Physics – Uspekhi **38** (7) 687 – 722 (1995)

CENTENARY OF I E TAMM'S BIRTH

PACS numbers: 98.80.Cq; 95.35.+d; 98.62.Ai

Large-scale structure of the Universe. Analytic theory

A V Gurevich, K P Zybin

Contents

Abstract. An analytical theory is developed of the nonlinear stage of the Jeans instability in a cold nondissipative gas in an expanding Universe. It is shown that the growth of this instability creates a giant dark-matter halo, of 200 kpc size, around galaxies. This halo has a density singularity described by $\rho \propto r^{-\alpha}$, where $\alpha \approx 1.7$–1.9. A comparison is made of this analytic theory with the results of numerical simulations and a good agreement between them is demonstrated. Analytic dynamic solutions are used to develop a statistical method for calculation of the correlation functions of galaxies and galaxy clusters. The theory is compared with the available observational data. The physical consequences of the proposed theory are considered.

A V Gurevich, K P Zybin P N Lebedev Institute, Russian Academy of Sciences, Leninskii pr. 53, 117924 Moscow
Tel. (095) 132-64-14, 132-60-50, 132-61-71
E-mail: alex@td.lpi.ac.ru; zybin@td.lpi.ac.ru

Received 14 March 1995
Uspekhi Fizicheskikh Nauk **165** (7) 723 – 758 (1995)
Translated by A Tybulewicz

1. Introduction

Observations show that, over distances of the order of the horizon radius, our Universe is homogeneous, isotropic, and expands uniformly. The expansion leads to a rapid cooling of matter. A cold gravitating gas is unstable because of the action of universal gravitational forces. Growth of the Jeans instability creates regions of strong compression with dimensions much smaller than the horizon radius. This is of decisive importance for the formation of the large-scale structure of matter in the Universe: galaxies, galactic clusters, superclusters, etc.

The main role in this process is played by the hidden mass, also called dark matter. This hidden mass manifests itself only in the gravitational interaction. It was first introduced in 1933 by Zwicky [1], who investigated galaxies in the Virgo cluster and observed that the masses of luminous galaxies are insufficient to account for the observed dynamics: it is necessary to introduce additionally the hidden mass [1]. Subsequently, dynamic hidden mass has been observed in many other clusters and galaxies, including our galaxy.

The existence of hidden mass and its special nonbaryonic nature are supported also by investigations of the process of nucleosynthesis. Heavy elements cannot form in the early Universe and their formation involves secondary nucleosynthesis in stars. The primordial nucleosynthesis yields, according to recent calculations, the following mass abundances X of some of the light elements [2]:

$$X(\mathrm{He}^4) \approx 0.24\,,\ X(\mathrm{D}) \sim X(\mathrm{He}^3) \sim 10^{-5}\,,\ X(\mathrm{Li}^7) \sim 10^{-10}\ .$$

These primordial abundances of light elements impose severe constraints on the ratio of the baryonic density of matter ρ_b to the critical density ρ_c, represented by the parameter $\Omega_\mathrm{b} = \rho_\mathrm{b}/\rho_\mathrm{c}$:

$$0.010h^{-2} < \Omega_\mathrm{b} < 0.015h^{-2}\ , \tag{1}$$

where h is the Hubble constant, normalised to 100 km s^{-1} Mpc^{-1}. On the other hand, a study of the dynamics of matter and, in particular, investigations of the gas in galactic groups and clusters give an estimate of the dynamic mass which is an order of magnitude larger [3, 4]:

$$\Omega = \frac{\rho}{\rho_\mathrm{c}} \geqslant 0.2\ . \tag{2}$$

Another important argument in support of the existence of the nonbaryonic dark matter follows from the anisotropy of the microwave background radiation. The argument is as follows. Fluctuations in the baryonic matter can grow only after the moment of recombination, which is characterised by a red shift $z_\mathrm{R} \sim 1000$. Up to that moment the density and temperature fluctuations are related by

$$\frac{\delta T_\mathrm{R}}{T} \approx \frac{\delta\rho_\mathrm{R}}{\rho_0}\ .$$

The very existence of structures in the Universe indicates that up to now the density fluctuations $\delta\rho_\mathrm{R}/\rho_0$ have been greater than or of the order of unity. Consequently, at the moment of recombination we should have

$$\frac{\delta\rho_\mathrm{R}}{\rho_0} \geqslant 10^{-3}\ .$$

However, observations show that the initial perturbations are considerably less than this value. According to the COBE (Cosmic Background Explorer) observations [5], the temperature fluctuations $\delta T/T$ amount to just 10^{-5}. The dark-matter concept makes it possible to solve this problem as well, because the dark-matter fluctuations grow much earlier than the moment of recombination. As a consequence, dark matter forms potential wells into which the baryonic matter drops after recombination.

Therefore, although we do not know yet the particles of which dark matter is composed, it is quite clear that these particles are of nonbaryonic nature and, consequently, interact very weakly with one another and with the baryonic matter. It is usual to assume that these particles are either a low-mass neutrino (known as hot dark matter — HDM [6]) or some hypothetical heavy particles, such photino, neutralino, etc. (known as cold dark matter — CDM [7–9]). In recent years a combination of both (CDM + HDM) has been considered.

From all this it follows that investigations of the dynamics of dark matter, representing over 90% of the mass of the Universe (when $\Omega = 1$), are of key importance for the understanding of the nonlinear structures that appear in the Universe.

The problem of the appearance of a large-scale structure in the Universe can be formulated as follows. Small initial perturbations grow linearly in the homogeneous and isotropic, uniformly expanding, Universe. It is natural to assume that correlations of linear perturbations of different scales are independent and are of Gaussian nature. The problem is then determined entirely by the form of the initial spectrum and by the parameter Ω. The nature of the spectrum of the initial fluctuations was considered on the basis of very general ideas by Zel'dovich [10] and Harrison [11]. The parameter Ω is usually taken as unity. In particular, a spectrum close to the Zel'dovich–Harrison spectrum (for $\Omega = 1$) follows from the inflation theory [12].

Large-scale structures grow as a result of nonlinear dynamics of the initial fluctuations. The problem has been tackled recently mainly by the method of direct numerical simulation. A nondissipative gas is replaced by an ensemble of identical particles interacting in accordance with the Newtonian law. The initial distribution of the particles is assumed to be uniform with random small perturbations. The initial spectrum is represented by a power law, quite close to that describing the Zel'dovich–Harrison spectrum. It is usual to vary the power exponent m of this spectrum and the parameter Ω in such calculations.

More complex spectra, resulting from a combination of CDM and HDM fluctuations, have also been analysed. In such cases the simulation is three-dimensional. Up to 3×10^6 particles have been considered recently [13–15]. It should be pointed out that such simulation is used only in the case of the nondissipative dark matter. Some investigators have also tackled the motion of the small fraction of the dissipative baryonic matter, described by hydrodynamic equations [16].

These results of numerical calculations are then compared with astronomical observations and particularly with the data on nonlinear structures. The conclusions of the current theories can be formulated as follows [17].

(1) After normalisation to the COBE data on the scale of 10^0, the CDM model is in agreement with the recent measurements of linear perturbations with other scales. The only possible exceptions are measurements on the scale of 1–5 arcmin in the RING experiments [18]. The latter are 3.3 times greater than the results of extrapolation of the COBE data in accordance with the linear Zel'dovich–Harrison spectrum. It should be pointed out that 1–5 arcmin corresponds to a scale of several megaparsecs.

(2) On scales of the order of $(10-15)\,h$ Mpc the galaxy–galaxy and cluster–cluster correlations are considerably higher than those deduced from the standard CDM scenario.

(3) A comparison of the galactic velocity distribution with observations shows that numerical calculations predict a much greater dispersion of the velocities, i.e. a more

chaotic motion. The observed motions include a correlated coherent component, which is considerably greater than the chaotic component.

(4) A preferred theory is that postulating the formation of galaxies already at the stage corresponding to the red shift $Z = 5-10$. Among the investigated numerical models this condition is obeyed by the pure CDM model only at low values of Ω, such as 0.2.

(5) An important role is played by direct collisions of galaxies, resulting in their coalescence. The influence of these processes on the distribution of galaxies is not yet clear.

(6) According to the current observations, the Hubble constant H_0 is closer to $60-80$ km s^{-1} Mpc^{-1}. The numerical calculations under discussion lead to the preferred value $H_0 \approx 50-30$ km s^{-1} Mpc^{-1} and also to small values $\Omega \approx 0.2-0.3$ (in the pure CDM model).

(7) Admission of a possible HDM + CDM combination eliminates a large number of these contradictions. At present this model seems to be the most promising [19].

On the whole, the remarkable success of numerical simulations is due to the feasibility of verifying the selection of the starting model of the initial spectrum and of the composition of dark matter on the basis of a detailed comparison with observational data. At present, it is most probable that the spectrum of the intial perturbations is close to the modified Zel'dovich – Harrison spectrum of a combination with 0.75 CDM + 0.25 HDM of dark matter with $\Omega = 1$. It should be stressed however that other versions of the models, particularly open models with $\Omega \approx 0.2$, cannot be excluded.

It is important to stress that very accurate answers cannot be obtained by numerical simulation, because of the fundamental limitations of the models. Even for the maximum possible number of particles considered at present (of the order of 10^6-10^7), the three-dimensional nature of the problem means that each measurement corresponds to $(1-2)\times 10^2$ particles. This means that the maximum range of the change in the perturbation spectrum considered in a model is only one to one-and-a-half orders of magnitude:

$$\frac{k_{\max}}{k_{\min}} \sim 10-30 \ .$$

This is insufficient to cover all the changes in the real spectrum, from the linear range (greater or of the order of 100 – 300 Mpc) to the subgalactic scales.

Moreover, the use of 'large' particles, whose mass is many orders of magnitude greater than the mass of dark-matter particles (in reality, their mass is between two and four orders of magnitude greater than the mass of a galaxy) in numerical calculations leads to a colossal overestimate of the role of the Coulomb scattering. This may have a significant influence on the particle velocity distribution and can distort the role of the dissipative processes. The need to use 'large' particles is one of the most serious shortcomings of the numerical methods. Moreover, it should be pointed out that the Poisson equation is usually solved by the fast Fourier transform method, which leads to an effective smoothing out of singularities.

These shortcomings do not apply to the analytic approach developed greatly in recent years. An analytic description of nonlinear dynamics of gravitating matter dates back to the paper of Chandrasekhar and Munch [20], who proposed to use statistical methods for the dynamics of continuous media, used extensively earlier in the theory of hydrodynamic turbulence. The theory has been utilised to the greatest extent in the well-known monograph of Peebles [21]. The underlying assumption of this theory is smoothness of the higher correlation functions, which makes it possible to decouple approximately a chain of coupled nonlinear equations.

Unfortunately, it is this assumption which is not satisfied in reality: special singular regions are selected where chains of correlation functions diverge. These regions are determined by the nonlinear dynamics of compression of cold gravitating matter and they play a fundamental role in the establishment of a nonlinear stochastic state. It should be pointed out that recent developments in the theory of hydrodynamic turbulence also support the decisive role of nonlinear singularities in higher correlation functions [22]. In the theory of a gravitating gas the role of singularities is considerably greater because of the Jeans instability of the homogeneous state of a system.

Zel'dovich [23] was the first to point out singular formations in the dynamics of a nondissipative gravitating gas. These are the well-known Zel'dovich pancakes, investigated also by others [24, 25]. It has been found since that a flat singularity of the pancake type is not the only one and not the main one. A three-dimensional compression singularity, which is of the highest order and is the basis of the steady-state stationary solution, has been discussed [25 – 29]: it is known as a nondissipative gravitational singularity (NGS). It is the NGS that determines the characteristic features of the pair correlation function in an advanced nonlinear stochastic state of dark matter [30]. Singularities of the Zel'dovich pancake type manifest themselves clearly only in the correlation functions of higher orders.

These singularities of the dynamics of a gravitating gas stand out clearly in the distributions of the density and velocity of matter, and in the behaviour of the correlation functions. This facilitates quite definite predictions, some of which (for example, the giant dark-matter halo around galaxies, the distribution of the velocities on the rotation curves, and the singular law of the behaviour of a pair correlation function) have already been confirmed by observations.

The present review is concerned with the state of the art of the analytical theory of the large-scale structure of the distribution of matter in the Universe. Section 2 gives the initial formulation of the problem and provides the equations describing the dynamics of the nondissipative cold matter, as well as the initial and boundary conditions for these equations. The important role of truncation of the Zel'dovich – Harrison spectrum in the short-wavelength range, due to the finite mass of dark-matter particles, is stressed.

Section 3 deals with the nonlinear growth of the initial fluctuations. The general solution is obtained and a growing unstable mode is identified. The relationships between the density and velocity in a growing mode, important for the development of a nonlinear theory, are considered. The nonlinear solution, obtained in the hydrodynamic approximation, is discussed in Section 4. In the initial perturbation there are singularities near which the nonlinearity plays a decisive role. It is shown that after a finite time a singularity

appears near these points in the initial hydrodynamic system of equations and then the system becomes invalid.

Section 5 deals with the solution after the appearance of a singularity. It is stressed that the initial singularity in the early flow is followed by regions of multistream flow. With time, the number of streams in such regions increases. Detailed studies have been made of the one-dimensional planar case. In the limit $t \to \infty$, the number of streams tends to infinity and in the vicinity of an initial singularity a stationary mixed kinetic NGS state is formed and it has a singularity of the average density $\bar{\rho} \propto x^{-4/7}$.

Section 6 is devoted to an NGS in the spherically symmetric case. It is shown that in this case the formation of an NGS is much faster than in the planar case. An infinite number of streams appears practically immediately after the initial singularity and a quasistationary density distribution $\rho \propto r^{-2} \ln^{1/3}(1/r)$ is established. The general case is considered in Section 7. It is stressed that, in general, a stationary spherically symmetric distribution appears near the initial singularity and at its centre there is a density singularity which can be extrapolated sufficiently accurately by the power law $\rho \propto r^{-\alpha}$, where $\alpha = 1.7-1.9$.

It is shown in Section 8 that the continuous growth of the Jeans instability creates NGSs of various scales and that large-scale singularities can capture small-scale ones, thus giving rise to a hierarchical structure in those regions where a multistream flow has developed and a quasisteady state has been established.

Section 9 discusses structures of the largest scale, which appear in transition regions of flow where kinetic mixing is not yet complete and separate caustic singularities are observed. The structure pattern is then cellular. It is stressed that in these regions the third-order and higher correlation functions should have singularities due to this cellular structure.

Section 10 deals with the correlation functions and presents a theory that takes account of the presence of strong correlations near singular points. It is shown that the observed singular structure of the pair correlation function $\xi \propto r^{-\alpha}$ is related uniquely to an NGS. Some astrophysical applications of the proposed theory are discussed briefly in Section 11.

It should be stressed that when speaking of an analytic theory of the hierarchical and large-scale structures, our aim is not to regard this theory as in any way opposite to numerical simulation, mentioned at the beginning of this review. Above all, it is clear that the term 'analytic theory' used by us is arbitrary, because in this theory the derivation and analysis of specific solutions rely significantly on numerical methods.

Moreover, the analytic theory concentrates mainly on the behaviour of matter in the vicinity of singularities. However, as shown below, the bulk of matter is concentrated at the periphery, i.e. effectively outside the singularities. Therefore, the general relationships applicable to the average behaviour of the relatively smooth part of the structure of matter are undoubtedly best described by direct numerical simulation.

On the other hand, the analytic theory is undoubtedly superior in the description of the behaviour of the density of matter near singularities of the distribution of the velocities and of the structure of the trapping regions, and also in the description of singular properties of the correlation functions. The two approaches are mutually complementary and obviously the fullest and most accurate description of the dynamics of the appearance and growth of the structural matter in the Universe requires a sensible combination of both approaches.

2. Initial equations. Initial and boundary conditions

2.1 Main simplifications

As pointed out above, it follows from the observational data that our Universe is on the whole homogeneous, isotropic, and expands uniformly. The maximum size of inhomogeneities, which govern the large-scale structure of matter, is $l < 300-500$ Mpc, i.e. the inhomogeneity scales are much less than the horizon radius $R \approx 5 \times 10^3$ Mpc. Consequently, in a theoretical investigation of this problem we need consider only such fluctuations whose scale l is small compared with the horizon radius R_{H}:

$$\frac{l}{R_{\mathrm{H}}} \ll 1 \,. \tag{3}$$

At the same time, the peculiar flow velocities V can be regarded as low compared with the velocity of light c:

$$\frac{V}{c} \ll 1 \,. \tag{4}$$

In fact, the peculiar velocities obey $V \leqslant lH$, where the Hubble constant H, governing the expansion of the universe, is

$$H \approx (50-100)\ \mathrm{km\ s^{-1}\ Mpc^{-1}} \,. \tag{5}$$

Hence and from inequality (3), bearing in mind that $R_{\mathrm{H}} H = c$, we obtain relationship (4). Under the conditions described by expressions (3) and (4), the relativistic effects are not important and the Newtonian dynamics is sufficient.

As pointed out above, according to current ideas, dark matter consists of noninteracting or very weakly interacting particles. The mean free path of such particles l_{f} is considerably greater than the scale of inhomogeneities. Then, the parameter

$$\frac{l}{l_{\mathrm{f}}} \tag{6}$$

makes it possible, at least in the first approximation, to ignore collisions of dark matter particles, i.e. to consider dark matter as a gas of noninteracting particles.

During the period of nonlinear growth of fluctuations, cooling in the course of expansion of the Universe ensures that the gas of particles forming dark matter always remains cold.

2.2 Kinetic and hydrodynamic equations

The dynamics of dark matter is that of a gas of noninteracting particles moving in its own self-consistent gravitational field. Under the conditions described by expressions (3), (4) and (6), this dynamics is described by the kinetic equation

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial \boldsymbol{r}} - \frac{\partial \psi}{\partial \boldsymbol{r}} \cdot \frac{\partial f}{\partial v} = 0 \,,$$
$$\nabla^2 \psi = \int f \, \mathrm{d}v \,. \tag{7}$$

Here, $f(\boldsymbol{r}, v, t)$ is the particle distribution function and $\psi(\boldsymbol{r}, t)$ is the potential of the gravitational field. In expressions (7) and later the system of units, selected for compactness, is such that $4\pi G = 1$, where G is the gravitational constant.

The specific form of the distribution function f may generally depend on the previous history of the appearance of fluctuations, on the masses of the particles forming dark matter, and on the interaction cross sections of these particles. However, in the epoch of interest to us, dark matter has already cooled (as pointed out above) and it has formed a cold collision-free gas with a negligible temperature. The initial distribution function can then be selected in the form

$$f(\boldsymbol{r}, v, t) = \rho(\boldsymbol{r}, t)\ \delta\big[v - \boldsymbol{V}(\boldsymbol{r}, t)\big]\ , \tag{8}$$

where ρ is the gas density and $\boldsymbol{V}$ is the hydrodynamic velocity.

Substitution of formula (8) into the set of expressions (7) and rewriting the resultant system of equations for the parameters ρ and $\boldsymbol{V}$, bearing in mind the cosmological expansion of the universe, i.e. adopting an expanding system of coordinates, yields the following hydrodynamic system of equations:

$$\begin{aligned} &\frac{\partial \boldsymbol{V}}{\partial t} + a^{-1}\boldsymbol{V}\cdot\nabla\boldsymbol{V} + a^{-1}\dot{a}\,\boldsymbol{V} + a^{-1}\,\nabla\varphi = \boldsymbol{0}\ , \\ &\frac{\partial \delta}{\partial t} + a^{-1}\,\nabla\cdot\big[(1+\delta)\boldsymbol{V}\big] = 0\ , \\ &\nabla^2\varphi = a^2\rho_0\delta\ . \end{aligned} \tag{9}$$

Here, $\boldsymbol{V} = \boldsymbol{V}(\boldsymbol{x}, t)$ is the peculiar velocity of matter, $\delta = \delta(\boldsymbol{x}, t) = [\rho(\boldsymbol{x}, t) - \rho_0(t)]/\rho_0(t)$ is the deviation of the density of the investigated gas from the average background density $\rho_0(t)$, and $a(t)$ is the scaling factor.

Transformation of the variables $\boldsymbol{x}$, $\boldsymbol{V}$, and t from an expanding system of coordinates to any fixed system can be made in accordance with the familiar relationships [21]:

$$\begin{aligned} &\boldsymbol{r} = a(t)\boldsymbol{x}\ , \qquad \boldsymbol{u} = \boldsymbol{V} + \boldsymbol{x}\dot{a}\ , \\ &\rho = \rho_0(1+\delta)\ , \qquad \varphi = \psi + \frac{a}{2}\,\ddot{a}x^2\ , \end{aligned} \tag{10}$$

where a dot represents differentiation with respect to time.

2.3 Initial and boundary conditions

The system of equations (9) can be solved by selecting the initial and boundary conditions for the functions $\boldsymbol{V}$ and δ. Under the conditions described by expression (4) the inhomogeneities of interest to us grow in an expanding homogeneous Universe. The boundary conditions are then unimportant, since the requirement of the vanishing of perturbations at infinity is readily satisfied by the expansion process.

Formally, the initial conditions for solving the system of equations (9) can be selected by specifying a set of four initial functions at some moment $t = t_i$:

$$\delta(\boldsymbol{x}, t_i) = \delta_i(\boldsymbol{x})\ , \qquad \boldsymbol{V}(\boldsymbol{x}, t_i) = \boldsymbol{V}_i(\boldsymbol{x})\ . \tag{11}$$

Which moment t_i should be selected as the initial depends on the physics of the interaction of the particles forming dark matter and the mass of these particles. In our case, the moment t_i corresponds to the beginning of that epoch of the evolution of the Universe when its general cosmological expansion is governed by dark matter. Usually, this is the moment when the dark-matter particles become non-relativistic [21]. Naturally, the initial fluctuations at this moment are small:

$$|\delta_i(\boldsymbol{x})| \ll 1\ . \tag{12}$$

The above inequality allows us to consider the initial stage of the growth of fluctuations in the linear approximation. In combination with the spatial distribution given by expression (11), we can now define the initial fluctuations in terms of their Fourier spectrum:

$$|\delta_i(\boldsymbol{k})|^2\ , \qquad \delta_i(\boldsymbol{k}) = \int_{-\infty}^{+\infty} \delta_i(\boldsymbol{x})\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\ . \tag{13}$$

2.4 Fourier spectrum of initial fluctuations

The linear theory is given in Section 3 of this review. Here, we shall consider some general properties of the initial spectrum. The form of the spectrum $|\delta_i(\boldsymbol{k})|^2$ depends on the previous history of the formation of fluctuations and their growth at the time when $0 < t < t_i$. We shall consider first the initial spectrum $|\delta_i^0(\boldsymbol{k})|^2$. The actual form of this spectrum is determined by the physical processes which have occurred in the early Universe. Fairly general ideas on this topic were put forward by Zel'dovich [10] and Harrison [11].

