation growth factors in O1 and L1 would agree. Because of our adopted normalization there are substantial initial fluctuations on small scales, so that many small objects collapse immediately after the start of a simulation. The steep initial spectrum then results in the rapid formation of larger structures as evolution proceeds. For models EdS1–5 the clustering scale approaches the box size for $a \geq 4$ and the experiments must be terminated.

III. THE EVOLUTION OF CLUSTERING

a) *Morphology of Clustering*

We begin by examining pictures of the models to get a qualitative feel for the results. Figure 1 shows four projections of model EdS1 and two of O1; the times plotted for O1 were chosen so that its linear-fluctuation amplitude corresponded to that of EdS1 at the two intermediate times. The initial conditions of the two models are almost indistinguishable, so only those of EdS1 are shown. Note that they have substantial large-scale fluctuations, and that careful examination reveals a trace of the underlying grid from which the particles were perturbed. As the simulation evolves, structure appears very rapidly, and filaments, superclusters of clumps, and large low-density regions develop. The range of structure present is much richer than in similar models evolved from white-noise initial conditions (Efstathiou and Eastwood 1981; Efstathiou and Barnes 1983). Because of the strong subclustering which is present, the filaments and voids in these models are less conspicuous than in simulations of neutrino-dominated universes. At times $a = 1.8$ in EdS1 and $a = 3.2$ in O1, the two-point correlation function is close to a power law, $\xi \propto r^{-\gamma}$, with $\gamma = 1.8$; under an appropriate scaling, these pictures might therefore be considered a reasonable statistical match to the observed galaxy distribution (however, see our more detailed analysis below). At later times much of the substructure is erased as clumps merge into larger and larger systems. By the end of the integration of EdS1, at $a = 4.5$, the distribution is dominated by a few large clumps with little apparent substructure. Note how closely the structure matches in the two models with different values of Ω. Nevertheless, it is clear that clusters are more concentrated in the open model. The simulation with a positive cosmological constant matches O1 very closely at the time when $\Omega = 0.2$ and the linear growth factor is 1.8 in both models.

b) *Power Spectra and Two-Point Correlation Functions*

In this section we consider the power spectrum $|\delta_k|^2$ of our simulations and its Fourier transform, the two-point correlation function $\xi(r)$. To obtain power spectra we derived a density field from the point distribution using a nearest-grid-point mass assignment on a 64^3 grid; we then Fourier transformed the result and binned the power density to the nearest integral wavenumber (see EDFW). The evolution of the power spectrum for EdS1–5 is shown in Figure 2*a*. At the initial time the spectrum flattens to the white noise level for $k > 16$, the Nyquist frequency corresponding to the initial particle grid, but is otherwise very close to the input theoretical power spectrum; remaining differences at small k are due to sampling noise. (Here and below we quote values of k in units such that $k = 1$ corresponds to the fundamental wave of our box.) The effective spectral index for this initial condition ranges from $n \approx -1$ for $1 < k < 6$ to $n \approx -2$ for $6 < k < 12$. At later times and at high wave number ($k > 16$), the power in excess of white noise is reduced somewhat as a result of the smoothing introduced by the mass assignment; it is also affected slightly by aliasing from wavenumbers with $k > 32$. In linear theory, all Fourier modes grow at the same rate, and $|\delta_k|^2 \propto a^2$ for $\Omega = 1$. We can check that the models remain linear on the largest scales by measuring the growth rate of the $k = 1$ mode and checking that it follows this scaling. In Figure 2*a* the growth expected is 12.2, slightly greater than the factor 10.5 found from the data; the discrepancy appears only at the end of the simulations. Once the fundamental mode becomes nonlinear, its growth rate is expected to decrease in this way (cf. EDFW). In the present models the fundamental mode saturates and grows very slowly for $a > 4$. At this stage, the periodic boundary conditions are becoming unrealistic, and the integration must be stopped. This reflects the fact that there is only a relatively small time window between the collapse of the smallest and largest structures in our models.

Figure 2*b* shows the evolution of the power spectrum of our ensemble of open models. The corresponding results from Figure 2*a* are superposed in order to facilitate comparison and have been multiplied by 0.9 to account for the slight difference between the theoretical input spectra (see § II). The growth of large-scale perturbations ($k < 8$) matches extremely well in the two sets of models, even though nonlinearities appear to have affected the evolution significantly for $k \geq 4$. At higher frequencies ($k > 8$) and at later times, there is an excess of power in the open models. This excess reflects the difference in cluster concentration visible in Figure 1 and is a consequence of the larger expansion of the open models after cluster collapse. When times with the same linear growth factor are compared, the power spectrum of the model with a positive cosmological constant is very similar to that of the corresponding open model but with slightly less power at high frequency.

An important test of the simulations comes from comparison of the power spectra for calculations in boxes of different size. Figure 3 shows the evolution of the power spectrum of the "large box" ensemble superposed on the appropriately shifted version of the data from Figure 2*a*. There are no free-fitting parameters in this comparison and, in the absence of spurious effects due to the finite particle number and the finite box size, the two power spectra should coincide. At large k there are slight differences due to the smoothing inherent in our method for calculating the power spectrum, but in the range $2 < k < 8$ the agreement is very good. This shows that the evolution of the Fourier components accessible to the "small box" calculation is not affected by its finite size. In principle, we could thus use linear theory to specify the evolution of longer wavelength components and so embed our models in an "infinite" universe. We have not bothered to do this because the neglected large-scale power is quite small at the times that interest us. The evolution of the power spectra in Figures 2 and

FIG. 1.—Two-dimensional projections of the particle positions in an Einstein–de Sitter and an open CDM simulation. The initial conditions for EdS1 are shown at top left. Subsequent plots show the model after expansion by factors, $a = 1.8$ (*center left*), $a = 2.4$ (*bottom left*), and $a = 4.5$ (*top right*). The remaining two plots show O1 at $a = 3.2$ when the linear growth factor from its initial conditions is 1.8 and $\Omega = 0.2$ (*center right*), and at $a = 8.4$ when the growth factor is 2.4 and $\Omega = 0.09$. Note that the side of the box corresponds to $32.5(\Omega h^2)^{-1}$ Mpc in these models, where Ω and h are the *present* values of the cosmological parameters.

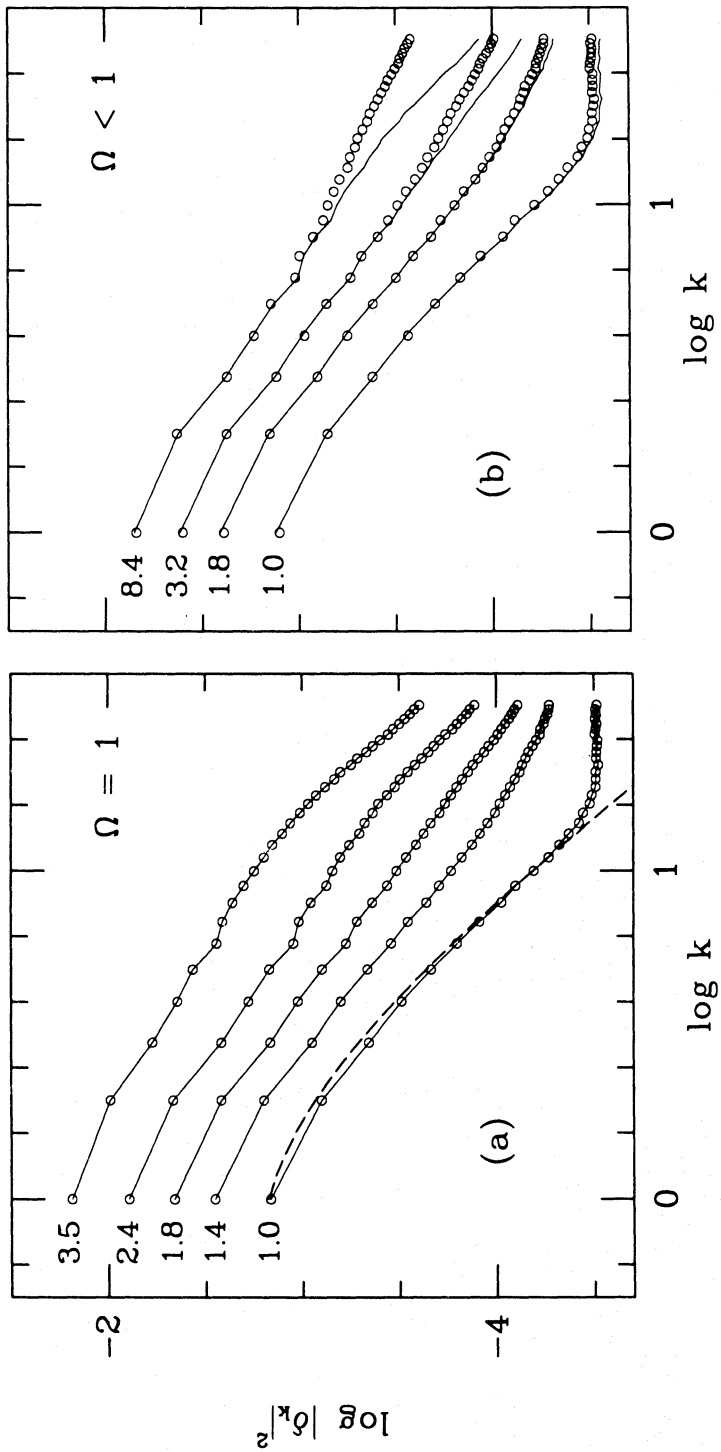


FIG. 2.—The evolution of the power spectrum for two ensembles of models, (*a*) EdS1–5 and (*b*) O1–4. The value of the expansion parameter *a* corresponding to each spectrum is given. The dashed line in (*a*) shows the theoretical spectrum (eq. [2]) used to generate the initial conditions. The solid lines in (*b*) repeat the results of (*a*) after multiplication by 0.9.

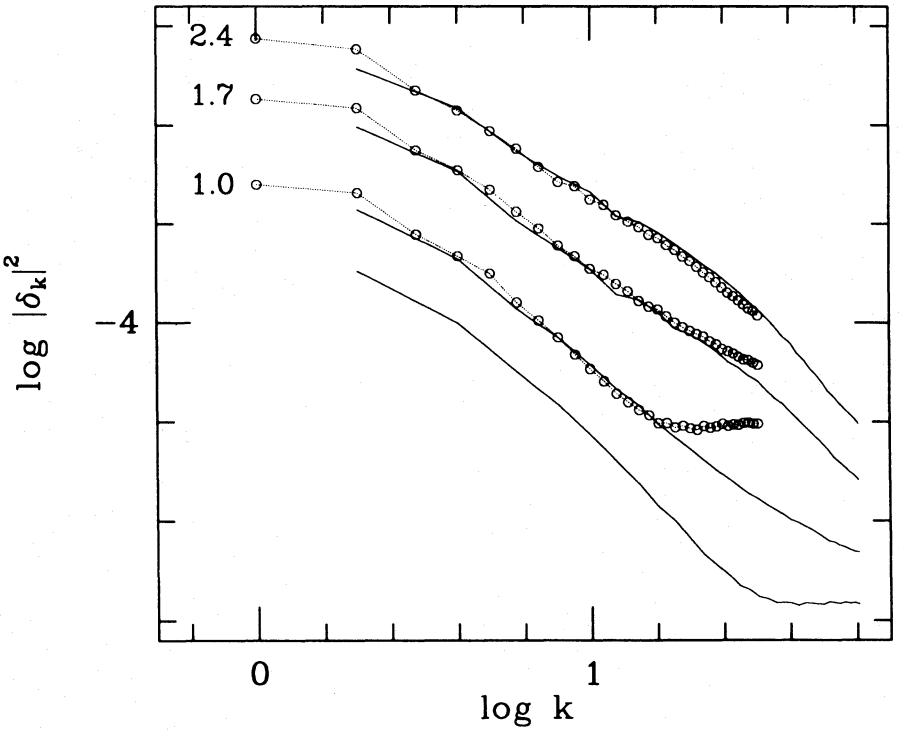


FIG. 3.—The evolution of the power spectrum of the "big box" models, EdS6–8, is compared with appropriately shifted results for our "standard" models, EdS1–5 (shown as solid lines). The agreement at small wavenumber shows that the evolution of large-scale structure in the standard models has not been adversely affected by the absence of wavelengths larger than the box size.

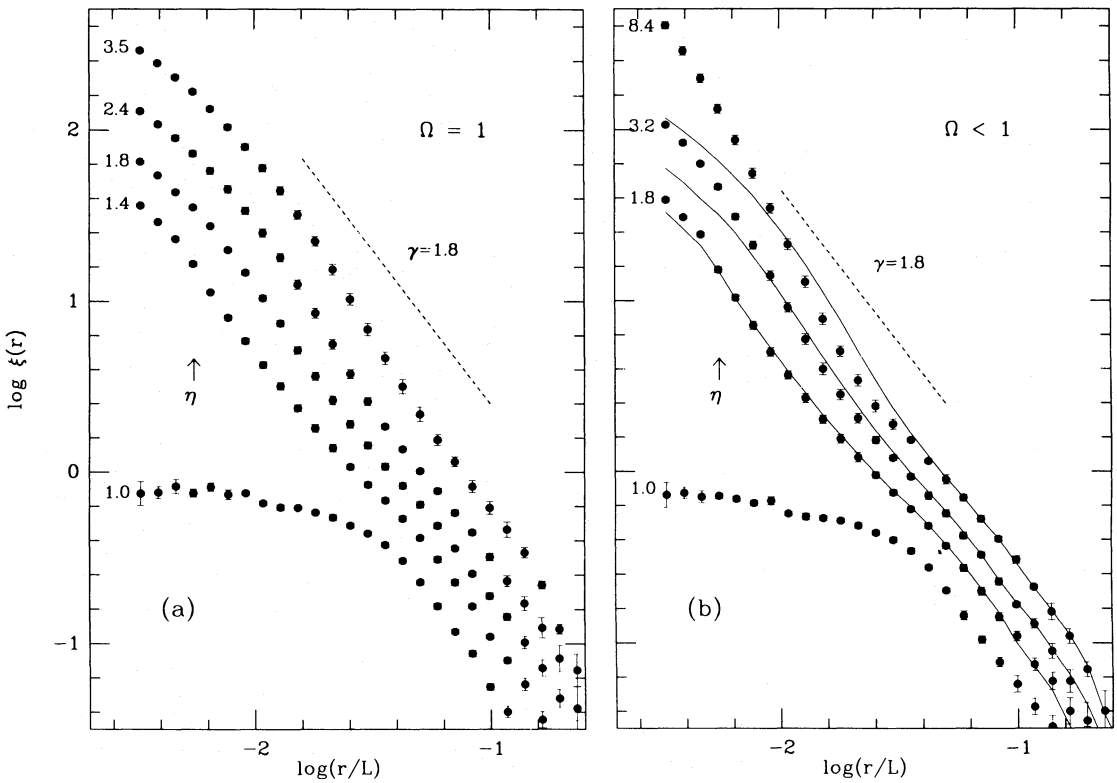


FIG. 4.—Two-point correlation functions $\xi(r)$ are shown for models (*a*) EdS1–5 and (*b*) O1–4, for various values of the expansion parameter *a*. The separation *r* is given in units of the side of the computational volume. The error bars give errors in the mean derived from the scatter within each ensemble. The dotted line is a power law with the index $\gamma = 1.8$, which fits the galaxy distribution. The solid lines in (*b*) repeat the results in (*a*) after multiplication by 0.9. An arrow in each panel marks the softening length of our simulations.

