# Evolution and clustering of rich clusters 

Nick Kaiser Institute of Astronomy, Madingley Road, Cambridge CB30HA

Accepted 1986 April 30. Received 1986 April 21


#### Abstract

Summary. In some hierarchical theories for galaxy formation the present evolution of the clustering pattern should approximate that which arises from scale-free initial conditions. The characteristic mass-scale of clustering is predicted to be evolving very rapidly, so we can hope to test these hypotheses by observations of clustering at quite accessible redshifts. We extend the well-known scaling laws for the dimensional quantities which characterize the clustering to give quantitative formulae for the evolution of general cluster catalogues. We show that X-ray selected clusters should display strong positive density evolution, with the characteristic luminosity, $L^{*}$, remaining roughly constant. The comoving number density of clusters selected according to a criterion like Abell's should, on the other hand, remain roughly constant. Core radii should be much smaller in the past, and the gravitational lensing power of these clusters should be greater than for an unevolving population of clusters. Consideration is given to processes which will modify the self-similar scaling results. A corollary of the rapid evolution in these models is that small-amplitude, long-wavelength density perturbations will spatially modulate the clustering process, resulting in an enhancement of the correlation function for optically selected clusters, $\xi_{\mathrm{c}}$, relative to the density correlation function $\xi_{\varrho}$. The enhancement factor $\xi_{\mathrm{c}} / \xi_{\rho}$ can be expressed in terms of the spectral index of the initial fluctuations and the present logarithmic slope of the mass distribution function of rich clusters. Estimates of the latter, combined with the observed clustering lengths of galaxies and rich clusters, require a spectral index $n \simeq-1.5$. A similar analysis applied to X-ray selected clusters suggests that these should be less strongly clustered than the optically selected clusters, though this result is sensitive to uncertainty in the slope of the X-ray luminosity function.


## 1 Introduction

An important constraint on theories for the formation of cosmological structure is that they should be able to account for the properties of rich clusters of galaxies. There are various currently popular theories for the type and spectrum of initial fluctuations, from which it should be possible, at least in principle, to make quantitative predictions for the form of present-day rich
clusters. If we ignore the gas content, progress can be made via $N$-body techniques, though the fact that rich clusters contain only a tiny fraction of all the matter in the Universe makes this a demanding task. In order to treat the gas content, which may make an important contribution to the total mass and is crucial for the X-ray properties, we would need full three-dimensional hydrodynamical simulations which are not at present a viable proposition.

In this paper we shall take a less direct approach to the problem: we shall not attempt to solve directly the problem of what non-linear structures form from general linear initial conditions; we will concentrate instead on how the population of clusters evolves with time. We shall show that, for an interesting class of initial conditions, those which are at least approximately scale-free, the clusters should evolve very rapidly with time, and in a manner which is readily predictable. Thus, while we will not be able to predict the form of present-day clusters, we shall make detailed quantitative predictions for the form of clusters at $z \geqslant 0$. A model will be constructed which incorporates both the collisionless dynamics of the 'dark-matter' and galaxies, but also the important process of gas infall. The model results in very simple analytic formulae for the evolution of the distribution of cluster masses and other properties. These formulae can readily be fitted to, and tested against, the data as these become available.

The layout of the paper is as follows: in Section 2 we will briefly review the characteristic observed properties of rich clusters which we will have cause to call upon throughout the paper.

In Section 3 we will discuss the theoretical framework within which we will discuss the data: we find the observed properties to be broadly consistent with those expected if clusters formed very recently by gravitational instability from primordial density fluctuations. We then focus attention on 'hierarchical' theories exemplified by the 'cold-dark-matter' hypothesis. We argue that cluster formation in these and similar models may be well approximated by that which follows from scale-free initial conditions with a suitable chosen spectral index $n$. This assumption results in a simple scaling solution for the non-linear structures which condense. We feel that this should provide a good description of the process by which clusters are assembled in this scenario.

In Section 4 we formulate this scaling law in terms of cluster catalogues: having specified some conditions which the selection criteria should fulfil in order that the scaling be manifest, we give the general result for the evolution of cluster properties, followed by some specific illustrative and practical examples: we consider core and halo X-ray emission; evolution of clusters as selected by Abell; optical core radii; and finally gravitational lensing.

In Section 5 we discuss those processes which break the self-similar scaling of our idealized model. These considerations constrain the range of masses and redshifts for which our various tests should be reliable and also suggest which of these tests should be most robust. These considerations are also important since we anticipate that the model will eventually be excluded by the data and, when this occurs, we would like to know at least in what direction the necessary modification of the model should proceed.

In Section 6 we change direction somewhat and consider the relationship between the rapid evolution of clusters in this scenario, and the clustering strength of these objects relative to that of the underlying matter. Using a simple model we obtain a formula which relates the clustering strength to the mass distribution of present-day clusters and the primordial spectral index $n$. We apply this to the observations of clusters in order to place constraints on the form and amplitude of matter fluctuations on large scales.

In Section 7 we summarize our main results.

## 2 Properties of rich clusters at the present epoch

In this section we will review some basic properties of rich clusters which we will call upon throughout the paper. This is by no means an exhaustive review of cluster properties, for which
the reader should refer to Sarazin (1986). Other useful reviews are those by Bahcall (1977) and Mushotzky (1984), the latter providing a concise introduction to the X-ray properties of clusters. For detailed discussion of the Einstein cluster observations see Abramopoulos \& Ku (1983) and Jones \& Forman (1984).

### 2.1 OPTICAL PROPERTIES

The most widely studied statistical catalogue of clusters of galaxies is that of Abell (1958). These clusters were selected and classified by Abell according to the number $N_{\mathrm{A}}$, this being the number of galaxies within two magnitudes of the third brightest which are contained within a cylinder of radius $R_{\mathrm{A}}=1.5 h^{-1} \mathrm{Mpc}$ centred on the cluster (after correction for foreground and background galaxies). Here and throughout, $h$ will denote Hubble's constant in units of $100 \mathrm{~km} \mathrm{~s}^{-1} \mathrm{Mpc}^{-1}$. Abell's primary classification was a discrete form of $N_{\mathrm{A}}$ known as 'Richness', and the intervals of $N_{\mathrm{A}}$ corresponding to the various richness classes can be found in his catalogue. For example, clusters of richness class $R=1$ have $50 \leqslant N_{\mathrm{A}} \leqslant 79$. The number $N_{\mathrm{A}}$ can be found tabulated in the paper by Kowalski et al. (1984).
Probably the best studied cluster of galaxies is the Coma cluster. Several hundred redshifts are known for galaxies in this cluster and virial analysis (see e.g. Kent \& Gunn 1982) gives a mass of $\simeq 10^{15} M_{\odot}$ within $R_{\mathrm{A}}$. Coma is a rather rich cluster [Abell classified it as $R=2$, but according to Dressler (1978) this may be an underestimate] and has a central velocity dispersion $\sigma_{0} \simeq 1000 \mathrm{~km} \mathrm{~s}^{-1}$. The data for velocity dispersions for $R=1$ clusters are less abundant, but a typical value for $\sigma_{0}$ may be $700 \mathrm{~km} \mathrm{~s}^{-1}$, in which case the mass interior to $R_{\mathrm{A}}$ would be $\simeq 5 \times 10^{14} m_{\odot}$.

The mass to light ratio for Coma is $M / L \simeq 400 h$, and this number is not atypical of rich clusters in general. This is much greater than the $M / L$ for any reasonable stellar population, for which $M / L \sim 2-10$, or indeed the mass to light ratios of galaxies including their dark haloes. Thus the stars in galaxies make only a small contribution to the mass of a rich cluster.

The density of material within radius $R_{\mathrm{A}}$ is, for Coma, about $300 \varrho_{\text {crit }}$, where $\varrho_{\text {crit }}=3 H^{2} / 8 \pi G$ is the density to close the Universe. For $R=1$ clusters the corresponding value is $\sim 150 \varrho_{\text {crit }}$ -

The galaxy density profile has been estimated for many clusters, but there is a severe problem as to how to assign cluster membership. A method which avoids this problem is to calculate the statistical average density profile which is given by the cluster-galaxy cross-correlation function. This has been estimated by Seldner \& Peebles (1977) and is well described by the two-power-law fit
$\xi_{\mathrm{cg}}=a r^{-1.7}+b r^{-2.5}$,
with $a=130$ and $b=70$ (for $h=1$ ). The two power laws cross over around $R_{\mathrm{A}}$ as it happens, and for radii of this order the effective logarithmic slope of the profile is about 2 .

The cluster profile may flatten off near the centre, and various attempts have been made to fit for a core radius $r_{\mathrm{c}}$. Bahcall (1975) finds that there is no correlation between $r_{\mathrm{c}}$ and richness and suggests that $r_{\mathrm{c}}$ provides a constant standard of length. However, other workers have found discrepant results when estimating core radii for even the same clusters (see e.g. Dressler 1978), so it seems that core radii, or at least those determined from galaxy counts, should be treated with some caution (see Beers \& Tonry 1986).
The space number density of clusters of richness $R \geqslant 1$ has been estimated by Abell to be $N \simeq 5 \times 10^{-6}\left(h^{-1} \mathrm{Mpc}\right)^{-3}$. Combining this with the density estimate we find that the material lying within an Abell radius of these clusters contributes about 1 per cent of closure density.
The luminosity function of rich clusters can be estimated from Abell's counts and is found to be rather steep: for instance, the number density of clusters of richness class $R \geqslant 2$ is about four times
smaller than that of $R \geqslant 1$, yet the number of galaxies within $R_{\mathrm{A}}$ is only about 60 per cent greater. It turns out that a more useful mass (or luminosity) estimate for our purposes is the number of galaxies within a radius within which the mean density of galaxies has some specific value, rather than within a fixed radius. If we assume that the number of galaxies per unit mass is reasonably constant from cluster to cluster, then this gives a reasonable measure of mass within a surface of given density contrast, which we shall denote by $M_{\Delta}$. If we make the assumption that individual clusters have a profile like the ensemble average it is straightforward to adjust Abell's richness classification to give $M_{\Delta}$. If we fit a power law through the counts for richness classes $R=1$ and $R=2$, we find $N\left(M_{\Delta}\right) \propto M_{\Delta}^{-1.8}$. For virialized clusters $M_{\Delta}$ should vary as $T^{3 / 2}$, so we can convert this mass function to a temperature distribution function: $N(T) \propto T^{-2.7}$, where $N(T)$ is the number density of clusters with virial temperature $\geqslant T$.

