PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 131–145

# The large scale structure of the universe: Dynamical and statistical aspects

VARUN SAHNI
Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkind, Pune 411 007, India

**Abstract.** The description of gravitational clustering is essentially statistical but its origin is dynamical. Hence both aspects of clustering: dynamical and statistical, must be understood in order to arrive at a proper appreciation of the subject of gravitational instability and the formation and evolution of the large scale structure of the Universe. Key dynamical aspects of gravitational clustering such as the Zeldovich approximation and its extension – the adhesion model are reviewed. Statistical indicators of clustering such as the correlation function and percolation theory, as applied to the large scale structure of the Universe have also been focussed on.

**Keywords.** Large scale structure; cosmology; correlation function; percolation theory.

**PACS No. 95.30**

## 1. Introduction

Observations of the large scale structure of the universe can be conveniently grouped under three main categories: (i) those that deal with the clustering properties of galaxies (these include statistical indicators such as: correlation functions, counts in cells, fractal analysis etc.); (ii) the peculiar velocities of galaxies after the smooth Hubble flow has been subtracted out; and (iii) the anisotropy of the cosmic microwave background radiation (CMB). (The latter provides a probe of the amplitude and spectrum of density inhomogeneities in the distant past, soon after the cosmological recombination of hydrogen.)

I shall briefly dwell on each of these observational aspects of large scale structure before proceeding to discuss related dynamical and statistical aspects of the formation and evolution of large scale structure in the universe.

### 1.1 *Clustering of galaxies and clusters*

The fact that galaxies are not randomly distributed on the celestial sphere can be easily discerned from projected two-dimensional catalogues. (Over 2 million galaxies have been mapped in 2D surveys, although the redshifts of only a small fraction of these is accurately known at present.) The clustering of galaxies is amply borne out in three dimensional redshift survey's such as the CfA survey and the southern sky redshift survey (SSRS), which, taken together contain over 15,000 galaxies. As figure 1 demonstrates,

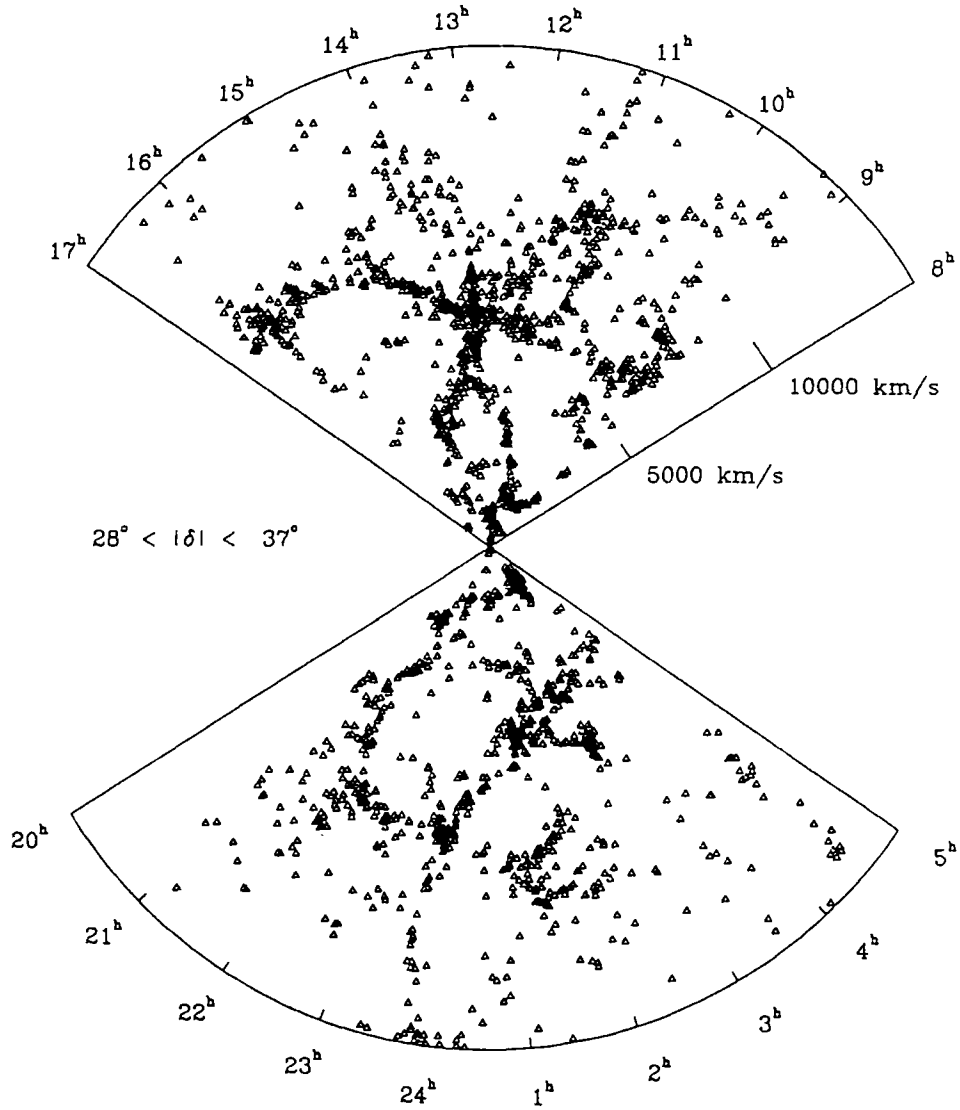


**Figure 1.** The right ascension is plotted against the radial velocity field for galaxies in the SSRS2-CfA2 redshift surveys. The combined surveys contain over 15,000 galaxies of which a fraction are shown in this figure [28]. The northern survey shows the coma cluster and the great wall. Several wall-like features and large voids are also present in the Southern survey.

galaxies appear to lie preferentially along sheets and filaments which are separated from each other by vast empty regions called voids. The typical length scale characterizing cellular structure in the universe is $\sim 30 - 50\,h^{-1}$ Mpc., h being the Hubble parameter in units of 100 km/sec/Mpc.