3 is such that as short-wavelength modes go nonlinear, they grow more rapidly than the linear mode, resulting in a progressive flattening of the power spectra. The power spectra do not evolve in a self-similar fashion.

A corresponding lack of self-similarity is shown by the evolution of the two-point correlation function $\xi(r)$, which we plot in Figure 4*a* for models EdS1–5. In order to achieve the highest possible resolution, correlation functions were computed by pair counts rather than by a Fourier transform of the power spectrum; all pairs were used for $\log(r/L) < -1$, and about 10% of pairs were used for larger separations. The effects of force softening are apparent as a slight flattening of $\xi(r)$ on the smallest scales shown [recall that $\log(\eta/L) = -2.3$], but generally the correlations have near power-law form on large scales. Note that $\xi(r)$ steepens with time and has a power-law index $\gamma = 1.8$ at $a = 1.8$. By the end of the run $\gamma = 2.1$, which is much steeper than the observed galaxy function. Power law fits to ξ over the range $30 > \xi > 0.3$ are given in Table 2 for all three of our ensembles. In Figure 4*a* the growth of $\xi(r)$ on large scale matches the linear expectation, $\xi \propto a^2$, while on small scale [$\log(r/L) = -2.5$], ξ is growing as $a^{2.2}$. This reflects the accelerated growth seen also in the power spectrum. As we discuss in the next section, the nonlinear clustering in these models does not match the predictions of earlier similarity arguments (Peebles 1974, 1980; Davis and Peebles 1977).

Two-point correlation functions for the open models are shown in Figure 4*b* and are compared with those of EdS1–5 at times chosen to have the same linear growth from the initial conditions. The correlations in the EdS models have been multiplied by 0.9 to account in the mean for the slight difference in the initial power spectra discussed in § II. As we found in earlier work (Frenk, White, and Davis 1983), the correlation functions agree well when times are matched in this way. For the open and positive Λ models, ξ is somewhat larger on small scales, and smaller on intermediate scales, than in the flat models, consistent with the excess power seen at high frequencies in Figure 2*b*. As Ω becomes small this disagreement becomes worse; $\xi(r)$ steepens on small scales but stops growing at large scale. At late times its shape is not well fit by a single power law, so that the exponents given in Table 2D must be treated with some caution.

TABLE 2

POWER-LAW FITS TO $\xi(r)$ OVER THE RANGE $0.3 \leq \xi \leq 30$

A. EdS1–EdS5

a	Δ	γ	$\log(r_0/L)$
1.4.....	1.4	1.70 ± 0.01	-1.55 ± 0.01
1.8.....	1.8	1.84 ± 0.01	-1.41 ± 0.01
2.4.....	2.4	1.96 ± 0.02	-1.29 ± 0.02
3.0.....	3.0	2.09 ± 0.02	-1.19 ± 0.02
4.0.....	4.0	2.14 ± 0.04	-1.06 ± 0.03

B. EdS6–EdS8

a	Δ	γ	$\log(r_0/L)$
1.3.....	1.8	1.65 ± 0.01	-1.76 ± 0.01
1.7.....	2.4	2.02 ± 0.01	-1.60 ± 0.02
2.1.....	3.0	2.14 ± 0.02	-1.50 ± 0.02
2.4.....	3.5	2.17 ± 0.03	-1.43 ± 0.03
2.8.....	4.0	2.21 ± 0.03	-1.37 ± 0.03

C. "Galaxies"

a	Δ	γ	$\log(r_0/L)$
1.4.....	1.4	1.96 ± 0.04	-1.17 ± 0.04
1.8.....	1.8	1.99 ± 0.05	-1.13 ± 0.04
2.4.....	2.4	1.92 ± 0.05	-1.07 ± 0.04
3.0.....	3.0	2.09 ± 0.07	-1.01 ± 0.05

D. O1–O4

a	Δ	γ	$\log(r_0/L)$	Ω
1.4.....	1.2	1.52 ± 0.01	-1.72 ± 0.02	0.37
1.8.....	1.4	1.73 ± 0.01	-1.59 ± 0.01	0.31
3.2.....	1.8	1.82 ± 0.02	-1.45 ± 0.02	0.20
5.1.....	2.1	1.75 ± 0.02	-1.39 ± 0.03	0.14
8.4.....	2.4	1.75 ± 0.02	-1.33 ± 0.02	0.09

NOTE.—Δ is the amplitude of mass fluctuations in the linear regime relative to that in the initial conditions of EdS1–5. The quoted errors come from least-squares fits to $\log[\xi(r)]$.

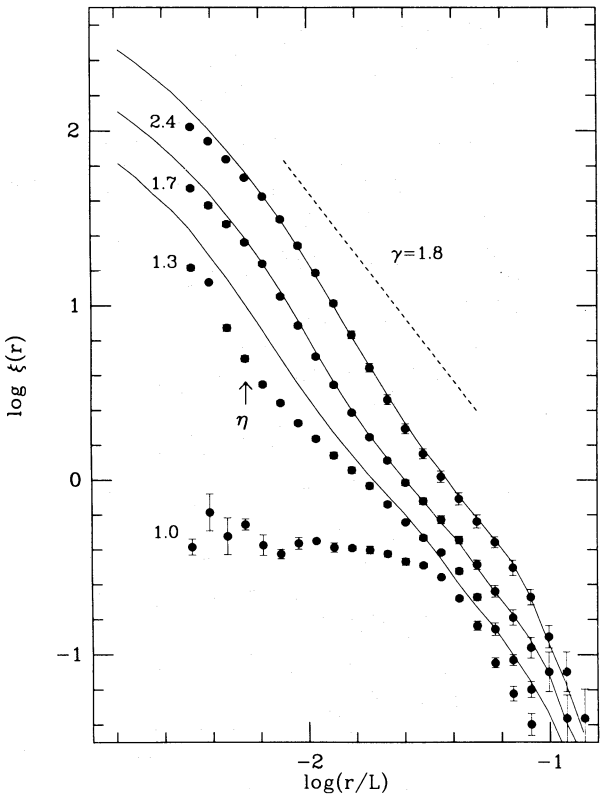


FIG. 5.—A test of boundary and discreteness effects. The points show ξ averaged over the ensemble EdS6–8 at various values of the expansion variable. The box length in these models is $L = 65h^{-2}$ Mpc. The correlation functions for the ensemble EdS1–5 at corresponding linear fluctuation amplitudes are shown as solid lines. These have been shifted by a factor of 2 to the left to enable a direct comparison with the points. Note the excellent agreement at late times and the size of the error bars (these are often smaller than the symbols plotted).

Figure 5 compares the two-point functions of our two ensembles with $\Omega = 1$ and boxes of differing size. The lines representing our "standard" models (EdS1–5) have been shifted to the left by a factor of 2 to account for the difference in box size, and the times plotted were chosen so that the linear fluctuation amplitudes on large scale should match. Once again we have no freedom to adjust the fit in this diagram; the curves should coincide, as indeed they do once initial transients have had time to subside. Slight differences on very small scale show the effect of the larger softening in the "big box" models. This test confirms the similar test carried out above for the power spectra and demonstrates that the rapid steepening of

$\xi(r)$ in the smaller simulations is an intrinsic property of the nonlinear clustering process rather than a consequence of their finite size. (This behavior is expected because of the non–self-similarity of the initial conditions.) This test also shows that the statistical properties of the small-scale structure are very similar in models where the particle masses differ by a factor of 8 and the softening lengths by a factor of 2. Thus, discreteness and two-body relaxation effects do not seem to be important on the scales which concern us (cf. Fall 1979). Figure 5 suggests that we can use periodic boxes of different size to extend our dynamic range; however, this technique is of limited use in practice, because we can never evolve a simulation beyond the time when the $k = 1$ modes begin to go nonlinear.

Under the hypothesis that the galaxies are an unbiased tracer of the mass distribution in the universe, we must require $\xi(r)$ in our models to have the form of the observed galaxy function. In addition, we must scale the models so that its amplitude is correct. The correlation length, r_0, defined by $\xi(r_0) = 1$, is measured to be $5 \pm 0.7h^{-1}$ Mpc for the galaxy distribution (Davis and Peebles 1983). At the time when the slope of the correlation function in our models has the observed power-law index, $\gamma = 1.8$, we find $r_0 = 1.27h^{-2}$ Mpc for the $\Omega = 1$ ensemble (at $a = 1.8$) and $r_0 = 5.8h^{-2}$ Mpc for the open ensemble (at $a = 3.2$, when $\Omega = 0.2$). Thus, for consistency with the observations, $h = 0.25$ is required if $\Omega = 1$ and $h = 1.1$ if $\Omega = 0.2$; both these values are outside the range covered by most modern determinations. For acceptable values of h, our models do not simultaneously reproduce the correct shape and the correct amplitude for $\xi(r)$. For example, when $a = 4$ in the EdS1–5 ensemble, we find $h = 0.57$ and $\gamma = 2.14$ (see Table 2A). The uncertainty as to which model time should be identified with the present is probably sufficient to allow large, but acceptable, values of h for the open models. Matching the correlation length of the positive Λ model to observation when $\Omega = 0.2$ also leads to $h = 1.1$, but in this case the age of the universe is 2.3 times the usual Einstein–de Sitter value, $\frac{2}{3}H_0^{-1}$.

A more direct test of whether the simulations match the observed shape of ξ is illustrated in Figure 6. Here we show the projected correlation function,

$$w(r_p) = \int_{-\infty}^{\infty} dy \xi[(r_p^2 + y^2)^{1/2}] , \quad (3)$$

for the $\Omega = 1$ models at $a = 1.8$ and for the open models at $a = 3.2$. In evaluating the integral in equation (3), we used the mean values of $\log \xi$ shown in Figures 4*a* and 4*b*. Note that the results are significantly affected by our force softening to the left of the vertical arrow. The filled circles in Figure 6 show observational estimates of $w(r_p)$ from the CfA redshift survey (Davis and Peebles 1983, Fig. 2). Our $\Omega = 1$ results fit the data very well, provided $h = 0.22$. The open arrows in the figure show how the model function shifts if it is scaled using $h = 0.3$. Clearly the value of h required for agreement at any given model time is determined to about 10%. The major uncertainty lies in deciding which model times could be consistent with the observational data. At $a = 3.5$ the amplitude of $w(r_p)$ matches the data if $h = 0.47$, but its dependence on r_p is much too steep to be acceptable. The amplitude of $w(r_p)$ shown for the $\Omega = 0.2$ ensemble in Figure 6 corresponds to the choice $h = 1.1$. In this case its shape deviates significantly from a power law, but the model still represents the CfA data fairly well. A value of Ω in the range $0.2 < \Omega < 1$ would clearly produce rough agreement with both the shape and the amplitude of the observed correlation function for any of the currently popular values of h. However, as we now discuss, such models would predict peculiar velocities for galaxies which are considerably larger than those observed.

c) *Velocity Moments*

The relative peculiar velocity distribution of pairs provides another important diagnostic for the simulations. Consider the moments of this distribution as a function of pair separation. The first moment points along the separation vector of the pairs and must cancel the Hubble flow if the pairs are to

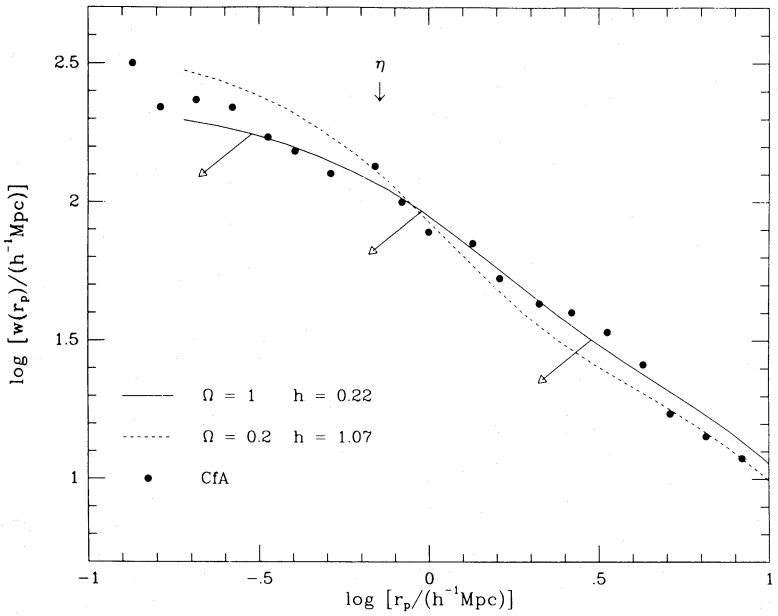


FIG. 6.—The projected correlation function $w(r_p)$ (eq. [3]). The filled circles show the results from the CfA redshift survey (Davis and Peebles 1983, Fig. 2), while the curves show results from our open and Einstein–de Sitter ensembles at $a = 3.2$ and $a = 1.8$ respectively. The vertical arrow marks the softening length of our simulations, and the diagonal arrows show how the results for a flat universe would change if they were scaled using $h = 0.3$ rather than $h = 0.22$.