For a higher richness class the data suggest an increasingly steep slope for the luminosity function, though the number of such clusters becomes very small so this result may not be very significant. A possible source of systematic errors in the luminosity function is discussed by Lucey (1983) who points out that foreground/background contamination requires a substantial correction and there is some danger that this may have been systematically misestimated by Abell. What is clear however, is the qualitative result that the mass function is very steep and consequently these clusters span only a very small range of masses.

## 2.2 x-RAY Properties

Rich clusters are conspicuous in the X-ray sky as extended sources of emission. There is now general agreement that the bulk of these X-ray photons are optically thin bremsstrahlung from hot intracluster gas. Most of the luminosity comes from the dense central regions of the cluster. Fits can be made to a profile with a core radius and, rather like the result for optical core radii, reveal a lack of correlation with luminosity (Abramopoulos \& Ku 1983). The X-ray emission is consistent with gas in equilibrium in the potential well of the cluster with temperature $T_{\mathrm{gas}} \sim$ $\sigma_{\text {galaxies. }}^{2}$. The fraction of the total cluster mass which is in gas is somewhat model dependent (Hendriksen 1985). The fraction of the mass in the core region which is in the form of diffuse gas is around $0.1 h^{-3 / 2}$. Dark matter is therefore necessary to bind the core, but if $h$ is small, the gas fraction may be appreciable. Iron lines are seen in the spectrum which indicate that the gas in the core region of the cluster is enriched with metals to about half the solar abundance.

Estimates of the cooling time for the gas show that the bulk of the gas cannot cool appreciably in the age of the Universe. However, in the dense inner core regions of the more massive clusters the cooling time is comparable to the age of the Universe (Abramopoulos \& Ku 1983; Jones \& Forman 1984) and cooling flows are found in many such clusters (Fabian, Nulsen \& Canizares 1984).

Various correlations between X-ray and optical properties are seen (Abramopoulos \& Ku 1983). There is a particularly tight correlation between total X-ray luminosity $L_{\mathrm{X}}$ and gas temperature, of the form $L_{\mathrm{X}} \propto T^{2.3 \pm 0.2}$. As discussed by Mushotzky (1984), this would be consistent with a model in which clusters of different mass have similar gas fractions and have core radii which are independent of luminosity or mass.

Various estimates have been made for the luminosity function for Abell clusters. Less data are available for X-ray selected clusters, but Piccinotti et al. (1982) have made such an estimate. They find $N \propto L_{\mathrm{X}}^{-1.15}$ where $N$ is the number density of clusters brighter than $L$. Using the $L-T$ correlation one can convert this to a temperature distribution: $N \propto T^{-2.7 \pm 0.3}$. Comparing this with the temperature function derived from the optical studies we find very good agreement. This gives us some confidence in this particular result which we shall call upon several times in the forthcoming sections.

## 3 Gravitational instability in the hierarchical picture

In this section we will discuss the theoretical framework within which we will attempt to interpret the properties of clusters described in the previous section.

### 3.1 HOMOGENEOUS SPHERICAL COLLAPSE

Probably the simplest model for the growth of non-linear structures by gravitational instability is provided by the homogeneous spherical collapse model. In this model we take a homogeneous, $\Omega=1$, background Universe, excise a sphere of matter, and replace it with a somewhat smaller concentric sphere of matter. We arrange that the age of this dust sphere is the same as that of the background Universe, but we let the sphere have negative total energy. The time evolution of such an inhomogeneity is well known: at early times the density is very close to that of the background and the fractional perturbation grows according to the usual Lifshitz growth rate $\Delta \varrho / \varrho \propto t^{2 / 3}$. As $\Delta \varrho / \varrho$ approaches unity the growth rate increases. If the perturbation starts at $t_{1}$ with amplitude, $\varepsilon$, then, when the background Universe has expanded by a factor $1.06 / \varepsilon$, the inner sphere is just turning around. At that time, the density of the inner sphere is about 5.5 times the density of the background Universe, whereas simply extrapolating the density perturbation by linear theory would give a density contrast of 1.06 . For a perfectly homogeneous perturbation, with infinite phase space density, the sphere will collapse to a singularity when the background Universe has expanded by a factor $1.68 / \varepsilon$. For a realistic collapse with any appreciable degree of inhomogeneity, we expect the collapsing perturbation to violently relax (Lynden-Bell 1967) and form a virialized sphere with virial radius roughly equal to one half of the radius at maximum expansion. A homogeneous sphere of this radius would have a density about 170 times that of the background Universe. Since the inner sphere is now stable, while the background Universe continues to expand, the density contrast grows thereafter like $t^{2}$.

This result can readily be generalized to an open background Universe. One finds that, in the limit as $\Omega \rightarrow 0$, the bound perturbation virializes at a density of $\simeq 80 \varrho_{\text {crit }}$, thus the density of a recently virialized sphere, in units of the critical density, is fairly insensitive to the value of $\Omega$ for the background Universe.

Comparing these results with the density of a typical rich cluster, which we estimated to be around $150 \varrho_{\text {crit }}$, we find the observed properties to be in good accord with this simple model, provided that the clusters have collapsed and virialized very recently.

### 3.2 THE HIERARCHICAL PICTURE

In the hierarchical picture we assume that, at some early epoch $t_{\mathrm{i}}$, there are small-amplitude density perturbations
$\Delta\left(\mathbf{r}, t_{\mathrm{i}}\right)=\frac{\varrho\left(\mathbf{r}, t_{\mathrm{i}}\right)-\bar{\varrho}\left(t_{\mathrm{i}}\right)}{\bar{\varrho}\left(t_{\mathrm{i}}\right)}$
which are assumed to take the form of some homogeneous and isotropic random process. This initial density field will give rise to a hierarchy of collapsing structures provided that there are fluctuations on all mass scales and that the amplitude of the fluctuations is a smoothly decreasing function of the mass scale, so that the smallest perturbations collapse first. The simplest example of such a field is the Gaussian field, which is constructed as a linear superposition of Fourier components with randomly chosen phases. Such a field is specified by a single function: the power spectrum, so there is considerable freedom in the choice of $\Delta(\mathbf{r})$. Gaussian fields are only one
possibility, and if non-Gaussian fields are allowed, the space of initial states multiplies enormously.

While the linear growth of fluctuations of arbitrary form can be solved analytically, the situation becomes very complicated when the fluctuations turn around and collapse. If we assume that the Universe is composed of collisionless particles, then, for any specific realization of any specific random process, we can integrate Newton's equations of motion for these particles into the non-linear regime (see e.g. Efstathiou et al. 1985). Various cases have now been investigated and one obtains from these results some suggestions as to how the general process of hierarchical clustering takes place. In these simulations one can identify, at any time, concentrations of matter with a range of masses. These concentrations progressively come together and merge into larger bound structures. While, for the simple isolated spherical perturbation considered in Section 3.1 the density contrast becomes arbitrarily large at late times, in a merging hierarchy this does not occur as the sub-structure seems to be rapidly erased as each new level of the hierarchy collapses. A detailed examination of these condensations (from simulations kindly made available by G. Efstathiou) reveals that, while there are no well-defined edges to these clusters, if we construct a sphere around the cluster centre such that the interior density is about 200 times the background density, then inside this sphere the particles are in a state of quasi-equilibrium, with little net infall or expansion, whereas outside this sphere the motion is predominantly an unequilibrated infall. Comparing this result with real clusters then one finds that, if these clusters form in a similar manner to those in the simulations, the Abell radius $R_{\mathrm{A}}$ roughly marks the boundary between infall and the virialized interior. As was the case for the simple homogeneous collapse, we are forced to the conclusion that rich clusters are forming at the present epoch, and are gaining mass at a great rate.

### 3.3 SELF-SIMILAR HIERARCHIES

For the general problem of non-linear evolution there is no alternative but to resort to numerical techniques. Even then, one can only treat the idealization of a Universe containing only collisionless material. There is, however, a special set of initial conditions which possess a particular symmetry: these are those in which the background Universe has closure density ( $\Omega=1$ ), so the Universe is scale-free, and in which the amplitude of the initial density fluctuations is a power-law function of the length scale, so these perturbations are also scale-free.

Even with these specially symmetrical initial conditions it seems that one cannot solve for the final state other than by numerical techniques. However, if one assumes that the present non-linear structure in our Universe did evolve from this type of initial condition, one can predict in detail the statistical nature of the clusters etc. at any other epoch, and thereby subject this hypothesis to a very powerful test. In this paper our goal is to develop the formalism of this test as appropriate to be applied to rich clusters. We will find that this method allows one to describe not only a collisionless dust of particles but also, to a useful approximation, the behaviour of the intracluster gas (Section 3.4). For the moment, however, we will consider only collisionless particles.