The simplest indicator of clustering is the two point correlation function $\xi(r)$, which when applied to galaxies estimates the probability (in excess of random), of finding a galaxy at a distance $r$ from a given galaxy. (For a uniform Poisson process $\xi(r) = 0$.) On scales $r < 10\,\mathrm{h}^{-1}$ Mpc the two point correlation function can be accurately described by a power-law [1]

$$\xi_{gg}(r) = \left(\frac{r}{r_{0,g}}\right)^{-1.8} \qquad r_{0,g} \simeq 5\,\mathrm{h}^{-1}\,\mathrm{Mpc}. \tag{1}$$

It is useful to express the density contrast in the universe as a Fourier transform

$$\delta(\mathbf{x}, t) = \frac{\rho - \bar{\rho}}{\bar{\rho}} = \int \mathrm{d}^3k \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \hat{\delta}(\mathbf{k}, t). \tag{2}$$

$\hat{\delta}(\mathbf{k}, t) = |\delta_k(t)| \exp(i\phi_k)$ (For a gaussian random field (GRF) the amplitudes $\delta_k$ have a Rayleigh distribution and the phases $\phi_k$ are independent and uniformly random on the interval $[0, 2\pi]$.) A GRF is completely specified by its power spectrum $P(k) = \langle |\delta_k|^2 \rangle$.

The correlation function is an important measure of clustering since $\xi(r)$ and $P(k)$ – the power spectrum of density fluctuations, form a Fourier transform pair

$$\xi(r) = \int \mathrm{d}^3\mathbf{k} \exp(i\mathbf{k}\cdot\mathbf{r}) P(k), \tag{3}$$

$$P(k) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{x} \exp(-i\mathbf{k}\cdot\mathbf{x}) \xi(x). \tag{4}$$

In standard models of structure formation based on gravitational instability the power spectrum $P(k)$ is related to (i) the spectrum of primordial fluctuations generated during inflation, and (ii) the nature of dark matter (whether 'hot' or 'cold'). The power spectrum in models with 'hot dark matter' shows a sharp cutoff on scales $\leq 40$ Mpc., (related to the free-streaming distance of the 'hot' light dark matter particle) so that super-cluster size objects are the first to form in such scenarios. Galaxies form later due to the fragmentation of primordial super-pancakes. This scenario is commonly referred to as the 'top-down' scenario for galaxy formation, and was first proposed by Zeldovich and co-workers in Moscow during the 1970s and 80s in connection with adiabatic baryonic models of structure formation.

The cold dark matter model (CDM) on the other hand, has power on all scales which is mainly related to the fact that the dark matter particle in this case is assumed to be 'cold' – having negligible free-streaming distance. As a result the first objects to form in this scenario have a low mass, and larger mass objects (galaxies and clusters) form due to the hierachical merging of smaller mass objects, leading to a 'bottom-up' scenario of structure formation.

It is also interesting that clusters of galaxies (which typically have $\geqslant 100$ members), have a correlation function which is similar in form to equation (1), but with a much larger correlation length: $r_{0,c} \simeq 12 - 25\,\mathrm{h}^{-1}$ Mpc. This trend is followed by richer (and rarer) objects which seem to display stronger clustering as shown in figure 2 where the amplitude of the two-point correlation function (scaled to 1 Mpc) between objects belonging to a given richness class (galaxy groups, Abell clusters, X-ray clusters, super-clusters etc.) is shown as a function of their mean separation. The reason for the increase

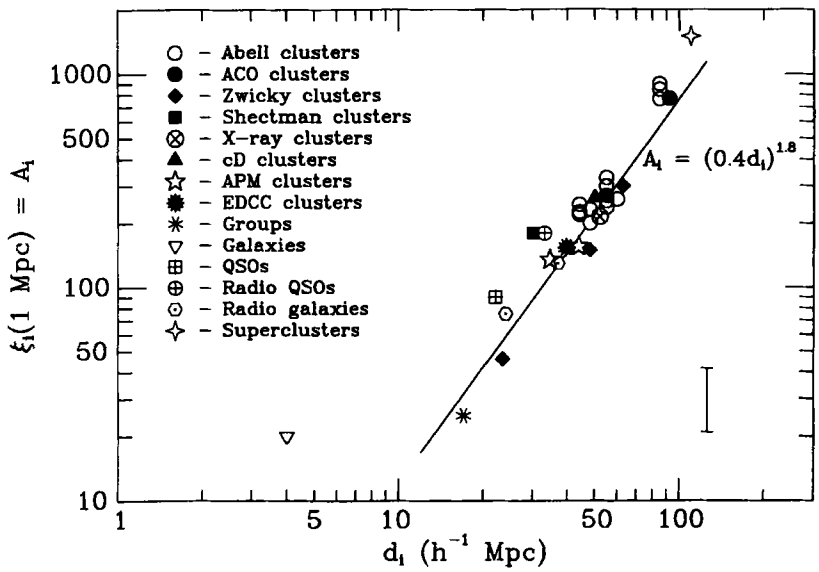


**Figure 2.** The two-point correlation function amplitude (scaled to 1 Mpc) between objects belonging to a given richness class is plotted against their mean separation $d_i$. Reproduced, with permission, from [29].

in clustering amplitude with richness is not well understood at present, although some arguments based on the statistics of Gaussian random fields have been presented [2, 3].

### 1.2 *Velocity fields of galaxies*

The existence of inhomogeneities in the distribution of galaxies results in a peculiar component to their motion over and above the smooth Hubble flow: $\vec{v}_{pec} = \mathbf{V} - H\mathbf{r}$ ($\mathbf{V}$ is the total galactic velocity and $H\mathbf{r}$ is the Hubble flow). Estimates of the peculiar velocity depend upon the sample depth, most surveys indicate $\vec{v}_{pec} \simeq 350\,\mathrm{km/sec}$, on scales $\sim 50\,\mathrm{Mpc}$. Direct evidence for the motion of our galaxy arises from the dipole anisotropy of the cosmic microwave background $\Delta T/T \simeq 1.2 \times 10^{-3}$, which is usually taken to indicate that the milky way is moving at a velocity $\vec{v}_{pec} \simeq 610\,\mathrm{km/sec}$ relative to the MBR. A study of large scale motions of elliptical galaxies made by Lynden–Bell *et al* [4] indicated that a sample of galaxies spanning a volume $\sim (50\,h^{-1})^3$ cubic megaparsec could be participating in a bulk flow directed towards Centaurus. The large amplitude of the flow ($\sim 570 \pm 60\,\mathrm{km/s}$ in the vicinity of the milky way) led some theorists to suggest that a large density fluctuation termed 'the great attractor' (GA) lying in the direction of Centaurus was responsible for this motion. Attempts to 'see' the great attractor, have not yet yielded definite results mainly because the GA region lies in a zone partly obscured by the plane of the milky way. More recent analysis of velocity data obtained from the IRAS satellite indicate that the peculiar velocity of the milky way may be due to the combined gravitational pull of a dozen or so clusters including Virgo, Hydra and Centaurus. If this is indeed the case, then the notion of a GA appears unnecessary.