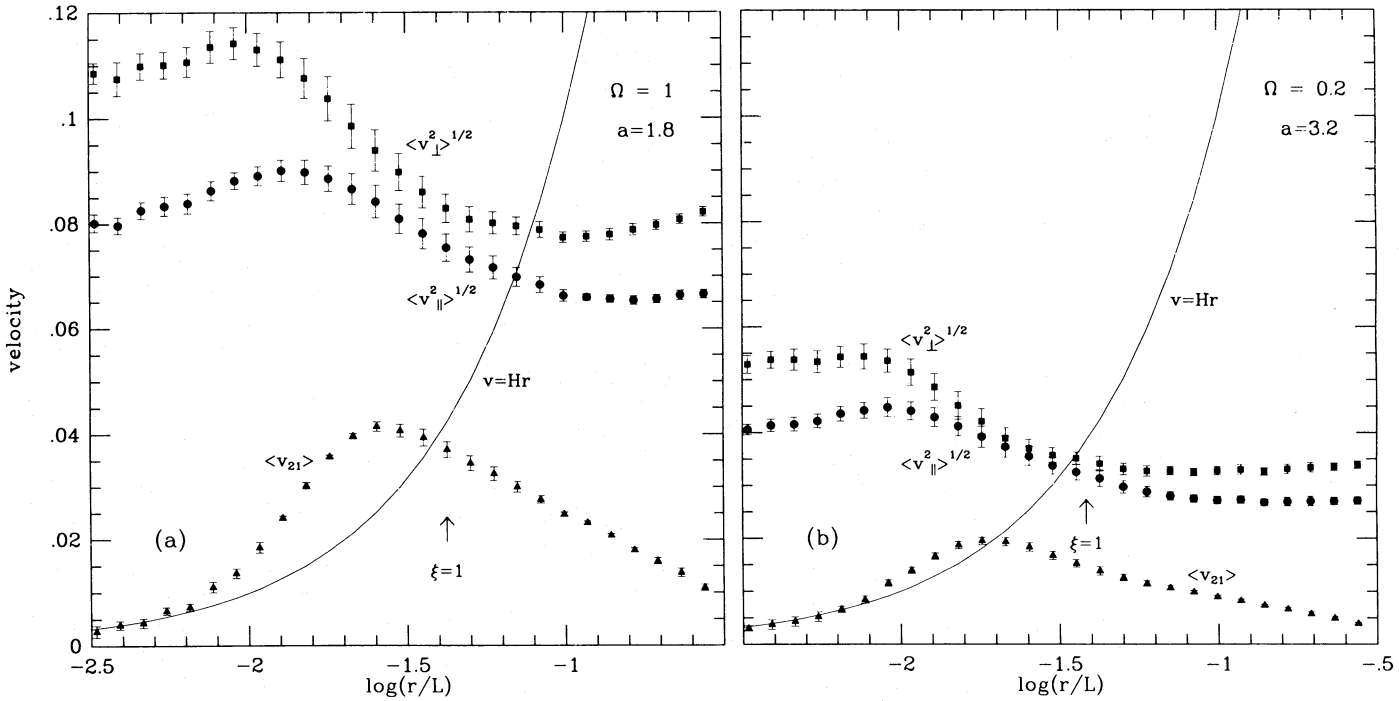


FIG. 7.—Velocity moments are shown in (*a*) for the $\Omega = 1$ ensemble EdS1–5 at $a = 1.8$, and in (*b*) for the open ensemble O1–4 at $a = 3.2$. The velocities have been scaled to units in which the Hubble velocity across the computational box is unity. A velocity of one thus corresponds to $3250(\Omega h)^{-1}$ km s^{-1}. The solid line shows the Hubble flow $v = Hr$. $\langle v_\parallel{}^2 \rangle^{1/2}$ and $\langle v_\perp{}^2 \rangle^{1/2}$ are the rms peculiar relative velocities of pairs in the radial and tangential directions respectively. A vertical arrow in each panel marks the point where $\xi = 1$.

remain, on average, at fixed physical separation. The second moments form a symmetric second-rank tensor; its two independent eigenvalues correspond to the radial and transverse velocity dispersions. In Figures 7*a* and 7*b* we plot these moments as a function of separation for models EdS1–5 (at $a = 1.8$) and for models O1–4 (at $a = 3.2$) respectively; the solid line is the Hubble line $v = Hr$, and arrows mark the correlation length r_0. As discussed above, $\xi(r)$ for both ensembles is approximately described by a power law of index $\gamma = 1.8$ at these times. Plots similar to Figure 7 have been given by Davis and Peebles (1977) for BBGKY models, and by Efstathiou and Eastwood (1981) for N-body simulations with white noise and subrandom initial conditions.

Our results are very similar to those of Efstathiou and Eastwood (1981) but differ substantially from the predictions of the BBGKY similarity solutions. The first moment, $\langle v_{21}(r) \rangle$, rises above the Hubble line for $\xi(r) > 1$ in both ensembles, and lies well above it for $\Omega = 1$, indicating that, on average, pairs on these scales are physically moving together. On the smallest scales $\langle v_{21}(r) \rangle = -Hr$, as expected for stable bound structures. The peculiar velocity dispersion tensor is slightly radially biased on all scales. (In accord with previous practice the radial dispersion plotted is not a central moment, and so should have the mean motion subtracted in quadrature when computing the anisotropy.) In our simulations, as in those of Efstathiou and Eastwood, the dispersions fall substantially with increasing separation, whereas the BBGKY solutions rise slowly. This discrepancy is probably the consequence of some of the simplifying assumptions which Davis and Peebles adopted in order to make the BBGKY equations tractable. The rms velocity difference drops slightly more rapidly with separation in the open models than in the closed models. This is yet another reflection of the fact that clusters are relatively more compact in the open models. The decrease of the dispersions at the smallest scales [$\log (r/L) < -2.3$] can be at least partly ascribed to the softening of the particles.

In order to test the effects of softening on the velocity distribution, we ran one simulation with $\Omega = 1$ and a severely softened force law. The potential was taken to vary logarithmically for separations less than $L/32$ and was further softened to approach a constant at separations smaller than $L/430$. It thus corresponded approximately to the interaction of truncated isothermal spheres. This model developed a correlation function with a pronounced bend at the outer softening length, and it had considerably reduced peculiar velocities on small scales. However, the qualitative nature of the velocity field remained unchanged; the dispersions decreased with separation for $\log (r/L) > -1.7$, and $\langle v_{21}(r) \rangle$ exceeded the Hubble flow in the region where $\xi \approx 1$. The velocity field at these separations is thus insensitive to the behavior of the interaction law on smaller scales. This model produced weakly concentrated clusters and a correlation function that is quite inconsistent with observations of galaxies.

A useful test that the velocities shown in Figure 7 are not unduly influenced by the discreteness of our particles (their individual mass exceeds that of an axion by a factor of order 10^{80}) is to compare them with the velocities found for the "big box" models EdS6–8. If this is done at times later than that plotted in Figure 7*a*, good agreement is found over the range $0.01 < r/L < 0.3$. However, $a = 1.8$ in EdS1–5 corresponds to $a = 1.27$ in EdS6–8, and at such an early time transients from the initial conditions are still visible in the velocity correlations of the "big box" models; corresponding effects may be seen in the two-point functions plotted in Figure 5. The good agreement at later times confirms that discreteness effects are indeed negligible in our models. This might also be inferred from the

fact that the dispersion on all scales in Figure 7 substantially exceed the Keplerian velocity of a binary. The dispersions must therefore result from virial motions within more massive aggregates.

To scale our models to physical units we set the Hubble velocity across the correlation length to its observed value of 500 km s^{-1}. (As discussed above, this corresponds to an unrealistically small value of h in EdS1–5.) For $\Omega = 1$, particles at small separation are then inferred to be moving with one-dimensional rms relative velocities up to 1100 km s^{-1}, while velocities up to 550 km s^{-1} are found in the open ensemble. As we discuss in § IV, values of 200–300 km s^{-1} are required to fit observations of galaxies, and the trend of the observed dispersions with separation disagrees with that seen in our models. Smaller velocities would be obtained if we considered models with lower Ω, but the scale discrepancy discussed in the previous section would then be worse. If one associates the observed galaxy distribution with the mass distribution of the simulations, then velocities in the Einstein–de Sitter models are clearly much too high to be consistent with observation. This disagreement is exacerbated if the present is identified with a larger expansion factor in EdS1–5 (the correlation amplitude of galaxies could then be matched for a larger value of h). It just reflects the well-known fact that the mass-to-light ratios obtained from dynamical analyses of groups and clusters of galaxies are much too small to close the universe. The open models give acceptable velocities at large separations ($\sim 5h^{-1}$ Mpc) but velocities which are still too large on smaller scales. The positive Λ model has very similar velocity correlations to O1 if the two models are compared when $\Omega = 0.2$, and so agrees no better with observation.

d) Three-Point Correlation Function

We have estimated the three-point correlation function $\zeta(r_1, r_2, r_3)$ for our models. This may be parameterized in terms of the usual "size" and "shape" parameters defined by

$$r = r_1 , \quad u = r_2/r_1 , \quad v = (r_3 - r_2)/r_1 , \quad (r_1 < r_2 < r_3)$$

(cf. Peebles and Groth 1975). Observational estimates from the galaxy distribution show that ζ is closely approximated by

$$\zeta(r_1, r_2, r_3) = Q[\xi(r_1)\xi(r_2) + \xi(r_2)\xi(r_3) + \xi(r_3)\xi(r_1)] , \quad (4)$$

where Q is a constant. Groth and Peebles (1977) find $Q = 1.3 \pm 0.2$ from an analysis of the angular correlation functions measured from the Zwicky and Lick catalogues. Efstathiou and Jedredjewski (1985) have estimated Q directly from the three-dimensional distribution using four independent redshift surveys; they find $Q = 0.8 \pm 0.2$, consistent with the Groth and Peebles' result. In our models ζ obeys the scaling with u and v implied by equation (4) very well. In Figure 8, we therefore plot Q, averaged over u and v, against the size parameter r for EdS1–5 and O1–4. In each case, the value of Q on small scales is significantly higher than that observed, and declines to $Q \approx 1$ where $\zeta \approx 1$. Thus the scale dependence of ζ in the N-body models disagrees slightly with observation. In the most extreme cases we find $Q > 2$ on small scales, well outside the observational limits. Note that as the models evolve, the small-scale values of Q decline. This behavior, and

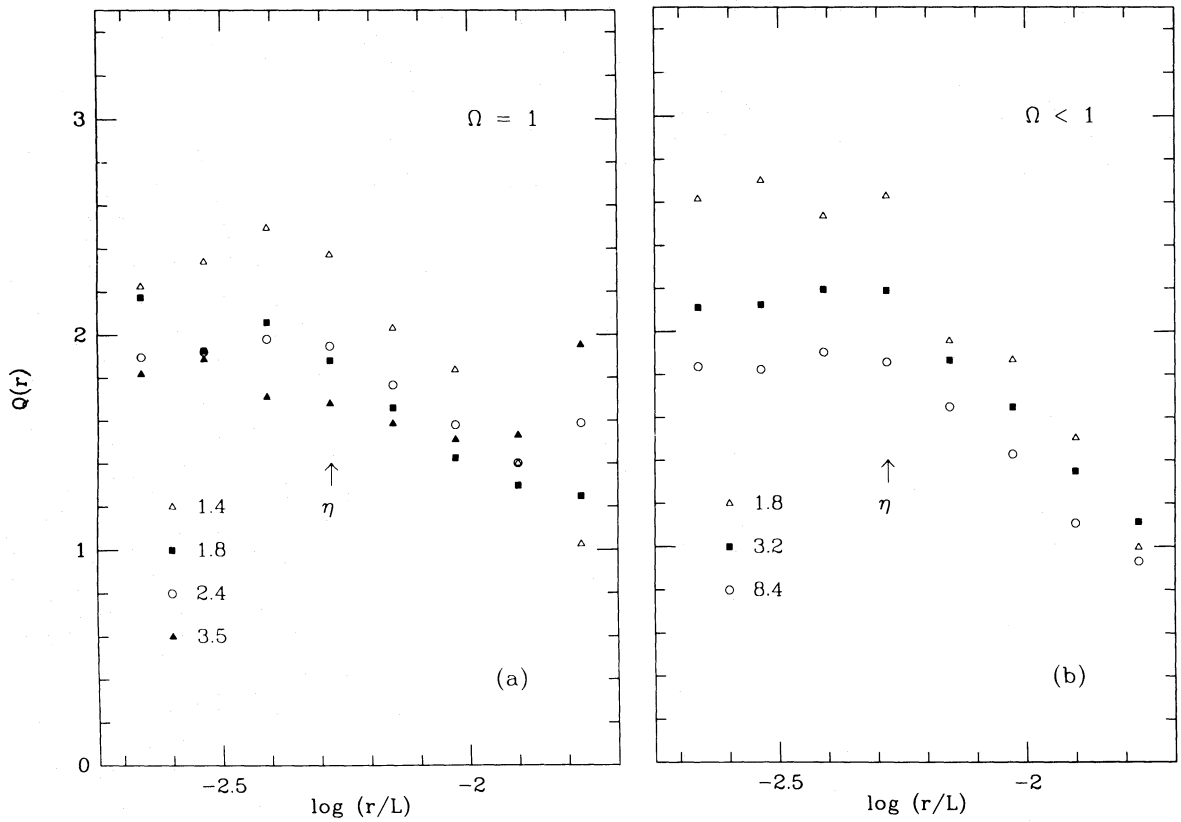


FIG. 8.—The dimensionless amplitude of the three-point correlation function Q (eq. [4]) plotted against the size parameter r. We have averaged Q over u and v because we find it to be almost independent of the triangle shape parameters. Results for the ensemble EdS1–5 are given in (*a*), and for O1–4 in (*b*). Values of Q for different expansion factors are plotted using different symbols as labeled in each panel. Arrows mark the softening length of our simulations.

the high initial values of Q, reflect the shape of the initial fluctuation spectrum. In the hierarchical model of equation (4), Q provides a measure of the dispersion in the number of neighbors found within a sphere of radius r centered on a randomly chosen particle. This dispersion is large compared to the mean number of neighbors if the initial fluctuation spectrum has considerable power on large scales (cf. the discussion of the multiplicity function below). Thus we expect Q to be larger in these models than the value $Q = 1.2 \pm 0.1$ found by Efstathiou and Eastwood for simulations with white noise initial conditions. Similarly, the gradual decrease in the value of Q with time is probably a result of the non–power-law initial conditions in our models. It is worth mentioning that Fry (1984) finds a dependence of Q on primordial spectral index n that is qualitatively similar to the results found here. Fry's calculation is based on second-order perturbation theory, and is not, however, directly applicable to the regime $\zeta > 1$.

e) Multiplicity Functions

Correlation studies of structure in our models may be complemented by an analysis of the distribution of cluster masses. This distribution, often known as the multiplicity function, is heavily influenced by correlations of very high order which cannot be estimated directly. We have found groups in our simulations by linking all particle pairs separated by less than a fraction b of the mean interparticle separation; each distinct subset of connected particles is then defined as a group. This method of cluster analysis has several advantages. In the first place it produces a unique group catalog for any chosen value of b. Second, these catalogs obey a nesting condition; all the members of any group defined for one particular b are also members of the same group in any catalog defined by a larger b (cf. Efstathiou, Fall, and Hogan 1979). Finally, the method makes no *a priori* assumption about the shapes of groups. We have found that if relatively large values of b are used ($b \approx 0.5$), the resulting groups are often quite irregular with several separate centers of concentration. Smaller values of b ($b \leq 0.25$) tend to pick out objects with a well-defined center and relatively regular structure. If the points had a fractal distribution, the shape of groups would be independent of b, so this behavior demonstrates a departure from spatial self-similarity in our models; further work will be required to determine whether this is real or is a consequence of our force softening on small scales ($b = 0.25$ corresponds to a separation of 1.7η in our models). Nevertheless it is interesting that the distribution of nearby galaxies appears to show a similar behavior (Einasto *et al.* 1984).