In fact, from the physical point of view, it is natural to expect that variations of the metric are not strongly divergent. In the Newtonian approximation the metric is

$$\mathrm{d}s^2 = \left(1 + \frac{2\varphi}{c^2}\right)c^2\,\mathrm{d}t^2 - \left(1 - \frac{2\varphi}{c^2}\right)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)\ ,$$

where φ is the gravitational field potential. Hence, the variations of the metric are

$$|\delta g_{00}| = |\delta g_{\alpha\beta}| = 2\varphi c^2\ .$$

Consequently,

$$|\delta g_{ik}| \propto \int_{-\infty}^{+\infty} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x})\,\varphi_i^{(0)}(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}\ . \tag{14}$$

The condition for a weak (logarithmic) rms divergence of expression (14) is

$$|\varphi_i^{(0)}(\boldsymbol{k})| \propto \frac{1}{k^{3/2}}\ .$$

Then, the relationship between the potential φ_i and the density deviation δ_i from the system of equations (9) leads to

$$|\delta_i^0(\boldsymbol{k})|^2 \propto k\ . \tag{15}$$

In the limit $k \to 0$, the spectrum described by expression (15) is truncated naturally by the horizon scale R_{H}^{-1}.

In the case of smaller scales ($k \to \infty$), the truncation of the spectrum is determined by evolution at times $0 < t < t_i$. This process can be described by some 'transition' function $C(k)$, which is multiplied by the initial spectrum $|\delta_i^0(k)|$ and determines the spectrum at a moment t_i:

$$|\delta_i(\boldsymbol{k})|^2 = |\delta_i^0(\boldsymbol{k})|^2\,C(k)\ . \tag{16}$$

The function $C(k)$ basically distorts the 'initial' spectrum on a small scale and truncates it quite abruptly. The truncation is determined by the nature of the dark-matter particles. Since at present these particles are not known, we do not know the exact form of the function $C(k)$.

We shall however limit ourselves to such transition functions $C(k)$ which fall sufficiently rapidly in the limit $k \to \infty$ and we shall thus identify the smallest scale $k_{\max}$. We shall show later (Section 4) that during the nonlinear stage of the Jeans instability the dominant role is not played by the spectrum $|\delta_i(\boldsymbol{k})|^2$, but by the specific realisation of the initial fluctuations $\delta_i(\boldsymbol{x})$. A sufficiently rapid fall of the spectrum in the limit $k \to \infty$ implies a sufficiently smooth function $\delta_i(\boldsymbol{x})$.

It will be shown in Section 4 that the function $\delta_i(\boldsymbol{x})$ should have a second derivative. It then follows from the Fourier spectrum (13) that

$$\frac{\partial^2 \delta_i}{\partial x^2} \propto \int_{-\infty}^{+\infty} k^2\, \delta_i\,(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x})\, \mathrm{d}\boldsymbol{k}\ , \tag{17}$$

and hence we obtain the following constraints on the spectrum $|\delta_i(\boldsymbol{k})|^2$ when $k \gg k_{\max}$:

$$|\delta_i(\boldsymbol{k})|^2 \leqslant k^{-10}\,, \quad k \to \infty\ . \tag{18}$$

Condition (18) is always satisfied by cold dark matter.

Therefore, at low values of $|\boldsymbol{k}\,|$ the initial spectrum is close to that given by expression (15) and it decays rapidly in the limit $k \to \infty$. The smallest scale $k = k_{\max}$, corresponding to the maximum of the initial spectrum, is governed by the mass of the dark-matter particles [31]. It should be stressed that, as shown later (Section 4) the presence of a maximum in the spectrum and the rapid fall of the spectrum for $k > k_{\max}$ is — according to condition (18) — of fundamental importance for an evolving nonlinear structure.

The set of the initial conditions (11) and of the conditions imposed on the initial spectrum, described by expressions (15) and (18), satisfy all the requirements necessary for the development of a theory.

3. Linear growth of initial perturbations

As pointed out above, according to the latest observational data (particularly on the anisotropy of the relic radiation [15]), the fluctuations in the early Universe are very small [see expression (12)], so that their initial growth can be described by a linear theory. Linearisation of the system of equations (9) gives [21]

$$\begin{aligned} &\frac{\partial \boldsymbol{V}}{\partial t} + a^{-1}\,\dot{a}\,\boldsymbol{V} + a^{-1}\,\nabla\varphi = \boldsymbol{0}\ , \\ &\frac{\partial \delta}{\partial t} + a^{-1}\,\nabla\cdot\boldsymbol{V} = 0\ , \\ &\nabla^2\varphi = a^2\,\rho_0(t)\,\delta\ . \end{aligned} \tag{19}$$

The system of equations (19) can be solved quite rapidly. In fact, it follows from the second equation of this system that the density changes only under the influence of the velocity divergence. The rotational component of the velocity deduced from the first equation of the system (19) is

$$\boldsymbol{V}^{\mathrm{rot}} = \boldsymbol{V}_i^{\mathrm{rot}}\,\frac{a_i(t_i)}{a(t)}\ . \tag{20}$$

We can see that this rotational (vortex) component of the velocity always decays because of the overall cosmological expansion. Calculation of the divergence from the first equation in the system (19) and its substitution into the second equation, gives the following equation for the eigenmodes:

$$\frac{\mathrm{d}^2 D}{\mathrm{d}t^2} + 2\,\frac{\dot{a}(t)}{a(t)}\,\frac{\mathrm{d}D}{\mathrm{d}t} = \rho_0(t)\,D\ . \tag{21}$$

Eqn (21) describes two modes: a growing one D_1 and a decaying one D_2. In the absence of expansion ($\dot{a} = 0$, $\rho_0 = \mathrm{const}$) these modes would have been exponential. However, expansion of the Universe alters in a radical manner the time dependences of the modes. For example, if $\Omega = 1$ (where Ω is the ratio of the density to the critical value) the time dependences of the modes are described by a power law. In fact, substitution of $a(t) \propto t^{3/2}$ and, correspondingly, of $\rho_0 = (2/3)t^{-2}$ into Eqn (21) gives

$$D_1 = \left(\frac{t}{t_i}\right)^{2/3}, \quad D_2 = \frac{t_i}{t}\ . \tag{22}$$

It is evident from relationships (20) and (22) that, as stated above, only one mode in the linear solution grows with time.

The eigenmodes D_1 and D_2 can then be used to find the complete solution of the linear problem that satisfies the initial conditions (11) [21]:

$$\begin{aligned} \delta &= \frac{\delta_i}{E}\left\{D_1(t)\,\dot{D}_2(i) - D_2(t)\,\dot{D}_1(i)\right\} \\ &\quad + \frac{\nabla\cdot\boldsymbol{V}_i}{a_i E}\left\{D_1(t)\,D_2(i) - D_2(t)\,D_1(i)\right\}\ , \\ \boldsymbol{V} &= \frac{a(t)}{4\pi E}\int \mathrm{d}\boldsymbol{x}'\,\delta_i(\boldsymbol{x}')\,\frac{\boldsymbol{x}'-\boldsymbol{x}}{|\boldsymbol{x}-\boldsymbol{x}'|^3}\left\{\dot{D}_1(t)\,\dot{D}_2(i) - \dot{D}_2(t)\,\dot{D}_1(i)\right\} \\ &\quad + \boldsymbol{V}_i^{\mathrm{div}}\,\frac{a(t)}{a_i E}\left\{\dot{D}_2(t)\,D_1(i) - \dot{D}_1(t)\,\dot{D}_2(i)\right\} + \boldsymbol{V}_i^{\mathrm{rot}}\,\frac{a_i}{a(t)}\ . \end{aligned} \tag{23}$$

Here, a_i is the scaling factor at the moment t_i; the quantities $\delta_i(\boldsymbol{x})$ and $\boldsymbol{V}_i(\boldsymbol{x})$ are determined by the initial conditions (11); V_i^{div} is the divergent component of the velocity. The normalisation constant E is defined by the relationship

$$E = D_1(i)\,\dot{D}_2(i) - D_2(i)\,\dot{D}_1(i)\ ,$$

where the index i in expressions $D_1(i)$ and $D_2(i)$ means that these functions are taken at the moment t_i.

We shall now consider some important properties of the linear solution of the system of equations (23). First, we shall separate a growing mode:

$$\begin{aligned} \delta &= \frac{\delta_i}{E}\,D_1(t)\,\dot{D}_2(i) + \frac{\nabla\cdot\boldsymbol{V}_i}{a_i E}\,D_1(t)\,D_2(i)\ , \\ \boldsymbol{V} &= \frac{a(t)}{4\pi E}\int \mathrm{d}^3x'\,\delta_i\,(\boldsymbol{x}')\,\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^3}\,\dot{D}_1(t)\,\dot{D}_2(i) \\ &\quad - \boldsymbol{V}_i^{\mathrm{div}}\,\frac{a(t)}{a_i E}\,D_2(i)\,\dot{D}_1(t)\ . \end{aligned} \tag{24}$$

We can see that during the linear stage all the scales in a growing mode increase identically, i.e. there is no distortion of the initial spectrum of fluctuations in this mode and the amplitudes of all the harmonics increase with time proportionately to the eigenfunction D_1. Consequently, the nonlinear effects begin to play a role first on a scale

corresponding to the maximum $k_{\max}$ of the spectrum (16) of the initial fluctuations.

Moreover, out of all the possible initial conditions set by selecting four arbitrary functions (representing the density and three components of the velocity), only one very definite combination increases with time. In fact, it follows from the first equation of the system (23) that the growing mode $D_1(t)$ is associated with the following combination of the initial functions:

$$\delta_i(\boldsymbol{x})\,\dot{D}_2(i) + \frac{\nabla\cdot\boldsymbol{V}_i}{a_i}\,D_2(i)\ . \tag{25}$$

A simple transformation of the second equation in the system (23) readily shows that indeed the combination (25) increases with time.

Therefore, out of an arbitrary set of four scalar initial functions a much narrower class of functions increases with time and this class corresponds to just one scalar initial function. This initial function can be selected to be, for example, the gravitational field potential φ. Then the initial density and velocity of the growing mode can be expressed in terms of this potential:

$$\delta_i(\boldsymbol{x}) = \frac{\nabla^2\varphi_i}{a_i^2\rho_0(i)}\,, \qquad \boldsymbol{V}_i = -\frac{\nabla\varphi_i}{a_i\rho_0(i)}\,\frac{\dot{D}_1(i)}{D_1(i)}\ . \tag{26}$$

We can also use a different, more convenient, method of specifying the initial conditions [27]. This can be done by the following selection at $t = t_i$:

$$\delta|_{t=t_i} = \delta_i(\boldsymbol{x})\,, \qquad \boldsymbol{V}|_{t=t_i} = 0\ . \tag{27}$$

In fact, it is evident from Eqns (24) and (25) that, instead of the usual density $\delta_i(\boldsymbol{x})$, we can always introduce a new 'effective' density

$$\tilde{\delta}_i(\boldsymbol{x}) = \frac{1}{E}\left[\delta_i(\boldsymbol{x})\,\dot{D}_2(i) + \frac{\nabla\cdot\boldsymbol{V}_i}{a(i)}\,D_2(i)\right]\ .$$

Then, apart from relatively unimportant decaying terms, the solution of the system of equations (24) becomes

$$\delta(\boldsymbol{x},\,t) = \tilde{\delta}_i(\boldsymbol{x})\,D_1\,, \qquad \boldsymbol{V}(\boldsymbol{x},\,t) = \frac{a(t)}{4\pi E}\int \mathrm{d}^3x'\tilde{\delta}_i(\boldsymbol{x}')$$
$$\times\,\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^3}\left[\dot{D}_1(t)\,\dot{D}_2(i) - \dot{D}_2(t)\,\dot{D}_1(i)\right]\ .$$

It should be noted that at the initial moment $t = t_i = i$ the term in the braces is identically equal to zero. The solution obtained is thus equivalent to the solution of Eqns (23) and (24) where the initial velocities are assumed to be zero.

It therefore follows that only one growing mode is obtained by solving the linear problem for arbitrary selected initial conditions. The initial conditions specified by expression (27) complete the solution of the problem for this mode. In future. when we shall develop a nonlinear theory, we shall begin with the initial conditions given by expression (27) and we shall study the nonlinear evolution of the growing unstable mode. The influence of the decaying mode on the nonlinear dynamics of the growing mode is of little importance. This particular point will be considered in the Appendix.

4. Nonlinear growth of the Jeans instability

4.1 Appearance of the initial singularity

In the preceding section we discussed the linear theory and demonstrated that in a cold nondissipative gas, considered in the linear approximation, the amplitude of the initial fluctuations in an unstable mode increases without distortion of the initial profile. This means that the linear approximation breaks down first and the nonlinear effects become important at those points where the initial density perturbations are largest, i.e. near each maximum of the initial distribution of the effective density $\delta_i(\boldsymbol{x})$. Therefore, in the investigation of the nonlinear stage of the Jeans instability it is important to consider first the dynamics of the system in the vicinity of a single maximum.

We shall solve the problem by selecting the system of coordinates at a point corresponding to some arbitrary initial density maximum. We shall bear in mind that the first to reach the nonlinear stage of growth are the inhomogeneities with the scale of $k_{\max}$. We shall begin our analysis from these inhomogeneities. A special feature of the dynamics of these inhomogeneities is that large scales ($k \ll k_{\max}$) have not yet been reached and smaller fluctuations are completely absent because of the rapid decay of the initial spectrum (16) for $k > k_{\max}$. Therefore, we can consider them as separate smooth maxima.

Transformation of the coordinates and velocities in accordance with expressions (10), and replacement of an expanding system of coordinates with one at rest and with its origin at the point of the density maximum δ_i, transform the system of equations (9) to

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial\boldsymbol{r}}\,(\rho\boldsymbol{u}) = 0\ ,$$
$$\frac{\partial\boldsymbol{u}}{\partial t} + \left(\boldsymbol{u}\cdot\frac{\partial}{\partial\boldsymbol{r}}\right)\boldsymbol{u} + \frac{\partial\psi}{\partial\boldsymbol{r}} = 0\ , \tag{28}$$
$$\nabla^2\psi = \rho\ .$$

In the vicinity of a density maximum when the conditions (15) and (18) are satisfied, the initial density distribution can be represented, apart from terms of higher orders of smallness, in the form

$$\rho(\boldsymbol{r}) = \rho_0(1-\xi^2)\,, \qquad \xi^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\ . \tag{29}$$

The coefficients a, b, and c in expressions (29) are, respectively, the terms of a Taylor expansion of the density near the point of the maximum $\boldsymbol{r} = \boldsymbol{0}$:

$$a^{-2} = \frac{1}{2}\,\frac{\partial^2\rho_i}{\partial x^2}\,, \qquad b^{-2} = \frac{1}{2}\,\frac{\partial^2\rho_i}{\partial y^2}\,, \quad c^{-2} = \frac{1}{2}\,\frac{\partial^2\rho_i}{\partial z^2}\ .$$

We can assume that the following inequalities are always satisfied:

$$a \geqslant b \geqslant c\ .$$

The condition specified by the first formula in expression (29), selected at the moment $t = t_i$, is one of the initial conditions for the system of equations (28). The other initial condition is selected, in accordance with expression (27), in the form

$$\boldsymbol{V}_i(\boldsymbol{r}) = \boldsymbol{0}\ . \tag{30}$$

4.2 One-dimensional flow

We shall now consider the solutions of the system of equations (28) subject to the initial conditions (29), (30). The relationships between the coefficients a, b, and c in expression (29) can be arbitrary. We shall begin with the limiting case

$$\frac{b}{a} \ll 1 , \quad \frac{c}{a} \ll 1 .$$

Then, in the first approximation with respect to the small parameters b/a and c/a, the problem can be regarded as one-dimensional. The system of equations (28) then becomes

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) = 0 ,$$
$$\frac{\partial u}{\partial t} + \left(u \frac{\partial}{\partial x}\right) u + \frac{\partial \Psi}{\partial x} = 0 , \tag{31}$$
$$\frac{\partial^2}{\partial x^2} \psi = \rho ,$$

and the initial conditions (29), (30) can be written in the form

$$t = 0 , \quad V_i = 0 , \quad \rho_i(x) = \rho_0 \left(1 - \frac{x^2}{a^2}\right) . \tag{32}$$

In the above expressions, time is measured from $t = t_i$.

The system of equations (31) can be integrated exactly (26). In fact, introducing $y = \partial\psi/\partial x$ and substituting the third equation into the first, we find that integration with respect to x gives

$$\frac{\partial y}{\partial t} + u \frac{\partial y}{\partial x} = c(t) , \tag{33}$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + y = 0 . \tag{34}$$

Eqn (33) describes the mass transport law since $y(x) = \int_0^x \rho(x_1)\,\mathrm{d}x_1 = m(x)$ is the mass of matter in the interval $(0, x)$. Therefore, if the total mass is conserved in the system, the constant $c(t) \equiv 0$.

Application of the hodograph transformation, i.e. the assumption that $t = t(u, y)$ and $x = (u, y)$, and substitution of the variables from expressions (33) and (34) gives

$$-\frac{\partial x}{\partial u} + u \frac{\partial t}{\partial u} = 0 , \quad \left(\frac{\partial x}{\partial y} - u \frac{\partial t}{\partial y}\right)\left(1 + y \frac{\partial t}{\partial u}\right) = 0 . \tag{35}$$

The second equation of the system (35) can be separated into two: either

$$\frac{\partial x}{\partial y} - u \frac{\partial t}{\partial y} = 0$$

or

$$\frac{\partial t}{\partial u} = -\frac{1}{y} .$$

It is easy to show that the first equation is degenerate and does not satisfy the initial conditions (26). It follows implicitly from the second equation that

$$x = -\frac{u^2}{2y} + H_1(y) , \quad t = -\frac{u}{y} + M_1(y) . \tag{36}$$

The functions $H_1(y)$ and $M_1(y)$ are determined by the initial conditions (29), (30):

$$M_1(y) = 0 , \quad x = H_1(y) . \tag{37}$$

Solutions of the system of equations (37) in the vicinity of a density maximum gives

$$H_1(y) = \frac{y}{\rho_0} + \frac{y^3}{3\rho_0^3 a^2} \tag{38}$$

[see Eqn (34)]. It therefore follows from relationships (36) and (38) that the solution is

$$x = \left(\frac{1}{\rho_0} - \frac{t^2}{2}\right) y + \frac{y^3}{3\rho_0^3 a^2} . \tag{39}$$

For short times $[t < (2/\rho_0)^{1/2}]$, it follows from expression (39) that the solution in the vicinity of a density maximum ($x = 0$) becomes

$$\rho = \frac{1}{1/\rho_0 - t^2/2} - \frac{1}{\rho_0^3 a^3} \frac{x^2}{(1/\rho_0 - t^2/2)^4} ,$$
$$u = -\frac{xt}{1/\rho_0 - t^2/2} + \frac{t}{3\rho_0^3 a^2} \frac{x^3}{(1/\rho_0 - t^2/2)^4} . \tag{40}$$

We can see that the density at the $x = 0$ maximum increases rapidly with time t, the peak becomes narrower, and flow towards the centre appears.

It is important to stress that the solution represented by expressions (36) and (39) exists only for a finite time. In fact, a flow singularity appears at time $t = t_c = (2/\rho_0)^{1/2}$:

$$\rho = \rho_0 \left(\frac{3x}{a}\right)^{-2/3} , \quad \psi = \frac{\rho_0 a^2}{4} \left(\frac{3x}{a}\right)^{4/3} ,$$
$$u = (2\rho_0)^{1/2} a \left(\frac{3x}{a}\right)^{1/3} , \quad t = t_c = \left(\frac{2}{\rho_0}\right)^{1/2} . \tag{41}$$

The density tends to infinity in the limit $x \to 0$. The velocity and the potential remain finite and the expression for the singularity contains only their derivatives. The appearance of a singularity in one-dimensional flow of a cold nondissipative gravitating gas was pointed out by Zel'dovich and Arnol'd [23, 24].

4.3 Spherically symmetric flow

We shall now consider another important limiting case, that of the spherically symmetric geometry when the parameters a, b, and c are all equal:

$$a = b = c , \quad \rho_i(r) = \rho_0 \left(1 - \frac{r^2}{a^2}\right) .$$

The system of equations (28) then becomes

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 \rho u) = 0 ,$$
$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial r} + \frac{\partial \psi}{\partial r} = 0 , \tag{42}$$
$$\frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) = \rho .$$

On introduction of $y = r^2 \partial\psi/\partial r$, we find from the system of equations (42) that

$$\frac{\partial y}{\partial t} + u\,\frac{\partial y}{\partial r} = c(t)\,, \qquad \frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial r} + \frac{y}{r^2} = 0\,. \tag{43}$$

Calculations analogous to those represented by the systems of equations (33)–(35) yield the following implicit solutions [26]:

$$t = (2y)^{-1/2} H^{-3/2}(y)\left(\arctan Z + \frac{Z}{1+Z^2}\right)\,,$$
$$r = \frac{1}{H(y)(1+Z^2)}\,, \qquad Z = -\frac{u}{\left[2yH(y)\right]^{1/2}}\,. \tag{44}$$

The expression for the function $H(y)$ in the vicinity of a maximum is found by analogy with expression (38):

$$H(y) = \left(\frac{y}{y_0}\right)^{-1/3}\left[1 - \frac{1}{5}\left(\frac{y}{y_0}\right)^{2/3}\right], \qquad y_0 = \frac{\rho_0 a^3}{3}\,. \tag{45}$$

An analysis of expressions (44) and (45) shows that a singularity again appears and this time this happens at a moment $t = t_c^{(1)} = \pi(3/8\rho_0)^{1/2}$ in the vicinity of the point $r = 0$:

$$\rho = \frac{3}{7}\left(\frac{40}{9\pi}\right)^{6/7} \rho_0 \left(\frac{r}{a}\right)^{-12/7}, \qquad \psi = \frac{7}{2}\,\beta\left(\frac{r}{a}\right)^{2/7},$$
$$u_r = -(2\beta)^{1/2}\left(\frac{r}{a}\right)^{1/7}, \qquad \beta = \frac{\rho_0 a^2}{3}\left(\frac{40}{9\pi}\right)^{6/7}. \tag{46}$$

The singularity described by expressions (46) is integrable. In fact, the potential ψ is finite and the mass function is $M(r) \propto r^{9/7}$, i.e. $M = 0$ in the limit $r \to 0$. This means that, in spite of the presence of a singularity of the density ρ at the centre, a black hole does not form since the mass does not rise sufficiently rapidly (in the case of a black hole, we should have $M \propto r$).

However, it should be stressed that, if we select a flatter initial density distribution $\rho_i(r)$

$$\rho_i(r) = \rho_0\left[1 - \left(\frac{r}{a_1}\right)^{2n}\right], \qquad n \geqslant 2\,, \tag{47}$$

a black hole forms at the same moment as a singularity at the centre. In fact, in the limit $r \to 0$, the density is

$$\rho(r) \propto r^{-12n/(3+4n)}\,. \tag{48}$$

This means that $M(r) \propto r^{9/(3+4n)}$ in the limit $r \to 0$, i.e. if $n \geqslant 2$, the mass at the centre rises more rapidly than in the case of a black hole. We can demonstrate that expressions (47) and (48) are valid also in the general relativity case.