With our assumed scale-free initial conditions and background Universe, there is only one physical scale in the problem; that being the mass- or length-scale of those fluctuations which are going non-linear, which will be a decreasing power-law function of redshift. It is conventional to parameterize this power law in terms of the spectral index, $n$, such that the non-linear mass-scale varies as $M_{\mathrm{NL}} \propto(1+z)^{-6 /(n+3)}$. Any dimensionless statistical property of the final density perturbations, such as the correlation function etc., must be a function only of the combination $M / M_{\mathrm{NL}}$; and the final state is temporally self-similar. The scaling for the density contrast field
$\Delta(\mathbf{r}, t)$, where $\mathbf{r}$ is a comoving spatial coordinate can be expressed as follows: if $\Delta_{1}(\mathbf{r})=\Delta\left(\mathbf{r}, t_{1}\right)$ and $\Delta_{2}(\mathbf{r})=\Delta\left(\mathbf{r}, t_{2}\right)$, then for any dimensionless statistic $S=S[\Delta]$ we have
$S\left[\Delta_{1}(\mathbf{r})\right]=S\left[\Delta_{2}(\alpha \mathbf{r})\right]$
where
$\alpha=\left(\frac{t_{2}}{t_{1}}\right)^{4 /(3 n+9)}$.
This result can be extended to the distribution function for peculiar velocities $\mathbf{v}$ which scale as $\mathbf{v}(\mathbf{r}) \rightarrow \beta \mathbf{v}(\alpha \mathbf{r})$, with $\beta=\left(t_{2} / t_{1}\right)^{(n-1) /(3 n+9)}$, this value for $\beta$ being dictated in order that the equation of continuity be satisfied.

This scaling solution is by no means novel. It is used by Efstathiou, Fall \& Hogan (1979), and more recently by Frenk et al. (in preparation) to evaluate the performance of $N$-body routines, and is also discussed by Peebles (1980). Our purpose here is to show how this scaling applies to rich clusters in order to test whether structure in our Universe is evolving according to equation (3.1). We will first make a few remarks concerning the applicability of these scale-invariant initial conditions.

Probably the most common example of these scale-free initial conditions is the Poisson distribution of point masses, where, for $M \gg M_{\text {point mass }}$ we obtain a Gaussian field with spectral index $n=0$. More general Gaussian fields can be constructed with arbitrary $n$. Non-Gaussian scale-free processes have also been postulated: one example is the 'cosmic string' hypothesis (see Turok \& Brandenberger 1985; Vilenkin 1985), in which the post-recombination density perturbation field takes the form of a scale-free distribution of mass points with a power-law spectrum of masses. This has spectral index $n=-1$ (at least for masses much less than the horizonal scale at the epoch of matter and radiation equality). Another example of a scale-free non-Gaussian process is provided by Peebles' (1983) prescription.

There is no particularly good reason to assume that the spectrum of fluctuations emerging from recombination should satisfy the scaling symmetry we have assumed. Indeed, in most models which have been considered, and in which the spectrum is assumed to be self-similar at very early times, the spectrum is modified as the fluctuations enter the horizon and the spectrum typically develops some feature around the horizon scale at $z_{\text {eq }}$. An example of this behaviour is provided by the 'cold-dark-matter' spectrum (Peebles 1984; Bond \& Efstathiou 1984; Blumenthal \& Primack 1984), where the spectrum asymptotes to $n=-3$ on very small scales and $n=+1$ on large scales. The transition between these asymptotic regimes is, however, very broad, taking several orders of magnitude in length scale. On the scale corresponding to rich clusters the effective (i.e. local) logarithmic slope of the power spectrum is $n_{\text {eff }} \simeq-1$ for the case of 'adiabatic' fluctuations, and is somewhat shallower, $n_{\text {eff }} \simeq-1.5$, for the 'isothermal' case. For cases such as these, provided we restrict ourselves to a small range of masses it should be a good approximation to model the spectrum by a scale-free spectrum with $n=n_{\text {eff }}$. The true spectrum will differ from the approximation only for masses very much greater than or very much less than the scale of interest. The physical reason that one can have reasonable confidence in this approximation is that the coupling between fluctuations on very different mass scales is very weak: the fluctuations on very small scales can be thought of as determining the graininess of the matter distribution; if one modifies the fluctuations on small scales in a physically allowable manner, i.e. conserving mass and momentum, then they will modify the fluctuations on larger scales, but only to a very minor extent. The reason for this is that the conservation laws effectively shield the monopole and dipole terms in the force due to the small-scale fluctuations, and the remaining quadrupole and higher order moments fall off very rapidly with distance. If one introduces small-scale fluctuations in an otherwise homogeneous Universe this will initiate a process of hierarchical clustering but with very steep spectral index $n=4$ (Zel'dovich 1965; Hogan, Kaiser \& Rees 1982).

For spectra of the type we are concerned with here which have $n \ll 4$ these fluctuations are negligible. There is, in fact, a more important interaction with small-scale fluctuations which arises because, as sub-structure is merged away, the binding energy must be absorbed by the higher levels of the hierarchical structure. The binding energy inherited from much lower levels of the hierarchy converges only if $n<1$, and we shall henceforth assume that this is the case. Fluctuations on very large scales can interact with smaller scale fluctuations via their tidal influence, but this is very small compared to the gravitational fields intrinsic to the scales of interest provided we have $n>-3$, which is just the condition that we have a hierarchical picture in any case. If we neglect the tidal forces then the only effect of long-wavelength perturbations is to cause the effective value of $\Omega$ to vary slightly from place to place. This will have only a very small influence on the distribution of cluster masses etc. but will be more important for the clustering of clusters. We will return to this point in Section 6.

It is difficult to quantify how strong the curvature of the spectrum would need to be before the approximation discussed above breaks down. The result must depend on the particular application, the most robust applications being those which concern phenomena spanning only a small range of masses. Since the range of masses of rich clusters is very small, it seems that for the present application the approximation should be rather good.

### 3.4 GAS INFALL

So far we have considered only collisionless particles. The X-ray properties, which increasingly dominate our empirical data base, are determined by the spatial distribution of diffuse intracluster gas. The cooling time for the bulk of this gas is much longer than the Hubble time in present epoch clusters, and the same will be true of the somewhat smaller units from which these clusters have been assembled in this scenario. The temperature of these sub-units will be typically somewhat lower that the virial temperature of the final cluster so the infall/merging process will be supersonic, and the gas will be heated to the virial temperature by shocking.

This process will no doubt be horrendously complicated. However, if we neglect dissipation, no additional physical scale is introduced into the problem and consequently the scaling laws we have written down must apply to an adiabatic gas component just as much as to the collisionless dust we have considered hitherto. This is not to say that the gas will end up distributed like the collisionless material. We do not expect that this will be the case since, while these two components end up sharing the same potential well, the processes by which they are heated to the virial temperature are very different. What we can assert is that whatever the distribution of collisionless matter and gas is in a present-day cluster, the distribution in a corresponding cluster at any earlier epoch will be precisely the same.

It is not difficult to enumerate several processes which break the self-similar scaling, and which must be important at some level, and we will discuss some of these in Section 5. However, we feel that by far the most important event in the history of any element of gas concerning its present state is the way it fell into the cluster in which it presently resides, and we feel that the assumption of adiabaticity should provide at least a good first approximation to this process. With this justification we will now explore the predictions that can be made on the basis of the processes we have discussed so far.

To recapitulate: we will assume:
(i) The Universe has closure density.
(ii) The initial density fluctuations are scale free.
(iii) The material content has two components: (a) collisionless particles (including galaxies) and (b) collisionless gas with $t_{\text {cool }} \gg t_{\text {dynamical }}$ at all times.

Under these conditions the density contrast and velocity field of the collisionless component has temporal self-similarity as described by equation (3.1), with analogous results for the density contrast and temperature of the gas.

In the next section we will translate these scaling laws into the language of cluster catalogues.

## 4 Evolution of cluster catalogues

### 4.1 FORMALISM

Equation (3.1) tells us how to relate the present epoch density contrast and peculiar velocity fields to those in the past. The prescription is simply to re-scale comoving distances as $r \rightarrow r \times(1+z)^{-2 /(n+3)}$ and peculiar velocities as $v \rightarrow v \times(1+z)^{(n-1) /(2 n+6)}$. If we apply this rescaling to a representative member of the ensemble of fields resulting from a particular random process then we will not obtain the density field etc. for the same realization at the earlier epoch $z \geqslant 0$, but we will get another representative member of the ensemble at that epoch.