In Figure 9 we plot the multiplicity functions for our ensemble of Einstein–de Sitter models at two times and for $b = 0.5$ and 0.25. We have followed White and Negroponte (1982) and Einasto *et al.* (1984) in plotting the number of particles as a function of the multiplicity of the group to which they belong; we have also chosen logarithmic bins such that the first contains singles, the second pairs, the third triplets and quadruples, the fourth quintuples to octuples, etc. The most striking aspect of the distributions in Figure 9 is that as the simulations evolve they become flat over the entire multiplicity range 2–1000; a particle is thus equally likely to find itself in a group in each logarithmic mass interval over this range. Multiplicity functions for our ensemble of open models resemble those shown in Figure 9; these seem to be characteristic of CDM models. Evolution from initial conditions with a large coherence length leads to U-shaped distributions in which particles are likely either to be singles or members of very large clusters. This may be seen clearly in the diagrams shown by Einasto *et al.* (1984) for models of a neutrino-dominated universe, and we have found similar results for the neutrino models of White, Frenk, and Davis (1983). Scale-free initial conditions with less power on large scales (e.g., white-noise models) produce configurations in which most particles tend to be in systems of intermediate multiplicity; they look rather like the non-dynamical hierarchical models illustrated by Einasto *et al.* Thus, as emphasized by Bhavsar, Gott, and Aarseth (1981), multiplicity functions seem to discriminate clearly between different kinds of linear initial condition. The flat multiplicity functions of Figure 9 are very similar to the histograms

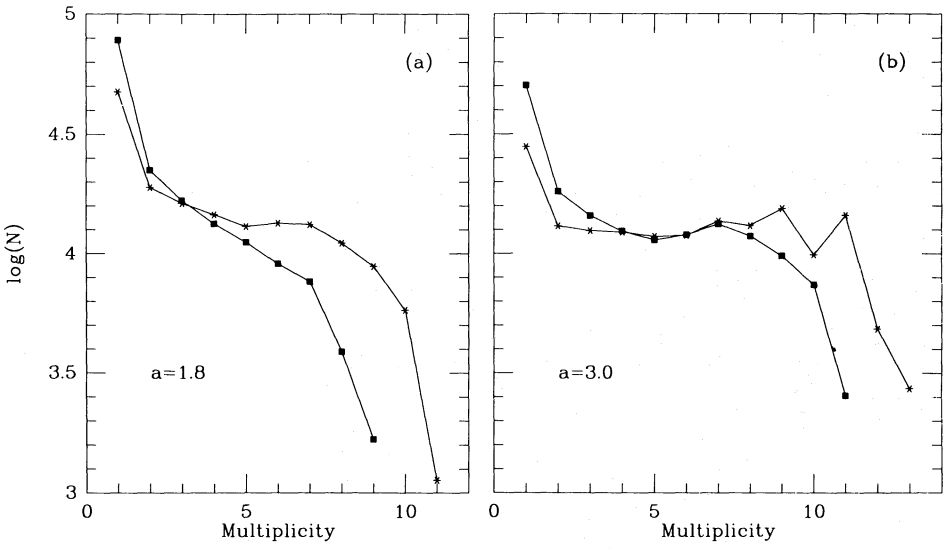


FIG. 9.—The number of particles in ensemble EdS1–5 in groups of a given multiplicity is plotted as a function of that multiplicity. The multiplicity parameter is defined to be the integral part of $\log_2 (N_g) + 1$, where N_g is the number of particles in a group. It is thus equal to 1 for singles, to 2 for pairs, to 3 for triplets and quadruples, and so on. Results are shown for $a = 1.8$ in (*a*) and for $a = 3.0$ in (*b*). In both plots filled squares give results for $b = 0.25$ and stars give results for $b = 0.5$, where b is the maximum separation at which particles are linked into a group; b is expressed in units of the mean interparticle separation.

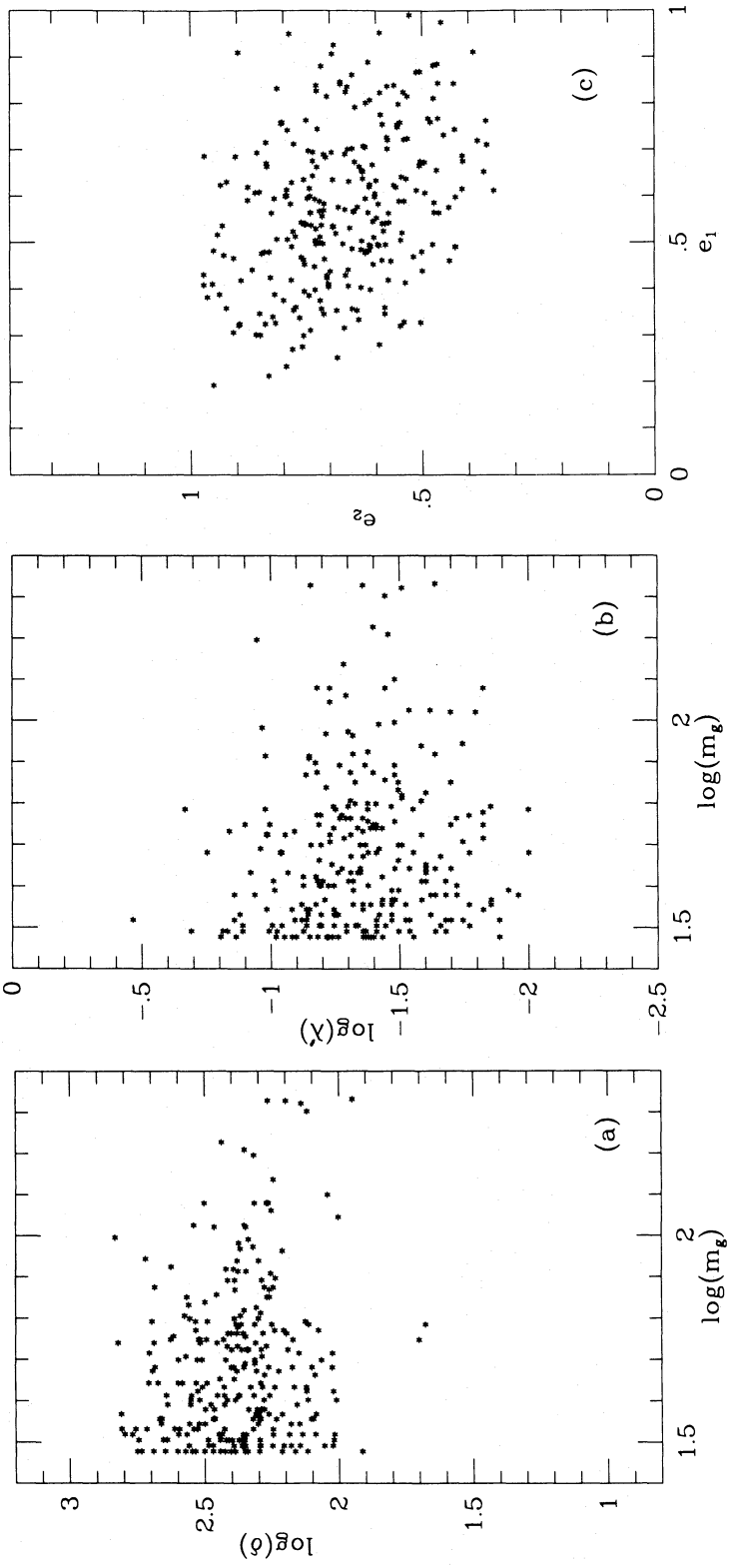


FIG. 10.—Properties of groups with more than 30 members picked out using $b = 0.25$ in the ensemble EdS1–5. In (*a*) we show the density of groups as a function of their mass: δ is the density ρ expressed in terms of the mean density within the simulation volume, and m_g is given in units of the mass of a single particle; (*b*) gives the spin parameter λ' as a function of group mass; and (*c*) shows the two axial ratios e_1 and e_2 plotted against each other. These quantities are all defined in eq. (5). Note that the results displayed here should be treated with caution because they are affected by the artificial force softening in our calculations.

obtained for the nearby galaxy distribution by Einasto *et al.* for corresponding values of b. We hope to come back to this question and make a detailed comparison with a larger sample of observational data in a later paper.

The groups in our simulations correspond to the bound clumps of cold dark matter which must make up the halos of individual galaxies and of groups and clusters of galaxies in the real universe. For small values of b our group-finding algorithm picks out single, well-defined clumps that can be identified with such structures. Although we do not have space here to present a thorough discussion of clump properties, we show some results in Figure 10 to give an idea of what is possible. For each group found by our algorithm we calculated a mass m_g, an inertia tensor with principal components $m_g a_1^2 > m_g a_2^2 > m_g a_3^2$, an internal velocity dispersion $\langle v^2 \rangle$, and an angular momentum vector $\boldsymbol{J}$. From these quantities we derive a mean density, two shape parameters, and a spin parameter using the formulae:

$$\rho = 3m_g/(4\pi 5^{3/2} a_1 a_2 a_3) ,$$
$$e_1 = a_2/a_1 ,$$
$$e_2 = a_3/a_2 ,$$
$$\lambda' = J(\langle v^2 \rangle/2)^{1/2}/Gm_g^2 . \quad (5)$$

In these equations ρ is the density of a uniform ellipsoid with the same mass and inertia tensor as the group, and λ' differs from the usual spin parameter in that it is defined using the kinetic energy of the group rather than its total energy. For an oblate system $e_1 = 1$, while for a prolate system $e_2 = 1$. Figure 10 shows the distribution of some of these quantities for the 273 groups with 30 or more members in models EdS1–5 at $a = 1.8$. We give ρ in units of the mean density of the calculation and m_g in units of the particle mass. For the value $b = 0.25$ used here, the algorithm clearly picks out groups at a fairly well-defined density contrast of about 200. There is no correlation of density with mass except for a slight tendency for the very largest groups to have low density. The spin parameter also shows at most a very slight anticorrelation with mass, and has a median value of about 0.045. The lack of trend agrees with the results of Efstathiou and Barnes (1983) for simulations from white noise initial conditions, but our characteristic value is 30% lower. This difference would be interesting if correct, but it is almost certainly due to the effects of softening in the present calculations; the mean separation of particles in our groups is approximately equal to η. The problem merits further careful investigation using higher resolution simulations. The groups tend to be quite strongly aspherical; axial ratios greater than 2 are quite common. Their generic shape is neither oblate nor prolate, although a clear majority of groups have $e_2 > e_1$, showing some preference for a prolate figure. Again this merits further study to investigate the influence of tides, of internal substructure, and of force-softening on the measured quantities. We have found very similar distributions for $a = 3$ in models EdS1–5, the main difference being that the mass distribution then extends to higher masses (see Fig. 9). The properties of groups in our open models are very similar to those shown in Figure 10, but since these groups are rather more compact, they are even more strongly affected by force softening.

A final problem which we have addressed with our group-finding algorithm is the derivation of the value of b for which linked structures extend across our computational volume and percolation occurs. Einasto *et al.* (1984) suggest this to be a useful measure of structure both in N-body models and in data catalogs; Dekel and West (1985), on the other hand, point out a number of difficulties in interpreting percolation properties. Models EdS1–5 percolate for values of b in the neighborhood of 0.7, both at $a = 1.8$ and at $a = 3$. Percolation is slightly harder at the later time and is also slightly harder in the open ensemble. This value of b is smaller than the value $b = 0.86$ required for percolation in a three-dimensional Poisson distribution. Thus by the criteria of Einasto *et al.* the present models appear filamentary rather than hierarchical, even though Figure 1 shows that clusters do grow in a hierarchical fashion. For the nearby galaxy distribution Einasto *et al.* found percolation to occur for $b = 0.65$, quite close to the value we obtain for our models.

f) Counts in Cells

Further information about galaxy clustering is contained in the distribution of counts in cells, and, in particular, in its departure from a Gaussian curve. For the N-body simulations this statistic is very easy to generate, but for the data it is essential to use volume-limited catalogs; this reduces the size of available surveys and makes comparison quite difficult. Figure 11 shows the distribution of counts in cubes of side $L/8$ and $L/4$ for EdS1–5 at $a = 1.8$. At this epoch $r_0/L = 1/26$, so these cubes should be compared with regions of the real universe which are $16h^{-1}$ and $32h^{-1}$ Mpc on a side. For the smaller cube the distribution is extremely non-Gaussian; there are an extended tail at high density and an excess of low density regions. For the larger cube the distribution is much closer to a Gaussian, but even here the departures are quite substantial. The expected rms fluctuation in the number in a cube is $(\delta\rho/\rho)_{\rm rms} = \bar{\xi}^{1/2}$, where $\bar{\xi}$ is the average of the correlation function over the cube. From our cell counts we find rms fluctuations of 70% and 35% for the small and large cubes respectively.

The dashed histogram in Figure 11 corresponds to the distribution of counts in cells of *redshift* space for cubes of side $L/4$. This curve was generated by displacing each particle in the z-direction by its z-component of peculiar velocity divided by the model Hubble constant and then using the same algorithm as before; this procedure approximates the construction of a volume-limited redshift catalog, the z-axis being the redshift direction. Observational catalogs of large-scale clusters and holes are, of course, measured in redshift space, not real space, and should be compared to this histogram. The difference between the redshift-space and the real-space histograms is seen mainly in the low-density cells. Because of the coherent outflowing of material from low-density regions, cells tend to appear emptier in redshift space than in real space. This effect can be substantial, particularly in a high-density universe.