It therefore follows that if the initial distribution of the density of nondissipative matter in the spherically symmetric case is sufficiently flat, nonlinear compression of this matter creates a black hole at the moment of appearance of the initial singularity. A black hole does not form in the case of the usual nondegenerate initial distribution described by expression (29).

4.4 General case

Finally, let us consider the solution of the system of equations (28) with the general initial conditions (29), (30). We shall introduce parameters ε_1 and ε_2, representing deviations from the spherical symmetry:

$$\varepsilon_1 = \frac{a-b}{a}\,, \qquad \varepsilon_2 = \frac{a-c}{a}\,.$$

The parameters ε_1 and ε_2 vary within the limits

$$0 \leqslant \varepsilon_1 \leqslant 1\,, \qquad 0 \leqslant \varepsilon_2 \leqslant 1\,.$$

It follows from the condition $a \geqslant b \geqslant c$ that we always have $\varepsilon_2 \geqslant \varepsilon_1$.

We shall consider first the case which does not differ too greatly from the spherically symmetric geometry [28]:

$$\varepsilon_1 \ll 1\,, \qquad \varepsilon_2 \ll 1\,.$$

The system of equations (28) can then be rewritten conveniently in the form

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial \boldsymbol{r}}(\rho\boldsymbol{u}) = 0\,,$$
$$\frac{\partial \boldsymbol{u}}{\partial t} + \left(\boldsymbol{u}\cdot\frac{\partial}{\partial \boldsymbol{r}}\right)\boldsymbol{u} + \boldsymbol{F} - \tilde{\boldsymbol{F}} = \boldsymbol{0}\,, \tag{49}$$
$$\operatorname{div}\boldsymbol{F} = \rho\,, \qquad \operatorname{div}\tilde{\boldsymbol{F}} = 0\,.$$

Here, the gradient of the potential $\nabla\psi$ is represented as the difference between the forces $\boldsymbol{F} - \tilde{\boldsymbol{F}}$. The forces $\tilde{\boldsymbol{F}}$ is proportional to the small parameters ε_1 and ε_2. In the spherically symmetric case, we have $\tilde{\boldsymbol{F}} = 0$.

Ignoring, in the first approximation, the force $\tilde{\boldsymbol{F}}$ and introducing the notation

$$\boldsymbol{u} = \frac{\boldsymbol{r}}{\xi}\,U(\xi, t)\,, \qquad \boldsymbol{F} = \frac{\boldsymbol{r}}{\xi}\,B(\xi, t)\,, \tag{50}$$

we find from the system of equations (49) that

$$\frac{\partial B}{\partial t} + \frac{U}{\xi^2}\frac{\partial}{\partial\xi}(\xi^2 B) = 0\,, \qquad \frac{\partial U}{\partial t} + U\,\frac{\partial U}{\partial\xi} + B = 0\,, \tag{51}$$

where ξ are the ellipsoidal coordinates defined in accordance with expression (29).

The system of equations (51) is analogous to the spherically symmetric system (43) with the radius r replaced by the variable ξ. It follows that the solution of the system (51) is known: it is exactly the same as the solution of the system (44) if the substitution $r \to \xi$ is made. The only difference is this: it follows from expression (50) that $\operatorname{curl}\boldsymbol{F}$ no longer vanishes because of the smallness of the parameters ε_1 and ε_2. The angular momentum is included by assuming that the force $\tilde{\boldsymbol{F}}$ in the system of equations (49) is such that $\operatorname{div}\tilde{\boldsymbol{F}} = 0$. The force $\tilde{\boldsymbol{F}}$ balances out exactly the rotational (vortex) component of the force $\boldsymbol{F}$.

We shall now take account of the presence of the force $\tilde{\boldsymbol{F}}$. The solution of the system of equations (51) makes it possible to represent the velocity in the form

$$\boldsymbol{u} = \frac{\boldsymbol{r}}{\xi}\,U(\xi, t) + v\,. \tag{52}$$

Substituting expression (52) into the system of equations (49), bearing in mind that $\boldsymbol{r}\times\nabla\xi \sim \varepsilon_1 \sim \varepsilon_2$, and taking the curl of the first equation, we find to within $O(\varepsilon_2^2)$ and $O(\varepsilon_2^2)$ that

$$\frac{\partial\boldsymbol{\omega}}{\partial t} = \operatorname{curl}\left(\frac{\boldsymbol{r}}{\xi}\,U \times v\right) - \operatorname{curl}\tilde{\boldsymbol{F}}\,, \tag{53}$$

where $\boldsymbol{\omega} = \operatorname{curl} v$, $\operatorname{curl}\tilde{\boldsymbol{F}} = \boldsymbol{r}\times\nabla(B/\xi)$.

The solution of the linear equation (53) — subject to the initial conditions (29), (30) — can be expressed in terms of the solution of the system (51). The new solution has the simple form

$$\boldsymbol{\omega} = -\boldsymbol{r} \times \nabla \frac{U}{\xi} , \tag{54}$$

which can easily be demonstrated by direct substitution of expression (54) into Eqn (53), and by several vector operations. It is clear from expression (54) that the value of $|\boldsymbol{\omega}|$ is of the order of $\varepsilon_1 \sim \varepsilon_2$. Therefore, at the moment of appearance of a singularity $t = t_g^{(1)}$ when — according to expression (46) — the radial velocity V_r is

$$V_r = -(2\beta)^{1/2} \left(\frac{r}{a}\right)^{1/7} ,$$

the velocities $\boldsymbol{u}_\perp$ transverse to the radius $\boldsymbol{r}$ remain small (of the same order of magnitude). In fact, it follows from expression (54) that

$$\begin{aligned} u_{x\perp} &= \frac{3}{2} V_r \frac{x}{r} (\varepsilon_1 y^2 + \varepsilon_2 z^2) , \\ u_{y\perp} &= \frac{3}{2} V_r \frac{y}{r} \varepsilon_1 , \quad u_{z\perp} = \frac{3}{2} V_r \frac{z}{r} \varepsilon_2 . \end{aligned} \tag{55}$$

It is evident from expressions (55) for these velocities that near a singularity the transverse components are not equivalent. It means that the compression is asymmetric. The velocity is lowest along the major semiaxis a of the initial ellipsoid. It follows from the above expressions that this velocity can be ignored.

Let us now identify the direction along which the compression rate is maximal. We shall do this by introducing the function

$$g = \frac{V^2}{V_r^2} = 1 + \frac{9}{4} \varepsilon_1^2 \frac{y^2}{r^2} + \varepsilon_2^2 \frac{z^2}{r^2} , \tag{56}$$

where $\varepsilon_1 > \varepsilon_2$ since $a > b > c$. Rewriting relationship (56) in polar coordinates, we shall find the points of extrema of the function $g(\theta, \varphi)$:

$$\theta = 0 , \quad \theta = \frac{\pi}{2} , \quad \varphi = 0 , \quad \cos^2 \varphi = \frac{\varepsilon_2^2}{\varepsilon_1^2} . \tag{57}$$

The last expression in formula (57) cannot be correct because $\varepsilon_2 > \varepsilon_1$. We can therefore easily see that $g_{\max} = 1 + \varepsilon_2^2$ is reached at $\theta = 0$. Therefore, the highest velocity is attained along the z axis, i.e. along the semiminor axis of the initial ellipsoid. This is in full agreement with the results of Zel'dovich [43] showing that the strongest compression occurs along the shortest of the axes of the initial ellipsoid.

The transverse velocity $\boldsymbol{u}_\perp$ determines the density $\boldsymbol{m}$ of the angular momentum of a gas relative to an inhomogeneity centre, because

$$\boldsymbol{m} = \boldsymbol{r} \times \boldsymbol{u} = \boldsymbol{r} \times \boldsymbol{u}_\perp .$$

It follows from the law of conservation of the angular momentum that the total (integral) momentum $\boldsymbol{M}$ naturally remains always zero. However, the mean square of the angular momentum is not conserved: at the moment when a singularity appears, it has a finite value proportional to ε^2. The density of the square of the angular momentum is then

$$m^2 = \frac{9}{4} V_r^2 r^2 \left[(\varepsilon_2 - \varepsilon_1)^2 \left(\frac{yz}{r^2}\right)^2 + \varepsilon_2^2 \left(\frac{xz}{r^2}\right)^2 + \varepsilon_1^2 \left(\frac{xy}{r^2}\right)^2 \right] .$$

4.5 Region near a singularity. General comments

It is evident from the above discussion that the evolution of the nonlinear solution of the system of equations (28) leads to the appearance of a singularity after a finite time. A description of the dynamics of the system during the subsequent stages can be provided if the system of equations (28) is supplemented by the rules of passage through a singularity. These rules are governed by the physical properties of the investigated system.

In the Euler approximation the system of equations (28) describes equally the hydrodynamics of ordinary (baryonic) matter and of nondissipative (dark) matter. However, after the appearance of a singularity the two cases differ fundamentally. In an ordinary gas the dominant role at a singularity and beyond it is played by dissipative processes. The dissipation, proportional to higher derivatives of the velocity is not important before the appearance of a singularity, but at the singularity the velocity gradient becomes $|\nabla \boldsymbol{V}| \to \infty$ and it begins to play the dominant role.

The dissipation means that the motion remains always of single-stream nature and is described by the system of equations (28). These equations should be supplemented by large discontinuities (shock waves) at which energy dissipation takes place, and also by a change in temperature which alters the pressure of the gas heated due to dissipation in the shock waves [32].

A different situation applies in the case of completely dissipation-free (dark) matter. There is now no dissipation so that the velocity can have any gradient. Moreover, special macroscopic flows are also possible when at the same point $\boldsymbol{r}$ in space there are several noninteracting streams travelling at different velocities $\boldsymbol{V}_i(\boldsymbol{r}, t)$. It is the passage of the initial singularity of the Jeans instability in nondissipative matter which is responsible for the appearance of multistream flows.

5. Multistream flows

It is natural to use the kinetic system of equations (7) in the description of multistream flows. Here $f(\boldsymbol{r}, v, t)$ is understood to be a function of the distribution of streams of matter in the velocity space. In other words, if there is one stream, the distribution function $f(\boldsymbol{r}, v, t)$ is defined by expression (8). If there are n streams, then

$$f(\boldsymbol{r}, v, t) = \sum_{i=1}^{n} \rho_i(\boldsymbol{r}, t)\, \delta\left[v - \boldsymbol{V}_i(\boldsymbol{r}, t)\right] . \tag{58}$$

It should be stressed that the adoption of the kinetic description represented by expressions (7) and (58) is justified by the equivalence of the gravitational and inertial masses: only because of this equivalence is the term $(\partial\psi/\partial\boldsymbol{r}) \cdot \partial f/\partial v$, describing the interaction with the field in expression (7), independent of the density of each of the streams ρ_i [26].

The gravitational forces acting on each of the streams are proportional to its density ρ_i (gravitational mass), but the inertial forces are also proportional to the density ρ_i (inertial mass). Consequently, the density ρ_i drops out of the system of kinetic equations (7). Consequently, the system (7) for a number of streams which have different densities ρ_i is exactly the same as the kinetic equation for a system of identical particles.

5.1 Solution near a singularity. Planar one-dimensional case

We shall now consider the behaviour of the solution near a singularity in the one-dimensional case. As pointed out already, in this case we should not use the system of equations (28), but the more general system of kinetic equations (7) and (58). Up to the moment of appearance of a singularity $t = t_c = (2/\rho_0)^{1/2}$, the flow is of single-stream nature and the solution (36) is still valid.

The solution of the kinetic system of equations (7) and (58) can be represented in the form of expression (8) with the parameters $\rho(x, t)$, $V(x, t)$ defined in accordance with formulas (36) and (41). Therefore, if we use expressions (8) and (41), we find that at the singularity $t = t_c$ we have

$$f(x, v, t)_{t=t_c} = \rho_1 x^{-2/3}\,\delta(v + v_1 x^{1/3}) ,$$
$$\rho_1 = \rho_0 \left(\frac{a}{3}\right)^{2/3} , \qquad v_1 = (2\rho_0)^{1/2}\,(3a^2)^{1/3} . \tag{59}$$

The solution described by the above expressions is naturally valid only if $x \ll a$.

Let us now consider the solution directly beyond the singularity i.e. in the case when $\tau > 0$, where $\tau = t - t_c$, $\tau \ll t_c$. We shall bear in mind that in the region of interest to us, $x \ll a$, the kinetic energy $K = v^2/2$ of a stream is considerably higher than its potential energy:

$$K = \frac{v^2}{2} = \frac{1}{2}\,v_1^2 x^{2/3} , \qquad \psi = \frac{1}{4}\,\rho_0 a^2 \left(\frac{3x}{a}\right)^{4/3} .$$

This means that the influence of the potential in the kinetic equation is not very great, so that in the first approximation we can ignore the potential. Then the solution of the kinetic system of equations (7) has the simple form

$$f(x, v, \tau) = f(x - v\tau, v, 0) . \tag{60}$$

It follows from expressions (59) and (60) that

$$f(x, v, \tau) = \rho_1 (x - v\tau)^{-2/3}\,\delta\left[v + v_1 (x - v\tau)^{1/3}\right] . \tag{61}$$

It is important to note that the argument of the δ function has one root for $|x| > x_c$, but for

$$|x| < x_c , \qquad x_c = \frac{2}{3\sqrt{3}}\,(v_1 \tau)^{3/2} \tag{62}$$

it has three roots. It then follows that three-stream flow appears in the vicinity of a singularity when $\tau > 0$. The distribution function (61) can then be rewritten in the form given by expression (58):

$$f(x, v, \tau) = \sum_{i=1}^{3} \rho_i(x, \tau)\,\delta\left[v + v_i(x, \tau)\right] , \tag{63}$$

where the quantity

$$\rho_i = \rho_1 \left[x - v_i(\tau)\right]^{-2/3} \left| \frac{\mathrm{d}\eta(v_i)}{\mathrm{d}v_i}\,(v_i) \right|^{-1}$$

represents the density of the ith stream and $v_i = v_i(x, \tau)$ are the roots of the equation

$$\eta(v_i) = v_i + v_1 (x - v_i \tau)^{1/3} = 0 . \tag{64}$$

Relationships (58) together with the system of kinetic equations (7) are together equivalent to introduction of multistream hydrodynamics. In fact, use of representation of the distribution function in the form of expression (58) leads to modification of the system of equations (7) to a new system

$$\frac{\partial \rho_i}{\partial t} + \frac{\partial}{\partial x}\,(\rho_i v_i) = 0 ,$$
$$\frac{\partial v_i}{\partial t} + \left(v_i \frac{\partial}{\partial x}\right) v_i + \frac{\partial \psi}{\partial x} = 0 , \tag{65}$$
$$\frac{\partial^2}{\partial x^2}\,\psi = \sum_i \rho_i .$$

The caustics, i.e. the surfaces on which streams [the roots of Eqn (64)] coalesce or multiply, are important in the systems of equations (65) and (7). The caustics correspond to the points x_c.

It in fact follows from expression (61) that if $|x| > x_c$, there are three streams, but only one stream if $|x| < x_c$. At the points x_c the derivatives of the coalescing velocities v_1 and v_2 have singularities:

$$\left.\frac{\partial v_1}{\partial x}\right|_{x=x_c} \to -\infty , \qquad \left.\frac{\partial v_2}{\partial x}\right|_{x=x_c} \to +\infty . \tag{66}$$

The stream densities ρ_1 and ρ_2, like the second derivative of the field $\partial^2\psi/\partial x^2$, also become infinite at the points $\pm x_c$. However, the potential itself and its first derivative remain continuous:

$$\rho_1\big|_{x=x_c} = \rho_2\big|_{x=x_c} = \frac{1}{2}\,\frac{\partial^2 \psi}{\partial x^2}\bigg|_{x=x_c}$$
$$= C(x_c, \tau)\left[x_c(\tau) - x\right]^{-1/2} ,$$
$$\left.\frac{\partial \psi}{\partial x}\right|_{x \to x_c - 0} = \left.\frac{\partial \psi}{\partial x}\right|_{x \to x_c + 0} , \qquad \psi(x_c + 0) = \psi(x_c - 0) . \tag{67}$$

The conditions described by expressions (66) and (67) apply at each point of coalescence of the streams. They represent a complete system of the boundary conditions for the system of equations (65) describing multistream hydrodynamics.

5.2 Numerical simulation

The system of equations (7) had been integrated numerically by the particle-in-a-cell method [33]. Numerical

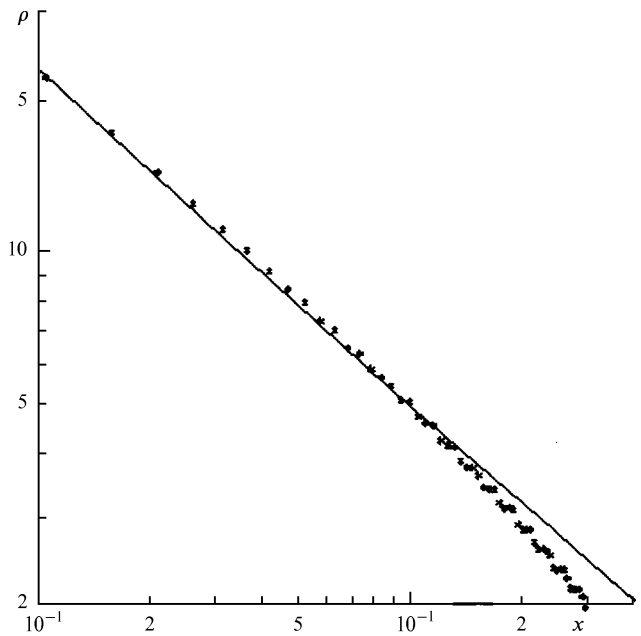

Figure 1. Distribution of the density $\rho(x)$ at the moment of appearance of a singularity.

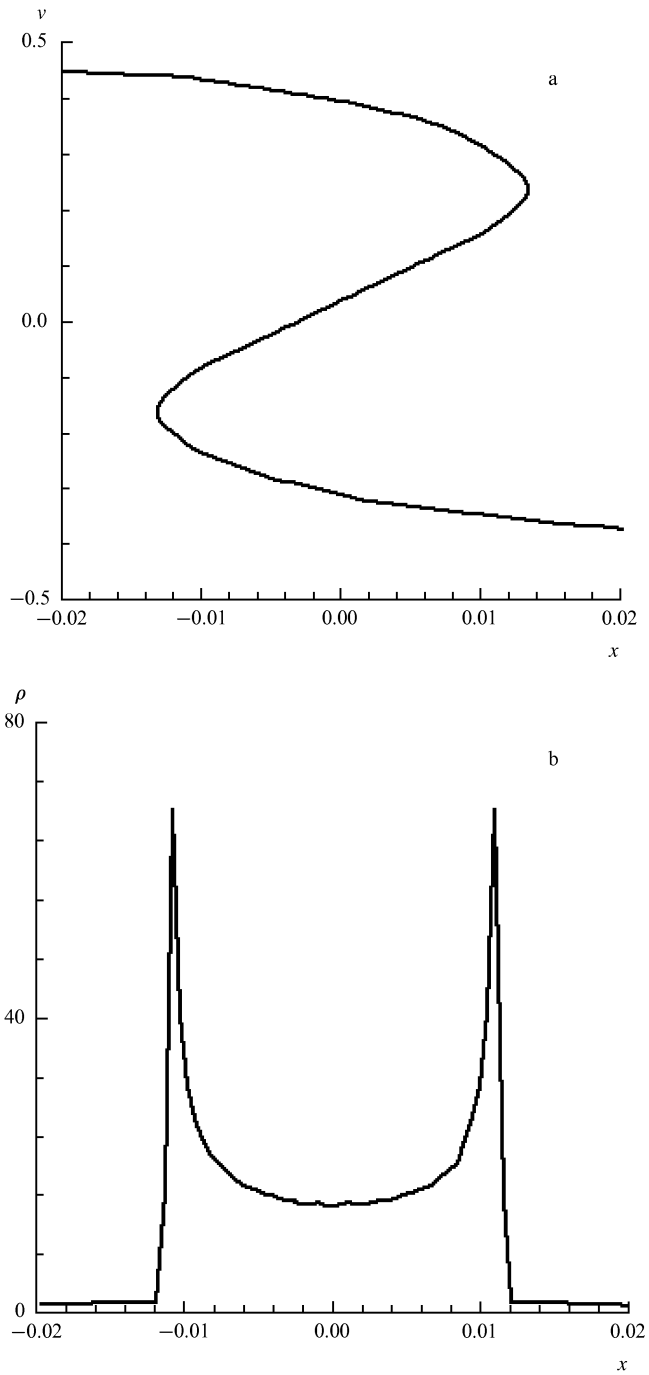

Figure 2. (a) Appearance of a region of three-stream flow. (b) Distribution of the density $\rho(x)$ in three-stream flow.

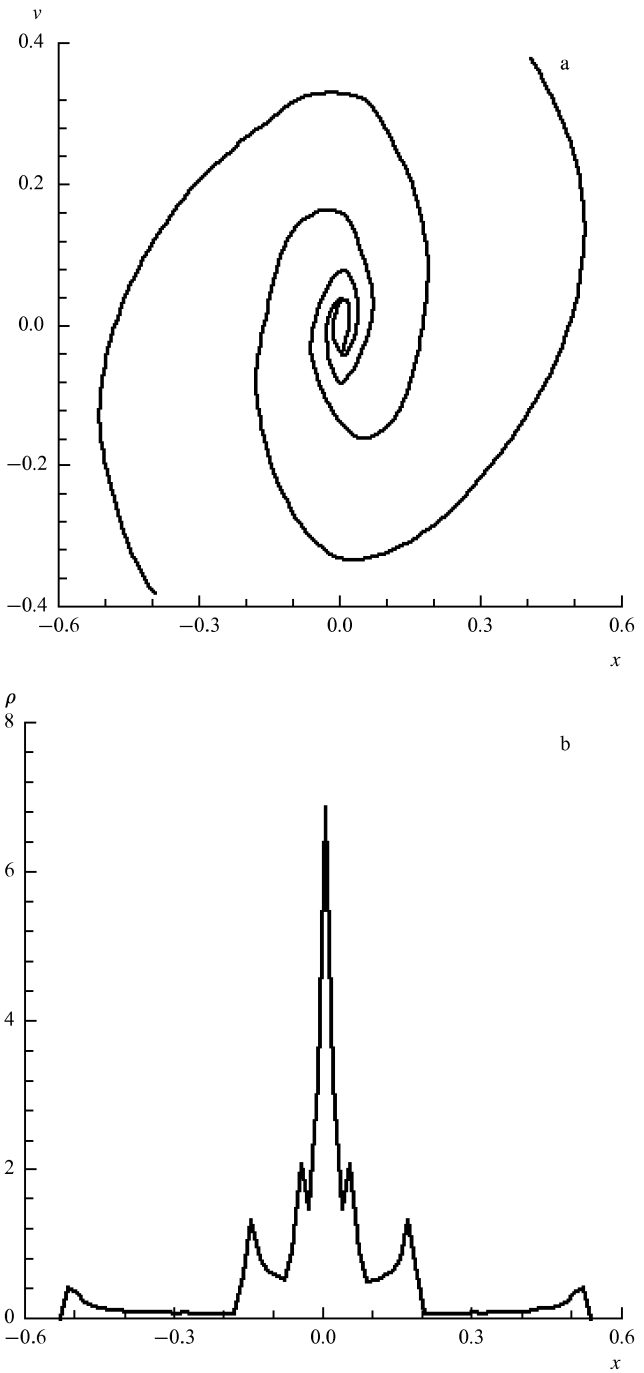

Figure 3. (a) Phase space of five-stream flow. (b) Density distribution $\rho(x)$ in five-stream flow.

calculations were carried out for $(1-2)\times 10^5$ particles. The solutions are presented in Figs 1–3. Fig. 1 shows the distribution of the particle density $\rho(x)$ at the moment $t = t_c$ when a singularity appears. We can readily see from Fig. 1 that the results of numerical calculations (points) are in agreement with the analytic expression (41).