In practice, we cannot readily measure the density contrast field or the peculiar velocity field at all points in space. For instance, it is very difficult to know the matter density contrast in a region which happens to be devoid of galaxies. What we can hope to determine is the density field around conspicuous condensations like rich clusters. These objects are conveniently described by a distribution function $n\left(X_{1}, X_{2}, \ldots, X_{N}\right)$, where the $X_{i}$ are a set of parameters describing the clusters (e.g. richness, X-ray luminosity etc.). We define $n\left(X_{1}, X_{2}, \ldots, X_{N}\right) d^{N} X$ to be the comoving number density of clusters with properties in the range $X_{i}$ to $X_{i}+d X_{i}$. For any given set of clusters, different group catalogues can be constructed according to different choices for parameters $X_{i}$. We have seen that with our scale-free assumptions for the initial state the evolution of the density field is almost trivially simple. However, this simplicity of the scaling will not necessarily be manifest in the cluster catalogue if the parameters are inappropriately chosen. If the scaling is to be manifest, we require that, if a particular cluster $C$ at the present epoch is assigned parameters $X_{i}$, then the parameters $X_{i}^{\prime}$ assigned to the scaled-down replica $C^{\prime}$ at some earlier time should be identical to the parameters $X_{i}$, aside from an appropriate scaling of any dimensionful parameters. In general a cluster catalogue will possess this desirable property provided that no physical scale is introduced in selecting and parameterizing the cluster. For instance, if we take $X_{1}$ to be the comoving radius within which the cluster has density contrast equal to 200 say, then we will have
$X_{1}^{\prime}=(1+z)^{-2 /(n+3)} X_{1}$
as desired. But if we set $X_{1}$ to be the radius at which the density has some particular value then no such simple scaling will apply. In the latter case, even though the clusters will satisfy a very simple symmetry of scaling, this will be obscured in the cluster catalogue by an awkward choice of parameter. Another suitable parameter, $X_{2}$ say, might be the central velocity dispersion $\sigma_{0}$, or perhaps the projected velocity dispersion at a radius $R=\sigma_{0} / H$ say. These parameters would scale according to
$X_{2}^{\prime}=(1+z)^{(n-1) /(2 n+6)} X_{2}$,
while an unsuitable parameter would be the projected velocity dispersion at a fixed physical distance of, say, $1.5 h^{-1} \mathrm{Mpc}$.
We now write down how some interesting dimensionful parameters will scale from some arbitrarily chosen cluster $C$ to its scaled replica $C^{\prime}$ at $z>0$. We will refer to these as the characteristic parameters. They will be used as a reference point against which to compare other
clusters. For concreteness we will take these to be the values corresponding to the material lying within a surface of given density contrast.
(i) The characteristic density $\varrho^{*}$ scales simply as
$\varrho^{*} \propto(1+z)^{3}$
i.e. like the background density.
(ii) The characteristic mass scales as
$M^{*} \propto(1+z)^{-6 /(n+3)}$
i.e. as the cube of the comoving length scale.
(iii) The characteristic physical radius scales as
$R^{*} \propto(1+z)^{-(5+n) /(n+3)}$.
(iv) The characteristic temperature scales as
$T^{*} \propto(1+z)^{(n-1) /(n+3)}$.
(v) Finally, the distance to a comparable neighbour cluster scales as $(1+z)^{-2 /(n+3)}$, so the characteristic number density will scale as
$N^{*} \propto(1+z)^{-6 /(n+3)}$.
For $n=-1$ for instance, we have $\varrho^{*} \propto(1+z)^{3} ; M^{*} \propto(1+z)^{-3} ; R^{*} \propto(1+z)^{-2} ; T^{*} \propto(1+z)^{-1}$ and $N^{*} \propto(1+z)^{3}$. We would therefore predict that clusters at $z \geqslant 0$ would be denser, less massive, smaller, cooler, and more numerous than those at the present epoch. Note that for this value of $n$, some of these quantities vary as quite high powers of $(1+z)$ so we have the hope of observing noticeable evolution even at quite moderate redshifts $z \sim 0.3$ or so.

With these preliminaries out of the way we can proceed to write down the scaling law for the distribution function $n\left(X_{1}, X_{2}, \ldots\right)$. Let us assume that we observe, for clusters at the present epoch, $n_{0}\left(X_{1}, X_{2}, \ldots\right)$. Then at redshift $z$ we have
$n_{z}\left(X_{1}, X_{2}, \ldots\right) d^{N} X=(1+z)^{6 / n+3} n_{0}\left(\frac{X_{1}^{*}(0)}{X_{1}^{*}(z)} X_{1}, \frac{X_{2}^{*}(0)}{X_{2}^{*}(z)} X_{2}, \ldots\right) \prod_{i} \frac{X_{i}^{*}(0)}{X_{i}^{*}(z)} d X_{i}$.
This equation follows simply from the fact that the comoving number density of clusters with dimensionless parameters $Y_{i}=X_{i} / X_{i}^{*}$ within some specified ranges must scale as $N^{*}$, i.e. as $(1+z)^{6 /(n+3)}$.

Since all of the factors $X_{i}^{*}(0) / X_{i}^{*}(z)$ are simple power laws of $1+z$, equation (4.1) provides a very simple expression which can be fitted to the data to find the single parameter $n$.

It is often convenient to work with cumulative distribution functions $N\left(X_{i}\right)$, often with just a single argument. We define the cumulative distribution in the parameter $X_{1}$ to be
$N\left(X_{1}\right)=\cdot \int_{X_{1}}^{\infty} d X_{1} \int_{-\infty}^{\infty} d X_{2} \ldots \int_{-\infty}^{\infty} d X_{N} n\left(X_{1}, X_{2}, \ldots\right)$.
This cumulative distribution function evolves according to
$N_{z}\left(X_{1}\right)=(1+z)^{6 /(n+3)} N_{0}\left(\frac{X_{1}^{*}(0)}{X_{1}^{*}(z)} X_{1}\right)$.
We will now examine some choices of $X_{i}$ for which we have, or can hope to gain, some knowledge of $n_{0}\left(X_{i}\right)$ or $N_{0}\left(X_{i}\right)$ in order to make predictions for the properties of clusters at moderate redshift.

### 4.2 X-RAY EMISSION

### 4.2.1 Total $X$-ray luminosity

Comparing the clusters $C$ and $C^{\prime}$, we find that the X-ray luminosity scales as
$L_{\mathrm{X}}^{*} \propto M^{*} \varrho^{*} T^{* 1 / 2}$.
So this characteristic luminosity scales as $L_{\mathrm{X}}^{*} \propto(1+z)^{(7 n+5) /(2 n+6)}$. For $n=-1$, for instance, we have $L_{\mathrm{X}}^{*} \propto(1+z)^{-1 / 2}$, so in this case $L^{*}$ varies quite slowly. $L^{*}$ would be constant if $n=5 / 7$.

By (4.1) we have
$n_{z}\left(L_{\mathrm{X}}\right) d L_{\mathrm{X}}=(1+z)^{(7-7 n) /(2 n+6)} n_{0}\left[(1+z)^{-(7 n+5) /(2 n+6)} L_{\mathrm{X}}\right]$.
The comoving emissivity is
$\varepsilon_{z} \propto \int d L_{\mathrm{X}} L_{\mathrm{X}} n_{z}\left(L_{\mathrm{X}}\right)$
SO
$\varepsilon_{z}=(1+z)^{(7 n+17) /(2 n+6)} \varepsilon_{0}$.
So, for $n=-1$, the total emissivity should have been larger in the past by a factor $(1+z)^{5 / 2}$. In models like 'cold dark matter' or 'cosmic string' then we predict strong positive density evolution (with $L^{*}$ staying roughly constant).

The constraints on evolution from the available data are rather weak. There is some evidence for positive evolution from the $H E A O$ counts (see figs 6, 7 of Kowalski et al. 1984), but this is not reflected in the $\log N-\log S$ relation from the Einstein medium sensitivity survey (Gioia et al. 1984). However, in both cases there are a very limited number of clusters. These complete surveys are limited to redshifts $z \leqslant 0.2$. Henry et al. (1982) have studied a number of clusters at higher redshift. While these data cannot be used to constrain the amplitude of the luminosity function since the selection criteria may vary in an unknown manner with redshift, it is interesting to look at the shape of the luminosity function they found which may be fairly representative of that for a complete survey at high redshift. They fitted various forms to the data and found, for a power-law fit, no change in the slope of the luminosity function, and, for a Schechter function fit, a slight increase in $L^{*}$ with $z$ (though the data also appear to be consistent with constant $L^{*}$ ). The results are therefore in good agreement with the prediction for $n \simeq-5 / 7$, since for $n=-5 / 7$ there would be no evolution of either of these quantities.

The predictions of our model are in sharp discord with those of Perrenod (1980). For instance, for $n=0$ we find positive evolution, so that $L^{*}$ should increase by about 50 per cent by $z=0.6$, whereas Perrenod finds a decrease of $L^{*}$ by about a factor 30 in the same interval. In Perrenod's model, as in that discussed here, the clusters are individually less massive at earlier epochs. In our model this does not lead to a great decrease in X-ray luminosity since this decrease in mass is compensated by an increase in the density so the product $M^{*} \varrho^{*}$ which appears in $L^{*}$ remains roughly constant. In Perrenod's model, the density of the clusters is assumed to be constant; this is inconsistent with the ideas that clusters formed by gravitational instability, in which case the clusters collapsing at any epoch will have a density which is some constant multiple of the background density.

### 4.2.2 Halo X-ray emission

The total X-ray flux comes predominantly from the material within a few core radii of the cluster centre. As we shall discuss in Section 5, gas dissipation may cause the evolution of the core regions
to depart from the idealized scaling laws. The X-ray emission from outside the core region should be less sensitive to this cooling and so should provide a more robust test.

The total X-ray surface brightness $\mu^{*}$ (in the rest frame of the cluster) at radius $r^{*}$ scales as $r^{* 2} \times(1+z)^{11 / 2}$, so the product $\mu^{*} / r^{* 2}$ is proportional to $(1+z)^{11 / 2}$, independent of $n$. This suggests the following test which may be performed given a detector with good spatial resolution. First find all the diffuse sources (i.e. clusters) that show up above a contour of given surface brightness in a chosen patch of the sky. Next, collect redshifts for these clusters by optical means. Finally, for each cluster calculate the comoving area $A=\pi r^{2}$ enclosed within the contour $\mu=\varkappa \times(1+z)^{11 / 2} r^{2}$, where the constant $\varkappa$ is chosen to place the contour at a convenient radius which is comfortably outside the inner core regions, but not so far that the surface brightness is immeasurably small. In the hierarchical models considered here, $A$ will scale simply as $r^{* 2}$, so we find, for instance

$$
\begin{equation*}
n_{z}(A)=(1+z)^{10 /(n+3)} n_{0}\left[(1+z)^{4 /(n+3)} A\right] \tag{4.6}
\end{equation*}
$$

by equation (4.1), and
$N_{z}(A)=(1+z)^{6 /(n+3)} N_{0}\left[(1+z)^{4 /(n+3)} A\right]$.
An analogous test can be performed with a low-energy detector. In this case the observed surface brightness at radius $r^{*}$ scales as
$\mu_{\mathrm{LED}}^{*} \propto(1+z)$.
So if we calculate the comoving area $A^{\prime}$ of the contour where $\mu_{\mathrm{LED}}=\varkappa \times(1+z)$, we should find evolution identical in form to equations (4.6) and (4.7), from which we can determine the spectral index $n$.