From the cumulative distribution corresponding to Figure 11, we find that 4% of cubes of side $L/4$ have $n/\bar{n} < 0.5$ in real space, versus 8% in redshift space. However, even in redshift space, the fraction of very empty cubes is very small; only 0.6% of cubes of side $L/4$ have $n/\bar{n} < 0.3$. A very similar result is found in models EdS6–8, if we analyze them at $a = 1.3$ when the correlations are similar to those of EdS1–5 at $a = 1.8$. In these models, 10% of $32h^{-1}$ Mpc cubes in redshift space have density less than half the mean. In the open models, the distribution of counts in large cells is virtually identical to that for the $\Omega = 1$ models, provided times with the same linear growth factor are compared. At late times, linear fluctuations cease to

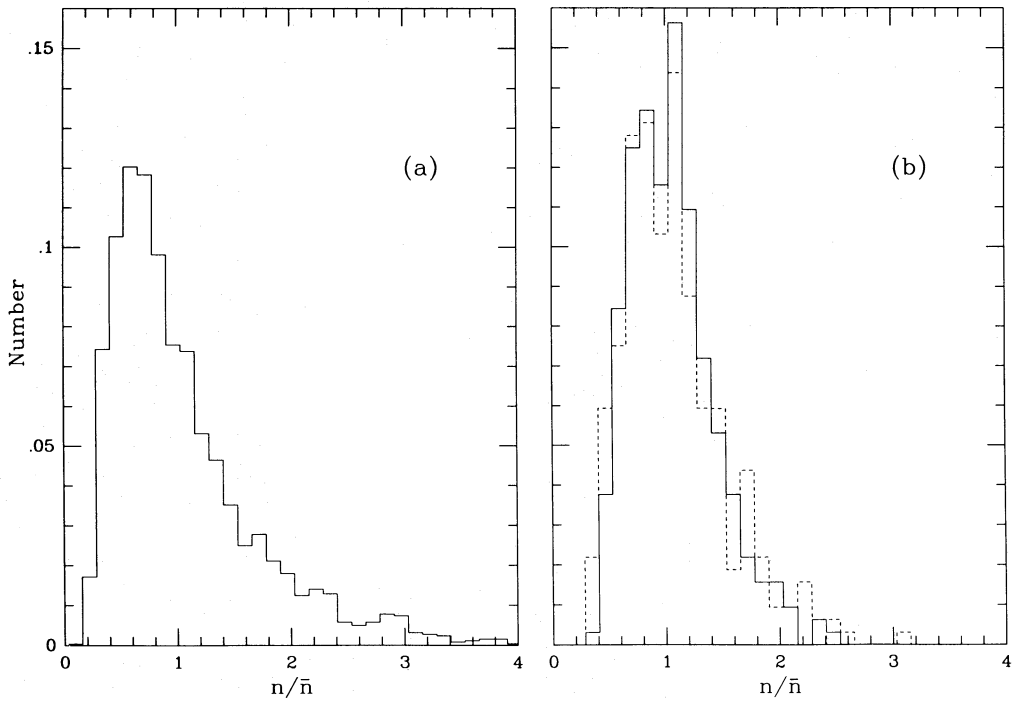


FIG. 11.—Histograms of counts in cells for the ensemble EdS1–5. The distribution of counts in cubes of side one-eighth that of the simulation volume is shown in (*a*), while (*b*) shows the corresponding distribution for cubes of side one-quarter that of the box. The dashed histogram in (*b*) shows the result of making the counts in "redshift" space rather than in real space.

grow in an open universe, and large holes would no longer expand in comoving coordinates.

It is difficult to find observational data to compare with these results. In a subset of the CfA redshift survey (Davis *et al.* 1982; Huchra *et al.* 1983) volume-limited to 6000 km s^{-1} and magnitude-limited to 7500 km s^{-1}, an analysis of counts in randomly placed spheres suggests that 20% of spheres of radius $20h^{-1}$ Mpc (i.e., with the same volume as a $32h^{-1}$ Mpc cube) have density less than half the mean. There is thus a discrepancy with the CDM prediction by about a factor of 2. Note, however, that the total volume of the survey is less than that of eight such spheres, so that the statistics are very poor. Nevertheless the number and sizes of voids within the CfA data set appears to be quite representative (cf. Oort 1983). If really large voids in the galaxy distribution such as the $30h^{-1}$ Mpc radius void in Boötes reported by Kirshner *et al.* (1981, 1983) turn out to be common, they may be difficult to explain within a CDM model. The Boötes region has not been fully surveyed, but Kirshner *et al.* claim the density of bright galaxies there to be less than 25% of the mean. This is particularly startling because the volume of the void is about 10% of all space that has been fully surveyed to date! Within our simulations we see no low-density regions of this size despite the fact that the total simulation "volume" (in units of r_0^3) is greater than the volume of "known" space.

IV. COMPARISON WITH A REDSHIFT CATALOG

In the previous section, we showed that CDM models with $\Omega = 1$ give excessively high velocities and two- and three-point correlation functions incompatible with observation. The amplitudes of the relative peculiar velocities, and of ξ, are better reproduced in models with $\Omega \approx 0.2$, but the discrepancy in the value of Q is worse. Since these open models (at $a \approx 3.2$) are the best approximation to the observed galaxy distribution we have found so far, we now compare some of them with the CfA redshift survey by generating "redshift" catalogs following the procedure of Davis *et al.* (1982). We situate the observer within the computational box and choose points at random in such a way as to mimic the selection function appropriate to the CfA northern catalog, when it is volume-limited to a distance of 4000 km s^{-1}, and magnitude-limited at $m_z = 14.5$ from 4000 km s^{-1} to 10,000 km s^{-1}. Further details are given by Davis and Huchra (1982). We recall that $\gamma = 1.8$ at $a = 3.2$ in O1–4, and that setting the model correlation length to its observed value then requires $h = 1.07$ and a physical box size of 142 Mpc.

Redshift catalogs constructed in this way from runs O2 and O3 (at $a = 3.2$) and projected onto the sky are shown in Figures 12*a* and 12*b*. The semi–volume-limited CfA northern catalog, which these samples model, is shown in Figure 12*c*. To illustrate the kind of fluctuation which may be expected, we have situated the origin near a prominent cluster in model O3, in much the same way as we happen to be near the Virgo supercluster. The selection criteria were unchanged between models O2 and O3, so this procedure results in the number of selected "galaxies" being considerably higher in O3. Wedge diagrams of these three catalogs are shown in Figures 13*a–c*. The simulation data have been plotted in redshift space, which includes the peculiar velocity distortion in the radial direction; the rms value of this velocity is 256 km s^{-1} for the catalog from O2 and 357 km s^{-1} for that from O3. A comparison of these three catalogs is quite instructive. "Fingers of God" can be seen in all three samples, as can long filamentary structures. The cluster we chose in O3 is clearly richer than Virgo, and there is as much difference between the two simulated catalogs as between them and the real data. The large-scale power in the perturbation spectrum produces significant fluctuations in volumes the size of the CfA catalog, so that it is difficult to

1985ApJ...292..371D

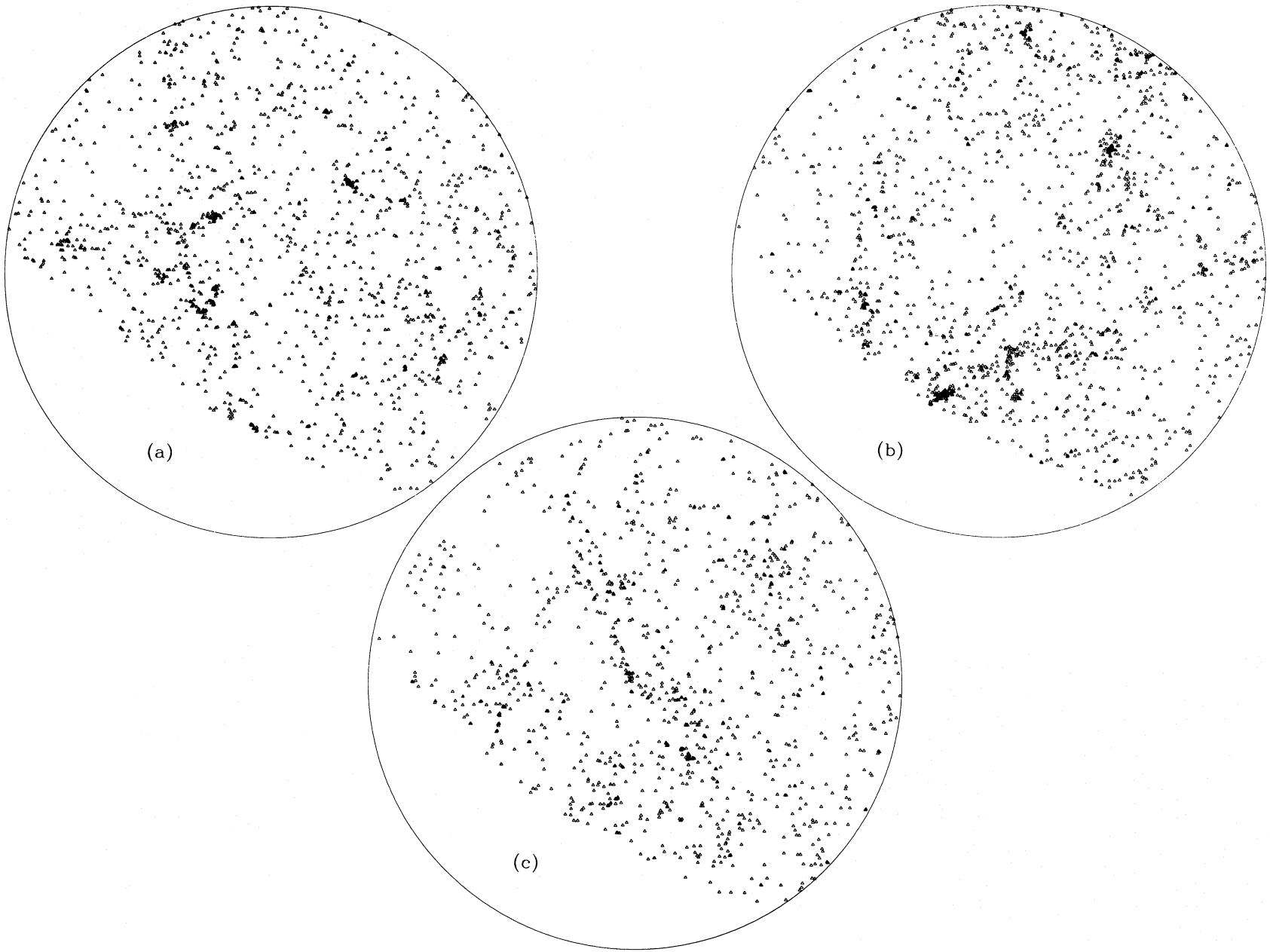


FIG. 12.—Redshift catalogs constructed from two open models (O2 and O3) are shown in (*a*) and (*b*) as projections onto the "sky." Particles were selected for inclusion in these catalogs in such a way as to mimic the northern CfA survey. The real data are shown in the same format in (*c*). These are equal area plots of the sky; the outer circle corresponds to Galactic latitude +40°, while the empty regions correspond to declinations below 0°. In constructing the catalog from O3 shown in (*b*), the "observer" was purposely sited near a prominent cluster.

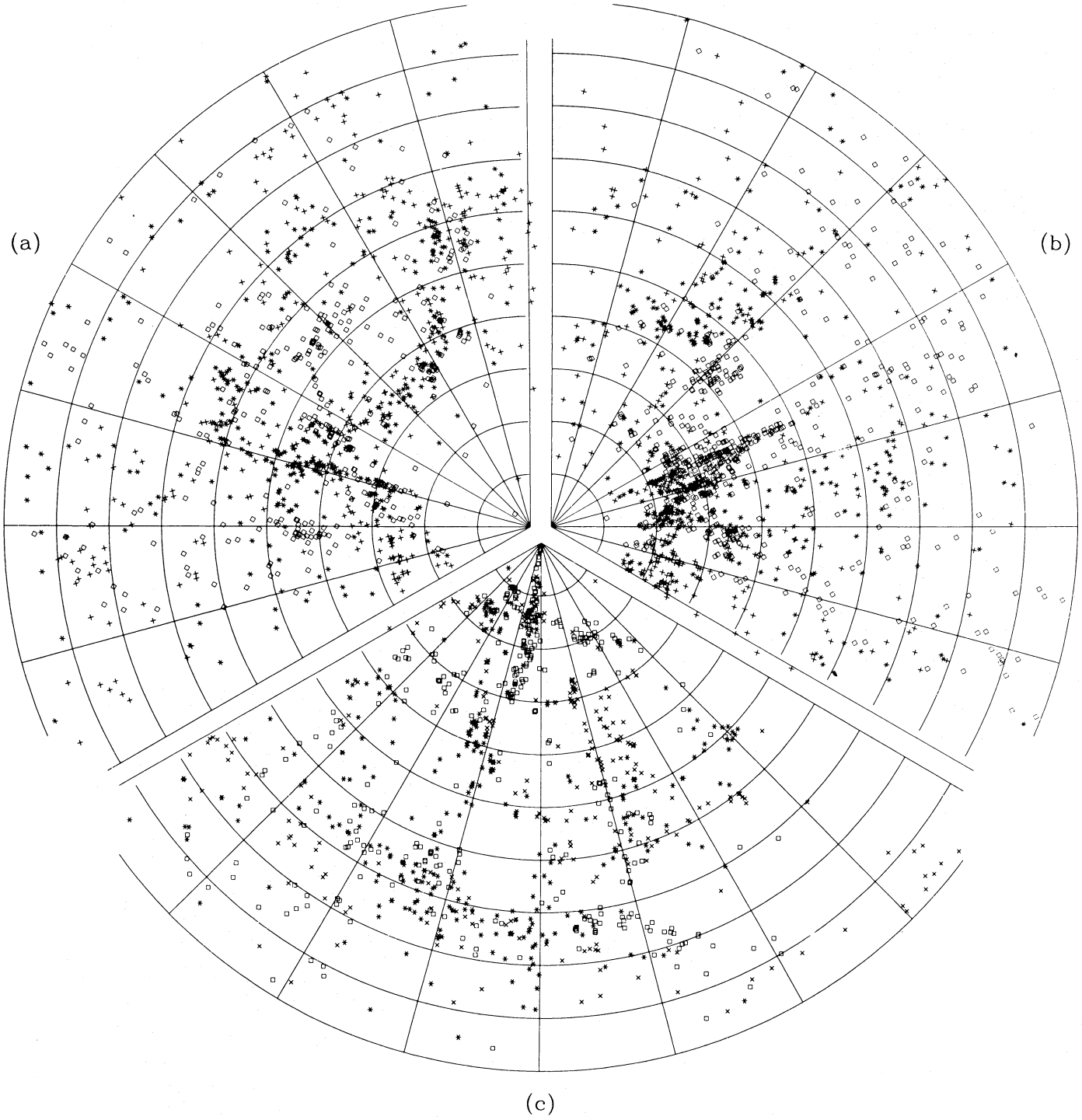


FIG. 13.—Wedge diagrams for the three catalogs illustrated in Fig. 12. The radial coordinate in these plots is recession velocity with successive circles corresponding to increments of 1000 km s^{-1}. The angular coordinate is right ascension with a line plotted for each hour. All "galaxies" in the declination range $0° < \delta < 45°$ are shown, with squares corresponding to $0° < \delta < 15°$, stars to $15° < \delta < 30°$, and crosses to $30° < \delta < 45°$. The three slices of this pie are labeled to correspond to Fig. 12. Note the large cluster in O3 and the Virgo Cluster in the CfA data.

judge the acceptability of a model on the basis of one representation. Such fluctuations also affect the observations, as may be seen by comparing galaxy distributions in the northern and southern hemispheres out to a distance of $50h^{-1}$ Mpc. Figures 12 and 13 give the overall impression that our open models match the observed galaxy distribution very well.