Since bulk flows in galaxies are directly related to fluctuations in the gravitational potential, considerable effort in the past few years has been paid to reconstruction methods (such as the 'potent' approach) which attempt to reconstruct the density field from a knowledge of peculiar velocities. Such an approach has the advantage that it maps thedensity field due to *all* matter – dark and visible, and is not biased towards luminous matter as are galaxy catalogues. (A detailed discussion of reconstruction methods can be found in [5].)

### 1.3 *Cosmic microwave background anisotropy*

A major astrophysical discovery of recent years has been the detection of large angle fluctuations ( $> 7°$) in the cosmic microwave background (CMB) made by the COBE satellite in 1992. (These fluctuations are thought to have a primordial origin unlike the much larger dipole fluctuation discussed earlier whose presence is attributed to the peculiar motion of our galaxy.) CMB fluctuations are related to fluctuations in matter at the time of the cosmological recombination of hydrogen (which also marks the decoupling of matter and radiation). The small amplitude of these fluctuations ($\Delta T/T \simeq 10^{-5}$) indicates that the universe was fairly homogeneous and isotropic on scales larger than the cosmological particle horizon, at the time of recombination. The spectrum of the anisotropy implies a primordial power spectrum for matter $P(k) = Ak^n$, $n = 1.1 \pm 0.32$ [6], which is in good agreement with predictions made by the inflationary scenario over a decade ago [7]. The COBE results lend further support to the 'standard model' of structure formation whereby the large scale structure in the universe forms due to the gravitational instability of density fluctuations which grow from initially small values $\delta\rho/\rho|_{\text{initial}} \sim \Delta T/T \simeq 10^{-5}$ at the time of recombination, to $\delta\rho/\rho|_{\text{final}} \gg 1$ today.

## 2. Nonlinear approximations to gravitational clustering

Gravitational instability in a collisionless self-gravitating fluid can be modelled by the coupled Euler–Poisson system of partial differential equations. These equations are written in terms of $\mathbf{x} = \mathbf{r}/a$ (the comoving spatial coordinate), $\mathbf{v} = \dot{\mathbf{r}} - H\mathbf{r} = a\dot{\mathbf{x}}$ (the peculiar velocity field), $\phi(\mathbf{x}, t)$ (the peculiar Newtonian gravitational potential) and $\rho(\mathbf{x}, t)$ (the matter density). Neglecting pressure terms we obtain [8]

*The Euler equation*

$$\frac{\partial(a\mathbf{v})}{\partial t} + (\mathbf{v} \cdot \nabla_x)\mathbf{v} = -\nabla_x \phi. \tag{5}$$

The term in the right hand side of eq. (5) represents the gravitational force which gives rise to acceleration in the fluid element.

*The continuity equation*

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_x[(1+\delta)\mathbf{v}] = 0, \tag{6}$$

where $\delta = \rho/\rho_0 - 1$ is the *density contrast* ($\rho_0$ is the mean background density), and

*The Poisson equation,*

$$\nabla_x^2\phi = 4\pi G a^2(\rho - \rho_0) = 4\pi G a^2 \rho_0 \delta. \tag{7}$$

The Euler–Poisson system forms a strongly coupled non-linear system of differential equations which is difficult to treat analytically. Several approximation schemes have been suggested which partially remedy this problem, we describe some of them below.

### 2.1 *Perturbation theory*

During the early stages of gravitational clustering the density contrast is small $\delta \ll 1$ and one can attempt to solve eq. (5)–(7) perturbatively by writing the density contrast as a series

$$\begin{aligned}\delta(\mathbf{x},t) &= \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x},t)\\ &= \sum_{i=1}^{\infty} D_+^n(t)\Delta^{(n)}(\mathbf{x}).\end{aligned} \tag{8}$$

The first term in the above expansion corresponds to the linear growth law: $\delta^{(1)} = D_+(t)\Delta^{(1)}(\mathbf{x})$. In a spatially flat matter dominated universe $D_+(t) \propto a(t) \propto t^{2/3}$ where $a(t)$ is the scale factor of the universe [9]. (The above expression tells us that left to themselves, density fluctuations would grow by a factor $\simeq 10^3$ between the epoch of decoupling and the present epoch.) Perturbation theory formally breaks down when $\delta \simeq 1$, to follow gravitational instability. Further one has to use more sophisticated non-linear approximations such as the Zel'dovich approximation described below.

### 2.2 *The Zel'dovich approximation*

In 1970, Zel'dovich proposed a remarkably simple approximation to follow the motion of fluid elements into the strongly non-linear regime. Zel'dovich suggested that instead of perturbing the density field one should instead perturb particle trajectories. The resulting Zel'dovich approximation (ZA) has the simple form [10, 11]:

$$\mathbf{r} = a\mathbf{x}(\mathbf{q},t) = a(t)[\mathbf{q} + D_+(t)\mathbf{u}(\mathbf{q})], \tag{9}$$

here $\mathbf{q}$ is the initial (Lagrangian) coordinate of a particle, and $\mathbf{x}$ is its final (Eulerian) coordinate. $\mathbf{u}(\mathbf{q})$ is the initial (comoving) velocity field of the particle at $\mathbf{q}$: $\mathbf{u}(\mathbf{q}) = \mathbf{v}/a\dot{D}_+$. Since any initial rotation is strongly damped during expansion we restrict our attention to potential flows so that $\mathbf{u}(\mathbf{q}) = -\nabla\Phi_0(\mathbf{q})$, $\Phi_0(\mathbf{q})$ being the linear velocity potential. It can be shown that $\Phi_0$ is related to the Newtonian gravitational potential $\phi$ by $\phi = A^{-1}\Phi_0$ where $A = 2/(3H^2a^3)$; $A$ is a constant for a spatially flat matter dominated universe.