We note that there should be no fundamental obstacle to detecting the rather numerous clusters predicted at $z \geqslant 0$ in these models, given a detector with sufficient resolution, since the contour whose radius we would like to determine in order to apply this test has surface brightness (as seen by the observer) which increases with redshift. The depth of such a survey will be limited by the angular resolution of the telescope.

### 4.3 OPTICAL PROPERTIES

### 4.3.1 Space density of Abell clusters

We now turn to the evolution of the optical properties of clusters. First, we predict the evolution of the comoving number density of clusters selected according to Abell's criterion. We are faced with the problem here that Abell's primary parameter $N_{\mathrm{A}}$ does not satisfy the scaling requirement discussed in Section 3, essentially because it makes reference to the physical scale $R_{\mathrm{A}}$. This means that for the two clusters $C$ and $C^{\prime}$ which are exact scaled replicas of each other, the parameter $N_{\mathrm{A}}$ will not necessarily scale in any sensible manner.

If we assume that the mass to light ratio of clusters does not vary appreciably with redshift, then $N_{\mathrm{A}}$ is essentially a measure of mass contained within a surface of given density. What we would prefer, for the present purpose, is a measure of $M_{\Delta}$, the mass within a surface of given density contrast. Then we would simply have
$n_{z}\left(M_{\Delta}\right)=(1+z)^{12 /(n+3)} n_{0}\left[(1+z)^{6 /(n+3)} M_{\Delta}\right]$.
We can relate the quantity $N_{\mathrm{A}}$ to $M_{\Delta}$ if we make the simplifying, but by no means rigorously justified assumption that individual clusters have a radially averged profile which varies like the
ensemble average, i.e. proportional to $\xi_{\mathrm{cg}} \sim r^{-2}$. This was the assumption we made in Section 2, where we determined $N\left(M_{\Delta}\right)$. At constant $\Delta$ we have $N_{\mathrm{A}}^{3} \propto M_{\Delta}^{2}(1+z)^{3}$, from which we infer that selecting clusters with constant richness is equivalent to selecting clusters according to $M_{\Delta} \propto(1+z)^{-3 / 2}$. Combining this with the scaling for $N\left(M_{\Delta}\right)$, we have
$N_{z}($ constant richness $) \propto(1+z)^{[6 /(n+3)-1.8(3-3 n) /(2 n+6)]}$
which, e.g. for $n=-1$ gives $N_{z} \propto(1+z)^{0.3 \pm 0.3}$. Abell plots the comoving number density of his clusters versus redshift (for $z \leqslant 0.2$ ) and finds consistency with $N_{z}=$ constant, and therefore with our ( $n=-1$ ) prediction.

It is interesting to note that, while the mass at fixed density contrast is varying very rapidly in these models, as is the number density of clusters selected according to density contrast, the number density of clusters selected by Abell should remain very nearly constant. Since Abell's $N_{\mathrm{A}}$ is essentially the mass within a fixed radius this should be proportional to $T$, provided $T$ is roughly constant with radius within each cluster. Thus, we find that the number density of clusters of given $T$ should be roughly constant in time. This result is relevant to the claim (Frenk et al. 1985) that, in the cold-dark-matter picture, galaxies form at redshift $z \sim 3$. This is based on the result that, at that epoch in the simulations, the comoving number density of condensations with virial temperature like that of galaxies was similar to the actual comoving number density of galaxies. The result we have obtained here shows that, if we applied the same method to 'date' the epoch of cluster formation, we would have a completely indeterminate result, and in a simulation, one would find the correct number density of objects with virial temperatures like those of clusters at essentially any epoch. What is happening here is that, if the spectral index is indeed close to -1 , individual clusters must be growing by accreting mass, but at a rate such that the virial temperature is barely changing with time. Since, in this picture, galaxies are quite analogous to clusters, in the sense that they apparently contain a similar fraction of the total mass, one must regard such estimates of the epoch of galaxy formation with great caution.

Returning to the matter in hand, the constant comoving number density of Abell clusters would disfavour models with spectral index $n$ significantly greater than or less than $n=-1$. However, the problem of foreground and background contamination may give a misleading impression here. Hopefully, with future cluster catalogues obtained by plate-scanning, this effect can be controlled.

### 4.3.2 Core radii

As we mentioned in Section 2, there is some evidence that core radii of clusters at the present epoch are uncorrelated with richness. If this is the case then we can make a clear prediction that core radii should vary with redshift as $R_{\mathrm{c}} \propto(1+z)^{5+n / 3+n}$. Or, for our fiducial $n=-1$ models, $R_{\mathrm{c}} \propto(1+z)^{-2}$. This $z$ dependence is much stronger than any sensible evolution due to the global cosmology. For instance, if we were to observe such a variation and attribute is to the effect of the deceleration parameter we would find $q_{0}=-4$. Bruzual \& Spinrad (1978) have fitted models to the galaxy counts for some very distant clusters and found them to be essentially independent of redshift. If this is really the case then we can immediately exclude essentially the whole class of hypotheses under consideration here. However, it is hard to know how seriously to take this discrepancy. Foreground/background contamination must be a serious problem for these distant clusters, and the systematic discrepancies between the core radii obtained by different workers even for clusters at low redshifts suggests that we take this result with a pinch of salt. None the less, the result is very interesting and provides the first suggestion that the models we are discussing may not describe reality. Since the prediction of the model is transparently simple, this aspect certainly warrants further investigation.

### 4.4 GRAVITATIONAL LENSING

Several instances of multiple images of distant QSOs are known and, if the common interpretation that these are lensed by intervening mass concentrations is correct, they provide an important probe of the mass distribution at moderate redshifts, $z \sim 1$.

A recent analysis (Turner, Ostriker \& Gott 1984) has shown that, while the small-angle splittings could be produced by intervening galaxies, the larger splittings require a more extended mass distribution. Turner et al. proposed that these cases result from the combination of a bright massive galaxy in a rich cluster core, the surface density of the core $\Sigma_{\mathrm{c}}$ being about one half of the critical value $\Sigma_{\text {crit }}$ required for the cluster to produce multiple images in its own right.

In the model of Turner et al. about 10 per cent of all the elliptical galaxies must inhabit cluster cores with core radii $\simeq 200 h^{-1} \mathrm{kpc}$ and velocity dispersion $\sigma_{0} \simeq 1200 \mathrm{~km} \mathrm{~s}^{-1}$. While this core radius is quite reasonable the required velocity dispersion is rather high, and the fraction of E galaxies which inhabit such cores is certainly much less than 10 per cent. Thus, the model proposed seems to be somewhat at odds with the observations if the clusters at $z \simeq 1$, which are responsible for the lensing, have characteristics like their present-day counterparts.

In the models under consideration here there is considerable evolution of cluster properties between $z \simeq 1$ and the present. The characteristic surface density varies as
$\Sigma^{*} \propto(1+z)^{(2 n+4) /(n+3)}$.
So, for $n=-1, \Sigma^{*} \propto(1+z)$. If we let $f_{z}(\Sigma)$ denote the fraction of material residing in regions of surface density $\geqslant \Sigma$ at redshift $z$, we have
$f_{z}(\Sigma)=f_{0}\left(\frac{\Sigma^{*}(0)}{\Sigma^{*}(z)} \Sigma\right)$.
So, for $n \simeq-1$, the cross-section for the galaxy/cluster lensing combinations proposed by Turner et al. would be greater than that provided by an unevolving population of clusters.

The presently available data do not permit a very accurate determination of $f_{0}(\Sigma)$, since this requires at least a knowledge of the bivariant distribution in binding energy and core radius for these clusters. We can get a crude estimate of the strength of this evolutionary effect if we assume that $T$ and $r_{\mathrm{c}}$ are uncorrelated (as is certainly allowed, and may well be required). If this is the case we must have
$f_{0}(\Sigma) \propto M_{\mathrm{c}}[T(\Sigma)] \times N[T(\Sigma)] \simeq T^{-1.7 \pm 0.2}$
with $T(\Sigma) \propto \Sigma$.
Thus, with $n=-1$, the fraction of galaxies resident in cores with $\Sigma \geqslant 0.5 \times \Sigma_{\text {crit }}$ would be about three to four times greater than the corresponding fraction for an unevolving population of clusters, and if the mass function of clusters steepens for the richest clusters as the optical data suggest, the gain would be even larger.
This assumes that the number of galaxies per unit mass in these high $z$ clusters is the same as those at the present epoch. While this remains to be tested empirically, we note that in some 'biasing' schemes (in which galaxy formation was more efficient in protoclusters) one would expect there to be more galaxies per unit mass in clusters at earlier epochs and this would further enhance the effect. Another point is that, if cluster cores are really involved in lensing events, we should be able to detect X-ray emission from the hot gas.
To summarize then, while the observations of present epoch clusters do not at present allow a precise prediction of the frequency of lensing events, we can at least assert that evolution of the clusters between the present and $z \sim 1$ is an important factor in some hierarchical models and the
evolutionary corrections to the predictions based on an unevolving population of clusters may go at least some way to resolving the problem of high frequency of large-angle splittings observed.