We produced the redshift catalogs for O2 and O3 by scaling the simulations so that the two-point correlation function of the open ensemble matches that of the observations. However, since the redshift catalogs contain only 3%–4% of the original particles, we can expect substantial fluctuations in estimates of $\xi(r)$ derived from them. In Figure 14 we plot $\xi(r)$ computed from the three redshift catalogs in a fashion that attempts to give equal weight to each volume element; this was the method used by Davis and Peebles (1983) for the CfA data. We give ξ as a function of the apparent spatial separation in redshift space; this is a biased estimator of the true $\xi(r)$. The catalog from O3 has correlations quite similar to the CfA data, but that from O2 has a smaller correlation length. Since the models have larger peculiar velocities on small scale than the real data (Fig. 7*b*), their correlation function in redshift space is shallower than that observed. This effect is evident in the O3 catalog for $r < 5h^{-1}$ Mpc. Note also that the CfA data become negative for $r > 28h^{-1}$ Mpc, that the same is true for the O3 catalog at $r > 20h^{-1}$, but that the O2 catalog is positive on almost all scales. The true $\xi(r)$ is actually positive over all these scales in models O2 and O3. The negative correlation seen in the O3 redshift catalog, and perhaps also that in the CfA data, results from the substantial inhomogeneity of these data sets and from the method of estimation. In the analysis of redshift catalogs, one solves for the mean density $\bar{n}$, as well as for $\xi(r)$. Despite the use of a density estimator that is minimally biased by inhomogeneity effects (see Davis and Huchra 1982 for details), the derived $\bar{n}$ for the O3 catalog is slightly higher than the true mean; this reduces the correlations on large scale. In addition, the foreground cluster in this catalog "beats" against a large low-density region just behind it (cf. Fig. 13*b*); this beating produces most of the anticorrelation seen on intermediate scales. One should clearly be cautious when interpreting reports of anticlustering in redshift surveys (Davis and Peebles 1983; Shanks *et al.* 1983).

The distribution of peculiar velocities provides the most stringent observational test of our models. In redshift surveys one can study $\xi(r_p, \Delta v)$, the two-point correlation of pairs as a function of their projected separation r_p and their line-of-sight velocity difference Δv. The rms distortion of ξ in the v direction as a function of r_p is termed $\sigma(r_p)$; its square is a linear combination of the radial and transverse components of $\langle v_{21}{}^2 \rangle$, suitably averaged over projection factors. There is excellent agreement between the estimates of $\sigma(r_p)$ from the CfA and AAT redshift surveys (Davis and Peebles 1983; Bean *et al.* 1983). Both groups report the velocity field to be smooth; the CfA data are best fit by a slowly rising $\sigma(r_p)$, while the AAT data are consistent with this, but are better fit by a constant $\sigma(r_p)$. We have applied the procedure of Davis and Peebles (1983) to our artificial "redshift" catalogs, and the results are plotted in Figure 15. The error bars are formal errors from fits to an exponential broadening function, which represents both the CfA data and the models quite well. The f parameter used by Davis and Peebles is here set to unity, which agrees with $\langle v_{21}(r) \rangle$ as plotted in Figure 7*b*. It was to be expected from the behavior of the full peculiar velocity-dispersion tensor that $\sigma(r_p)$ would fall with radius; Figure 15 demonstrates that it is indeed possible to see this trend in a redshift catalog, and that it is inconsistent with the CfA and AAT data. We showed in § IIIc that the falling dispersions of our models are not a consequence of their experimental limitations; they must therefore be considered a serious disagreement with observation. The substantial difference between $\sigma(r_p)$ in the two simulations is caused by the large cluster in the foreground of the O3 catalog.

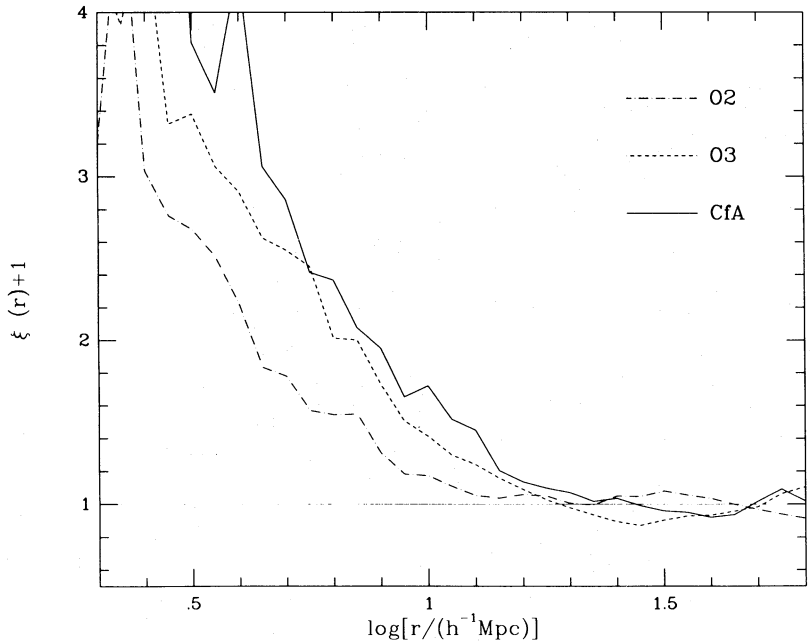


FIG. 14.—Correlation functions calculated in redshift space for the three catalogs of Fig. 12. Note that r is not a true physical distance in this plot since it includes distortions caused by peculiar velocities. The scaling of the simulations assumes $h = 1.1$, the value required to get $\xi(r)$ for the ensemble of open models to match observation (see Fig. 6).

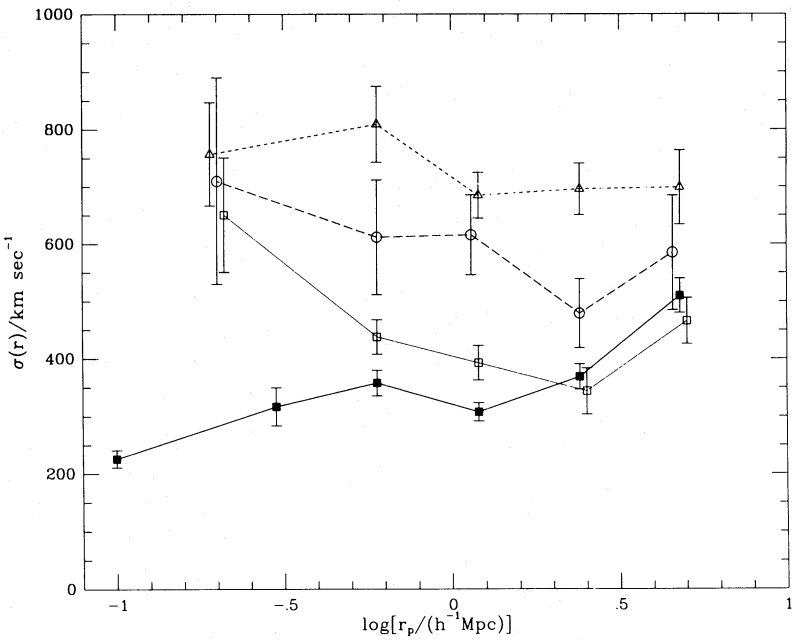


FIG. 15.—The rms line-of-sight peculiar relative velocity as a function of projected separation for catalogs made from O2 (*open squares*), from O3 (*open triangles*), from the CfA redshift survey (*filled squares*; from Davis and Peebles 1983), and from the "galaxies" found in EdS1–5 by the algorithm described in § V. The two open models have been scaled using $h = 1.1$, while the biased formation ensemble has been scaled using $h = 0.44$.

The velocity moments are computed on a pair-weighted basis, and the presence of this cluster skews the results. This effect is undoubtedly larger in the O3 catalog than in the real data, because of the richness of the cluster we chose to sit next to.

We have found that no CDM model is in complete agreement with observation if galaxies are distributed in the same way as the mass; the open models are a better fit than models with $\Omega = 1$, but are still not fully acceptable. To try and circumvent these problems, we now explore the effect of relaxing the assumption that galaxies trace the mass. This is clearly necessary if we wish to find a viable model with $\Omega = 1$.

V. BIASED GALAXY FORMATION

In a universe dominated by CDM, galaxy formation presumably occurs at peaks of the matter-density field. Now suppose that galaxy formation is a process that can go to completion only under favorable circumstances and can be suppressed if the environment changes. This might occur, for example, if the first generation of stars photoionizes, or shock heats to high temperature, those regions that were somewhat slower to collapse, cool, and fragment. It is then reasonable to expect that the highest peaks (which have the shortest cooling and collapse times) will be the most likely to complete the process of galaxy formation successfully. A similar biasing of the bright-galaxy distribution might occur if the stellar initial mass function is sensitive to local environment, or perhaps simply if galaxy formation is sufficiently sensitive to protogalactic cooling time.

To simulate this kind of situation, we applied a low-pass filter to the initial random-density field of our models and tabulated the result on a 64^3 grid. We then identified all peaks of amplitude at least ν times the rms density fluctuation and tagged the particle whose Lagrangian position lay nearest each one as a "galaxy." Only one "galaxy" was allowed for each connected region above the threshold. The tagged particles were followed over the course of the simulation. This identification procedure involves two parameters, the dispersion r_s of the Gaussian filter $\exp(-r^2/2r_s^2)$ used to smooth the density field, and the threshold ν. The scale parameter defines the size of the regions from which a galaxy forms, while the threshold is some measure of the "difficulty" of galaxy formation. Because there is so much power on large scales in the CDM models, it is significantly more probable for a local peak to rise above a fixed threshold if it lies in an incipient protocluster than if it lies in an incipient void. As a result our "galaxies" are a biased subset and are more strongly clustered than the simulation as a whole. We find the strength of this bias to increase strongly with ν, but only rather weakly with r_s; the number of particles tagged as galaxies depends strongly on both parameters. A similar effect has recently been explored by Kaiser (1984*a*) as a possible explanation for the large correlation length of the distribution of rich galaxy clusters. Analytic discussions of biasing in the present context are given by Bardeen (1984) and Kaiser (1984*b*). With $r_s = L/192$ and $\nu = 2.5$, we obtain about 1000 tagged points, or about 3% of the original. After adopting a suitable scaling, this turns out to be comparable to, or somewhat less than, the expected number of bright galaxies in a volume the size of our box (i.e., to a density $\sim 0.003h^3$ Mpc^{-3}).

The result of applying this procedure to run EdS1 is illustrated in Figure 16, which compares the distribution of "galaxies" with that of the dark matter at $a = 1.4$. The "galaxies" follow the clustering structure of the underlying matter field but with greatly enhanced amplitude. The effect of this bias on $\xi(r)$ is shown in Figure 17*a*, where we plot the evolution of the "galaxy" correlation function in models EdS1–5 and compare it with the unbiased results obtained when all particles are used; power-law fits to these correlation functions are given in Table 2C. The "galaxy" correlations substantially exceed the mass correlations on all scales, but the effect is most marked at small separations and produces an increase in the effective power-law index γ. Much of the clustering of the "galaxies" is due to the pattern imposed by the

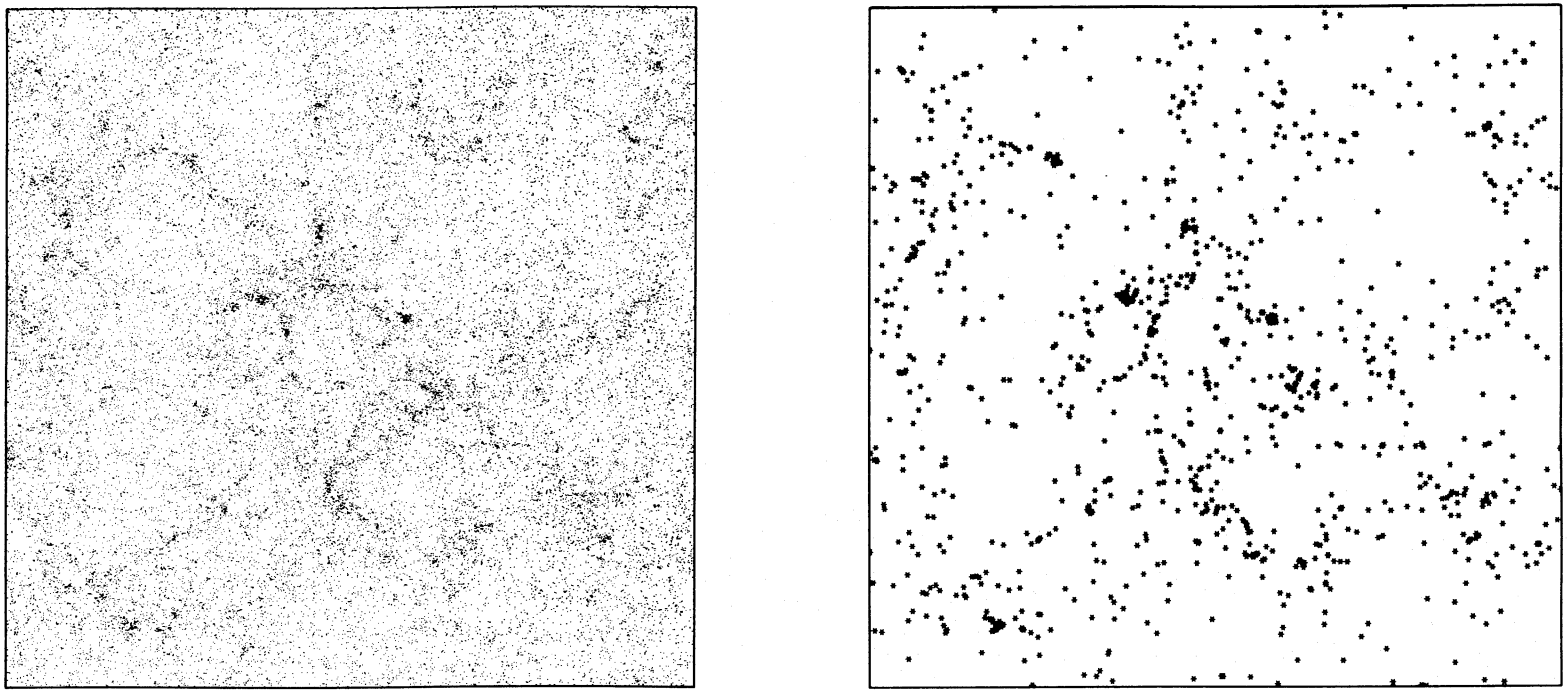


FIG. 16.—The projected distribution of all particles (*left*) and of the "galaxies" (*right*) in EdS1 at $a = 1.4$. The side of the box is $32.5h^{-1}$ Mpc. "Galaxies" are assumed to form only at the 2.5 σ peaks of the linear density distribution.