A generic feature of the Zel'dovich approximation (for random initial distributions of $\Phi_0$) is the intersection of particle trajectories (shell crossing) which formally leads to singularities in the density field. As Zeldovich pointed out, the motion described by

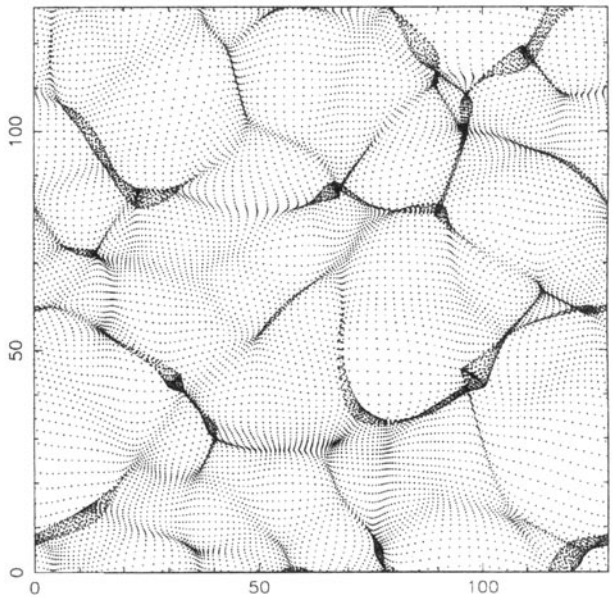


**Figure 3.** Caustics arising in a two dimensional realisation of the Zel'dovich approximation are shown for random (Gaussian) initial conditions. Reproduced, with permission, from [8].

eq. (9) is in many respects very similar to the equation describing the propagation of light rays in geometrical optics [11]. In optics, caustics – describing alternating bright and dark regions in the intensity of light – form from generic initial conditions (all of us are accustomed to seeing such features at the bottom of a shallow pool of water). Similarly gravitational instability too – as predicted by the Zel'dovich approximation – will result in the formation of caustics in the density field, as shown in figure 3.

The prediction of *planar caustics* follows immediately from eq. (9) if we note that the condition for mass conservation

$$\mathrm{d}M = \rho_0 \mathrm{d}^3\mathbf{q} = \rho(\mathbf{x}, t)\mathrm{d}^3\mathbf{x} \tag{10}$$

leads to the following expression for the density in terms of the Jacobian of the transformation from $\mathbf{q}$ to $\mathbf{x}$:

$$\begin{aligned}\frac{\mathrm{d}V_e}{\mathrm{d}V_l} &= J\left[\frac{\partial \mathbf{x}}{\partial \mathbf{q}}\right] \\ &= \left|\delta_{ij} - D_+(t)\frac{\partial^2\Phi_0}{\partial q_i \partial q_j}\right| \\ &= [1 - D_+(t)\lambda_1(\mathbf{q})][1 - D_+(t)\lambda_2(\mathbf{q})][1 - D_+(t)\lambda_3(\mathbf{q})],\end{aligned} \tag{11}$$

where $\mathrm{d}V_e \equiv \mathrm{d}^3\mathbf{x}$ and $\mathrm{d}V_l \equiv \mathrm{d}^3\mathbf{q}$, are the Eulerian and Lagrangian volume elements, respectively, and $\lambda_1, \lambda_2$ and $\lambda_3$ are the eigenvalues of the deformation tensor $d_{ij} \equiv \partial^2\Phi_0/\partial q_i \partial q_j$.

The resulting expression for the density distribution is

$$\rho(\mathbf{r},t)=\frac{\rho_0}{a^3}[1-D_+(t)\lambda_1(\mathbf{q})]^{-1}[1-D_+(t)\lambda_2(\mathbf{q})]^{-1}[1-D_+(t)\lambda_3(\mathbf{q})]^{-1}. \quad (12)$$

For a Gaussian random field the probability for the largest eigenvalue to be positive is 92%. A volume element satisfying this criterion will collapse along at least one direction, resulting in the formation of two dimensional pancake-type singularities in the density field, which arise when $1-D_+(t)\lambda_1=0$; (the eigenvalues are assumed to be ordered $\lambda_1>\lambda_2>\lambda_3$) see figure 3. Thus Zel'dovich showed that the first singularities in the density field are likely to be planar, a development which led to the 'pancake' model of structure formation pioneered by Zel'dovich and his collaborator's in Moscow in the 1970's and 1980's.

An interesting feature of ZA is that caustics percolate, even though they occupy a negligible fraction of the total volume. For a Gaussian random field a phase has to occupy at least 16% of the total volume in order to percolate. The reason for percolation in ZA is simple, regions which collapse in ZA occupy 92% of the Lagrangian volume, guaranteeing that such regions will percolate in Lagrangian space. Since the Zel'dovich mapping (9) preserves topology, regions which percolate in Lagrangian space continue to do so after they are mapped into Eulerian space (even though they now occupy a much smaller volume). One sees a similar property in the large scale distribution of galaxies, which also percolates at a small value of the 'filling factor'. Thus the Zel'dovich approximation could be on the right track in describing the gravitational clustering process that led to the formation of large scale structure in the universe.