## 5 Processes which disrupt the self-similar scaling

We have argued in Section 3 that our simple two-component model (collisionless matter+adiabatic gas) should provide a reasonable first approximation to the process of cluster formation. However, there are processes, either operating at present, or which will have been important in the past, which will introduce physical scales into the problem (other than the mass scale of non-linearity) and will disrupt the perfect self-similarity of our model. We will now consider some of these processes. We will find constraints on the range of mass scales and epochs over which we can expect the various scaling laws to provide a valid description. We will discover which of these results will be least sensitive to self-similarity breaking processes, and which therefore provide the most robust tests of the hierarchical hypothesis.

## 5.1 gas cooling

The cooling time for a typical element of gas in a rich cluster is much longer than the Hubble time. In the core regions of the most massive clusters these time-scales become comparable. While our adiabatic model may still provide a reasonable approximation for the initial assembly of the cluster, after a Hubble time the gas in the centre will lose energy and will start to sink to the centre. Evidence for this is provided by the observations of cooling flows (Fabian et al. 1984). Mass rates calculated for these flows suggest that up to $\sim 10^{12} M_{\odot}$ of gas has cooled out of the cores of these clusters. This is only a small fraction of the total core gas mass, so we conclude that overall core properties such as the total X-ray luminosity will only have been marginally affected. However, as we go back to earlier epochs the ratio of cooling time to Hubble time decreases quite rapidly and we should see the cooling flows becoming more important and influencing the core gas properties to a greater extent. To some extent this nuisance factor can be avoided if we restrict attention to the less massive clusters for which cooling appears to be less important today, but eventually, and at not too large a redshift, our predictions for the total luminosity must become unreliable.
Even when a considerable fraction of the core gas has participated in cooling this will not seriously affect the dynamics of merging except in so far as a small fraction of the intracluster gas will have been converted to dark matter. The predictions for the halo properties of the cluster should be very insensitive to these changes taking place in the core region, so this should provide a robust test of the theory. Similarly, global properties of the mass distribution should still be reliably predicted by the model.

### 5.2 STRIPPING

Stripping of gas is probably relatively unimportant for the bulk of the gas in clusters at the present epoch, simply because the amount of gas that can reasonably be associated with the infalling galaxies is small compared to the mass of intracluster gas. However, the high metal content of this gas, or at least that gas near the core which is luminous in X-rays, suggests that this material has at some time in the past been processed through galaxies. If galaxies formed at quite an early time, when the scale of clustering was around $M_{\text {gal }} \sim 10^{12} M_{\odot}$, as is usually assumed to be the case in the hierarchical picture, any particular galaxy will have passed through several levels of the hierarchy before finding itself in a present-day cluster, and enriched gas may have been returned to the intracluster gas during these earlier phases. In any case, it is hard to imagine that stripping would
have had a drastic effect on the X-ray luminosities since the infalling galaxies would deliver the gas with a total energy quite similar to that gained by infalling diffuse gas. This means that the final density, and X-ray luminosity, of the stripped gas would be very similar to that of the gas it displaces.

### 5.3 WINDS AND OTHER ENERGY INPUT FROM GALAXIES

We have assumed that the energy of the gas was obtained by shocking as it fell in, and the present state of the gas seems to be compatible with this hypothesis. There cannot have been much larger energy input from the galaxies since this would have driven the intracluster gas out. It would seem rather coincidental, and extravagant, if the energy input of the gas was comparable to its present thermal energy, so we conclude that again, such energy input is in all likelihood only a small perturbative effect today.

In the models considered here though, the virial temperature of the cluster varies as $T \propto(1+z)^{-1}$, for the case $n=-1$, so the effect of winds may have been more important in the past. However, since the evolution of the virial temperature of the clusters is rather slow, any predictions for clusters at observable redshifts should be essentially unaffected.

### 5.4 DISCRETENESS OF GALAXIES

We have envisaged that the bulk of the collisionless component is composed of particles that are microscopic compared to the mass of the cluster so any dynamical evolution is negligible. The galaxies, however, may have suffered some dynamical friction near the core of the cluster. It is somewhat uncertain as to how strong this effect should be since we do not know how much mass is attached to individual galaxies. We can, however, be reasonably confident that, as far as the halo of the cluster is concerned, such effects are not very important.

### 5.5 CURVATURE OF THE SPECTRUM

We argued in Section 3 that, provided that the curvature of the spectrum is small, as in the 'cold dark matter' spectrum, it should be a good approximation to take the initial spectrum to be scale-free with some appropriate $n_{\text {eff }}$. The proviso here was that one restricts attention to phenomena spanning only a small mass range. Present-day global properties of clusters certainly satisfy this criterion. Less certain is whether the core properties are sensitive to the high-frequency part of the spectrum (for instance, the phase-space density in the core of the cluster must be limited by the phase-space density of the infalling sub-clumps and this will depend on the initial fluctuations on scales considerably less than that of the final cluster as a whole). The safest tests in this regard should again be those which depend on global properties.

As we look back in earlier epochs we may find (if $n$ is close to -3 ) that even the global properties of clusters change considerably. If this turns out to be the case then we may become sensitive to, and hence be able to measure, the curvature, if any, of the initial spectrum. An obvious higher order approximation would be to patch together self-similar solutions at different epochs with continuously varying $n_{\text {eff }}$. We will not develop that theme in the present paper.
A more extreme form of curvature arises in the 'pancake' picture, in which the initial spectrum is more drastically affected prior to recombination, with all of the power above some cut-off frequency being damped. This gives rise to a sequence of cosmogony which is radically different from the hierarchical picture. The distinction between the final structures forming in these different scenarios has been discussed extensively elsewhere. Here we merely wish to note that the evolution of clusters would be easily distinguishable too. In the pancake picture we would
expect individual clusters to be gaining mass much as in the hierarchical picture, so at earlier epochs one would expect the clusters to be less massive. The distinguishing feature is that, while, in the hierarchical scheme, these clusters would be much more numerous in the past, the opposite would hold in the pancake picture. A possible exception to this behaviour would arise if the damping cut-off were on a very small scale and these mini-pancakes formed quite early. In this case, at later times, these pancakes would merge together and the more massive clusters produced would evolve much as in the way we have discussed above.

## $5.6 \Omega \ll 1.0$

If the density parameter $\Omega$ is significantly less than unity then a physical scale is introduced into the background Universe; the time when $\Omega$ started to depart from unity, and one would not obtain a self-similar final state even if the initial spectrum of fluctuations were perfectly self-similar. Generally speaking, for the same initial spectrum, one would expect the evolution to slow down as the Universe becomes open, since the long-wavelength linear modes grow more slowly, so the mass scale of non-linearity will not grow as fast. If $\Omega$ has become very small then we should notice that those clusters which did manage to collapse would have become well separated from the Universal expansion and should have steeper profiles than would otherwise be the case. However, since these profiles also depend on the slope of the initial spectrum, it would be difficult to disentangle these effects. In order to make more quantitative statements it will probably be necessary to perform numerical calculations.

## 6 Clustering and evolution

It is well established that the length scale of Abell clusters is much larger than that of galaxies (Hauser \& Peebles 1973; Bahcall \& Soneira 1983; Klypin \& Kopylov 1983), and is also much larger than the scale of clustering of the mass distribution. The author has argued (Kaiser 1984) that this situation would arise quite naturally in a hierarchical picture with Gaussian initial fluctuations if the power spectrum had more power on large scales than a 'white-noise' spectrum (i.e. if the spectrum had negative spectral index $n$ ). The important steps in this argument are:
(i) The sites of formation of rare massive condensations (like clusters) of mass $M$ should be identified with the high peaks above some threshold $v \sigma(M)$ of the initial density perturbation field when this has been filtered with a window containing mass $M$. Here $\sigma(M)$ is the rms of the fluctuations of the smoothed density field, and $v$ is a dimensionless parameter which must be adjusted to give the correct space number density of clusters.
(ii) The initial field can be decomposed into a 'noise' component, containing the high spatial frequency modes which determine the nature of the clusters which form, and a 'signal' component containing only low-frequency modes with wavelength much greater than the scale of the clusters.
(iii) The final step is to argue that, in the absence of the signal, the clusters which form would, like the noise component from which they derive, have negligible correlations on large scales. The effect of adding the 'signal' to reconstitute the entire initial spectrum is to modulate the space density of the clusters which form. This gives, at large scales, a cluster correlation function with the same form as that of the underlying noise, but with an amplitude which depends on the threshold parameter $v$.

This analysis (discussed in more detail by Bardeen et al. 1985) provides a simple connection between the initial spectrum and the final clustering pattern. Now the initial spectrum of fluctuations also determines the mass spectrum $n(M)$ of the clusters in the final state and all other properties of these clusters. However, it is rather difficult to determine how these are related.

What we shall show in the present section is that, using the same idea of signal/noise decomposition, and the assumption that the initial fluctuations are approximately scaleinvariant, one can obtain an expression for the linear relationship between the cluster correlation function $\xi_{\mathrm{c}}$ and that of the matter $\xi_{\varrho}$ in terms of the final mass spectrum $n(M)$. This relation is useful for two purposes: first, the quantities which appear in this equation can readily be calculated in numerical simulations for which enhancement of the clustering is seen. This provides a test of the reliability of the signal/noise decomposition approach. Secondly, we can apply this to the real world, and for an assumed initial spectral index $n$, infer the amplitude of $\xi_{\varrho}$ on large scales. Combining this with independent measures of matter fluctuations on smaller scales, we can limit the range of viable values of $n$.

### 6.1 ASSUMPTIONS AND APPROXIMATIONS

We will assume that the initial random density fluctuation field $F_{\mathrm{i}}(\mathbf{r})=\Delta\left(\mathbf{r}, t_{\mathrm{i}}\right)$ can be decomposed into two statistically independent fields $F_{\mathrm{i}}(\mathbf{r})=S_{\mathrm{i}}(\mathbf{r})+N_{\mathrm{i}}(\mathbf{r})$, where $S$ determines the coherence of the field $F_{\mathrm{i}}$ on scales much greater than the scale of individual clusters. For Gaussian fields this decomposition is always possible since such a field has Fourier components which are statistically independent. For non-Gaussian fields this decomposition is not always strictly allowed. We shall restrict attention to fields for which this decomposition is allowed.