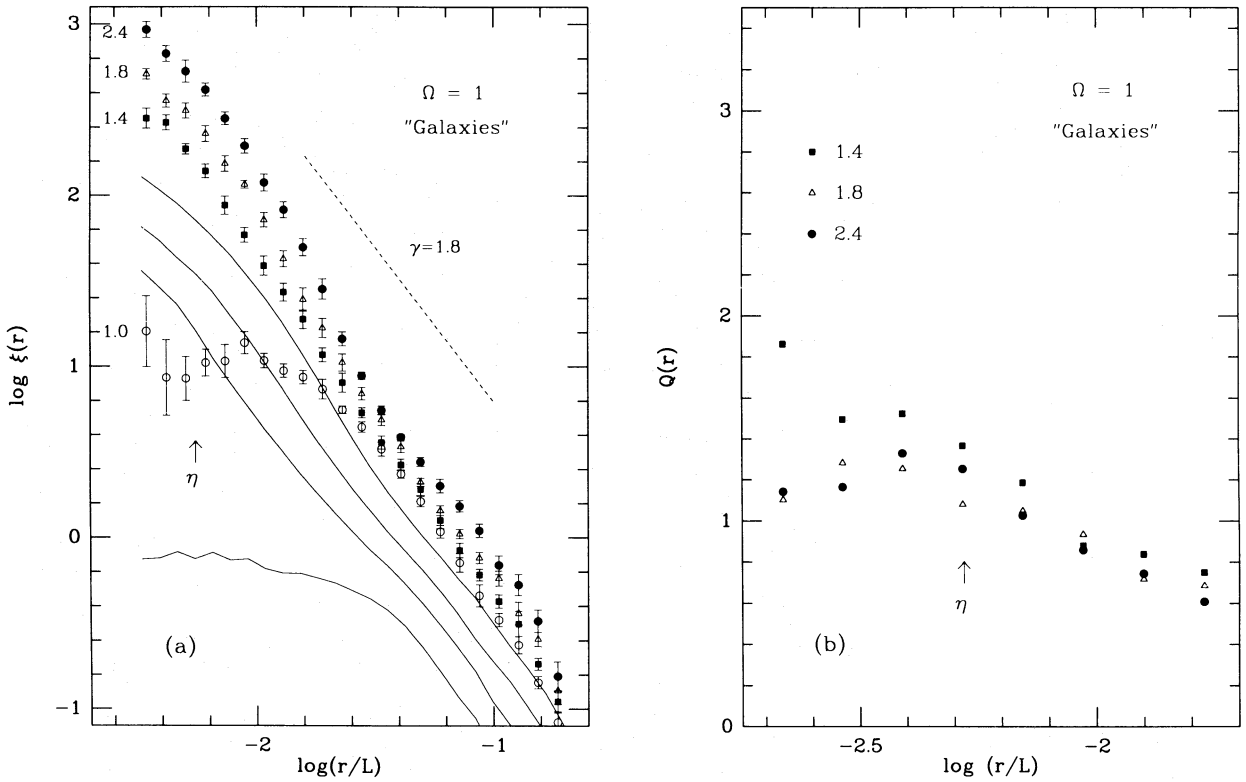


FIG. 17.—Correlation properties of "galaxies" in EdS1–5 on the biased galaxy formation model. (*a*) compares the evolution of $\xi(r)$ for the "galaxies" (*symbols with error bars*) with that of the mass distribution as a whole (*solid lines*). (*b*) shows Q values for the "galaxies" in the same format as Fig. 8.

statistical properties of the linear density field rather than to the effects of gravity. This is particularly evident on large scales, where correlations grow very little until the underlying matter distribution has become significantly clustered. As a result, $\xi(r)$ for the "galaxies" steepens even more rapidly than for the simulation as a whole. By $a = 1.4$, our models have an effective γ which exceeds 1.8, and by $a = 1.8$, the galaxy correlation function is clearly too steep to match observation. The correlation length r_0 for the "galaxies" exceeds that of the underlying mass distribution by a factor of 2.4 at $a = 1.4$. The two- and three-point correlations of the "galaxies" obey the empirical formula of equation (4) at least as well as those of the mass, and, as shown in Figure 17*b*, they imply Q values which agree better with observation. The relative peculiar velocities of the "galaxies" behave in a similar way to those of the simulations as a whole; the bias is evident only as a slight increase in the dispersions. This may be seen in Figure 18, as may the fact that the dispersion profiles for the "galaxies" are somewhat flatter than those of Figure 7*a*.

To scale these "galaxy" catalogs to the real world, we again set the correlation length r_0 to $5h^{-1}$ Mpc. Unfortunately, while the correlation exponent γ is about 1.8 at $a = 1.2$, the peculiar velocity distribution shows the effects of initial transients until about $a = 1.8$. As a compromise we identify the present with $a = 1.4$ when ξ is only slightly steeper than the observations and velocity transients have already subsided to a relatively low level. At this time $r_0 = 0.07L$, leading to a value of 0.44 for h and to a box size $L = 168$ Mpc. With this scaling the velocity unit in Figure 18 is 7400 km s^{-1}, and the one-dimensional dispersion has a maximum of 575 km s^{-1}, from which it drops to values of order 400 km s^{-1}. When redshift catalogs are made from the "galaxy" distributions, these dispersions are found to lead to the values of $\sigma(r_p)$ shown in Figure 15, which are in as good agreement with observation as those for unbiased models with $\Omega = 0.2$. The redshift catalogs for our "galaxies" are too sparse for visual comparison with the CfA survey (as in Figs. 12 and 13), but we can get some information about the probability of large low-density regions from the distribution of counts in cells. In EdS1–5, we find only one cube out of 40 with side $7r_0$ ($35h^{-1}$ Mpc) to have a density less than half the mean. Holes appear to be relatively less abundant in these "galaxy" samples than in the overall mass distribution (cf. § III*f*). Thus voids as large as the one in Boötes may be hard to accommodate in our model. Simulations of larger scale are required to obtain a definitive answer on this point.

Although our prescription for galaxy formation is very crude, it is clear that the biases which arise if there is an effective threshold for galaxy formation can be very strong. Furthermore, their effects are of the kind required to bring an Einstein–de Sitter CDM universe into agreement with observation. The particular models which we discuss above have two-point correlations and peculiar velocity distributions which are quite similar to those that we obtained earlier for our open models. They are more successful than the open models in reproducing the observed three-point correlations, and they imply an age for the universe which is not uncomfortably short. Furthermore, they satisfy the theoretical desire for a flat universe without contradicting dynamical studies of galaxy clustering. Finally, they lead to predicted fluctuations in the microwave background which are far below the current observational upper limits (Bond and Efstathiou 1984; Vittorio and Silk 1984). On the negative side, the biased models seem less able to produce large-scale voids than the open models, and they predict that galaxies formed at quite recent epochs—this

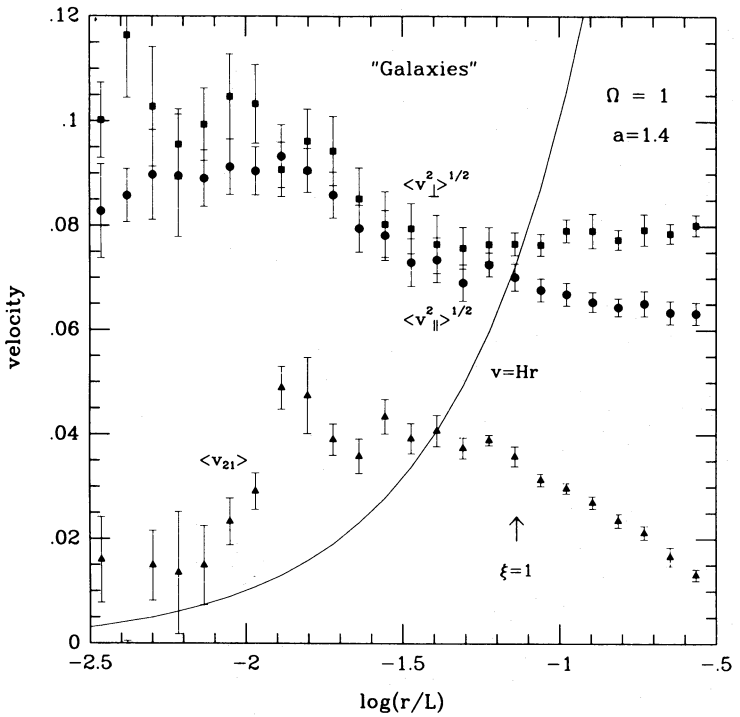


FIG. 18.—Moments of the relative peculiar velocity distribution of pairs of "galaxies" in EdS1–5 at $a = 1.4$ in the same format as Fig. 7.

may conflict with observation. Clearly, what is now required is a proper physical model for galaxy formation which can be grafted onto simulations to see if the distribution of our "galaxies" is indeed realistic. These biased galaxy-formation models are in many ways the closest we have come to matching the observed galaxy distribution, and they involve the minimum gravitational interaction!

VI. CONCLUSIONS

The numerical simulations discussed in this paper were designed to study nonlinear gravitational clustering in a universe dominated by cold dark matter. We assumed adiabatic primordial density fluctuations with a constant-curvature spectrum. We followed the evolution of two ensembles of simulations with $\Omega = 1$, another ensemble with $\Omega < 1$, and a single simulation of a flat universe with a nonzero cosmological constant. Our main conclusions are as follows:

1. Initial conditions of this kind lead to a very rapid growth in the characteristic scale of clustering. As a result, nonlinear structures on a wide range of scales are present at any given moment. This is reflected in the filamentary structures, the superclusters of clumps, and the large low-density regions which are visible in pictures of the particle distribution. This large-scale morphology is clearly a very general consequence of nonlinear evolution from such initial conditions. While its scale and its evolution rate depend on Ω, Λ, and h, its form is almost independent of them.

2. The particle two-point correlation function $\xi(r)$ steepens with time in all the models, showing that their evolution is not self-similar. The good agreement we find between ensembles of different scale argues strongly that this result is not a consequence of the limited spatial and temporal resolution of our experiments, but rather reflects the lack of spatial self-similarity in the theoretical initial conditions. In each ensemble of simulations, a time interval can be chosen so that ξ provides an acceptable match to the shape of the observed galaxy two-point function over a limited range of scales. The amplitude of the observed function can then be matched by choosing h appropriately. For Einstein–de Sitter models we require $h \approx 0.22$, which is unacceptably low. On the other hand, for models in which ξ has the right shape when $\Omega = 0.2$, we need $h \approx 1.1$; this seems rather high. Our model with zero spatial curvature and positive cosmological constant has very similar properties to open models with the same value of Ω; however, it implies a considerably greater age for the universe.

3. The peculiar velocity distribution in our simulations does not give a good match to observations of galaxies. For $\Omega = 1$ we find rms one-dimensional relative peculiar velocities of $\sigma \approx 1000$ km s^{-1} at separations $r \approx 1h^{-1}$ Mpc, much in excess of the observational estimates, $\sigma \approx 200$–300 km s^{-1}; this is a reflection of the well-known fact that the mass-to-light ratios of small groups and clusters appear much smaller than that needed to close the universe. For $\Omega = 0.2$ we find $\sigma \approx 500$ km s^{-1} on small scales, which is still rather too big. At $r \approx 5h^{-1}$ Mpc, however, peculiar velocities in the open models are quite similar to those observed. In the simulation with a nonzero cosmological constant, the peculiar velocity distribution resembles that in an open model with the same value of Ω. In all the models σ is found to decline markedly with increasing pair separation, whereas the observations show it to be constant, or to increase slightly with scale.

4. If we write the irreducible three-point correlation function as $\zeta = Q\xi^2$, we find that, although Q is independent of triangle shape, it decreases slowly with triangle size in all our models. On scales $r \leq 5h^{-1}$ Mpc, we find values of Q for the mass distribution which are considerably greater than those estimated for galaxies. This discrepancy gets smaller as our models evolve, but it always remains significant.

5. The distribution of cluster masses is very broad in our

models and is quite different from that found either in simulations with white-noise initial conditions or in models where the initial density field has a large coherence length. Despite the fact that structure grows hierarchically in the present models, percolation across them is easier than across a Poisson distribution of points; in this sense our models show filamentary structure. These properties correspond quite closely to the results of Einasto *et al.* (1984) for the galaxy distribution in the neighborhood of the Local Supercluster.

6. The distribution of particle counts in cells is found to have significantly non-Gaussian tails, even for cubic cells with side exceeding $6r_0$ (where $\xi(r_0) = 1$). 4% of such cubes are underdense by a factor of 2. This frequency doubles if the distribution is analyzed in "redshift space." The observed frequency of such large low-density regions is very poorly known but appears to be about a factor of 2 larger than in our models. None of our simulations contains an underdense region as extreme as the apparent void in Boötes.