### 2.3 *The adhesion model*

Tests of the Zel'dovich approximation against the results of $N$-body simulations show that the approximation works exceedingly well until the formation of the first pancakes. The Zel'dovich approximation attempts to describe particle motion 'inertially' – entirely through the initial conditions. There is therefore no inbuilt feedback mechanism which can modify particle trajectories especially when they pass through high density regions such as caustics. This drawback in ZA results in its breakdown soon after the formation of caustics which are shown in figure 3. If one applies ZA naively to later epochs one will find that the thickness of pancakes grows rapidly as particles sail through them oblivious of the backreaction of strong gravitational fields which they encounter enroute. Contrary to this one finds that the thickness of pancakes soon stabilises in $N$-body simulations. A successful attempt to incorporate the 'adhesive' effects of gravity on small scales into the Zel'dovich approximation has resulted in the adhesion model of structure formation (AM) [11, 12]. AM allows the study of large scale structure to be carried out in the strongly nonlinear regime after the first pancakes have formed, it reduces to ZA at early times. In AM particles are made to stick together once they enter pancakes, so that the total velocity flow is conserved. To appreciate the formal structure of AM let us recast the Euler equation in terms of the comoving velocity field $\mathbf{u}=(\mathbf{v}/\dot{a}a)=\mathrm{d}\mathbf{x}/\mathrm{d}a$ and the comoving velocity potential $\mathbf{u}=-\nabla\Phi$. As a result equation (5) reduces to

$$\frac{\partial\mathbf{u}}{\partial a}+(\mathbf{u}\nabla_x)\mathbf{u}=-\frac{3}{2a}[A\nabla_x\phi+\mathbf{u}], \quad (13)$$

$$\frac{\partial \Phi}{\partial a} - \frac{1}{2}(\nabla_x \Phi)^2 = \frac{3}{2a}(A\phi - \Phi). \tag{14}$$

In the linear regime of gravitational instability $\Phi = A\phi$, leading to

$$\frac{\partial \Phi}{\partial a} - \frac{1}{2}(\nabla_x \Phi)^2 = 0 \tag{15}$$

which has the analytical solution

$$\Phi(\mathbf{x}, a) = \Phi_0(\mathbf{q}) - \frac{(\mathbf{x} - \mathbf{q})^2}{2a}. \tag{16}$$

Differentiating eq. (16) alternately with respect to $\mathbf{x}$ and $\mathbf{q}$, we get the Zel'dovich approximation

$$\begin{aligned} \mathbf{x} &= \mathbf{q} - a(t)\nabla_q \Phi \\ \mathbf{u} &= -\nabla_x \Phi(\mathbf{x}, a) = -\nabla_q \Phi_0(\mathbf{q}). \end{aligned} \tag{17}$$

We thus see that the Zel'dovich approximation was inspired by the linear relationship $\Phi = A\phi$ which was successfully extended into the non-linear regime. (As a consequence of this relationship the velocity vector remains parallel to the acceleration vector so that particles move along straight line trajectories in ZA.) The adhesion approximation simplifies equation (14) by assuming that the right hand side in that equation can be replaced by a mock viscosity term intended to mimick the 'adhesive' effects of gravity on small scales. The resulting Euler equation becomes [12, 13]

$$\frac{\partial \mathbf{u}}{\partial a} + (\mathbf{u}\nabla)\mathbf{u} = \nu \nabla^2 \mathbf{u}. \tag{18}$$

This equation is the three dimensional generalization of Burgers' equation [14, 15]. If the motion happens to be potential $\mathbf{u} = -\nabla\Phi$, eq. (18) can be solved analytically by means of the Hopf–Cole substitution $\Phi(\mathbf{x}, a) = -2\nu \log U(\mathbf{x}, a)$ which leads to the diffusion equation

$$\frac{\partial U}{\partial a} = \nu \nabla^2 U \tag{19}$$

which can be solved exactly. We therefore get

$$\mathbf{u}(\mathbf{x}, a) = \frac{\int \mathrm{d}^3\mathbf{q}((\mathbf{x} - \mathbf{q})/a) \exp[-S(\mathbf{x}, a; \mathbf{q})/2\nu]}{\int \mathrm{d}^3\mathbf{q} \exp[-S(\mathbf{x}, a; \mathbf{q})/2\nu]}, \tag{20}$$

where

$$S(\mathbf{x}, a; \mathbf{q}) = -\Phi_0(\mathbf{q}) + \frac{(\mathbf{x} - \mathbf{q})^2}{2a(t)}. \tag{21}$$

The trajectory of a particle can be determined from the integral equation [16, 17]

$$\mathbf{x}(\mathbf{q}, a) = \mathbf{q} + \int_0^a \mathrm{d}a' \mathbf{u}[\mathbf{x}(\mathbf{q}, a'), a']. \tag{22}$$

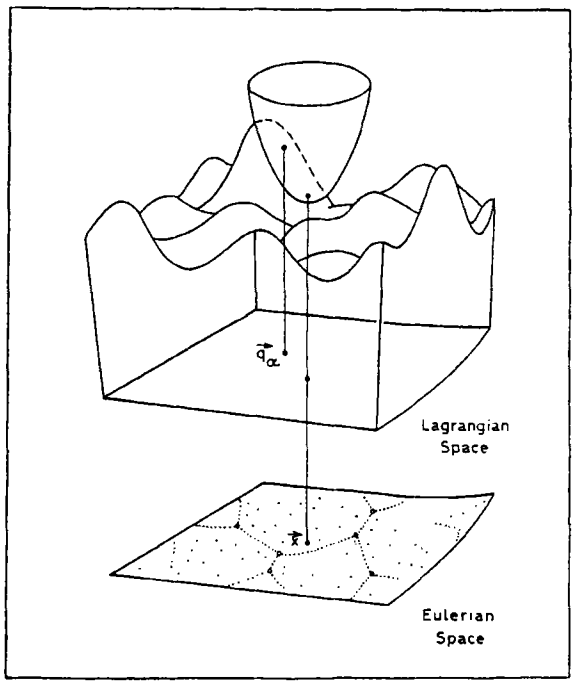


**Figure 4.** The geometrical construction of the adhesion model. A paraboloid descends onto the linear gravitational potential and defines sites of structure formation in two dimensions. The paraboloid is tangent to the potential in *Lagrangian* space, its apex gives the location of two dimensional caustics (filaments and clumps) in *Eulerian* space. Reproduced, with permission, from [8].