Fig. 2 demonstrates the formation of three-stream flow. We can see clearly the caustic singularities described by expression (66) and the behaviour of the density near the caustics described by the system of equations (67). During the subsequent evolution the field gradually slows down the central stream, which stops and begins to move in the opposite direction; at a certain moment this stream is reversed, i.e. a singularity analogous to that described by expression (41) again appears at the centre of the distribution. This is followed by the formation of a five-stream flow zone near the centre.

Naturally, caustics separating the five-stream and three-stream flow zones also appear. In the $(x-v)$ space the flow is in the form of a spiral twisted towards the centre (Fig. 3a). On each of the caustics the density still has a singularity described by expression (67) (Fig. 3b). The spiral subsequently twists tighter and tighter, and the number of the streams near the centre increases. Regions with different numbers of streams are separated by caustics. The result is that the density function is strongly jagged, but

the integral of the density m and the potential ψ still remain fairly smooth functions described by expression (67).

We shall now consider the state established as a result of prolonged mixing in the limit $t \to \infty$. As pointed out above, at the moment of appearance of the initial singularity the kinetic energy near the singularity is much higher than the potential energy. Consequently, there is no local trapping of matter, so that immediately after the initial reversal the singularity at the centre disappears (Fig. 2b) and the density dis-tribution becomes smooth in the vicinity of the point $x = 0$.

It would be natural to expect that also after prolonged mixing a similar smooth density distribution would also appear:

$$\rho = \rho_* \left[1 - \left(\frac{x}{a_*} \right)^2 \right] , \tag{68}$$

which would differ from the initial density distribution only by a change in the constants ρ_* and a_*. However, this is incorrect.

A numerical calculation shows that after a long time there is accumulation of matter at the centre. The density $\rho(x)$ fluctuates strongly in view of relationships (67). Therefore, such accumulation can be seen more clearly in the case of the integral density curve, i.e. when the mass of matter is

$$m(x) = \int_0^x \rho(x_1)\, \mathrm{d}x_1 ,$$

as shown in Fig. 4. For values $t \geqslant 10t_c$ the curve plotted in Fig. 4 practically ceases to vary: it thus shows the steady-state distribution of the integral density. We can see that it differs considerably from the dependence $m(x) = \rho_* x$ represented by formula (68).

It is also evident from Fig. 4 that in this steady state the average density $\bar{\rho}(x) = d\bar{m}/\mathrm{d}x$ rises strongly on approach to the centre ($x = 0$). Moreover, in this region the fluctuations of the density ρ becomes stronger, indicating accumulation of the caustic singularities and a considerable deviation of the mean field potential from the law $\psi \propto x^2$ in the limit $x \to 0$ [this law follows from formula (68)].

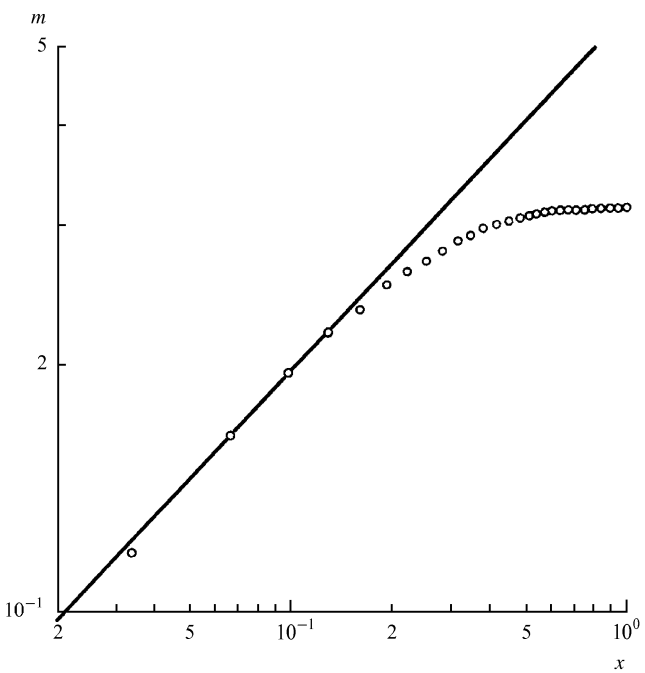

Figure 4. Integral density $m(x)$ at times $t > 10t_c$.

5.3 Adiabatic approximation

The nature of the singularities predicted by the numerical solution can be identified by analytic consideration of the process of multistream mixing after a long time t. Then, instead of the variable v, it is convenient to introduce a new variable into the system of kinetic equations (7); this variable is the adiabatic invariant

$$I = \int v\, \mathrm{d}x = \int_{-x_m}^{x_m} \left[2(E - \psi) \right]^{1/2} \mathrm{d}x . \tag{69}$$

Here, $E = v^2/2 + \psi$ is the energy of a stream and the reflection points $\pm x_m$ are defined by the condition $E = \psi(x_m)$. For a given potential, the definition (69) establishes a unique relationship between the energy E and the invariant I:

$$E = E(I) .$$

If, in accordance with expression (69), we adopt new variables in the system of equations (7), the result is

$$\begin{gathered} \frac{\partial f}{\partial t} \pm \left[2(E - \psi) \right]^{1/2} \frac{\partial f}{\partial x} \\ + \int_{-x_m}^{x_m} \frac{\partial\psi(x, t)/\partial t - \partial\psi(x_1, t)/\partial t}{\left\{ 2\left[E - \psi(x_1, t) \right] \right\}^{1/2}} \, \mathrm{d}x_1 \frac{\partial f}{\partial I} = 0 , \\ \frac{\partial^2 \psi}{\partial x^2} = \int_{\psi}^{E_m} \frac{f}{\left[2(E - \psi) \right]^{1/2}} \, \mathrm{d}E . \end{gathered} \tag{70}$$

We shall now consider the approximation which will be called the adiabatic model. This model can be described as follows. The process of mixing involves oscillations of the streams and the appearance of multistream flows. The adiabatic invariant I is conserved in these oscillations if they appear against the background of a sufficiently slowly varying potential. The distribution function of the adiabatic invariants $f(I)$ is then also conserved, which follows directly from the system of kinetic equations (70).

The boundary of the multistream region (mixing region) is the first caustic which appears immediately after the initial singularity. In the adiabatic model it is assumed that sufficient mixing, accompanied by conservation of the adiabatic invariant I, occurs immediately after the passage of the first caustic.

It therefore follows that the distribution function $f(I)$ is conserved in the region bounded by the first caustic. It follows from the law of conservation of the mass of matter crossing the first caustic that

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \rho(x_c) \frac{\mathrm{d}x}{\mathrm{d}t} = f(I_c) \frac{\mathrm{d}I_c}{\mathrm{d}t} , \tag{71}$$

where x_c is the coordinate of the caustic and I_c is the adiabatic invariant on the caustic. Relationships (41) and (69) then give

$$I_c = \int_0^{x_c} u\, \mathrm{d}x = \frac{3}{4} V_1 x_c^{4/3} . \tag{72}$$

If x_c is found from formula (72), then expressions (71) and (41) can be used to determine the distribution function $f(I_c)$:

$$f(I_c) = \rho(x_c)\,\frac{dx_c}{dI_c} = \left(\frac{3}{4}\right)^{3/4} \frac{\rho_1}{V_1}\left(\frac{I_c}{V_1}\right)^{-3/4} . \tag{73}$$

In view of the conservation of the distribution function of the adiabatic invariance, it follows that for any value of I, we have

$$f(I) = f(I_c)|_{I_c=I} . \tag{74}$$

The equation for the field potential can then be written in the form

$$\frac{\partial^2\psi}{\partial x^2} = \int_{\psi}^{E_m} \frac{f[I(E)]}{[2(E-\psi)]^{1/2}}\,dE , \tag{75}$$

where the dependence $I(E)$ is defined by expression (69).

The system of equations (70) considered in the adiabatic approximation thus assumes the form of expressions (69), (73)–(75). Its solution in the limit $x \to 0$ should be sought naturally in the form of the power law:

$$\psi = \psi_1 x^{\alpha} .$$

It then follows from expression (69) that

$$I = C_0\,\frac{E^{1/2+1/\alpha}}{\psi_1^{1/\alpha}} , \qquad C_0 = 2^{1/2}\int_0^1 (1-y^{\alpha})^{1/2}\,dy . \tag{76}$$

We can see from expressions (73) and (74) that

$$f = \left(\frac{3}{4C_0}\right)^{3/4} \frac{\rho_1}{V_1}\,\frac{\psi_1^{1/2}}{V_1}\left(\frac{E}{\psi_1}\right)^{-(3\alpha+6)/8\alpha} . \tag{77}$$

Substitution of expression (77) in the equation for the potential (75) gives finally the required steady-state solution:

$$\alpha = \frac{10}{7} , \qquad \psi_1 = V_1^2\left[\frac{49}{30}\,C_1\left(\frac{3}{4C_0}\right)^{3/4}\frac{\rho_1}{V_1^2}\right]^{8/7} , \tag{78}$$

where

$$C_1 = 2^{-1/2}\int_1^{\infty} y^{-0.9}\,(y-1)^{-1/2}\,dy ,$$

$$C_0 = 2^{1/2}\int_0^1 (1-y^{10/7})^{1/2}\,dy .$$

The solution described by expressions (76)–(78) is fully mixed and steady-state, since the distribution function then depends only on the stream energies E.

It is thus evident from the adiabatic model that multistream mixing establishes a steady-state distribution with a singularity of the average density at the point $x = 0$. In fact, if $\alpha = 10/7$, we have

$$\bar{\rho} = \frac{30}{49}\,\psi_1 x^{-4/7} . \tag{79}$$

A singularity appears also in the steady-state distribution function (77):

$$f \propto E^{-9/10} . \tag{80}$$

As pointed out above, relationship (79) is best examined by considering the curve $m(x)$ shown in Fig. 4. We can see that, in the limit $x \to 0$, relationship (79) agrees well with the results of a numerical simulation. It follows from the distribution described by expression (58) that the real distribution function consists of a set of δ functions and, therefore, it fluctuates strongly. However, once again the agreement is satisfactory between numerical calculations and relationship (80) [29].

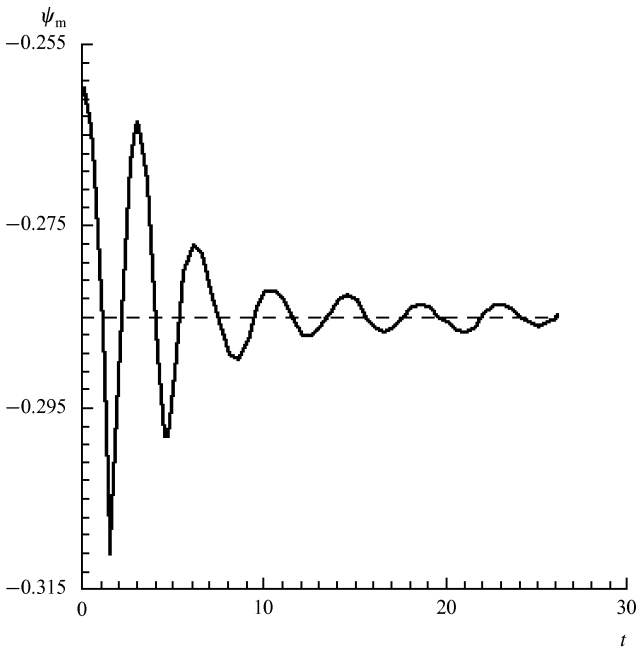

Figure 5. Time dependence of the minimum of the potential.

Let us now consider our adiabatic approximation. We can see that the solution described by expressions (79) and (80) is of scaling nature, that the real fluctuations of the potential are small (Fig. 5), and that moreover the period of the oscillations of the streams

$$T(E) = \frac{\partial T}{\partial E} = \frac{6}{5}\,C_0\psi_1^{10/7}E^{1/5} \tag{81}$$

tends to zero for $E \to 0$. Therefore, if $E \to 0\,(x \to 0)$, the adiabatic solution is asymptotically correct. Consequently, it is also valid for all energies $E \ll 1$, as confirmed by a comparison with numerical results.

Relationship (81) and the scaling nature of the solution described by expressions (79) and (80) provide essentially the justification for the use of the adiabatic approximation.

5.4 Accumulation of caustics

The reduction in the oscillation period with energy described by expression (81), means that in the limit $E \to 0$, i.e. near the bottom of a potential well, the caustics form more and more frequently. We shall now consider the law governing such accumulation of the caustics. According to expressions (61) and (63), the process of multiplication of the caustics is equivalent to the appearance of new roots of the δ function. At any moment t the distribution function retains its general form:

$$f = \rho(x, t)\,\delta[v - V(x_0, v, t)] ,$$

where $\tau = t - t_0$, t_0 is the initial moment of time and x_0 is the initial point of a path which is an integral of motion.

We shall assume that this motion occurs asymptotically in a mixed potential. We then have

$$\tau = \int_{x_0}^{x} \left(\tilde{v}^2 - 2\psi_1 x_1^{10/7}\right)^{-1/2} dx_1 , \qquad \tilde{v}^2 = 2E . \tag{82}$$

At $\tau = 0$, it follows from expression (41) that

$$V = -V_1 x_0^{1/3} .$$

Replacement of the variable x in expression (82) with a new variable $\cos\theta = (2\psi_1/\tilde{v}^2)^{1/2} x^{5/7}$ gives

$$\tilde{v}\tau = \left(\frac{\tilde{v}^2}{2\psi_1}\right)^{7/10} \int_b^a \cos^{2/5}\theta \, \mathrm{d}\theta , \tag{83}$$

$a = \arccos\left[(2\psi_1)^{1/2} x_0^{5/7}/\tilde{v}\right]$ and $b = \arccos\left[(2\psi_1)^{1/2} x^{5/7}/\tilde{v}\right]$.

In view of the symmetry of the pattern relative to the variables $v-x$, the appearance of the new roots, i.e. of the new turns of the spiral, can be observed more conveniently along the v axis. The points v_n can be found by substituting $x = 0$, i.e. $b = \pi/2$, in Eqn (83). Then after a long time in the limit $\tau \to \infty$ the solution of Eqn (83) becomes

$$\frac{\pi}{2} + (2\psi_1)^{7/10} \frac{\tau}{q_0 v^{2/5}} = \arccos\left[(2\psi_1)^{1/2} \frac{x_0^{5/7}}{v}\right] , \tag{84}$$

where

$$q_0 = \frac{1}{2\pi} \int_0^{2\pi} \cos^{2/5}\theta \, \mathrm{d}\theta .$$

If x_0 is expressed in terms of v by means of relationships (82) and (84), the result is the following transcendental equation:

$$v^{8/5} = V_1^3 (2\psi_1)^{-7/10} \sin^{7/5}\left[(2\psi_1)^{7/10} \frac{\tau}{q_0 v^{2/5}}\right] . \tag{85}$$

After a long time ($t \to \infty$) the solution of Eqn (85) is practically coincident with zeros of the sine:

$$(2\psi_1)^{7/10} \frac{\tau}{q_0 v_n^{2/5}} = \pi n .$$

Therefore, the sequence of roots v_n obeys

$$v_n \propto \left(\frac{\tau}{n}\right)^{5/2} , \tag{86}$$

where $n = 1, 2, \ldots, N$.

The maximum number N of the caustics can be estimated from Eqn (85). Since the sine is always less than unity, we find that

$$N \geqslant \tau (2\psi_1)^{7/10} V_1^{3/4} . \tag{87}$$

According to inequality (87), the number of caustics increases proportionately to time τ.

The relative separation between the caustics $\Delta v_n/v_n$ decreases with increase in N. In fact, we find from formula (86) that

$$\frac{\Delta v_n}{v_n} = \frac{5}{2n} , \tag{88}$$

i.e. according to equality (88), the reduction in the scale (and, consequently, the reduction in the smallest distance between the caustics) obeys

$$\frac{\Delta v_N}{v_N} = \frac{5}{2N} \propto \frac{1}{\tau} . \tag{89}$$

In view of the above relationship, similar relationships (86)–(89) are obeyed also by the coordinates x_n of the caustics. We recall that on each caustic the density becomes infinite [see the system of equations (67)]. Therefore, with the passing of time, the actual pattern of the density distribution becomes more and more jagged and the degree of jaggedness increases on approach to the centre of a singularity in the region where $x \to 0$.

5.5 Natural oscillations

Only small-scale caustic waves have been considered so far. The steady kinetic state of expression (80) can be regarded as an equilibrium dynamic system. A deviation from the position of equilibrium, such as (in particular) the initial state described by the conditions (32), should result in the excitation of natural oscillations of the system.

These natural oscillations can be described by representing the distribution function and the field potential as follows:

$$f = f_0(E) + f_1 , \quad \psi = \psi_0(x) + \varphi_1 , \tag{90}$$

where f_0 and ψ_0 are the steady-state distribution function and the field potential; f_1 and φ_1 are the deviations of the distribution function and of the potential from the equilibrium values.

Substitution of expression (90) into the system of equations (7) gives

$$\begin{gathered} \frac{\partial f_1}{\partial t} + v \frac{\partial f_1}{\partial x} + \frac{\partial \psi_0}{\partial x} \frac{\partial f_1}{\partial v} + \frac{\partial \varphi_1}{\partial x} \frac{\partial f_0}{\partial v} + \frac{\partial \varphi_1}{\partial x} \frac{\partial f_1}{\partial v} = 0 , \\ \frac{\partial^2 \varphi_1}{\partial x^2} = \int_{-\infty}^{+\infty} f_1 \, \mathrm{d}v . \end{gathered} \tag{91}$$

The linear part can be separated in a natural manner from the system of equations (91):

$$\begin{gathered} \frac{\partial f_1}{\partial t} + v \frac{\partial f_1}{\partial x} + \frac{\partial \psi_0}{\partial x} \frac{\partial f_1}{\partial v} + \frac{\partial \varphi_1}{\partial x} \frac{\partial f_0}{\partial v} = 0 , \\ \frac{\partial^2 \varphi_1}{\partial x^2} = \int_{-\infty}^{+\infty} f_1 \, \mathrm{d}v . \end{gathered} \tag{92}$$

The system (92) describes eigenmodes of linear oscillations. The first (fundamental) mode can be seen in the oscillations of the field potential shown in Fig. 5. The period and, consequently, the angular frequency of this first mode, are

$$T = 4.2 , \quad \omega = 1.50 .$$

Since the investigated system is nondissipative, linear oscillations should naturally be undamped. However, Fig. 5 demonstrates clearly that the amplitude of the oscillations decays: this is due to the fact that the oscillation amplitude is finite. The decay of this amplitude is a consequence of the nonlinear interaction of the eigenmodes.

We shall describe this process by expanding the function f_1 in terms of the ket eigenvectors $|g_n(x, v)\rangle$ of the system of equations (92):

$$f_1 = \sum_n \exp(\lambda_n t) A_n(t) \, |g_n(x, v)\rangle , \tag{93}$$

where λ_n are the eigenvalues of the system (92); $A_n(t)$ is the amplitude corresponding to the eigenmode $|g_n(x, v)\rangle$. Substitution of expression (93) into the system of equations (91) and multiplication by the corresponding bra

vector of the system (92) gives the following equation for the amplitude of the lth-mode:

$$\frac{\mathrm{d}A_l}{\mathrm{d}t} = \sum_{m,n} \exp\left[(\lambda_n + \lambda_m - \lambda_l)t\right] A_n A_m K_{l,m,n} \,, \tag{94}$$

where

$$K_{l,m,n} = \langle g_l \Big| \int_0^x \mathrm{d}x_1 \int_{-\infty}^{+\infty} \mathrm{d}v_1\, g_n(x_1, v_1) \frac{\partial}{\partial v} \Big| g_m \rangle \,.$$

Eqn (94) for A_l can be used to find the asymptotic law of decay of the amplitudes after a long time t. Bearing in mind that all the eigenvalues λ_n of the system are purely imaginary, Eqn (94) considered in the limit $t \to \infty$ becomes

$$\frac{\mathrm{d}A_l}{\mathrm{d}t} = -\sum_m A_{l-m} A_m K_{l,m,l-m} \,. \tag{95}$$

The form of Eqn (95) for the mode amplitudes is a natural consequence of the quadratic nature of the nonlinearity in the system of equations (91).

Eqn (95) has the obvious solution

$$A_l = \frac{C_l}{t} \,, \tag{96}$$

where the constants C_l satisfy the following system of algebraic equations:

$$\sum_m \left\{ C_{l-m} C_m K_{l,m,l-m} + \delta_{m,l} C_m \right\} = 0 \,.$$

We can thus see that after a long time the amplitudes of the eigenmodes decrease as $1/t$.

The correspondence between the decay law of the eigenmodes (96) and the law describing the reduction in the scale by the caustic waves (89) is quite natural. The decrease in the mode amplitudes is due to continuous transfer of energy to higher harmonics, which occurs precisely because of the caustic-induced reduction in the scale. In this sense the process of nonlinear relaxation is fully analogous to the nonlinear Landau damping of plasma oscillations [34]. The first (fundamental) mode decays most slowly and the dynamics of this mode can be judged on the basis of the time dependence of the minimum of the potential shown in Fig. 5. We can see that the $1/t$ law is supported quite well by the numerical calculation.