Let us imagine that we evolve this field into the non-linear regime and identify clusters by some density contrast algorithm such that, like real clusters, these objects are recently virialized, i.e. they have a density contrast $\simeq 170$. For each of these clusters we can calculate the mean initial density contrast
$\Delta_{\mathrm{i}}=\frac{\int d M F_{\mathrm{i}}}{\int d M}$
where the integrals are over those mass elements which end up in the cluster. We know that $\Delta_{\mathrm{i}} \sim 1 /\left(1+z_{\mathrm{i}}\right)$, because the clusters have only recently collapsed. We postulate that, at least for clusters of similar mass, the scatter of the values of $\Delta_{\mathrm{i}}$ is relatively small and so we can write $\Delta_{\mathrm{i}}=\varkappa /(1+z)$, where $\varkappa$ is a constant of order unity. In principle, this assumption could be tested by suitable numerical simulations for specific choices of initial fluctuations, and the constant $\varkappa$ determined. In the absence of suitable simulations from which to determine $x$, we simply assume that it is equal to the value appropriate for spherical collapse: $x \simeq 1.7$. As discussed by Kaiser \& Davis (1985) one expects this spherical approximation to hold best for the most massive clusters.

### 6.2 MODULATION OF CLUSTER FORMATION BY LONG-WAVELENGTH FLUCTUATIONS

We now ask: what is the effect on the population of clusters which grow from the initial field $N(\mathbf{r})$ when we add in the small-amplitude long-wavelength fluctuations $S(\mathbf{r})$ ? Neglecting the tidal influence of $S$, since the amplitude is small, the only effect locally (around the position $r$ ) is the same as if we add a small constant perturbation $\varepsilon=S(\mathbf{r})$.

The effect of this spatially constant perturbation is simply to hasten the collapse of a cluster so that it reaches the same density contrast as in the unperturbed Universe at a slightly earlier epoch: $1+z^{\prime}=1+\varepsilon / \varkappa$. Thus the clustering pattern in the perturbed region will, at least in terms of the density contrast, appear essentially identical to that in the unperturbed region, but at a slightly
earlier time, assuming that $\varepsilon$ is positive. For a negative perturbation, the growth of clustering will be retarded somewhat.

With our assumed scale-free initial conditions though, this small spatially varying time shift introduced by $S(\mathbf{r})$ is just equivalent to a modulation of the characteristic mass scale $M^{*}$ from place to place by an amount $\Delta M^{*}=M^{*} \times 6 S / \varkappa(n+3)$. It is then quite straightforward to show that this leads to a modulation of the number density of clusters of fractional amplitude
$\frac{\Delta N}{N}(\mathbf{r})=\left[1-\frac{6}{\varkappa(n+3)} \frac{d \ln f}{d \ln M_{\Delta}}\right] \times S(\mathbf{r})$
where
$f(M)=\int_{M}^{\infty} d M_{\Delta} M_{\Delta} n\left(M_{\Delta}\right)$.
The first term in the square brackets is the change in space density of the clusters that would occur due to $S(\mathbf{r})$ if there were no modulation of $M^{*}$, while the second term arises entirely from this modulation.

Since, by equation (6.1), the perturbation in the number density is just a linear multiple of $S$, and the correlation function of the density field is just equal to the correlation function of $S$, the ratio of the correlation function of the clusters to that of the underlying density field is given by the square of the term in square brackets
$\xi_{\mathrm{c}}=\left[1-\frac{6}{\varkappa(n+3)} \frac{d \ln f}{d \ln M_{\Delta}}\right]^{2} \times \xi_{\varrho}$.
This equation can be tested against the numerical experiments of Barnes et al. (1985) who applied cluster-finding algorithms to a variety of simulations and estimated the enhancement of clustering strength. One of their simulations (case B) meets the requirements we have described (scale-free initial conditions etc.) and they also give some information from which one can estimate $d \ln f / d \ln M$. This quantity can only be estimated rather crudely but it does seem to become increasingly negative with increasing $M_{\Delta}$, and equation (6.2) gives $\xi_{c}$ which increases from $\simeq 5 \xi_{\varrho}$ to about $17 \xi_{\varrho}$ for the most massive clusters. For the lower mass clusters the enhancement of $\xi_{\mathrm{c}}$ is similar to that predicted but the more massive clusters do not show as large an enhancement as we predict. It is possible that this discrepancy is due to a poor estimate of $f(M)$ from the limited data given. A second possibility is that the simulation is not evolving realistically (Barnes et al. note that the evolution does not seem to be very self-similar, possibly due to the late start time $z=0.6$ used). A further source of concern is the correlation function of the matter which does not show the $\xi \propto r^{-1}$ tail at large $r$ which, according to linear theory, must result from a spectral index with $n=-2$. In fact, $\xi_{\rho}$ falls much more rapidly. Such a departure of $\xi_{\rho}$ from the ensemble average may be quite reasonable, given the limited volume of the simulation, but if that is the case, little confidence can be attached to the estimates of $\xi_{\mathrm{c}}$ at large scales either. A third possibility is that equation (6.2) may give seriously misleading results. It would be nice to have simulations for which one could eliminate the first two possibilities.
We now proceed to apply equation (6.2) to real clusters: from our simple power-law approximation to $N(M)$ (Section 2) we estimate $d \ln f / d \ln M=-0.8 \pm 0.2$. In Table 1, we have estimated the 'amplification factor' $A=\xi_{\mathrm{c}} / \xi_{Q}$ for a variety of values of $n$. The value of $n$ which is relevant here is the effective spectral index around the scale of the clusters, not of the large-scale clustering. We do not need to assume that the power-law initial spectrum extends accurately to large scales: if the spectrum on the large scale is greater than or smaller than the extrapolation of the power law, then the correlation function will be correspondingly larger or smaller.

Table 1. Correlation function enhancement factor.

| $n$ | $A$ | $r_{0, \varrho} / h^{-1} \mathrm{Mpc}$ |
| :--- | ---: | :---: |
| -1.0 | $5.8 \pm 1.7$ | $10.0 \pm 1.5$ |
| -1.5 | $8.4 \pm 2.7$ | $6.1 \pm 1.3$ |
| -2.0 | $14.6 \pm 5.9$ | $1.7 \pm 0.6$ |

From the values of $A$ in Table 1 , one can infer, for any assumed value of $n$, the amplitude of $\xi_{Q}$ at separation $r \simeq 25 h^{-1} \mathrm{Mpc}$ where $\xi_{\mathrm{c}}$ is observed to be around unity. The formulae are not valid for smaller separations because, on those scales, the linearity assumption has broken down, while on larger scales the estimates of $\xi$ rapidly cease to be significant. It is interesting to compare these amplitudes with the upper limits to the amplitude of the galaxy correlation function on these scales. It is not easy to quantify the errors in the estimates of $\xi_{\mathrm{g}}$, but it is probably safe to assert that $\xi_{\mathrm{g}}$ is not greater than about 0.2 at these separations. If we assume that $\xi_{\varrho}$ is roughly equal to $\xi_{\mathrm{g}}$ on these scales then any of these values of $n$ would be allowed.

If we require that the spectral index remain roughly constant on scales $\sim 5-25 h^{-1} \mathrm{Mpc}$, we can calculate the radius $r_{0, \varrho}$ for which the matter correlation has unit amplitude. These values are given in column 3 of Table 1 . If we require that this agree with the radius at which the galaxies have unit correlation, this singles out a value of $n$ : we find we need $n \simeq-1.5$. This is somewhat more negative than the spectral index for the cold-dark-matter spectrum in this range of scales and suggests that we need a spectrum with rather more power at large scales. If we require that the correlation length for the matter be about one half that of the galaxies, as in the simplest 'biasing' picture (see e.g. Bardeen et al. 1985), then we are forced to a spectrum with an even more negative value for $n$, say $n \simeq-1.8$.

An important assumption used in deriving equation (6.2) was that the small-scale and large-scale fluctuations $N$ and $S$ be statistically independent. This is always valid for Gaussian fluctuations, but for some non-Gaussian processes this may be violated. For example, in the 'cosmic string' picture the clusters are thought to form from the 'daughter loops' which form, approximately 10 at a time, from larger single loops which enter the horizon and self-intersect several times. In this case the clustering we have described above will be augmented by correlations between daughter loops due to their common ancestry. As discussed by Turok (1985) the level of clustering produced by this self-intersection seems quite able to account for the observed properties of clusters.