For cosmological applications it is necessary to be able to solve eqs (20), (22) in the limit $\nu \to 0$. This can be done using the method of steepest descents which indicates that the maximum contribution to the integral in eq. (20) will come from those points where $\nabla_q S = 0$. This results in an interesting geometrical interpretation which we describe below (see [18, 19]). We notice that the condition $\nabla_q S = 0$ defines a paraboloid $P = (\mathbf{x} - \mathbf{q})^2/2a(t)$ tangential to the velocity potential $\Phi_0$ at $\mathbf{q}$. The large scale structure of the universe is obtained by descending a paraboloid $P$ onto the (initial) potential $\Phi_0$ ($P$ and $\Phi_0$ are not permitted to intersect). The apex of $P$ defines the final location of particles which are initially located at $\mathbf{q}$ (the point of contact between $P$ and $\Phi_0$). The mapping between the initial position $\mathbf{q}$ of a particle (also called its Lagrangian coordinate), and its final location $\mathbf{x}$ (Eulerian coordinate) is single valued at early times when $a(t)$ (the radius of curvature of $P$) is small. It is multi-valued at late times when $a(t)$ is large. This means that the velocity flow is *single stream* at early times (i.e. caustics have not yet formed), but *multistream* at late times (after the formation of caustics). It is easy to see that for small $a$ the geometrical approach reduces to the Zel'dovich approximation. However the full power of the adhesion approach is felt at late times when the Zel'dovich approximation is no longer valid. The adhesion model at such times tesselates the universe into pancakes (or sheets), filaments and clumps

complemented by large empty regions – voids (see figure 4). Sheets form at the apex of the paraboloid $P$ when it touches $\Phi_0$ in exacly two points; filaments form when $P$ touches $\Phi_0$ in three points and clumps when the points of contact are exactly four. (More than four contact points are not generic in three dimensions.) According to the adhesion model gravitational instability proceeds in several distinct stages: during the first particles move 'inertially' to form two-dimensional 'pancakes', within pancakes flows become multistream and matter aggregates in regions of higher densities – filaments (which form at the intersections of pancakes). Finally matter flows along filaments to collect in clumps (the intersections of filaments). Thereafter neighboring clumps attract and merge and gravitational clustering proceeds in a hierarchical manner.

The adhesion model permits us to view the formation of large scale structure in the universe in a geometrical manner. The picture of clustering at different epochs is obtained by varying the radius of curvature of the paraboloid which is simply the scale factor of the universe. Although AM does not provide details of fluid motion within caustics it does give a fairly accurate description of structure on large scales as verified by detailed tests against $N$-body simulations [17, 20].

## 3. Percolation and the shape of large scale structure

Dynamical methods which describe how structures form must be supplemented by statistical methods describing different aspects of gravitational clustering. The simplest statistic to be applied to galaxy catalogues is the two point correlation function $\xi$ which has been described earlier. Since $\xi$ is the Fourier transform of the power spectrum $P(k) = |\delta_k|^2$ it ignores information contained in the phases of Fourier components $\delta_k$. This is a major limitation since it is well-known that beyond the linear regime of gravitational instability coupling between modes becomes important leading to the build up of phase correlations. (Initially the phases of the Fourier components of the density field are uncorrelated and random.) A consequence of mode coupling is the departure of the galaxy distribution from Gaussianity and the increasing coherence length of structures. (In the universe one sees 'patterns' in the distribution of galaxies on scales of 50–100 Mpc. i.e. on scales on which $\xi(R)$ is effectively zero.) Large scale structure in the universe appears to be concentrated in large objects (such as the *great wall*) occupying a relatively small volume of space (i.e. the filling factor of such objects is very small), see figure 1. These 'superclusters' of galaxies are separated by voids – vast regions of space with little if any matter. Zel'dovich first drew attention to the smallness of the filling factor of the galaxy distribution in the universe and suggested using percolation theory to understand its topological properties [33].

It is well known that a random (Poisson) distribution of points percolates at a filling factor (FF) of $\sim 0.319$. In other words the fraction of volume occupied by the percolating phase is $\sim 32\%$. On the other hand systems evolving under gravitational instability show a much smaller FF at percolation $\sim 0.02$–$0.07\%$ [30]. This is also what one finds from galaxy surveys leading one to believe that the galaxy distribution is strongly non-Gaussian. In a comprehensive study we have analysed the percolation properties of $N$-body simulations intending to model the universe at different stages of its evolution.

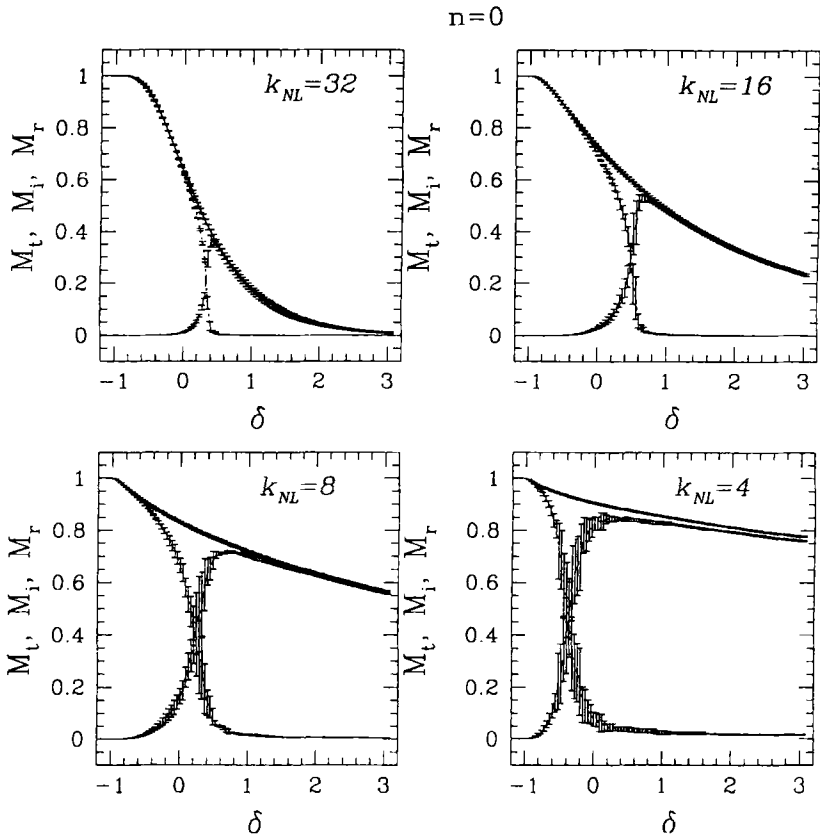


**Figure 5.** Percolation in $n = 0$ model at different epochs measured by the scale of nonlinearity $k_{\rm NL}^{-1}$ which increases from top to bottom. The filling factor of the largest cluster $M_i$ (dotted line), the filling factor of all but the largest cluster $M_r$ (dashed line) and the total filling factor $M_t$ (solid line) are plotted against the density contrast. Reproduced, with permission, from [21].