5.6 Change in the entropy

Since the investigated system is conservative, both the decay of the modes and the caustic-induced reduction in the scale imply simply the transfer of energy to smaller scales. In view of this, it is useful to consider the total entropy of the system

$$S = \int f \ln \frac{e}{f} \,\mathrm{d}v\,\mathrm{d}x \,. \tag{97}$$

It follows from Eqns (7), (58), and (97) that in the initial state the total entropy is $-\infty$. The dynamic process described by the system of equations (7) cannot change the entropy. Therefore, at any moment the entropy should be $-\infty$. The validity of this conclusion is readily confirmed if we bear in mind that the distribution function can always be represented in the form given by expression (58). However, if we consider a mixed state described by formula (80), we find that in this state — in accordance with expression (97) — the entropy has the finite value

$$S_{\mathrm{f}} = \frac{Q_0 \psi_m^{\beta+1}}{\beta+1} \left(\ln Q_0 + \beta \ln \psi_m - \frac{\beta}{\beta+1} \right) \,.$$

Here, according to formula (80), $\beta = 9/10$, ψ_m is the depth of a potential well, and

$$Q_0 = \left(\frac{3}{4C_0} \right)^{3/4} \frac{\rho_1}{V_1} \frac{\psi_1^{1/2}}{V_1} \psi_1^{9/10} \,.$$

The whole entropy of the mixed state is therefore concentrated not in the mean average distribution, but in giant small-scale fluctuations governed by the presence of an infinite set of caustics. In this sense we can say that relaxation of the investigated system to a mixed kinetic state by the purely dynamic forces discussed above is incomplete: in the absence of dissipation the system always conserves nonequilibrium (giant) fluctuations.

6. Spherically symmetric singularity

6.1 Spherically symmetric case

We shall now consider the process of compression of a spherically symmetric bunch after the appearance of a singularity. We shall analyse first the qualitative features of flow. According to expression (46), near a singularity the stream velocity is

$$V_r = -(2\beta)^{1/2} \left(\frac{r}{a} \right)^{1/7}$$

and the potential of the gravitational field is given by

$$\psi = \frac{7}{2} \beta \left(\frac{r}{a} \right)^{2/7} \,.$$

It is thus clear that the kinetic energy of a stream is less than the potential energy:

$$|\psi| > \frac{V_r^2}{2} \,.$$

This means that the gravitational field has the dominant influence on the flow of a gas near a singularity right up to the moment of reversal at $t = t_{\mathrm{c}}^{(1)}$.

The form of the gravitational potential after a reversal at $t > t_{\mathrm{c}}^{(1)}$ can be found if, instead of the usual density ρ, we introduce the effective density $\tilde{\rho} = r^2 \rho$, which takes account of the spherical accumulation of flow. The system of equations (42) can be rewritten in the form

$$\frac{\partial \tilde{\rho}}{\partial t} + \frac{\partial}{\partial r} (\tilde{\rho} u) = 0 \,,$$

$$\frac{\partial u}{\partial t} + \left(u \frac{\partial}{\partial r} \right) u + \frac{\partial \psi}{\partial r} = 0 \,,$$

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) = \tilde{\rho} \,, \tag{98}$$

i.e. the hydrodynamic equations have exactly the same form as in the planar case and only the Poisson equation is modified.

If we assume that after reversal the density $\tilde{\rho}$ at the centre (at $r = 0$) is — as in the planar case — finite (see

Fig. 2b), i.e. if $\tilde{\rho} = \tilde{\rho}_0$, it follows from the system of equations (98) that in the limit $r \to 0$ the potential is of the form

$$\psi = \tilde{\rho}_0 \ln r \,. \tag{99}$$

In other words, an infinitely deep potential well appears after reversal. However, the velocity v is low. It is therefore clear that the motion of the gas and the particle density near a singularity in the limit $r \to 0$ are governed entirely by the action of the field.

It is also important to note that a logarithmic potential well which appears at a time t close to $t_c^{(1)}$ is very narrow. This means that the frequency of the oscillations of the streams trapped in the well is high and, as shown below, it increases rapidly with the depth in the well, i.e. on approach to the limit $r \to 0$. Consequently, in the vicinity of a centre an infinite set of caustic singularities appears immediately at $t > t_c^{(1)}$ and these singularities converge towards the point $r = 0$. Consequently, an infinite number of streams also appears. A structure of this kind will be called a non-dissipative gravitational singularity (NGS).

6.2 Adiabatic theory

It is natural to describe a gravitational singularity with the aid of the theory of adiabatic trapping [35], since the parameters of a potential well vary slowly with time compared with the frequency of the oscillations of the streams trapped in the well [see inequality (106) given later]. This means that, in the first approximation, the distribution function can be represented in the form

$$f = f(E, t) = f(I) \,,$$

$$I = \int_0^{r_1} V \, \mathrm{d}r = 2^{2/3} \int_0^{r_1} (E - \psi)^{1/2} \, \mathrm{d}r \,, \qquad \psi(r_1, t) = E \,.$$

Here, $E = V^2/2 + \psi$ is the energy, I is the adiabatic invariant, and $r_1 = r_1(E, t)$ is a reflection point.

Further calculations, fully analogous to those represented by expressions (70)–(75) in the planar case, yield a nonlinear integrodifferential equation for the potential [26]:

$$\begin{gathered} \frac{\mathrm{d}^2\tilde{\psi}}{\mathrm{d}x^2} + \frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}x} = \int_{\tilde{\psi}}^{0} F^{1/8}(y)\,(y - \tilde{\psi})^{-1/2} \, \mathrm{d}y \,, \\ F(y) = \int_{-\infty}^{x_1(y)} \left[y - \tilde{\psi}(x)\right]^{1/2} e^x \, \mathrm{d}x \,, \qquad \tilde{\psi}(x_1) = y \,. \end{gathered} \tag{100}$$

Here, $x = \ln\left[r/r_0(t)\right]$, $r_0(t)$ is the coordinate of the first caustic, and $F(y)$ is the distribution function of the stream energies,

$$y = \frac{E}{\psi_1 \tau^{1/3}} \,, \qquad \tau = \frac{t - t_c^{(1)}}{t_c^{(1)}} \,, \qquad \psi(x, \tau) = \psi_1 \tau^{1/3} \tilde{\psi}(x) \,.$$

Let us consider the asymptotic behaviour of the function $\tilde{\psi}$ in the limit $x \to \infty$ or $r \to 0$. Let us assume (this will be justified later) that the function $F(y)$ falls quite rapidly in the limit $y \to \infty$, so that the following integral converges:

$$\int_{-\infty}^{0} F^{1/8}(y) \, \mathrm{d}y = B_0 \,.$$

Then, the first equation in the system (100) assumes the following form in the limit $x \to -\infty$:

$$\frac{\mathrm{d}^2\tilde{\psi}}{\mathrm{d}x^2} + \frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}x} = \frac{B_0}{(-\tilde{\psi})^{1/2}} \,. \tag{101}$$

The solution of Eqn (101), which becomes $-\infty$ in the limit $x \to -\infty$, is

$$\tilde{\psi}(x) = -\left(-\frac{3}{2} B_0 x\right)^{2/3} \left[1 + \frac{2}{9} \frac{\ln(-x)}{-x} + \cdots\right] . \tag{102}$$

Therefore, in the limit $r \to 0$ the gravitational field potential is

$$\tilde{\psi}(x) \propto \ln^{2/3} \frac{1}{r} \,. \tag{103}$$

This potential $\tilde{\psi}$ tends to $-\infty$ when $r \to 0$. The effective density $\tilde{\rho}$ vanishes on approach to $r \to 0$ in accordance with the following logarithmic law:

$$\tilde{\rho} \propto (-x)^{-1/3} \propto \ln^{-1/3} \frac{1}{r} \,. \tag{104}$$

A comparison of expressions (103) and (104) with expression (99) demonstrates near a singularity that the distributions of the field $\tilde{f}$ and of the density $\tilde{\psi}$ are governed entirely by the action of the gravitational field.

The asymptote of the energy distribution function for $E \to -\infty$ is also readily obtained from expressions (100) and (102):

$$f(E, \tau) = B_1 \tau^{1/6} (-y)^{-1/32} \exp\left[-\frac{1}{8}(-y)^{3/2}\right] . \tag{105}$$

The above distribution function falls exponentially in the limit $E \to 0$, which is in particular the justification for Eqn (101). It should be pointed out that, in spite of its exponential nature, the distribution function (105) is very far from a Boltzmann equilibrium.

Finally, let us consider the validity of the adiabatic approximation, i.e. the slowness of the change in the potential $\tilde{\psi}$ during one oscillation period T:

$$p = \left|\frac{T}{2\pi\psi} \frac{\partial\psi}{\partial t}\right| \ll 1 \,. \tag{106}$$

Direct calculations, carried out by the technique of adiabats of an invariant, readily show that the parameter p is given by [26]

$$p = p_0 (-y)^{1/4} \exp\left[-(-y)^{3/2}\right] , \tag{107}$$

where p_0 is a constant of the order of unity and y is defined by system (100).

It therefore follows that the condition of validity of the adiabatic approximation, given by inequality (106) at sufficiently high values of $|y|$, is always well satisfied, i.e. in the vicinity of the bottom of a potential well the solutions (103)–(105) are asymptotically exact.

We shall now consider some properties of the solution obtained. The average density obeys

$$\rho \propto r^{-2} \ln^{-1/3} \frac{1}{r} \,, \tag{108}$$

and, by analogy with the planar case described by expression (79), it has a singularity at $r \to 0$. The actual (not average) density is strongly jagged by the caustics (in the adiabatic limit the jaggedness is everywhere dense), but the first integral

$$m(r) = 4\pi \int_0^r r^2 \rho \, \mathrm{d}r \,,$$

is, as in the planar case (Fig. 4), smooth and tends to zero in the limit $r \to 0$.

In contrast to the planar case, in the spherical case the singularity at $r \to 0$ also has the field potential given by expres-sion (103). Consequently, local trapping and multi-stream flow occur immediately after the formation of the initial singularity described by formula (46) and, in view of the narrowness of the wall, i.e. because relationships (106) and (107) are obeyed, the processes of mixing and transition to a kinetic state occur practically immediately at $t > t_c^{(1)}$, whereas in the planar case such a state appears only after $t \gg t_c$.

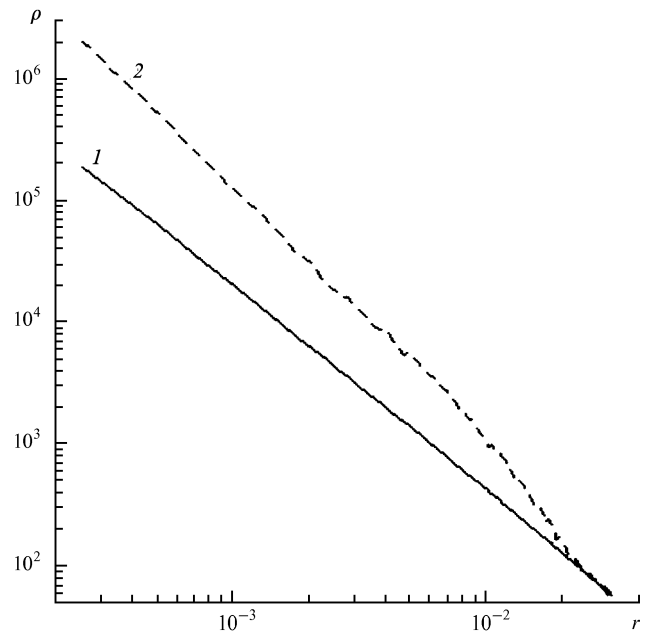

Figure 6. Distribution of the density $\rho(r)$ at the moment of appearance of a singularity.

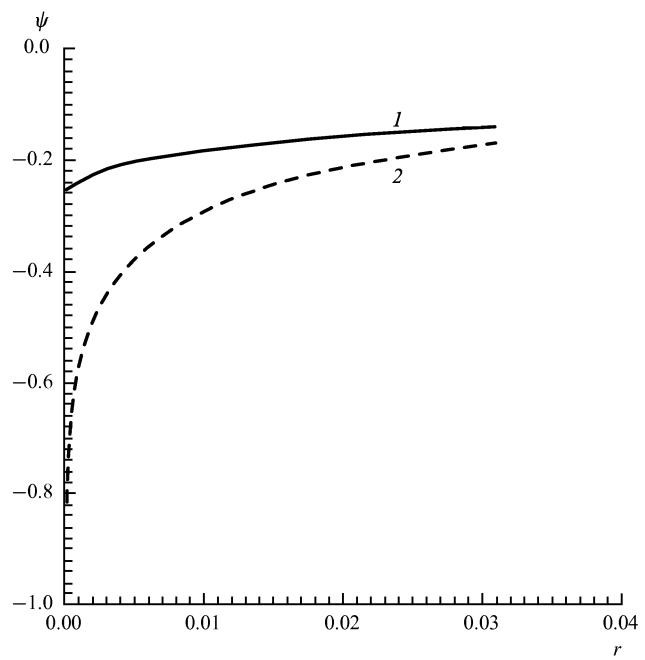

Figure 7. Potential $\psi(r)$ at the moment of appearance of a singularity.

V_r / V														
	4	0	0	0	0	0	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5	24	0	0	0	0	0	0	0	0	0	0	0	0	0
	52	1	0	0	0	0	0	0	0	0	0	0	0	0
	117	8	0	0	0	0	0	0	0	0	0	0	0	0
	183	29	0	0	0	0	0	0	0	0	0	0	0	0
	239	73	4	0	0	0	0	0	0	0	0	0	0	0
1.0	313	174	52	7	1	0	0	0	0	0	0	0	0	0
	360	258	141	56	15	9	3	0	0	0	0	0	0	0
	372	306	240	165	85	62	30	4	0	0	0	0	0	0
	391	351	292	241	197	152	69	16	1	0	0	0	0	0
	412	410	342	310	257	220	131	43	6	0	0	0	0	0
0.5	363	416	359	310	283	276	185	80	15	2	0	0	0	0
	360	409	416	344	316	314	224	93	15	1	0	0	0	0
	345	376	423	370	337	353	224	69	3	0	0	0	0	0
	329	371	438	380	384	350	151	12	0	0	0	0	0	0
	330	389	426	415	373	207	29	0	0	0	0	0	0	0
0.0	320	366	427	414	273	42	0	0	0	0	0	0	0	0
	331	368	430	323	68	0	0	0	0	0	0	0	0	0
	350	378	420	133	2	0	0	0	0	0	0	0	0	0
	318	403	244	12	0	0	0	61	443	874123	815491	81420	422238	
	311	337	66	0	258	113019	482	4912267	184414	741147	855	593	357	
−0.5	356	217	119	119920	621430	688	132	0	0	0	0	0	0	0
	366	232	131	11081	127	0	0	0	0	0	0	0	0	0
	326	806	619	9	1	0	0	0	0	0	0	0	0	0
	475	612	18	7	0	0	0	0	0	0	0	0	0	0
	514	197	0	0	0	0	0	0	0	0	0	0	0	0
−1.0	364	58	0	0	0	0	0	0	0	0	0	0	0	0
	221	16	0	0	0	0	0	0	0	0	0	0	0	0
	119	4	0	0	0	0	0	0	0	0	0	0	0	0
	67	1	0	0	0	0	0	0	0	0	0	0	0	0
	31	0	0	0	0	0	0	0	0	0	0	0	0	0
−1.5	14	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.000000					0.041667				0.083333				r

Figure 8. Phase plane in the case of spherically symmetric compression.

6.3 Numerical simulation

In addition to the adiabatic theory, the problem has been tackled by numerical simulation of the process of kinetic mixing in the spherically symmetric case [36]. The results of such simulation are presented in Figs 6–8. We can see that a density singularity (curve *1* in Fig. 6) appears at the moment $t = t_c^{(1)}$. However, the potential at this moment still has a finite value (curve *1* in Fig. 7). The results obtained are in good agreement with the exact analytic solution given by formula (46).

According to the adiabatic theory predictions, a sharp dip of the potential appears at the centre of the initial distribution of matter at $t > 0$ (curve *2* in Fig. 7). Even at very low values of τ there are multiple reflections of particles (in the limit $r \to 0$) from the self-consistent field potential and a density distribution close to that described by the law (108) is established (curve *2* in Fig. 6).

In the spherically symmetric case the behaviour of the system differs considerably from that in the planar case. Immediately after the appearance of a singularity a mixed multistream state begins to evolve and this state is

characterised by a distribution function in which it is impossible to identify δ-like peaks typical of multistream flow (Fig. 8).

We recall that in the planar case the distribution function is formed by consecutive tripling of the streams at the centre of the distribution and quite a long time is required for the appearance of a fairly smooth average function.

It therefore follows that the analytic adiabatic theory predictions are well supported by the numerical simulation results. The spherically symmetric flow seems to be a unique system in which the dynamic process establishes a kinetic mixed state after a singularity in a time interval which can be as short as we please. It is evident from Fig. 8 that the incoming δ-like hydrodynamic stream is rapidly converted into a stream with a wide distribution function.

The planar and spherically symmetric cases represent the limits of the initial general problem described by the system of equations (28). In both cases a density singularity appears under steady-state conditions: the differences are solely in the order of the singularity and in the time taken to reach the steady state. It is therefore natural to expect that the main features of this limit will be manifested also in the general case.

7. Nondissipative gravitational singularity

7.1 General case

The spherically symmetric case is degenerate in the sense that all the particle motion occurs only along the radius r. We shall now study the evolution in time of a general singularity described by expressions (55), when — apart from the radial motion — there are also transverse components of the velocity. By analogy with expressions (49)–(55), the ellipticity parameters near the maximum of the initial density will be regarded as small:

$$\varepsilon_1 = \frac{a-b}{a} \ll 1 , \qquad \varepsilon_2 = \frac{a-c}{a} \ll 1 .$$

We shall deal with this problem by rewriting the kinetic equation (7) in spherical coordinates:

$$\frac{\partial f}{\partial t} + p_r \frac{\partial f}{\partial r} + \frac{p_\theta}{r^2}\frac{\partial f}{\partial \theta} + \frac{p_\varphi}{r^2 \cos^2\theta}\frac{\partial f}{\partial \varphi} + \left\{\frac{\partial \psi}{\partial t} - \frac{p_\varphi}{r^2\cos^2\theta}\frac{\partial \psi}{\partial \varphi}\right\}\frac{\partial f}{\partial E} - 2\left(p_\theta \frac{\partial \psi}{\partial \theta} + \frac{p_\varphi}{\cos^2\theta}\right)\frac{\partial f}{\partial m^2} - \frac{\partial \psi}{\partial \varphi}\frac{\partial f}{\partial p_\varphi} = 0 ,$$

$$\nabla^2 \psi = \int f \, \mathrm{d}p_\varphi \, \mathrm{d}p_\theta \, \mathrm{d}p_r , \tag{109}$$

where the momenta p_r, p_θ, and p_φ are related to the velocities v_r, v_θ, and v_φ by

$$v_r = p_r , \qquad v_\theta = \frac{p_\theta}{r} , \qquad v_\varphi = \frac{p_\varphi}{r\cos\theta} ,$$

where $E = p_r^2/2 + p_\theta^2/2r^2 + p_\varphi^2/(2r^2\cos^2\theta) + \psi$ is the energy of the system, $m^2 = p_\theta^2 + p_\varphi^2/\cos^2\theta$ is the square of the angular momentum, and $p_\varphi = m_\varphi$ has the physical meaning of the angular momentum relative to the z axis. As usual, the relationships applicable to the initial singularity [expressions (46) and (55)] should be regarded as the initial conditions for the system of equations (109).

We shall now consider a system of expressions (55). Two velocity components, u_φ and u_θ, are transverse to the radial velocity v_r. Consequently, the components $u_{\perp x}$, $u_{\perp y}$, $u_{\perp z}$ are not independent and, as can be seen from expressions (55), the component $u_{\perp x}$ can be ignored near a singularity. This means that one of the transverse components, for example u_φ, can be assumed to be zero.

Without limitations on the generality, the expression for the transverse component u_θ is described by the following expression, accurate to within ε^2:

$$u_\theta = \frac{3}{2} v_r \varepsilon , \tag{110}$$

where $\varepsilon^2 = \varepsilon_1^2 + \varepsilon_2^2 - \varepsilon_1\varepsilon_2$. Consequently, the system of equations (109) can be simplified. In fact, if the distribution function is represented in the form

$$f = \tilde{f}\delta(p_\varphi) ,$$

the result is

$$\frac{\partial \tilde{f}}{\partial t} + p_r \frac{\partial \tilde{f}}{\partial r} + \frac{m}{r^2}\frac{\partial \tilde{f}}{\partial \theta} + \frac{\partial \psi}{\partial t}\frac{\partial \tilde{f}}{\partial E} - \frac{\partial \psi}{\partial \theta}\frac{\partial \tilde{f}}{\partial m} = 0 ,$$

$$p_r = \left[2\left(E - \psi - \frac{m^2}{2r^2}\right)\right]^{1/2} . \tag{111}$$

For simplicity, we shall omit the tilde above f.

An analysis of the solution of the systems of equations (51) and (55) shows that

$$m \sim \varepsilon , \qquad \frac{\partial \psi}{\partial \theta} \sim \varepsilon ,$$

i.e. the angular part of the distribution function is small. Moreover, this angular part decays with time because of energy transfer to higher harmonics, in full analogy with the process discussed in Section 5.

Averaging over the angle θ in Eqns (111) and bearing in mind that the quantities E and m^2 are integrals of motion, we find that the spherical part of the distribution function is described by the following kinetic equation:

$$\frac{\partial f}{\partial t} \pm \left[2\left(E - \psi - \frac{m^2}{2r^3}\right)\right]^{1/2}\frac{\partial f}{\partial r} + \frac{\partial \psi}{\partial t}\frac{\partial f}{\partial E} = 0 . \tag{112}$$

The Poisson equation then becomes

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) = 2^{1/2}\int_0^\infty \mathrm{d}m^2 \times \int_{\psi + m^2/2r^2}^0 f(E, m^2, r, t)\left(E - \psi - \frac{m^2}{2r^2}\right)^{-1/2} \mathrm{d}E . \tag{113}$$

7.2 Adiabatic theory

We shall use the adiabatic approximation (see Sections 5 and 6) to find the solution of the system of equations (112), (113) [28]. In this approximation the function f depends on the square of the angular momentum m^2 and on the adiabatic invariant:

$$I = 2^{1/2}\int_{r_{\min}(E)}^{r_{\max}(E)} \left[E - \psi(r_1, t) - \frac{m^2}{2r_1^2}\right]^{1/2} \mathrm{d}r_1 , \tag{114}$$

where $r_{\min}$ and $r_{\max}$ are zeros of the radicand. The energy E in Eqns (113) and (114) is measured from the caustic. It follows from the hydrodynamic solution that

$$m^2 = m_0^2 r_0^{16/7}, \quad E(r_0) = E,$$
$$E(r_0) = \frac{9}{7}\psi_0 r_0^{9/7} + o(\varepsilon), \tag{115}$$
$$m_0^2 = 0.0881\tilde{\rho}_0\varepsilon^2, \quad \tilde{\rho}_0 = \rho_0 a^{12/7}.$$

The condition of continuity of flow on the caustic gives

$$\rho(r_0)\,\mathrm{d}r_0 = f(I_0)\,\mathrm{d}I_0. \tag{116}$$

According to the adiabatic theory, the form of the distribution function $f(I)$ is governed by the conditions at the trapping boundary, i.e. on the caustic at $I = I_0(t)$. Conservation of the adiabatic invariant means that the distribution function is

$$f(I) = f(I_m)|_{I_m = I}. \tag{117}$$

It follows from relationships (114)–(117) and from the law of conservation of the angular momentum that

$$f(I) = f_0 I^{1/8}(E)\,\delta(m^2 - l_0^2 I^2), \tag{118}$$

where

$$f_0 = \frac{49}{72}\frac{\tilde{\rho}_0}{\psi_0}\left(\frac{7}{9\psi_0}\right)^{7/2} C_1^{-9/8}, \quad l_0^2 = \left(\frac{7}{9\psi_0}\right)^8 \frac{m_0^2}{C_1^2},$$
$$C_1 = 2^{1/2}\int_{\kappa_{\min}}^{\kappa_{\max}}\left[E - \psi_0\kappa^{2/7} - \frac{m_0^2}{2\kappa^2}\left(\frac{7}{9\psi_0}\right)^8\right]^{1/2}\mathrm{d}\kappa.$$

Substitution of relationship (118) into Eqns (113) and (114) gives

$$\frac{\partial}{\partial r} r^2 \frac{\partial\psi}{\partial r} = 2^{1/2} f_0 \int_0^\infty \mathrm{d}m^2$$
$$\times \int_{\psi + m^2/2r^2}^0 I^{1/8}(E)\,\delta(m^2 - l_0^2 I^2)\left(E - \psi - \frac{m^2}{2r^2}\right)^{-1/2}\mathrm{d}E,$$
$$I = 2^{1/2}\int_{r_{\min}}^{r_{\max}}\left[E - \psi(r_1, t) - \frac{l_0^2}{2r_1^2} I^2\right]^{1/2}\mathrm{d}r_1. \tag{119}$$

When l_0^2 tends to zero, the system of equations (119) reduces exactly to the system (100) derived in the absence of the momentum. In this limit, the field potential is

$$\psi = -\left[\frac{3}{2} C_0 \ln\frac{r_0(t)}{r}\right]^{2/3} \tag{120}$$

where r_0 is the position of the caustic separating the regions of single-stream and multistream flow, and C_0 is a normalisation constant.