7. Artificial redshift catalogs of points, chosen from our simulations in a way which mimics the selection of galaxies for the CfA redshift survey, show considerable resemblance to the real data. These samples also show that substantial fluctuations are to be expected in estimates of clustering statistics derived from samples of this size. In particular it is quite likely that the anticorrelation seen in the CfA survey at separations of order $20h^{-1}$ Mpc is an artifact of the relatively small volume surveyed.

8. In all the above comparisons with observation, we assume that the distribution of galaxies may be identified with the mass distribution in our models. Observed galaxies may, however, be a biased reflection of the real underlying mass distribution. We have investigated the kind of bias to be expected if bright galaxies form only at relatively high peaks of the linear density distribution. If this is the case, their large-scale clustering follows that of the dark matter, but with strongly enhanced amplitude. For plausible parameters we find that the correlation length of galaxies in the present universe may exceed that of the dark matter by a factor of about 2.5. The peculiar velocities of galaxies are then smaller than those expected for an unbiased distribution by a similar factor. Applying a simple model of biased galaxy formation to our Einstein–de Sitter simulations, we find a "galaxy" distribution which fits the observed two- and three-point correlations and the observed peculiar velocity distribution quite well. This kind of model may therefore reconcile the theoretical imperative of a flat universe with the low density values implied by observations of galaxy motions.

The evolution of a universe dominated by cold dark matter is difficult to simulate because of the speed with which the characteristic scale of clustering grows. Even with a very large increase in computing resources, it would be possible to follow clustering at times before the start of our models over expansion by at most a factor of 2. Nevertheless, models with higher spatial resolution than those of the present paper are desirable in order to study the formation of galactic halos in greater detail. This is clearly a prerequisite for any more detailed theory of galaxy formation, and it would also be useful for a more thorough investigation of the consequences of our own very simple prescription. Models with a larger dynamic range are necessary to study the large-scale morphology of the "galaxy" distribution because small-scale resolution must be retained to locate the initial density peaks on which "galaxies" form. Until we have made such models, we cannot be sure that we can reproduce not only the correlation functions and other small-scale statistics of galaxies but also the filaments and voids which characterize their large-scale distribution. Higher resolution models would also help us to pin down the epoch of galaxy formation. This is important because it appears to be quite recent, particularly in the biased formation models, and so may be in conflict with observation.

Structure forms in a universe dominated by CDM in a well-specified way, and it is worth mentioning several aspects of its evolution which may be amenable to direct observational testing. The rapid increase in the characteristic scale of clustering might be observable on Mpc scales if $\Omega = 1$ and galaxies trace the mass distribution; however, such models are already ruled out by the kinematics of nearby galaxies. If the universe is open, or if our biased galaxy formation model is appropriate, the correlation function is a much weaker function of lookback time. Existing data on angular correlations of galaxies at $z \approx 0.3$ (Koo and Szalay 1984) may put significant constraints on these possibilities. The biased formation model implies that "failed" galaxies should be visible in our local region of space. The voids should contain some baryons either in low-density gas clouds or in low-surface-brightness galaxies. Space Telescope will be an excellent tool to search for such clouds. One would look for low-redshift Lyα absorption clouds in the spectra of distant QSOs. It is known that the Lyα clouds observed at $z = 2$ are more weakly clustered than bright galaxies (Sargent *et al.* 1980); this may already be evidence for biased galaxy formation. At low redshift the distribution of clouds found by Space Telescope could be correlated with the galaxy distribution measured by redshift surveys; if the clouds clustered only weakly with galaxies, and lay preferentially in "voids," it would be clear evidence that galaxy formation is biased on large scales.

The distribution of rich clusters of galaxies and of large voids presumably gives us information about initial fluctuations on large scale. The correlation function of Abell clusters can be understood as reflecting the enhanced clustering of rare peaks in the matter distribution if the underlying mass distribution is positively correlated on the relevant scales (Kaiser 1984*a*; Barnes *et al.* 1985). The linear correlation function corresponding to equation (2) is negative for $r > 18(\Omega h^2)^{-1}$ Mpc, so that if Abell clusters are indeed positively correlated at $r = 100h^{-1}$ Mpc (Bahcall and Soneira 1983), then we must conclude that $\Omega h < 0.18$. The Boötes void appears incompatible with CDM models in which galaxies are assumed to follow the mass distribution, as well as with the recipe for biased galaxy formation which we have considered. It is not clear whether an alternative recipe could explain the low density of galaxies in such a large region of space.

Further constraints on the CDM model come from observational limits on the anisotropy of the microwave background radiation. Recent calculations have shown that the stringent limit of Uson and Wilkinson (1984) is able to rule out CDM models with $\Omega h^{1.3} < 0.2$ at the 95% confidence level unless the universe was reionized at $z \geq 50$ (Bond and Efstathiou 1984; Vittorio and Silk 1984). The energy requirements for reionization are quite extreme, and if this loophole is discounted, the microwave background limits exclude all unbiased models for which we find even rough agreement with the observed peculiar velocity distribution of galaxies. The microwave limits are, however, very far from constraining the kind of biased formation model we have considered.

A great virtue of the general theoretical framework investi-

gated in this article is the fact that it makes very specific predictions for the initial conditions from which structure must form. It may therefore be possible to disprove it, in the same way as we found the standard neutrino-dominated universe to be untenable once its observational consequences were analyzed sufficiently carefully (White, Frenk, and Davis 1983; White, Davis, and Frenk 1984). Conversely, if the properties of a universe filled with CDM are found to agree with observation, this must be considered a significant success, since there is very little freedom to adjust the theoretical predictions. The direct detection of most of the elementary-particle candidates for CDM remains a remote possibility, so their cosmological implications may be the strongest indication of their existence. The major uncertainty in these implications comes from the fact that while we can predict the distribution of mass, what we see is the distribution of galaxies. As we have shown, the latter may be strongly influenced by the details of galaxy formation, a process which we are unlikely to understand fully in the foreseeable future. Nevertheless, it is remarkable how many aspects of the observed galaxy distribution are reflected quite faithfully by the distribution of CDM, and our crude attempts to model galaxy formation suggest that $\Omega = 1$ might even be *required* to get good agreement with observation. This seems too good to be true, but perhaps it hints that we are at last approaching a correct resolution of the missing mass problem.

This research has been carried out over a period of more than two years and has benefited greatly from discussions with many people. We would like to extend particular thanks to all participants in the 1984 program of the Evolution of Structure in the Universe at the Institute for Theoretical Physics. We are also very grateful to the director and to his staff for their warm hospitality during our stay. Various stages of this work were supported by the UK SERC, by the NSF under grants nos. AST 81-14715, AST 81-18557, PHY 77-27084 (supplemented by funds from the National Aeronautics and Space Administration), and by Department of Energy contract AT03-82ER1240069. Considerable amounts of computer time were provided by the ITP, by the Royal Greenwich Observatory, and by the Institute of Astronomy, Cambridge.

REFERENCES

Bahcall, N. A., and Soneira, R. M. 1983, *Ap. J.*, **270**, 20.
Bardeen, J. M. 1984, in *Inner Space/Outer Space* (Chicago: University of Chicago Press), in press.
Bardeen, J. M., Steinhardt, P. J., and Turner, M. S. 1983, *Phys. Rev. D*, **28**, 679.
Barnes, J., Dekel, A., Efstathiou, G., and Frenk, C. S. 1985, *Ap. J.*, submitted.
Bean, A. J., Efstathiou, G., Ellis, R. S., Peterson, B. A., and Shanks, T. 1983, *M.N.R.A.S.*, **205**, 605.
Bhavsar, S. P., Gott, J. R., and Aarseth, S. J. 1981, *Ap. J.*, **246**, 65.
Blumenthal, G. R., Faber, S. M., Primack, J. R., and Rees, M. J. 1984, *Nature*, **311**, 517.
Blumenthal, G. R., and Primack, J. R. 1983, in *Fourth Workshop on Grand Unification*, ed. H. A. Weldon, P. Langacker, and P. J. Steinhardt (Boston: Birkhauser), p. 256.
Bond, J. R., Centrella, J., Szalay, A. S., and Wilson, J. R. 1983, in *Formation and Evolution of Galaxies and Large Structures in the Universe*, ed. J. Audouze and J. Tran Thanh Van (Dordrecht: Reidel), p. 87.
Bond, J. R., and Efstathiou, G. 1984, *Ap. J. (Letters)*, **285**, L45.
Bond, J. R., and Szalay, A. S. 1983, *Ap. J.*, **274**, 443.
Bond, J. R., Szalay, A. S., and Turner, M. S. 1982, *Phys. Rev. Letters*, **48**, 1636.
Centrella, J., and Melott, A. L. 1983, *Nature*, **305**, 196.
Davis, M. 1984, in *Inner Space/Outer Space* (Chicago: University of Chicago Press), in press.
Davis, M., and Huchra, J. 1982, *Ap. J.*, **254**, 425.
Davis, H., Huchra, J., Latham, D. W., and Tonry, J. 1982, *Ap. J.*, **253**, 423.
Davis, M., and Peebles, P. J. E. 1977, *Ap. J. Suppl.*, **34**, 425.
———. 1983, *Ap. J.*, **267**, 465.
Dekel, A., and West, M. J. 1985, *Ap. J.*, **288**, 411.
Efstathiou, G., and Barnes, J. 1983, in *Formation and Evolution of Galaxies and Large Structures in the Universe*, ed. J. Audouze and J. Tran Thanh Van (Dordrecht: Reidel), p. 361.
Efstathiou, G., Davis, M., Frenk, C. S., and White, S. D. M. 1985, *Ap. J. Suppl.*, **57**, 241 (EDFW).
Efstathiou, G., and Eastwood, J. W. 1981, *M.N.R.A.S.*, **194**, 503.
Efstathiou, G., and Jedredjewski, R. 1985, in preparation.
Efstathiou, G., Fall, S. M., and Hogan, C. 1979, *M.N.R.A.S.*, **189**, 203.
Einasto, J., Klypin, A. A., Saar, E., and Shandarin, S. F. 1984, *M.N.R.A.S.*, **206**, 529.
Fall, S. M. 1979, *Rev. Mod. Phys.*, **51**, 21.
Frenk, C. S., White, S. D. M., and Davis, M. 1983, *Ap. J.*, **271**, 417.
Fry, J. 1984, *Ap. J.*, **279**, 499.
Groth, E. J., and Peebles, P. J. E. 1977, *Ap. J.*, **217**, 385.
Guth, A. 1981, *Phys. Rev. D*, **23**, 347.
Guth, A., and Pi, S.-Y. 1982, *Phys. Rev. Letters*, **49**, 1110.
Harrison, E. R. 1970, *Phys. Rev. D*, **1**, 2726.
Hawking, S. 1982, *Phys. Letters*, **115B**, 295.
Hockney, R. W., and Eastwood, J. W. 1981, *Computer Simulations using Particles* (New York: McGraw-Hill).
Huchra, J., Davis, M., Latham, D. W., and Tonry, J. 1983, *Ap. J. Suppl.*, **53**, 89.
Kaiser, N. 1984*a*, *Ap. J. (Letters)*, **284**, L9.
———. 1984*b*, in *Inner Space/Outer Space* (Chicago: University of Chicago Press), in press.
Kibble, T. W. B. 1976, *J. Phys. A*, **9**, 1387.
Kirshner, R. P., Oemler, A., Schechter, P. L., and Shectman, S. A. 1981, *Ap. J. (Letters)*, **248**, L57.
———. 1983, in *IAU Symposium 104, Early Evolution of the Universe and its Present Structure*, ed. G. O. Abell and G. Chincarini (Dordrecht: Reidel), p. 197.
Klypin, A. A., and Shandarin, S. F. 1983, *M.N.R.A.S.*, **204**, 891.
Koo, D., and Szalay, A. S. 1984, *Ap. J.*, **282**, 390.
Melott, A. L., Einasto, J., Saar, E., Suisalu, I., Klypin, A. A., and Shandarin, S. F. 1983, *Phys. Rev. Letters*, **51**, 935.
Oort, J. 1983, *Ann. Rev. Astr. Ap.*, **21**, 373.
Pagels, H. R., and Primack, J. R. 1982, *Phys. Rev. Letters*, **48**, 223.
Peebles, P. J. E. 1974, *Ap. J. (Letters)*, **189**, L51.
———. 1980, *The Large Scale Structure of the Universe* (Princeton: Princeton University Press).
———. 1982, *Ap. J.*, **258**, 415.
———. 1984, *Ap. J.*, **277**, 470.
Peebles, P. J. E., and Groth, E. J. 1975, *Ap. J.*, **196**, 1.
Preskill, J., Wise, M. B., and Wilczek, F. 1983, *Phys. Letters*, **120B**, 127.
Press, W. H. 1980, *Phys. Scripta*, **21**, 702.
Sargent, W. L. W., Young, P. J., Boksenberg, A., and Tytler, D. 1980, *Ap. J. Suppl.*, **42**, 41.
Shanks, T., Bean, A. J., Efstathiou, G., Ellis, R. S., and Peterson, B. A. 1983, *Ap. J.*, **274**, 529.
Shapiro, P. R., Struck-Marcell, C., and Melott, A. L. 1983, *Ap. J.*, **275**, 473.
Starobinskii, A. 1982, *Phys. Letters*, **117B**, 175.
Turner, M., and Schramm, D. 1978, *Nature*, **279**, 303.
Uson, J. M., and Wilkinson, D. T. 1984, *Ap. J. (Letters)*, **277**, L1.
Vittorio, N., and Silk, J. 1984, *Ap. J. (Letters)*, **285**, L39.
White, S. D. M., Davis, M., and Frenk, C. S. 1984, *M.N.R.A.S.*, **209**. 27P.
White, S. D. M., Frenk, C. S., and Davis, M. 1983, *Ap. J. (Letters)*, **274**, L1.
White, S. D. M., and Negroponte, J. 1982, *M.N.R.A.S.*, **201**, 401.
Weinberg, S. 1979, *Phys. Rev. Letters*, **42**, 850.
Wilson, M. L., and Silk, J. 1981, *Ap. J.*, **243**, 14.
Yang, J., Turner, M. S., Steigman, G., Schramm, D. N., and Olive, K. A. 1984, *Ap. J.*, **281**, 493.
Zel'dovich, Ya. B. 1970, *Astr. Ap.*, **5**, 84.
———. 1972, *M.N.R.A.S.*, **160**, 1P.

MARC DAVIS: Department of Astronomy, University of California, Berkeley, CA 94720

GEORGE EFSTATHIOU: Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, England

CARLOS FRENK: Astronomy Centre, Physics Building, University of Sussex, Falmer, Brighton BN1 9QH, England

SIMON WHITE: Steward Observatory, University of Arizona, Tucson, AZ 85721