In our simulations we first define a density threshold, then identify clusters at this threshold using a nearest neighbor (friends-of-friends) algorithm. Having done this we determine the mass fraction in clusters. The models considered are $N$-body simulations with power spectrum $P(k) \equiv \langle |\delta_k|^2 \rangle \propto k^n$, for $k < k_N$, and $P(k) = 0$ for $k \geq k_N$, $n = -3, -2, -1, 0$ and $+1$, where $k_N$ is the Nyquist cutoff. $\langle\ \rangle$ denotes the average over an ensemble. The epochs we consider are characterized by the scale of nonlinearity $k_{\rm NL}^{-1}$ at that epoch determined by

$$\sigma \equiv \langle \delta^2 \rangle^{1/2} = D_+(t) \left( 4\pi \int_0^{k_{\rm NL}} P(k) k^2 {\rm d}k \right)^{1/2} = 1, \tag{23}$$

$D_+(t)$ being the linear growing-mode of density fluctuations.

Our results are shown in figure 5 for the spectrum $n = 0$, corresponding to an equal distribution of power amongst all modes in the initial conditions (white noise) (a more

comprehensive discussion of our work may be found in [21, 22]). We have plotted: (1) $M_i$ – the mass contained in the largest cluster divided by the total mass in clusters (the largest cluster is formally 'infinite' after percolation sets in), $M_i$ is also called the mass fraction or filling factor (FF) of the largest cluster; (2) $M_r$ – the mass fraction in the rest of the clusters (other than the largest); (3) $M_t = M_i + M_r$ – the total mass fraction in clusters (i.e. the total filling factor). $M_i, M_r, M_t$ are plotted as functions of the density threshold for different expansion epochs in figure 5. We find that $M_i$ decreases with increasing $\delta$, for $\delta > \delta_c \simeq 0.5$, $M_i$ is very small and clusters do not percolate. On the other hand for $\delta < \delta_c$, $M_i$ increases sharply indicating that virtually all the mass (above the threshold) is in the largest (percolating) cluster. Thus $\delta = \delta_c$ defines the percolation threshold which appears to be sensitive both to the spectral index and the amount of evolution the system has undergone. A related issue of considerable importance concerns the *topology* of large scale structure, i.e. whether the distribution of large scale structure is 'meatball-like', 'sponge-like' or some other. High values of the filling factor at percolation (relative to a Gaussian distribution) determine a 'meatball-like' topology, whereas low values indicate a 'bubble-like' or a 'sponge-like' topology [24].

We should note that $M_r \gg M_i$ for $\delta \gg \delta_c$ indicating that most of the mass is in small unconnected clusters at high densities. As we decrease the threshold $\delta$ we find that $M_r$ peaks, then rapidly decreases. The fall in $M_r$ is compensated by the increase in $M_i$, caused by the fact that clusters, as they grow in size, tend to connect with each other to form superclusters which eventually percolate. (The error bars in figure 5 indicate the rms dispersion in four different $N$-body realisations of the same spectrum.)

The smallness of the filling factor in our simulations indicates that clustered matter is likely to be distributed in either sheets or filaments (since both sheets and filaments occupy less space than spheres). To check whether this is true we employed a statistic sensitive to shape originally suggested by Babul and Starkman [23]. This 'shape statistic' is constructed by determining the moment of inertia (deformation) tensor of matter distributed within a radius $R$ of a fiducial point. The tensor is then diagonalized and a triplet of shape functions $S_i$ introduced by the relations

$$S_1 \equiv \sin\left[\frac{\pi}{2}(1-\mu)^p\right], \tag{24}$$

$$S_2 \equiv \sin\left[\frac{\pi}{2}a(\mu,\nu)\right], \tag{25}$$

$$S_3 \equiv \sin\left[\frac{\pi\nu}{2}\right], \tag{26}$$

where $\mu \equiv \sqrt{\lambda_2/\lambda_1}$, $\mu \equiv \sqrt{\lambda_3/\lambda_1}$. (The three eigenvalues of the deformation tensor $\lambda_1$, $\lambda_2$, $\lambda_3$ are arranged in decreasing order, so that $\lambda_1 > \lambda_2 > \lambda_3$.) The function $a(\mu,\nu)$ is implicitly defined by the equation

$$\frac{\mu^2}{a^2} - \frac{\nu^2}{a^2(1-\alpha a^{1/3}+\beta a^{2/3})} = 1. \tag{27}$$

The parameters $p$, $\alpha$ and $\beta$ are chosen to normalize $S_n$ (for details see [23, 8, 21, 22]).

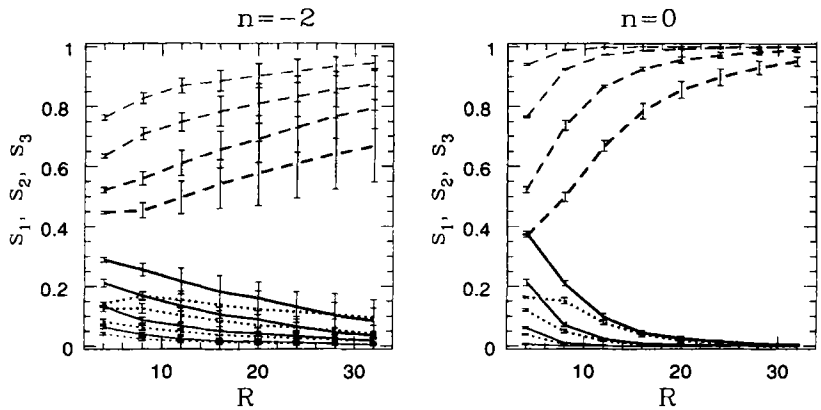


**Figure 6.** Structure functions averaged over randomly chosen high density regions are plotted as functions of the window radius $R$ (measured in units of the grid size). Solid lines plot the average linearity $\bar{S}_1$, dotted lines the average planarity $\bar{S}_2$, and dashed lines the average sphericity $\bar{S}_3$. Heavier lines correspond to later epochs. The y-axis extends from 0 to 1. Reproduced, with permission, from [22].