An analysis of the system of equations (119) by a procedure analogous to that described above shows that in a region of radius

$$0 \leqslant r \leqslant r_\varepsilon, \quad r_\varepsilon = 0.0731\left[\frac{\chi_0(t)}{\psi_0}\right]^{7/2}\varepsilon^{46/49},$$

there is a power-law solution of the system (119):

$$\psi = -\chi_0(t) + \psi_1 r^{2/7}, \quad \psi_1 = 5.742\,\psi_0\varepsilon^{-32/49}. \tag{121}$$

Outside the region defined by the above expressions the momenta are unimportant and, therefore, the solution described by expression (120) is valid. The depth of the potential well $\chi_0(t)$ in the solution (121) is found by assuming continuity of the potential at a point $r = r_\varepsilon$. With logarithmic accuracy, we have

$$\chi_0 = \left\{\frac{3}{2} C_0 \ln\left[\frac{r_0}{0.07}\left(\frac{3}{2}C_0\right)^{-7/3}\psi_0^{-7/2}\varepsilon^{-46/49}\right]\right\}^{1/2}. \tag{122}$$

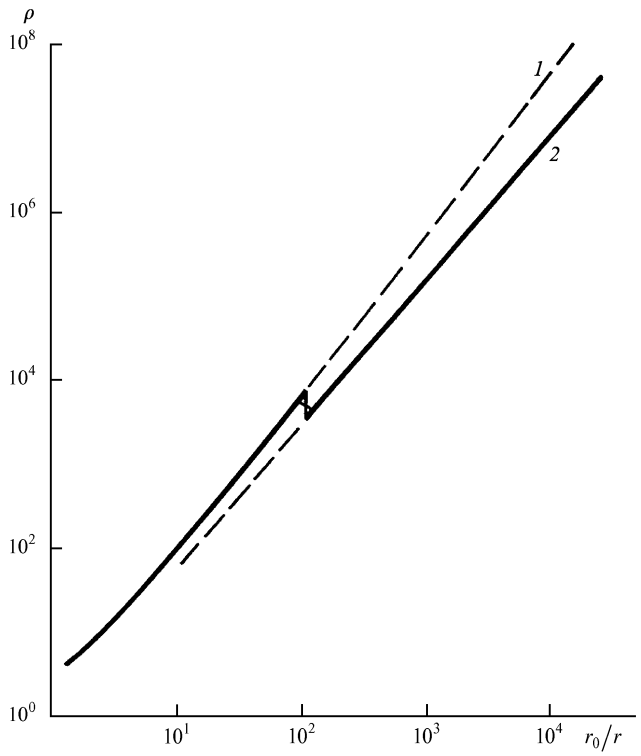

Figure 9. Radial distribution ($r_\varepsilon = 10^{-2}$) of the density in a mixed state: (*1*) in accordance with expression (120); (*2*) in accordance with expression (121).

Relationships (120)–(122) solve, in the adiabatic approximation, the problem of the appearance of a non-dissipative gravitational singularity in the presence of momentum. The radial distribution of the density near this singularity is shown in Fig. 9. The dashed line in Fig. 9 is the behaviour of the density described by expressions (120) and (121) (curves *1* and *2*, respectively). The difference between these two curves is slight. For example, if they are represented by the power low $r^{-\alpha}$, then for the value $r_\varepsilon = 10^{-2}$ adopted in this figure, we have $\alpha = 1.87$ and 1.72 for curves *1* and *2*, respectively.

It therefore follows that the steady-state distribution of the density near a singularity can be described approximately by a power law:

$$\rho = Kr^{-\alpha}, \quad \alpha \approx 1.7 - 1.9, \tag{123}$$

where with an error of about 5% the value of α can be regarded as constant.

7.3 Numerical simulation

In numerical simulation of the situation after the moment of appearance of a singularity we solved Eqns (112) and (113) subject to the initial conditions (110) at $t = t_c$. The simulation results show that immediately after the singularity a steep dip of the potential appears in the vicinity of the centre at $r = 0$, but the dip is shallower than in the spherical case. The solution then obtained has the

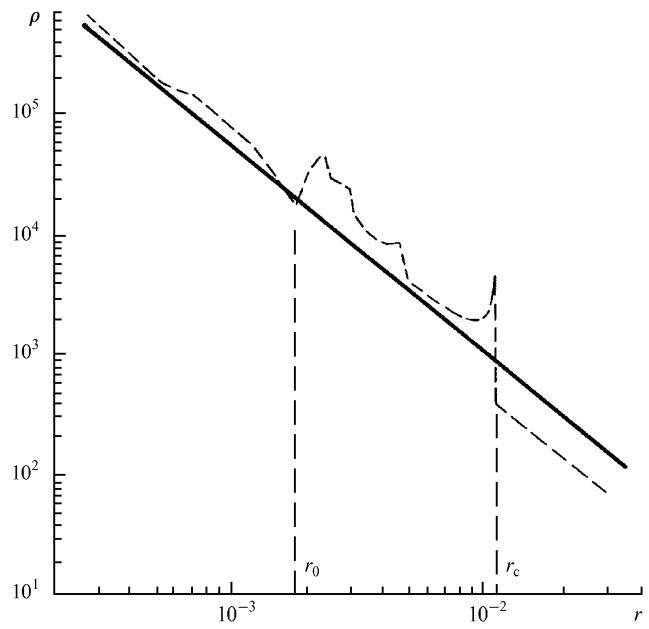

Figure 10. Spatial distribution of the density in the case when $\varepsilon^2 = 0.1$.

properties of both spherically symmetric and planar solutions.

We can see in fact three clearly separate zones in Fig. 10: (1) $0 < r < r_0$ is the zone where intense mixing of the streams takes place; (2) $r > r_c$ is the region of hydrodynamic single-stream flow, separated from the third zone by a caustic located at the point r_c; (3) $r_0 < r < r_c$ is an intermediate transition zone, the structure of which resembles the multicaustic pattern of one-dimensional flow (in this zone the mixing is not complete and the separate caustics can be seen). The continuous line in Fig. 10 is the distribution of the average density formed after the appearance of the singularity. We can see that the results of the numerical calculation agree well with the adiabatic theory predictions represented by expression (123). This prediction [25, 27] and the calculations [35] are supported by the results of numerical calculations carried out recently by other authors [58].

It therefore follows that during the development of the Jeans instability in cold nondissipative matter in the vicinity of the initial maximum of the effective density we can expect, after a sufficiently long time, a density singularity of the type described by expression (123) and this singularity is practically independent of the ellipticity ε of the initial maximum. However, the quantity ε influences significantly the time taken to reach the steady-state solution: the larger the value of ε, the longer the time needed for a system to assume a mixed state.

8. Hierarchical structure

8.1 General qualitative pattern

We have considered above the nonlinear dynamics of a single bunch which appears in the vicinity of the effective density maximum described by expression (29). Beginning from this section, we shall analyse the behaviour of a random distribution of the density of a nondissipative self-gravitating gas which has a wide spectrum of the initial fluctuations. It is important to stress that this is the Zel'dovich–Harrison spectrum that terminates abruptly at the wave number $k = k_{\max}$, which corresponds to the mass of dark matter particles (see Section 2). In the region of $k = k_{\max}$ the spectrum therefore has a clear maximum, so that inhomo-geneities which are specifically of the $R \approx R_m = k_{\max}^{-1}$ scale are the first to reach the nonlinear evolution stage.

Under the conditions specified by expression (18) there are no small-scale fluctuations and the large-scale fluctuations are still weak: they have not reached the nonlinear stage. This means that the nonlinear growth of the first inhomogeneities is of purely dynamic nature and, consequently, it can be described accurately by the solutions given above.

In a time exceeding the Jeans time $[t_g = (4\pi G\rho_0)^{-1/2}]$ the first steady-state dynamic NGS structures form on a scale of R_m. They are characterised by a definite scaling law of the distributions of the density, velocity, and field potential, described by expression (121), which depend little on the actual form of the initial maxima.

It is very important to note that during the period of the linear growth of inhomogeneities their scales increase with time in the course of the Hubble expansion of the universe, proportionately to the scaling factor. After the separation of a nonlinear bound object (NGS), the scale and structure of this object become fixed. During the subsequent Hubble expansion such NGSs participate as separate elements or objects separated by distances which increase, but the objects themselves do not change.

Moreover, the process of expansion of the Universe reduces the average density of matter, but the density in an NGS remains constant. The latter density is governed by the characteristic scale of the NGS and by the moment at which it forms. At a given moment the observed size of the object R' in comoving coordinates is related to its size R at the moment of formation by

$$R' = \frac{R}{1 + z_R}\,, \tag{124}$$

where z_R is the cosmological red shift associated with the moment of formation of the NGS.

The monotonic nature of the fluctuation spectrum corresponding to $k < k_{\max}$ and the continuity of the growth of the fluctuation amplitude with time [we are speaking here naturally of an unstable mode described by expression (24)] imply that the initial scale R_m of the nonlinear stage of the evolution is followed by larger scales ($R = 1/k > R_m$). Then, NGSs of scale $R_m < R$ may occur as separate elements of an NGS scale R. Then, after completion of nonlinear relaxation, NGSs of scale R also become separate objects in the expanding universe, so that the Hubble expansion inside them ceases.

This is followed by the nonlinear stage of the evolution of NGSs of scale $R_1 > R$. They may include NGS elements of scale R and of scale R_m. This is how a hierarchical structure of NGSs of various scales, embedded in one another, gradually grows. The present section is intended to give a description of this process.

It should be stressed that, since the initial fluctuation spectrum is random, different numbers of NGSs with different scales may combine to form separate objects of a hierarchical structure, depending on the statistics of the maxima in the initial spectrum. Some of these objects may

not combine with others at all and they retain their independence.

A study of various possible combinations of these objects should identify their statistics and the correlation properties. The statistics is dealt with in Section 10. Finally, we should mention that we are discussing here only completely formed NGSs. Some structures appear also in the course of their formation. Moreover, these structures are the largest-scale objects that separate out in the Universe. They are dealt with in Section 9.

8.2 Hierarchy of scales

We shall now consider a random distribution of the initial density of a gravitating gas characterised by a wide spectrum of fluctuations which are assumed to be homogeneous and isotropic. The dynamics of this gas is still described by the system of equations (7) subject to the initial conditions (27), where $\delta_i(\boldsymbol{r})$ is a random function with a wide spectrum of scale R. Let us introduce a distribution function averaged over the scale R:

$$f_R(\boldsymbol{r}, v, t) = \int f(\boldsymbol{r}+\boldsymbol{s}, v, t)\, W(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} \ . \tag{125}$$

Here $W(\boldsymbol{s})$ is a smoothed-out function, which is normalised and falls rapidly at infinity:

$$\int_{-\infty}^{+\infty} W(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} = 1 \ .$$

Suitable averaging in the system of equations (7) gives

$$\frac{\partial f_R}{\partial t} + v\cdot\frac{\partial f_R}{\partial \boldsymbol{r}} - \frac{\partial \psi}{\partial \boldsymbol{r}}\cdot\frac{\partial f_R}{\partial v} + S_R = 0 \ , \tag{126}$$

where

$$\nabla^2\psi = \int f\, \mathrm{d}v \ ,$$

$$S_R = \int \nabla(\delta\psi)\cdot\frac{\partial}{\partial v} f(\boldsymbol{r}+\boldsymbol{s})\, W(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} \ ,$$

$$\delta\psi = \psi(\boldsymbol{r}+\boldsymbol{s}) - \psi(\boldsymbol{r}) \ .$$

We shall show that the correlation integral S_R is generally small, so that in the first approximation it can be ignored. In the same approximation the system of equations (126) is practically identical with the initial system of equations (7). The only difference is that the initial density distribution $\delta_i(\boldsymbol{r})$ is averaged:

$$\delta(\boldsymbol{r}, R, t_i) = \int_{\infty}^{+\infty} W\left(|\boldsymbol{r}-\boldsymbol{y}|\right) \delta_i(\boldsymbol{y}, t_i)\, \mathrm{d}\boldsymbol{y} \ . \tag{127}$$

This means that all the maxima of the initial density with a scale smaller than R are smoothed out. Therefore, the solution of the system of equations (126) with the distribution (127) gives rise to an NGS of scale R or greater. It is understood that the singular law of the density distribution in an NGS, described by expression (123), applies also to scales of order R.

Let us now consider a scale $R_1 \gg R$. The equations for f_{R1} are of the same form as the system of equations (126), but the initial inhomogeneities may contain only the density maxima of scales $\tilde{R} \geqslant R_1$ because of the averaging described by expression (127). The dynamic evolution of such systems leads to the formation of stationary gravitational singularities of scale R_1 which contain singularities described by the function f_R of scale R, which behave as small elements trapped in the field of an NGS of scale R_1. This applies also to singularities of scale $R_2 \ll R$ trapped in the field of an NGS of scale R and described by a function $f_{R\,2}$.

In accordance with expression (124), it is important to stress that the size of each object R, R_1, or R_2 is governed by the red shift z_R which corresponds to its formation. The bulk of the objects of size R is formed when

$$\langle \delta^2(\boldsymbol{r}, R, t)\rangle = \delta_0^2 \sim 1 \ . \tag{128}$$

The quantity δ_0 is not known exactly, but in the case of a power-law spectrum it is independent of the smoothing-out scale R_0 and it is a universal constant.

Extrapolation of the linear law representing the increase in the density up to the moment t, described by expression (128), readily yields a relationship between z_R and $\boldsymbol{R}$ for a power-law spectrum of the initial perturbations:

$$1 + z_R = \left(\frac{R_{\mathrm{f}}}{R}\right)^{(m+3)/2} \ , \tag{129}$$

where m is the power exponent of the spectrum of the initial density perturbations (for the Zel'dovich–Harrison spectrum we have $m = 1$) and R_{f} is the maximum scale which at a given moment begins the nonlinear stage of its evolution described by expression (128) ($z_{R_{\mathrm{f}}} = 0$). From relationships (124) and (129) we can now obtain an expression relating the sizes of the objects in the comoving coordinate system in the present epoch and at the moment of their formation

$$R' = R^{(m+5)/2} R_{\mathrm{f}}^{-(m+3)/2} \ . \tag{130}$$

It therefore follows that a complete solution for the established range of scales $R < R_{\mathrm{f}}$ represents a hierarchical structure consisting of NGSs of different scales embedded in one another and moving along finite paths, and also of separate NGSs moving independently. Typical dimensions of NGSs are then described by relationships (128) and (130).

Scaling relationships (123) for the dark matter density, for the potential, and for the velocity are independent of the initial spectrum and remain the same for all the scales where multistream flow has already been established. On the other hand, the number of growing NGSs of different sizes depends strongly on the distribution of the initial fluctuations.

8.3 Cloud-in-cloud parameter and estimate of the correlation integral

The main parameter that characterises the degree of embedding or inclusion of smaller objects of a hierarchical struc-ture in larger ones is the cloud-in-cloud parameter $\epsilon(\boldsymbol{R})$, i.e. the probability that an object of size R is inside some other object. This probability is governed by the dependence of the 'concentration' of the maxima of the initial density distribution on their size $n(R)$ and by the spatial correlations of these maxima.

If in the long-wavelength range the initial spectrum obeys a power law $|\delta(k)|^2 \propto k^m$ (see Section 2), then in the case of sufficiently large values of R the scaling invariance of the spectrum leads to

$$n(R)\, \mathrm{d}R = 3\beta R^{-4}\, \mathrm{d}R \ , \tag{131}$$

where β is a dimensionless parameter which depends on the power exponent m of the spectrum. The concentration of objects of size exceeding R is $n(>R) = \beta(R^{-3} - R_{\rm f}^{-3})$, where $R_{\rm f}$ is the size of the largest objects that have formed up to a given moment.

Ignoring the correlations in the distributions of the objects, we can estimate the cloud-in-cloud parameter from

$$\epsilon(R) = \int_R^{R_f} \frac{4}{3}\pi R^3 n(R)\, {\rm d}R = 4\pi\beta \ln \frac{R_{\rm f}}{R} . \tag{132}$$

An estimate of the dimensionless parameter β deduced from observational data on clusters and groups of galaxies [37] gives

$$\beta \approx 6.4 \times 10^{-3} .$$

This empirical estimate may be somewhat underestimated because the contribution of dark matter to the masses of the objects is not taken fully into account. A theoretical estimate of the upper limit of β for a spectrum with the power exponent m also gives a low value [38]:

$$\beta < 0.016 \left(\frac{6}{m+5}\right)^{3/2} .$$

Therefore, the degree of embedding of objects in a real hierarchical structure is not very large. The evolution pattern of a hierarchical structure depends on time t. At any given moment t a hierarchical structure evolves from a scale $R_{\rm m}$ to a scale $R(t)$, which is given by expression (128). The maximum turbulence scale is of the order of $R_{\rm f}$: it is governed by the range of scales that have reached the nonlinear stage up to the moment in question.

We shall now estimate the correlation integral S_R in the system of equations (126). Such an integral for an NGS of smaller scale R trapped in an NGS of larger scale R_1 differs from an integral for untrapped NGSs, i.e. for those that move independently. In view of the smallness of the parameter β, the interaction between NGSs can be estimated from the Coulomb law.

The order of magnitude of a correlation integral is

$$S_R \sim \frac{f}{t_R} , \qquad t_R = p t_{\rm g} , \tag{133}$$

where the parameter p is the ratio of the 'mean travel time' to the Jeans time. For a freely moving NGS this parameter is

$$p \sim (\beta\Lambda)^{-1} , \tag{134}$$

where Λ is the Coulomb logarithm.

We shall now consider the case of trapped NGSs. In view of the smallness of the dimensionless parameter β, the probability of trapping an NGS of smaller size R by one of larger size R_1 is low: it is of the order of β. If we estimate the parameter p for this case and bear in mind the density distribution described by expression (123), we obtain

$$p \sim (\Lambda\beta^2)^{-1} \left(\frac{R_1}{R}\right)^{18/7} \left(\frac{r}{R_1}\right)^{6/7} . \tag{135}$$

We can see that the parameter p is always large: this is true both in the case of a freely moving NGS, described by formula (134), and in the case of trapped NGSs described by formula (135). Therefore, in the first approximation, we can ignore S_R in the system of equations (126).

In the case of a freely moving NGS a collision occurs mainly at the 'edges', which does not affect the structure of a singularity. However, in the case of trapped NGSs when $t > t_R$ a distortion occurs in the central part of the singularity. However, it should be pointed out that the estimate given by formula (135) is obtained on the assumption that in the course of a motion of an NGS of smaller scale R in an NGS of larger scale R_1 the former does not lose its mass. On the other hand, according to expression (123) the bulk of the mass is concentrated at $r \sim R$. Consequently, during the motion of an NGS of scale R in a potential of scale R_1 there is a loss of the mass of the NGS in a region of high gradients.

The process of inelastic head-on collisions may also be important. For example, it should be noted that, apart from the scattering of NGSs by one another, there is also scattering of dark-matter particles by an NGS. In view of the large mass of an NGS, the scattering of dark-matter particles changes only the angle but not the energy. Such scattering accelerates the change of the distribution function to the isotropic form by averaging this function over the angles.

In principle, the collisional processes discussed here lead to 'thermalisation' of matter. As is well known, under thermal equilibrium conditions, we have

$$\nabla\psi = -T\,\frac{\nabla\rho}{\rho} , \qquad \nabla^2\psi = 4\pi G\rho , \tag{136}$$

where T is the temperature of a gas in equilibrium. The solution of the system of equations (136) yields a singular density distribution (an isothermal sphere):

$$\rho = \frac{T}{2\mu\pi G}\, r^{-2} . \tag{137}$$

The law given by expression (137) is close to expression (123), but the distribution function described by Eqns (118) and (119) differs very greatly from the Maxwell–Boltzmann function. This is because the law described by expression (123) reflects mainly accumulation of streams directed towards the centre, whereas the equilibrium distribution described by expression (137) is dominated by the particles trapped inside a potential well. Therefore, a long period is necessary to convert the distribution function to its equilibrium form. Such an equilibrium is never reached in cold nondissipative matter.

In fact, the attainment of an equilibrium described by expression (137) requires a time much longer than the mean free time

$$t_{\rm f} \sim (n\sigma V)^{-1} , \tag{138}$$

where n is the dark-matter density, σ is the collision cross section of dark-matter particles, and V is their characteristic velocity, found from the distribution function defined by Eqns (118) and (119). Even in the case of a light neutrino with mass $m_\nu \approx 2$ eV the time $t_{\rm f}$ is between eight and nine orders of magnitude longer than the lifetime of the Universe. In the case of heavy dark-matter particles the time $t_{\rm f}$ is obviously even longer.

We have considered so far only the nondissipative dark matter. It is the absence of dissipation in cold self-gravitating dark matter that gives rise to NGSs. However, the dynamics of the baryonic matter is largely determined by dissipative processes. Such matter loses

energy by emission of radiation and drops to the bottom of the potential wells formed by the cold dark matter. The luminous baryonic matter then acts as an indicator of the structure of dark matter, identifying in particular the position of the centre of an NGS. This produces a unique object: baryonic matter with a halo of dark matter.