We can also estimate the enhancement of the clustering strength of clusters selected according to their X-ray properties. As before, we must calculate how the properties of the clusters change under the influence of the long-wavelength 'swell' $S(\mathbf{r})$. The characteristic mass varies as $\Delta M^{*}=6 M^{*} S / \varkappa(n+3)$. The characteristic temperature varies as $T \propto M^{2 / 3}$, so we have
$\Delta T^{*} / T^{*}=(2 / 3) \Delta M^{*} / M^{*}$.
Our approximation is that, aside from this small shift of physical scale, the form of the clustering is unchanged. This means that the core density of the characteristic $M^{*}$ cluster is unchanged, and the characteristic total X-ray luminosity $L_{\mathrm{X}}^{*} \propto M^{*}\left(T^{*}\right)^{1 / 2}$ will change according to
$\Delta L_{\mathrm{X}}^{*} / L_{\mathrm{X}}^{*}=(4 / 3) \Delta M^{*} / M^{*}$.
The comoving number density of these clusters satisfies $N^{*} \propto\left(M^{*}\right)^{-1}$. Now, recalling from Section 2 that
$N\left(L_{\mathrm{X}}\right) \simeq N^{*}\left(L_{\mathrm{X}} / L_{\mathrm{X}}^{*}\right)^{-1.15}$
we can find the fluctuation induced in the comoving number density of clusters brighter than some threshold $L$ :
$(\Delta N / N)_{L} \simeq\left(\Delta N^{*} / N^{*}\right)+1.15\left(\Delta L^{*} / L^{*}\right)=[-1+(4 / 3) \times 1.15] \Delta M^{*} / M^{*}$.
To get the actual fluctuation in number density we must add to this the underlying density fluctuation $S(\mathbf{r})$. This gives us, analogous to equation (6.2),
$\xi_{\mathrm{c}} \simeq\left[1+\frac{6}{\chi(n+3)} \times(0.5)\right]^{2} \xi_{\rho}$.
This results in an enhancement which is somewhat smaller than that obtained for the optically selected clusters. For instance, with $n=-1.5$, equation (6.2) gives $\xi_{\mathrm{c}} \simeq 8 \xi_{Q}$ whereas these X-ray selected clusters have $\xi_{\mathrm{c}} \simeq 5 \xi_{Q}$. This result is, however, quite sensitive to the assumed slope of the X-ray luminosity function. An analogous calculation can be performed for clusters selected according to their flux as measured by a low-energy detector. The result is a still smaller enhancement for $\xi$, essentially because one of the effects of the perturbation $S$ is to increase the characteristic temperature of the clusters, and this gives a greater boost to the total flux than to the low-energy flux. Again, however, the result is quite sensitive to the assumed slope of the luminosity function. While neither the clustering properties of X-ray selected clusters nor their luminosity function are very well determined at present, it will be interesting as more extensive surveys become available to see how well, if at all, these formulae are obeyed.

The basic result we have tried to get across in this section is that there is an intimate relationship between the rapid evolution of clusters in the models under consideration and their clustering properties. In the Universe as a whole there is a rapid increase in the characteristic mass of clusters, and the effect of small-amplitude density fluctuations on scales larger than that of individual clusters is to advance or retard this process slightly, depending on the sign of the density fluctuations. When this modulation is combined with the steep slope of the luminosity function for the most massive clusters, this can result in very strong modulation of the number density of clusters selected according to mass or some other selection criterion.

## 7 Conclusions

Comparison of the observed velocity dispersions and profiles of rich clusters with theoretical and numerical models suggests that these objects have virialized very recently, and that beyond about $1.5-2 h^{-1} \mathrm{Mpc}$ the material is still falling in. Only a small fraction of the matter in the Universe resides in the virialized regions of these clusters, and the number density of these objects is a steeply falling function of virial temperature [ $N(T) \propto T^{-\gamma}$ with $\gamma \simeq 2.7$ ], so we are dealing with a well-defined upper end to the distribution of bound objects in the Universe. The X-ray luminosity is determined by the density reached in the core, at a radius typically 5-10 times smaller than the virial radius, so, not surprisingly, there is considerable scatter in the correlation between X-ray luminosity and richness. It is reassuring, however, that the $N(T)$ inferred from X-ray observations is in good accord with that derived from optical studies, since it is the potential well of the cluster that links the gas and the galaxies.
In hierarchical theories such as the 'cold dark matter' or 'cosmic string' pictures, the density fluctuations which give rise to rich clusters are approximately scale-free and, to the extent that dissipative processes can be neglected, result in a rapid self-similar evolution of the cluster population.

The formulae for the time evolution of the dimensional quantities that characterize the clustering are well known. Provided that no additional dimensional quantities are introduced in
the selection and classification of clusters to form a general cluster catalogue, the time evolution of the distribution function of such a catalogue can also be written down.

X-ray selected clusters are predicted to display dramatic positive density evolution, but with $L^{*}$ remaining roughly constant. Failure to observe the very numerous clusters predicted here would be a severe embarrassment for these theories. The comoving number density of clusters selected according to a criterion like Abell's should, on the other hand, remain roughly constant in time. Core radii should have been much smaller in the past, and the gravitational lensing power of these cluster cores should be substantially greater than for an unevolving population of clusters.

Various physical processes, and also any curvature of the initial spectrum of fluctuations, will act to disrupt the self-similar scaling, particularly as regards the core properties of the clusters. While this may complicate the evolution of the total X-ray luminosities, one can check for this by observing the halo properties which should be quite insensitive to dissipation and other disruptive effects.

A corollary of the rapid evolution of the cluster population in these models is an enhancement of the clustering strength of clusters relative to that of the matter. At large separations one finds a simple linear amplification of $\xi$ which depends on the present logarithmic slope of the mass function $N(M)$, and on the spectral index of the linear density fluctuations from which the clusters form. Estimates of $N(M)$, combined with estimates of the clustering strengths of galaxies and of rich clusters, seem to require a spectral index $n \simeq-1.5$ or even more negative, whereas, for reasonable values of $h$, the 'cold dark matter' spectrum gives $n \simeq-1.0$, so there is apparently a conflict here. The clustering of X-ray selected clusters can also be predicted in a similar fashion, but the result is rather sensitive to uncertainty in the luminosity function. Available estimates of $N\left(L_{\mathrm{X}}\right)$ suggest a weaker clustering amplitude than we predict for optically selected clusters.

## Acknowledgments

The author acknowledges the support of the SERC and the hospitality of the Institute of Astronomy at Cambridge, where this work was performed.

## References

Abell, G. O., 1958. Astrophys. J. Suppl., 3, 211.
Abramopoulos, F. \& Ku, W. H. M., 1983. Astrophys. J., 271, 446.
Bahcall, N. A., 1975. Astrophys. J., 198, 249.
Bahcall, N. A., 1977. Ann. Rev. Astr. Astrophys., 15, 505.
Bahcall, N. A. \& Soneira, R. M., 1983. Astrophys. J., 270, 20.
Bardeen, J. M., Bond, J. R., Kaiser, N. \& Szalay, A. S., 1985. Astrophys. J., 304, 15.
Barnes, J., Dekel, A., Efstathiou, G. \& Frenk, C. S., 1985. Astrophys. J., 295, 368.
Beers, T. C. \& Tonry, J. L., 1986. Astrophys. J., 300, 557.
Blumenthal, G. R. \& Primack, J. R., 1984. Preprint.
Bond, J. R. \& Efstathiou, G., 1984. Astrophys. J., 285, L45.
Bruzual, G. \& Spinrad, H., 1978. Astrophys. J., 220, 1.
Dressler, A., 1978. Astrophys. J., 226, 55.
Efstathiou, G., Fall, S. M. \& Hogan, C. J., 1979. Mon. Not. R. astr. Soc., 189, 203.
Efstathiou, G., Davis, M., Frrenk, C. S. \& White, C. D. M., 1985. Astrophys. J. Suppl. Ser., 57, 241.
Fabian, A. C., Nulsen, P. E. J. \& Canizares, C. R., 1984. Nature, 310, 733.
Frenk, C. S., White, S. D. M., Efstathiou, G. \& Davis, M., 1985. Nature, 317, 595.
Gioia, I. M., Maccacaro, T., Schild, R. E., Stocke, J. T., Liebert, J. W., Danziger, I. J., Kunth, D. \& Lub, J., 1984. Astrophys. J., 283, 459.
Hauser, M. G. \& Peebles, P. J. E., 1973. Astrophys. J., 185, 757.
Hendriksen, M. J., 1985. PhD thesis, University of Maryland.
Henry, J., Soltan, A., Briel, U. \& Gunn, J. E., 1982. Astrophys. J., 262, 1.

Hogan, C. J., Kaiser, N. \& Rees, M. J., 1982. Phil. Trans. R. Soc. London, A307, 97.
Jones, C. \& Forman, W., 1984. Astrophys. J., 276, 38.
Kaiser, N., 1984. Astrophys. J., 284, L9.
Kaiser, N. \& Davis, M., 1985. Astrophys. J., 297, 365.
Kent, S. M. \& Gunn, J. E., 1982. Astr. J., 87, 945.
Klypin, A. A. \& Kopylov, A. A., 1983. Sov. Astr. Lett., 9, 41.
Kowalski, M. P., Ulmer, M. P., Cruddace, R. G. \& Wood, K. S., 1984. Astrophys. J. Suppl., 56, 403.
Lucey, J. R., 1983. Mon. Not. R. astr. Soc., 204, 33.
Lynden-Bell, D., 1967. Mon. Not. R. astr. Soc., 136, 101.
Mushotzky, R. F., 1984. Phys. Scripta, T7, 157.
Peebles, P. J. E., 1980. The Large-Scale Structure of the Universe, Princeton University Press.
Peebles, P. J. E., 1983. Astrophys. J., 274, 1.
Peebles, P. J. E., 1984. Astrophys. J., 277, 470.
Perrenod, S. C., 1980. Astrophys. J., 236, 373.
Piccinotti, G., Mushotzky, R. F., Boldt, E. A., Holt, S. S., Marshall, F. E., Serlemitsos, P. J. \& Shafer, R. A., 1982. Astrophys. J., 253, 485.
Sarazin, C. L., 1986. Rev. Mod. Phys., 56, 1.
Seldner, M. \& Peebles, P. J. E., 1977. Astrophys. J., 215, 703.
Turner, E. L., Ostriker, J. P. \& Gott, J. R., 1984. Astrophys. J., 284, 1.
Turok, N., 1985. Preprint.
Turok, N. \& Brandenberger, R. H., 1985. Preprint.
Vilenkin, A., 1985. Phys. Rep., 121, 263.
Zel'dovich, Ya. B., 1965. Adv. Astr. Astrophys., 3, 241.