Defined in this manner $S_i$ can be evaluated for a distribution of points (as originally suggested in [23]) or for a density field (which is the approach we employ). In the latter case it is sufficient to evaluate $S_i(R)$ within a radius $R$ of fiducial points lying above a given density threshold $\delta_i$. The location of the fiducial points is chosen randomly (the sole criterion being that they lie in regions where $\delta > \delta_i$). One can define the average shape functions $\bar{S}_i = 1/N \sum_{i=1}^{N} S_i$ defined for the ensemble as a whole, and study its properties. $S_i$ can be viewed as a vector, defined in the first octant: $S_i = (1,0,0)$ describes a one dimensional filament, $S_i = (0,1,0)$ – a two dimensional 'panacake', and $S_i = (0,0,1)$ – a three dimensional sphere. In figure 6 we plot $\bar{S}_i$ as functions of both the radius $R$ and the time (characterized by the scale of non-linearity $k_{\rm NL}^{-1}$). We find that as the $N$-body simulation evolves structures within it proceed to become more and more filamentary. This is true for all spectra we consider and appears to be an important generic property of systems evolving under gravitational instability [31, 32].

## 4. Conclusions

Both dynamical and statistical aspects are crucial to a deep understanding of large scale structure in the universe. Whereas statistical indicators such as the two point correlation function are useful in measuring clustering, they must be complemented by topological descriptors in order to understand the global properties of large scale structure such as its connectedness and topology. We show that gravitationally clustered systems in general percolate at lower filling factor's than a Gaussian random field, this property indicates that the large scale structure in such systems is closer to being 'sponge-like' or 'bubble-like' than 'meatball-like'. We also argue that whereas the first gravitationally bound objects are likely to be pancakes, gravitational clustering thereafter leads to the increasing filamentarity of structures. Whether the degree of filamentarity seen in

simulations matches that in redshift surveys of galaxies (such as IRAS), is presently being studied [27].

## Acknowledgement

The author is grateful to B S Sathyaprakash and Sergei Shandarin for stimulating discussions and many fruitful collaborations.

## References

[1] M Davis and P J E Peebles, *Astrophys. J.* **267**, 465 (1983)
[2] N Kaiser, *Astrophys. J.* **284**, L9 (1984)
[3] J M Bardeen, J R Bond, N Kaiser and A S Szalay, *Astrophys. J.* **304**, 15 (1986)
[4] D Lynden-Bell, S M Faber, D Burstein, R L Davies, A Dressler, R J Terlevich and G Wegner, *Astrophys. J.* **326**, 19 (1988)
[5] A Dekel, *Ann. Rev. Astron. Astrophys.* **32**, 371 (1994)
[6] M White, D Scott and J Silk, *Ann. Rev. Astron. Astrophys.* **32**, 319 (1994)
[7] A R Liddle and D H Lyth, *Phys. Rep.* **231**, 1 (1993)
[8] V Sahni and P Coles, *Phys. Rep.* **262**, 1 (1995)
[9] In a static universe the growth of $\delta$ is much faster $D_+(t) \propto \exp\sqrt{4\pi G\rho t}$
[10] Ya B Zel'dovich, *Astron. Astrophys.* **4**, 84 (1970)
[11] S F Shandarin and Ya B Zel'dovich, *Rev. Mod. Phys.* **61**, 185 (1989)
[12] S N Gurbatov, A I Saichev and S F Shandarin, *Mon. Not. R. Astron. Soc.* **236**, 385 (1989)
[13] S N Gurbatov, A I Saichev and S F Shandarin, *Sov. Phys. Dokl.* **30**, 921 (1985)
[14] J M Burgers, *The non-linear diffusion equation* (Reidel, Dordrecht, 1974)
[15] G B Whitham, *Linear and non-linear waves* (Wiley, New York, 1974)
[16] A Nusser and A Dekel, *Astrophys. J.* **362**, 14 (1990)
[17] D H Weinberg and J E Gunn, *Mon. Not. R. Astron. Soc.* **247**, 260 (1990)
[18] L A Kofman, D Yu Pogosyan, S F Shandarin and A L Melott, *Astrophys. J.* **393**, 437 (1992)
[19] V Sahni, B S Sathyaprakash and S F Shandarin, *Astrophys. J.* **431**, 20 (1994)
[20] B S Sathyaprakash, V Sahni, D Munshi, D Pogosyan and A L Melott, *Mon. Not. R. Astron. Soc.* **275**, 463 (1995)
[21] B S Sathyaprakash, V Sahni and S F Shandarin, *Mon. Not. R. Astron Soc.* (submitted)
[22] B S Sathyaprakash, V Sahni, S F Shandarin, *Astrophy. J. Lett.* **462**, L5 (1996)
[23] A Babul and G D Starkman, *Astrophys. J.* **401**, 28 (1992)
[24] In a meatball topology underdense regions percolate easily whereas overdense do not. In a bubble-topology the underdense phase is enclosed by overdense 'bubbles', as a result the overdense phase percolates whereas the underdense one does not. In a sponge topology both phases are symmetric and both can percolate with equal ease; (see [8, 25] for a detailed discussion of these issues). Other statistical indicators used to investigate the global properties of large scale structure include the 'genus characteristic' and the minimal spanning tree [8, 26].
[25] C Yess and S F Shandarin, *Astrophys. J.* **465**, 2 (1996)
[26] A L Melott, *Phys. Rep.* **193**, 1 (1990)
[27] B S Sathyaprakash, V Sahni, S F Shandarin and K Fisher, in preparation
[28] L N da Costa, in *Cosmic velocity fields*, Proceedings of the ninth IAP Astrophysics meeting edited by F R Bouchet and M Lachieze-Rey (Editions Frontieres 1994) p. 475
[29] N A Bahcall and M J West, *Astrophys. J.* **392**, 419 (1992)
[30] A A Klypin and S F Shandarin, *Astrophys. J.* **413**, 48 (1993)
[31] R C Pearson and P Coles, *Mon. Not. R. Astron. Soc.* **272**, 231 (1995)
[32] B S Sathyaprakash, V Sahni, S F Shandarin and K Fisher, in preparation
[33] Ya B Zel'dovich, *Sov. Astron. Lett.* **8**, 102 (1982)