Examples of such objects are galaxies: in this case the evidence for the presence of a dark-matter halo is provided by flat rotation curves [39]. Other examples of such objects are clusters which contain a trapped hot gas. It should be pointed out that as the baryonic matter drops to the bottom of the potential wells, its density rises and, in the region where the average densities of the baryonic and dark matter become equal, there may be considerable distortions in the canonical distribution of dark matter described by Eqns (118) and (119).

9. Large-scale structure

In the preceding section we discussed a mixed kinetic state of nondissipative dark matter. This state appears in those regions where matter has oscillated many times in a self-consistent gravitational field. Kinetic mixing establishes a steady-state distribution function of matter. This steady state, which we call here an NGS, forms at the minima of the potential of the initial spectrum of fluctuations. It is characterised by a definite distribution function (118) and by a density singularity described by expression (123). A system of NGSs of different scales forms a hierarchical structure.

However, a general pattern of the structures that form in the Universe changes with time and depends on the initial fluctuation spectrum. For example, for the Zel'dovich–Harrison spectra a hierarchical structure of NGSs appears first of all on a small scale. Up to the present epoch, NGSs have evidently formed on scales $R \leqslant 5-10$ Mpc (galaxies, clusters). On the other hand, in the range of very large scales ($R > 100-200$ Mpc) the fluctuations are small and their growth is described by the linear theory. The distribution of matter in this range of scales is essentially homogeneous: there are no clear structures.

The range of scales 10 Mpc $\leqslant R \leqslant$ 100 Mpc represents the transition between the two cases. Within this range we can identify a subrange of intermediate scales $R \approx 50-100$ Mpc where perturbations just reach the nonlinear stage ($\delta \sim 1$). In this subrange the flow is of potential nature and it is described by single-stream hydrodynamics. The density of matter grows rapidly near a minimum of the potential. The distributions of the density and velocity are described by the system of equations (51) and by the set of expressions (55): they are characterised, apart from the scale R, by the dimensionless parameters ε_1 nd ε_2. Subranges of intermediate scales correspond in the Universe to nonlinear struc-tures of the 'Great Attractor' type [40] and in these structures the density contrast at the maximum reaches $\delta \sim 2-5$.

In the subrange $R \approx 20-50$ Mpc the flow is no longer of single-stream nature: the first caustics appear, but the number of streams is still small. The caustics represent bent flat objects of the Zel'dovich pancake type. The caustics intersect, forming filamentary objects and 'nodes', and a cellular structure is established.

This section deals with the cellular structure. The main structure elements are the caustics (Section 4) which are planar objects separating the zones of single-stream and triple-stream flow at which two streams merge [see expressions (66) and (67)].

We shall consider particularly the solution of the system of equations (49) near a caustic. An initial singularity, described by the set of expressions (41), is followed by the formation of a pair of caustics. Expansion of the solution (36) near a caustic for $\tau \ll 1$ gives the following expressions for the stream densities ρ_1, ρ_2, and ρ_3:

$$\rho_1 = \frac{\rho_0}{1+\tau}\,\frac{1-\zeta_1^2}{\zeta_1^2-\tau/(1+\tau)}\,,$$

$$\rho_2 = -\frac{\rho_0}{1+\tau}\,\frac{1-\zeta_2^2}{\zeta_2^2-\tau/(1+\tau)}\,,$$

$$\rho_3 = \frac{\rho_0}{1+\tau}\,\frac{1-\zeta_3^2}{\zeta_3^2-\tau/(1+\tau)}\,,$$

where

$$\zeta_1 = 2\left(\frac{\tau}{1+\tau}\right)^{1/2} - \frac{x-x_c}{3a\tau}\,,$$

$$\zeta_{2,3} = -\left(\frac{\tau}{1+\tau}\right)^{1/2} \pm\left\{\frac{x-x_c}{a}\,(\tau+\tau^2)^{-1/2}\right\}^{1/2}\,,$$

$$x_c = \frac{2}{3}\,\tau a\left(\frac{\tau}{1+\tau}\right)^{1/2}\,.$$

On approach to a caustic the densities ρ_2 and ρ_3 increase without limit proportionately to $(x-x_c)^{-1/2}$. The total density of matter

$$\rho = \sum_{i=1}^{3} \rho_i(x,\ \tau)\,, \tag{139}$$

plotted in Fig. 2b, also increases as $(x-x_c)^{-1/2}$ on approach to a caustic. Differentiation of the expression for x_c gives the velocity and acceleration of a caustic:

$$v_c = \frac{\mathrm{d}x_c}{\mathrm{d}\tau} = \frac{1}{1+\tau}\left(\frac{\tau}{1+\tau}\right)^{1/2}\left(1+\frac{2}{3}\tau\right)\,,$$

$$w_c = \frac{\mathrm{d}^2x_c}{\mathrm{d}\tau^2} = \frac{1}{2}\,\tau^{-1/2}\,(1+\tau)^{-5/2}\,.$$

A reference system moving together with the caustic is accelerated. Therefore, the following inertial force appears in the system:

$$W_c = 2\rho_0 a\, w_c\,.$$

If we use this circumstance and also the system of equations (65), we can find the potential in the vicinity of a caustic:

$$\psi = \psi_0 + \frac{\rho_0 a^2}{1+\tau}\left\{-\left[\tau^{-1/2}(1+\tau)^{-3/2} + 2\left(\frac{\tau}{1+\tau}\right)^{1/2}\left(1-\frac{\tau}{3}\right)\right]\frac{x-x_c}{a} - \frac{1}{6\tau}\,(1-3\tau)\left(\frac{x-x_c}{a}\right)^2 - \frac{4}{3}\,\frac{\theta(x-x_c)}{\tau^{1/4}(1+\tau)^{1/4}}\left(\frac{x-x_c}{a}\right)^{3/2}\right\}\,. \tag{140}$$

It is evident from the solution (140) that the potential does not form a well. Therefore, matter flows across the caustic and a trapped state of matter does not appear in its vicinity.

The solution (140) is valid at times close to the time $t = t_c$ of appearance of the initial singularity (i.e. at $\tau \ll 1$). After a sufficiently long time ($\tau \gg 1$) the form of the solution (140) depends on how fast the density of the initial distribution of matter decreases. If it decreases in accordance with the power law, i.e. if

$$\rho_0(x) = \rho_0 \frac{\mu}{3A^{1/3}} \left(\frac{x}{a}\right)^{-1-\mu/3} \qquad \text{for} \quad x \gg a\,,$$

where A and μ are arbitrary positive constants, the potential of the field near the caustic front is given by

$$\psi = \psi_0 + \rho_0 a \left\{ -\left[\frac{\pi}{2} - A^{-1/3}(\pi\tau)^{-\mu/3}\right](x - x_c) + \frac{2\sqrt{2}}{3a} C(\mu, A)\, \tau^{-(1/2+\mu/3)/(1+\mu/3)} (x - x_c)^{3/2}\, \theta(x - x_c) + \frac{\mu}{12aA^{1/3}} (\pi\tau)^{-(1+\mu/3)} (x - x_c)^2\, \theta(x_c - x) \right\}, \qquad (141)$$

where

$$C(\mu, A) = 2A^{-1/2(\mu+3)} \left(1 + \frac{\mu}{3}\right)^{-1/2} \left(\frac{\mu}{3}\right)^{1/2(1+\mu/3)}.$$

Therefore, the solution (141) derived for $\tau \gg 1$ is qualitatively similar to the solution (140) derived for $\tau \ll 1$. Only the time dependences of the coefficients are different.

We shall now consider filamentary structures that form as a result of crossing of caustic surfaces. We shall assume that caustic crossing does not create a deep potential well, capable of trapping streams. (This assumption will be justified later.) Then, in the first approximation, near a point of intersection all the streams move freely and do not influence one another. This means that the total potential of the gravitational field is a superposition of the potentials of two caustics.

It follows from the solution (141) that

$$\psi = \psi_0 + \sum_{j=1,2} \rho_0^{(j)} a_j \left\{ -\left[\frac{\pi}{2} - A^{-1/3}(\pi\tau_j)^{-\mu/3}\right] d_j + \frac{2\sqrt{2}}{3a_j} C(\mu, A)\, \tau_j^{-(1/2+\mu/3)/(1+\mu/3)}\, d_j^{3/2}\, \theta(d_j) + \frac{\mu}{12a_j A^{1/3}} (\pi\tau_j)^{-(1+\mu/3)}\, d_j^2\, \theta(-d_j) \right\}. \qquad (142)$$

Here, the index j numbers the caustics, d_j denotes the distance from the point X where the potential up to the jth caustic is calculated; $d_j > 0$ if the point X is in the region of three streams of the jth caustic, but $d_j < 0$ if the point X lies outside this region. The shading of the caustic shown in Fig. 11 is directed towards the region where there are three streams for a given caustic.

It is evident from Fig. 11 that the expressions for d_j ($j = 1, 2$) are given by

$$d_1 = x\,\cos\varphi - y\,\sin\varphi\,, \quad d_2 = -x\,\cos\varphi - y\,\sin\varphi\,, \qquad (143)$$

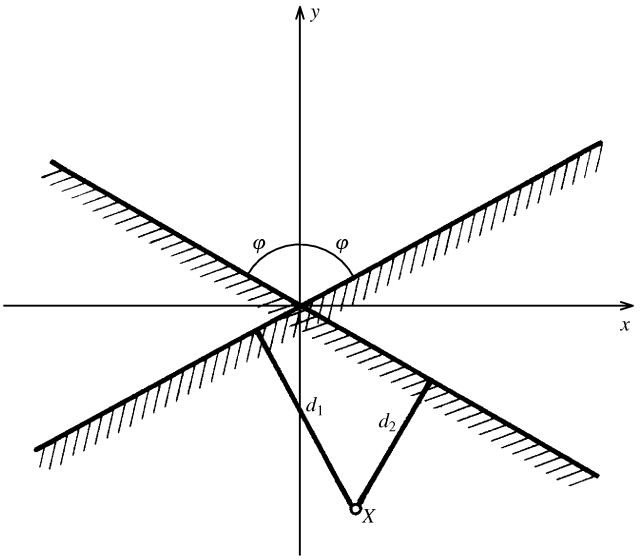

Figure 11. Pattern of intersecting caustics (X is the point at which the potential is calculated).

where (x, y) are the coordinates of the point X. Substitution of the formulas for d_j into expression (142) gives the potential near the point of intersection of two caustics. The general expression is very cumbersome.

The behaviour of the potential can be made qualitatively clear by considering the simplest case when the intersecting caustics have the same parameters, i.e. when

$$\rho_0^{(1)} = \rho_0^{(2)} = \rho_0^{(3)}, \qquad a_1 = a_2 = a_3\,, \qquad \tau_1 = \tau_2 = \tau\,.$$

Then,

$$\psi(x, y) = \psi_0 + 2\rho_0 a \left\{ \left[\frac{\pi}{2} - A^{-1/3}(\pi\tau)^{-\mu/3}\right] y\,\sin\varphi + \frac{2\sqrt{2}}{3a} C(\mu, A)\, \tau^{-(1/2+\mu/3)/1+\mu/3)} \times \left[(x\cos\varphi - y\sin\varphi)^{3/2}\theta\,(x\cos\varphi - y\sin\varphi) + (-x\cos\varphi - y\sin\varphi)^{3/2}\theta\,(-x\cos\varphi - y\sin\varphi)\right] + \frac{\mu}{12aA^{1/3}} (\pi\tau)^{-(1+\mu/3)} (x^2\cos^2\varphi + y^2\sin^2\varphi) \right\}. \qquad (144)$$

All the terms, apart from the linear, in expression (144) are positive and they increase with distance away from the point of intersection of the caustics. In other words, the potential $\psi(x, y)$ has the shape of a bent 'dish' tilted in the direction of negative values of y. It follows from expression (144) that the tilt of this dish is so large that the potential has no minimum in the vicinity of a filamentary singularity. The edges of the dish are only slightly bent since $\tau^{-(1/2+\mu/3)/(1+\mu/3)} \ll 1$, $\tau^{-(1+\mu/3)} \ll 1$.

The potential behaves similarly also for caustics with any set of parameters. It is evident from expression (144) that the linear term in the potential vanishes only in the degenerate case when $\varphi = 0$, i.e. when the caustic fronts collide.

It therefore follows that intersection of the caustics creates filamentary objects which also have a purely kinematic structure: their self-consistent gravitational field is incapable of trapping streams of matter, which justi-

fies — in particular — the hypothesis on the superposition of the potential described by expression (141).

In addition to filamentary objects, which are formed as a result of intersection of a pair of caustics, 'nodes' may form as a result of intersection of a filament with a plane. Let us consider the structure of the potential near objects of the node type. We shall do this by adopting a reference system in which the point of intersection is at rest. We shall describe the spatial orientation of a caustic by the vector

$$\lambda_j = (\alpha_j, \beta_j, \gamma_j), \quad |\lambda_j| = 1, \quad j = 1, 2, 3,$$

which is oriented along the normal to the caustic front and directed to the region of three streams.

The distance from the point $\boldsymbol{X}$ to the jth caustic is

$$d_j = \lambda_j \boldsymbol{X}.$$

Calculations similar to those made in the derivation of expression (144) yield the potential:

$$\psi(x, y, z) = \psi_0 + \sum_{j=1}^{3} \rho_0^{(j)} a_j \left\{ -\left[\frac{\pi}{2} - A^{-1/3}(\pi\tau_j)^{-\mu/3}\right] d_j + \frac{2\sqrt{2}}{3a_j} C(\mu, A)\, \tau_j^{-(1/2+\mu/3)/(1+\mu/3)}\, d_j^{3/2}\, \theta(d_j) + \frac{\mu}{12 a_j A^{1/3}} (\pi\tau_j)^{-(1+\mu/3)}\, d_j^2\, \theta(-d_j) \right\}. \quad (145)$$

Expression (145) is fully analogous to (144). As in the case of a filament, in the vicinity of a node there is in general no potential well capable of trapping and thus forming a steady kinetic state.

It therefore follows that the range of scales in which a large-scale cellular structure (pancakes, filaments, nodes) appears is — in a mathematical sense — an intermediate asymptote between the region of linear growth of perturbations and the kinetic region of multistream flow. Consequently, such large-scale structures cannot reach a steady state.

The whole pattern evolves in time because, on the one hand, all the larger scales approach the nonlinear stage and, on the other, the number of caustics, filaments, and nodes increases in the case of small scales. The strength of these singularities decreases with time (Fig. 5) and the range of scales gradually shifts to a mixed kinetic region where a hierarchical structure of NGSs, discussed above, is formed.

It should also be pointed out that singularities of the density of nondissipative matter in cellular structures are weaker than NGSs. In fact, it follows from expressions (141), (144), and (145) that the density singularities near pancakes, filaments, and nodes can be described by

$$\rho \propto d^{-1/2},$$

where d is either a coordinate along the normal to a caustic, a radius in a plane perpendicular to a filament, or a radius vector in the case of a node. In contrast to NGSs, the contribution of the singularities to the pair correlation function is not the main one. Filamentary singularities can make a singular contribution to a three-point correlation function, whereas singularities of the pancake type can make a contribution only to the fourth-order correlation function.

10. Correlation functions

10.1 General comments. Definition of an object

In the preceding sections we have considered the structure of nonlinear formations of dark matter. This section deals with investigations of the statistical properties of such formations. As shown in Section 7, in the vicinity of an initial density maximum the nonlinear gravitational self-compression creates an NGS, which is a steady-state self-trapped spherically symmetric distribution of the dark matter density $\rho(r)$ with a singularity at its centre described by expression (123):

$$\rho(r) = \tilde{\rho} \left(\frac{r}{R}\right)^{-\alpha}. \quad (146)$$

According to expression (123), the quantity α lies in the range 1.7 – 1.9. We shall use the average value: $\alpha \approx 1.8$. We can see that the power exponent α is constant under steady-state conditions and independent of the profile and scale of the initial maximum. This is in fact the main scaling property of three-dimensional nonlinear gravitational compression of cold nondissipative matter.

It therefore follows that each NGS represents a spherically symmetric formation with the intensity distribution described by expression (146) and this singularity is characterised by just two parameters: R and $\tilde{\rho}$. The radius R is a comoving size of the density maximum at a moment t_R [see expression (128)], when $\tilde{\delta}(\bar{0}, t_R) = \delta_0 \sim 1$. The moment t_R is a characteristic moment in time when a given density peak becomes separated by a nonlinear process from the overall cosmological expansion and the physical (intrinsic) size of the peak becomes fixed. The red shift corresponding to the moment t_R will be denoted by z_R.

It follows from the definition of t_R that at $t > t_R$ a peak passes through a nonlinear evolution stage and is transformed into an NGS of size R. It then exists as an independent object with a constant physical size and a steady-state distribution of the density inside it. The size R' of an object, expressed in comoving coordinates, observed during the present epoch, is related by expression (124) to its comoving size R at the moment of formation.

In this section an object is understood to be a fully formed NGS, which is a gravitationally bound dark-matter formation, in which the density distribution is given by expression (146) and whose size is R. Such objects form a hierarchical structure (Section 8). The degree of clusterisation of the structure is slight [see expression (132)].

The main task in this section is the development of a statistical theory of such objects. This will be done following Ref. [30]. We shall use the language of the statistics of N points representing the positions in space of all the objects formed with all possible sizes. The starting point in our theory is combined distribution of the probabilities $P(\bar{y}_i, R_i)$ of the formation of N objects of given dimensions R_i with centres at given points $\bar{y}_i$.

The influence of nonlinear dynamics on the statistical properties of a system of objects is usually described by an infinite chain of coupled equations for the moments of an N particle distribution function, known as the Bogolyubov – Born – Green – Kirkwood – Yvon (BBGKY) chain. This problem has been discussed on many occasions for the case of self-gravitating matter and, in particular, it is dealt with in Peeble's monograph [21].

Such a chain of equations is usually solved by truncating it and neglecting higher momenta. However, as pointed out above, a strongly correlated singular distribution of matter appears in regions of the effective density maxima, where a pair correlation function and higher momenta are characterised by $\xi \gg 1$ and where, consequently, such a truncation procedure is incorrect. The formation of strongly correlated distributions should be taken into account properly and this aspect underlies the method considered below.

The appearance of strongly correlated distributions can be described by introducing the transition probability determined on the basis of the nonlinear solution described by expression (146). This approach makes it possible to find directly the relationship between the final values of the correlation functions and their initial values, and it is the main new feature of our method. It makes it possible to avoid solving a chain of coupled equations and it can effectively take care of the whole BBGKY hierarchy in the region defined by $\xi \gg 1$.

10.2 Model of the formation of correlations

As pointed out above, an NGS is formed fairly rapidly after the effective density, given by expression (29), reaches $\delta \sim 1$ at a maximum. In our statistical model we shall assume that this occurs instantaneously. If some smaller object is trapped by a given NGS, the probability $f(r)$ of finding it at present in a unit volume at a distance r from its centre can be regarded as proportional to the density of matter given by expression (146):

$$f(r, R) = \frac{3-\alpha}{4\pi} R'^{-3} \theta(R'-r) \left(\frac{r}{R'}\right)^{-\alpha} . \tag{147}$$

The constant in the above expression is found by postulating mass conservation in a volume $r < R$ during the formation of an NGS when the comoving size R' of an object during the present epoch is given by formula (124).

The density distribution inside an NGS, which is governed by the laws of nonlinear compression, is given by expression (146) only in the asymptotic limit $r \to 0$ and the error contributed by it is of the order of 1 at $r \sim R'$. Consequently, at distances $r \geqslant R'$ we can expect the corrections to the probability distribution (147) to be significant.

We shall now determine the conditions for trapping one object by another. An object of size R_2 with its centre at a point $\bar{y}_2$ traps an object of size R_1 with its centre at the point $\bar{y}_1$ if the following conditions are satisfied:

$$R_1 < R_2 , \quad |\bar{y}_1 - \bar{y}_2| < R_2 . \tag{148}$$

If the above conditions are not satisfied by any objects ($\bar{y}_2$, R_2), we can regard an object ($\bar{y}_1$ R_1) as remaining in place.

In calculation of diagrams of higher order in terms of the number of participating objects the important aspect is the distribution of the times of formation of objects of a given size (red shift). Strictly speaking, the times of formation are subject to a scatter, but for the sake of simplification let us assume that all the objects with a given size R are formed simultaneously and the corresponding red shift is z_R. Here, z_R corresponds to the moment when the majority of the objects of size R forms. The value of z_R is determined by the spectrum of the initial inhomogeneities $|\delta_i(k)|^2$ (Section 8).

Every object in our model is thus characterised by the coordinates of its centre $\bar{x}_i$ and by its size R_i. The structure of the distribution of matter is described by the statistics of a large number N of points in the physical space $\bar{x}$ and in the space of sizes R. The nonlinear evolution is described by a random process of consecutive trapping of smaller objects by larger ones. The rules for finding the appropriate transition probability in the case of a single trapping event are discussed below. In calculation of the final j-point correlation functions it is essential to know also the initial correlations of all the N objects and the rules for calculation of the transition probability for multiple trapping. Let us now consider these topics.

10.3 Calculation of j-point correlation functions

We shall determine the transition probability for a test object which is trapped by several other objects. Let the radius of the first object, which has captured the test object, by R_1. This is followed by the formation of an object of size $R_2 > R_1$, which traps the first object. The probability $W(\bar{r})$ of a transition of the test object to the final state, given by the vector $\bar{r}$ relative to the centre of the second object, can be represented in the form

$$W(\bar{r}) = \int f(|\bar{r}+\bar{x}|, R_1) f(|\bar{x}, R_2|) \, \mathrm{d}\bar{x} ,$$

if in accordance with our assumption the internal structure of the trapped object is not altered by collisions and tidal forces because of insufficient time.

For the sake of brevity, let us adopt the term the 'density of the probability of finding objects (of a given size) at such points' to mean the density of the probability of finding objects per unit volume (per unit interval of sizes) in the vicinity of such points.

Since in the process of mixing the various trapped objects quickly 'forget' their initial correlations, the nominal densities of the probability of a transition involving several objects within one other object can be calculated independently and multiplied. For example, if there are two objects of sizes R_1 and R_2 and they are trapped by an object of size R_3 which is at a point $\bar{x}_3$, the density of the probability of detecting them after mixing at points $\bar{x}_1$ and $\bar{x}_2$ is quite simply

$$W(\bar{x}_1, \bar{x}_2) = f(|\bar{x}_1 - \bar{x}_3|, R_3) f(|\bar{x}_2 - \bar{x}_3|, R_3) .$$

The density of the probability of a transition in general is calculated in a similar manner. We shall use $W_a^n(\bar{x}_i, \bar{y}_i, R_i)$ to denote the density of the probability of finding n objects with dimensions R_i at points $\bar{x}_i$ on the assumption that they are formed at points $\bar{y}_i$. Then the density of the probability of finding, during the present epoch, j objects with dimensions $R_1, \ldots, R_j$ at points $\bar{x}_1, \ldots, \bar{x}_j$ can be represented by

$$\begin{aligned} P(\bar{x}_1, \ldots, \bar{x}_j, R_1, \ldots, R_j) &= \sum_a P_a(\bar{x}_1, \ldots, \bar{x}_j, R_1, \ldots, R_j) \\ &= \sum_{n=j}^{\infty} \sum_a \int \ldots \int W_a^n(\bar{x}_i, \bar{y}_i, R_i) P_a^n(\bar{y}_i, R_i) \\ &\times \prod_{i=1}^{n} \mathrm{d}\bar{y}_i \prod_{i=j+1}^{n} \mathrm{d}\bar{x}_i \, \mathrm{d}R_i . \end{aligned} \tag{149}$$

Here, the index a is used to number all possible variants of the spatial embedding of n objects in one another (in other words, all the different trapping configurations). For example, if $j=2$ and $n=2$ or 3, all possible variants of the embedding of objects *1*, *2*, and *3* have the form shown schematically in Fig. 12a.

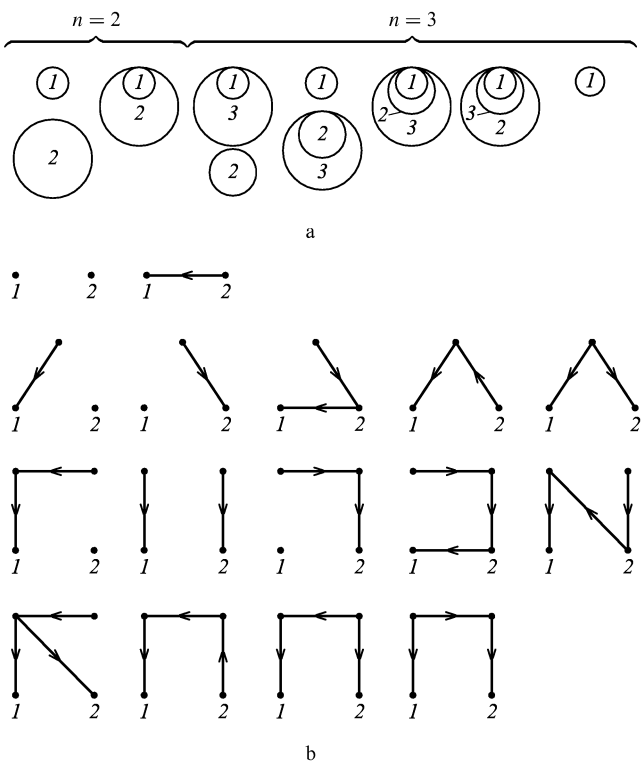

Figure 12. (a) Schematic representation of all the variants of spatial embedding of n objects in one another, corresponding to Eqn (149) with $j=2$. (b) All permissible diagrams with $n \leqslant 4$ that contribute to the pair ($j=2$) correlation function of the participating objects.

It is important to note that in the enumeration of the trapping configuration in expression (149) only those objects are included which actually trap at least one of the objects $1, \ldots, j$. The objects which are trapped, but themselves do not trap any other object are ignored. Finally, the summation in expression (149) over n up to ∞ is meaningful if the Universe contains any number of large objects. In reality, because of the limited evolution time of galaxies it is sufficient to consider only, for example, $n < (2-3)j$. This is justified further by the circumstance that the terms with large values of n should be small because of the smallness of the cloud-in-cloud parameter ϵ.

In expression (149) the quantity $P_a^n(\bar{y}_i, R_i)$ is the density of the probability of the formation of given n objects at points $\bar{y}_i$ in the configuration shown in Fig. 12a, so that none of these objects traps any other external objects. This probability density can be described by

$$P_a^n(\bar{y}_i, R_i) = \Theta_a^n(\bar{y}_i, R_i)\, P^n(\bar{y}_i, R_i) \,.$$

Here, $P^n(\bar{y}_i, R_i)$ is the density of the probability of the formation of some n objects of dimensions R_i at points $\bar{y}_i$, and of the formation of all the other objects wherever possible, but in such a way that they do not trap the selected objects. The function $\Theta_a^n(\bar{y}_i, R_i)$ is equal to unity if the objects are in the configuration shown in Fig. 12a and zero in the opposite case. This function ensures that the conditions of expression (148) are satisfied for each pair of objects. Thus, in the case when $j=2$ and $n=2$ or 3 the above configurations are described by the following functions:

$$\Theta_0^2 = \theta(|\bar{y}_1 - \bar{y}_2| - R_2) \,,$$

$$\Theta_1^2 = \theta(R_2 - |\bar{y}_1 - \bar{y}_2|) \,,$$

$$\begin{aligned}\Theta_2^3 = {} & \theta(R_3 - R_2)\,\theta(R_3 - |\bar{y}_1 - \bar{y}_3|)\,\theta(|\bar{y}_2 - \bar{y}_3| - R_3) \\ & \times \theta(|\bar{y}_1 - \bar{y}_2| - R_2) + \theta(R_3 - R_1)\,\theta(R_2 - R_3) \\ & \times \theta(R_3 - |\bar{y}_1 - \bar{y}_3|)\,\theta(|\bar{y}_2 - \bar{y}_3| - R_2) \,,\end{aligned}$$

$$\begin{aligned}\Theta_3^3 = {} & \theta(R_3 - R_2)\,\theta(R_3 - |\bar{y}_2 - \bar{y}_3|)\,\theta(|\bar{y}_1 - \bar{y}_3| - R_3) \\ & \times \theta(|\bar{y}_1 - \bar{y}_2| - R_2) \,,\end{aligned}$$

$$\Theta_4^3 = \theta(R_3 - R_2)\,\theta(R_3 - |\bar{y}_2 - \bar{y}_3|)\,\theta(R_2 - |\bar{y}_1 - \bar{y}_2|) \,,$$

$$\begin{aligned}\Theta_5^3 = {} & \theta(R_2 - R_3)\,\theta(R_3 - R_1)\,\theta(R_3 - |\bar{y}_1 - \bar{y}_3|) \\ & \times \theta(R_2 - |\bar{y}_2 - \bar{y}_3|) \,,\end{aligned}$$

$$\begin{aligned}\Theta_6^3 = {} & \theta(R_3 - R_2)\,\theta(R_3 - |\bar{y}_2 - \bar{y}_3|)\,\theta(R_3 - |\bar{y}_1 - \bar{y}_3|) \\ & \times \theta(|\bar{y}_1 - \bar{y}_2| - R_2) \,.\end{aligned}$$

The corresponding densities of the transition probabilities are of the form

$$W_0^2 = \delta(\bar{x}_1 - \bar{y}_1)\,\delta(\bar{x}_2 - \bar{y}_2) \,,$$

$$W_1^2 = f(|\bar{x}_1 - \bar{x}_2|, R_2)\,\delta(\bar{x}_2 - \bar{y}_2) \,,$$

$$W_2^3 = f(|\bar{x}_1 - \bar{x}_3|, R_3)\,\delta(\bar{x}_2 - \bar{y}_2)\,\delta(\bar{x}_3 - \bar{y}_3) \,,$$

$$W_3^3 = f(|\bar{x}_2 - \bar{x}_3|, R_3)\,\delta(\bar{x}_1 - \bar{y}_1)\,\delta(\bar{x}_3 - \bar{y}_3) \,,$$

$$W_4^3 = f(|\bar{x}_1 - \bar{x}_2|, R_2) f(|\bar{x}_2 - \bar{x}_3|, R_3)\,\delta(\bar{x}_3 - \bar{y}_3) \,,$$

$$W_5^3 = f(|\bar{x}_1 - \bar{x}_3|, R_3) f(|\bar{x}_2 - \bar{x}_3|, R_2)\,\delta(\bar{x}_2 - \bar{y}_2) \,,$$

$$W_6^3 = f(|\bar{x}_1 - \bar{x}_3|, R_3) f(|\bar{x}_2 - \bar{x}_3|, R_3)\,\delta(\bar{x}_3 - \bar{y}_3) \,.$$

Each term in the sum (149) can be assigned a diagram in accordance with the following rules:

(1) each arrow $i \longrightarrow k$ corresponds to a factor $f(|\bar{x}_i - \bar{x}_k|, R_i)$;

(2) a vertex with $i \neq 1, \ldots, j$ which is not approached by any arrow ($i \cdot \rightarrow$) corresponds to a factor $\delta(\bar{x}_i - \bar{y}_i)$;

(3) the whole diagram with n vertices is assigned additionally a factor $P_a^n(\bar{y}_i, R_i)$.

Integration is carried out over all the values of $\bar{x}_i$, $\bar{y}_i$, and R_i apart from $\bar{x}_1, \ldots, \bar{x}_j$ and $\bar{R}_1, \ldots, \bar{R}_j$. The sum (149) contains all possible diagrams that include $n \geqslant j$ vertices and which are characterised by the following properties:

(4) not more than one arrow enters any vertex;

(5) the only vertices from which not even one arrow emerges are those labelled $1, \ldots, j$.

In this way the sum contains only the diagrams that do not contain loops or parts which are not associated with the vertices $1, \ldots, j$. All permissible diagrams $n \leqslant 4$, which contribute to the pair ($j=2$) correlation function are plotted in Fig. 12b.

Summation of all the possible diagrams for a given j yields the density of the probability of finding, during the present epoch, j objects of given dimensions at given points. Therefore, expression (149) and the diagram procedure

described above solve, in principle, the problem of calculating the final j-point correlation function on condition that the probability density $P^n(\bar{y}_i, R_i)$ is known for any initial distribution of an arbitrary number n of untrapped objects at the moment of their formation.

The quantity $P^n(\bar{y}_i, R_i)$ represents the density of the probability of finding a given relative configuration of the density maxima of various dimensions during the linear stage and, as assumed by us, it is governed entirely by the initial spectrum of inhomogeneities $|\delta_i(k)|^2$. Calculation of this quantity for an arbitrary spectrum is a separate task, which is not considered here.

We shall use later $P^n(\bar{y}_i, R_i)$ in that simple case when there are no initial correlations between the positions of all the objects formed. We can show that this simplification does not alter the form of the pair correlation function ξ in the strong correlation range ($\xi \gg 1$). In this case, on condition that the total number of the objects formed is large, we obtain

$$P^n(\bar{y}_i, R_i) = n(R_1)\ldots n(R_n)\left\{1 - \epsilon(v) + \frac{1}{2}\epsilon^2(v) + \mathrm{O}(\epsilon^3)\right\}, \tag{150}$$

where

$$\epsilon(v) = \int_{v(\bar{y}_i, R_i)} n(R)\,\mathrm{d}R\,\mathrm{d}\bar{y}\,.$$

Integration in the above expression is carried out over the full range of v in the space of $(\bar{y}, R)$ which contains objects that have trapped at least one of the objects $(\bar{y}_i, R_i)$; $n(R)\,\mathrm{d}R$ is the concentration of the objects with dimensions ranging from R to $R + \mathrm{d}R$. In the case of a scaling-invariant object this concentration is given by formula (131).

10.4 Pair correlation function

We shall now consider in greater detail the most interesting, from the point of view of observations, pair correlation function. We shall calculate this function from expression (149) for $j = 2$. We shall use expression (150) to describe P^n. We note that in the case of small values of n of interest to us, the quantity $\epsilon(v)$ is a small parameter proportional to the cloud-in-cloud parameter (132):

$$\epsilon(v) \leqslant \sum_{i=1}^{n} \int_{R_i}^{R_f} \frac{4}{3}\pi R^3 n(R)\,\mathrm{d}R = 4\pi\beta \sum_{i=1}^{n} \ln\frac{R_f}{R_i}\,.$$

All the diagrams that contribute to the pair correlation function consist of one or two connected pieces and are of the form shown in Fig. 13a. It can readily be demonstrated that any of these diagrams containing a 'tail' of k links (or two tails having a total of k links) represents a small correction of the order of ϵ^k to the corresponding diagram without a tail. Thus, all the leading terms in the expansion in terms of ϵ have no tails. We shall number these diagrams, beginning from those most important in terms of the small parameter ϵ, exactly as shown in Fig. 13b.

In view of the homogeneity and isotropy of the statistics of the initial perturbations, the required probability density (and the pair correlation function) depend only on $r = |\bar{x}_1 - \bar{x}_2|$. The zeroth diagram describes the contribution of the initial correlations of the objects not trapped by any other objects. Since in this approximate calculation the

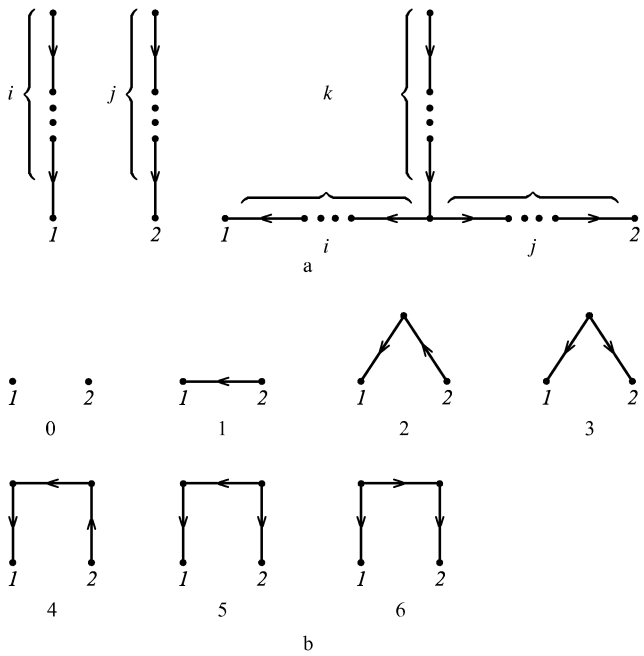

Figure 13. (a) General form of all the diagrams that contribute to the pair correlation function ($i, j, k = 0, 1, 2, \ldots$). (b) Diagrams for $j = 2$, $n \leqslant 4$, which are the leading diagrams in terms of the small parameter ϵ.

initial correlations are ignored, the contribution of the zeroth diagram is trivial:

$$P_0(r, R_1, R_2) = n_1 n_2\,\theta(r - R_2)\left\{1 + \mathrm{O}(\epsilon)\right\},$$

$$n_1 \equiv n(R_1)\,, \qquad n_2 \equiv n(R_2)\,.$$

The first diagram describes the contribution of pairs:

$$P_1(r, R_1, R_2) = n_1 n_2\,\theta(R_2' - r)\,\frac{3-\alpha}{3}\,(1 + z_2)^3\left(\frac{r}{R_2'}\right)^{-\alpha}, \tag{151}$$

where $z_2 = R_2/R_2' - 1 = (R_f/R_2')^{(m+3)/(m+5)} - 1$ [see expression (29)] is the red shift at the moment of formation of the larger object in a pair. If $r < R_2'$ and $R_1 \neq R_2$, it is this first diagram that dominates the pair correlation function.

The next diagrams describe the contribution of three-member systems:

$$P_2(r, R_1, R_2) = n_1 n_2\,\frac{2}{3}\,\frac{3-\alpha}{m+5}\,A\left[\left(\frac{R_2'}{R_1'}\right)^{3-\alpha} - 1\right] \times \beta(1 + z_2)^3\left(\frac{r}{R_2'}\right)^{3-2\alpha}, \quad r \ll R_2'\,, \tag{152}$$

$$P_3(r, R_1, R_2) = n_1 n_2\,\frac{3-\alpha}{16\gamma}\,A\left[\frac{17}{6} + \frac{22}{\gamma - 6}\right] \times \beta(1 + z_2)^3\left(\frac{r}{R_2'}\right)^{3-2\alpha}, \quad r \ll R_2'\,, \tag{153}$$

where $\gamma = (m+5)(3-\alpha) + 3$ and A is a constant of the order of 1.

Expressions (151)–(153) contain the leading terms of the expansion in ϵ. We can show that there are no other contributions, diverging in the limit $r \to 0$, to the correla-

tion function. In calculation of the dominant diagram described by expression (151) no use has been made of the relationship (130) between the moment of formation of an object and its size. In view of this, the main contribution to the correlation function can be considered ignoring the relationship (130) and assuming that R_i and z_i are independent characteristics of an object.

Moreover, we need not use the assumption represented by expression (150) that there are no correlations in the positions of the objects formed and we can consider the case of an arbitrary probability $P^n(\bar{y}_i, R_i)$. Then, the contribution of the main diagram becomes

$$P_1(r, R_1, R_2) = C(R_1, R_2)\, n_1 n_2\, \theta(R_2' - r) \times \frac{3-\alpha}{3}(1+z_2)^3 \left(\frac{r}{R_2'}\right)^{-\alpha} ,$$

where

$$C(R_1, R_2) = \frac{3}{4\pi n_1 n_2 R_2^3} \int_{|\bar{y}_1|<R_2} P^2(\bar{y}_1, \bar{0}, R_1, R_2)\, \mathrm{d}\bar{y}_1$$

is a factor which is independent of r and takes account of the initial correlations. Therefore, in the most general case, the main contribution to the probability has the form given by expression (155) below, accurate apart from the factor C which is independent of r.

The observed objects (such as clusters of certain richness, etc.) correspond in reality to a whole range of sizes from R_{i1} to R_{i2} (or, correspondingly, of masses from M_{i1} to M_{i2}), and also possibly to a range of the moments of formation. Therefore, a more correct, i.e. corresponding closer to reality, expression for (for example) the pair autocorrelation function is

$$\xi_i(r) = \frac{1}{n_i^2} \int_{R_{i1}}^{R_{i2}} \int_{R_{i1}}^{R_{i2}} P(r, R_1, R_2)\, \mathrm{d}R_1\, \mathrm{d}R_2 - 1 ,$$
$$n_i = \int_{R_{i1}}^{R_{i2}} n(R)\, \mathrm{d}R . \tag{154}$$

In the case of strong correlations ($\xi_i \gg 1$) the leading term found from expressions (151) and (154) is

$$\xi_i(r) = \left(\frac{a d_i}{r}\right)^{\alpha} , \tag{155}$$

where

$$a = \beta^{1/3}\left(1 - \frac{M_{i1}}{M_{i2}}\right)^{1/3} (3-\alpha)^{1/\alpha} (1+z_i)^{(3-\alpha)/\alpha} , \tag{156}$$

on the condition that the ranges of the masses $M_{i2} - M_{i1}$ and of the times of formation $z_{i1} - z_{i2}$ are fairly narrow. Here, $d_i = n_i^{-1/3}$ is the average distance between the objects, given by

$$d_i = R_{i1}\beta^{-1/3}\left(1 - \frac{M_{i1}}{M_{i2}}\right)^{-1/3} . \tag{157}$$

The quantity z_i denotes the average moment of formation of the investigated objects: $M_{i1,2} = (4/3)\,\pi\rho_0 R_{i1,2}^3$. The ratio of the baryonic and dark components of the masses of the objects is, on average, constant, so that the baryonic masses (156) and (157) can be substituted in expressions M_{i1} and M_{i2}.

11. Conclusions

We shall conclude by considering briefly some of the astrophysical manifestations of the processes discussed above.

11.1 Giant halo of galaxies

Nonlinear structures which appear in the dark matter distribution are discussed above. However, it is the baryonic matter which is observable. It is natural to assume that in a homogeneous Universe the baryonic dark components are mixed uniformly. The ratio of their densities is given by the parameter

$$P = \frac{\rho_{\mathrm{b}}}{\rho_{\mathrm{d}}} , \tag{158}$$

which is a universal constant.

It should be noted that this parameter P is related to other universal constants by

$$P = \frac{8\pi}{3}\frac{G\rho_{\mathrm{b}}}{\Omega H^2} . \tag{159}$$

If $\Omega = 1$ and $50\ \mathrm{km\ s^{-1}\ Mpc^{-1}} \leqslant \mathrm{H} \leqslant 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, and since $\rho_{\mathrm{b}} \approx 3 \times 10^{-31}\ \mathrm{g\ cm^{-3}}$, we obtain

$$0.02 \leqslant P \leqslant 0.08 , \tag{160}$$

which is in agreement with the current ideas [2].

During the linear stage of the evolution of inhomogeneous structures in the universe, after recombination, the ratio of the densities of the baryonic and dark matter components (158) is conserved. During the nonlinear stage, stable spherically symmetric NGSs form in the nondissipative matter distribution. This is accompanied by the appearance of gravitational potential wells. The baryonic gas is heated during the nonlinear stage by compression and explosion of supernovas. It emits radiation, losing energy, and gradually drops to the bottom of a potential well, forming a galactic system.

Luminous galaxies become distributed in the central regions of spherically symmetric dark matter structures (NGSs), which appear as a giant halo of the galactic systems. This halo was predicted in Refs [25, 41].

Let us determine the size of the giant halo [42]. We shall bear in mind that, according to expression (123), the distribution law of the density of the nondissipative matter in an NGS is

$$\rho = K r^{-\alpha} , \quad \alpha \approx 1.8 .$$

Consequently, the total dark matter mass in a galactic NGS is

$$M = \frac{4\pi}{3-\alpha} K R_{\mathrm{g}}^{3-\alpha} , \tag{161}$$

where R_{g} is the effective size of the dark-matter halo. Hence,

$$R_{\mathrm{g}} = \left[\frac{(3-\alpha)M}{4\pi K}\right]^{1/(3-\alpha)} = \left[\frac{(3-\alpha)M_{\mathrm{G}}}{4\pi K P}\right]^{1/(3-\alpha)} . \tag{162}$$

Here, M_{G} is the observed baryonic mass of a galaxy; P is the universal parameter, defined above [expressions (159) and (160)], and the constant K is a typical parameter which can be different for different galaxies: it is governed by the scale of the galaxies and by the moment of appearance of an NGS.

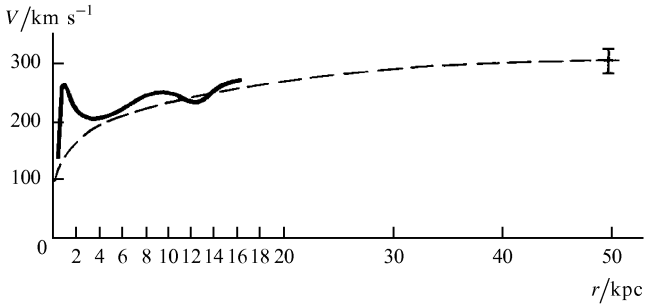

Figure 14. Rotation of our galaxy. Dashed curve is the theoretical dependence $V(r)$.

The constant K can be determined from the rotation curves. The observation of these curves has made it possible to detect the latent (dark) matter in the vicinity of galaxies. Fig. 14 shows the rotation curve of our galaxy. The dashed curve is the theoretical dependence $V(r)$, which is plotted — in accordance with the virial theorem — for the density distribution given by expression (123):

$$V^2 = (2-\alpha)\psi = \frac{4\pi}{3-\alpha} KGr^{2-\alpha} .$$

The constant K is then

$$K \approx 1.7 \times 10^{16} \text{ g cm}^{-1/2} . \tag{163}$$

It follows from formula (163) that if the distance of the Sun from the centre of the galaxy is taken to be $r_\odot = 8$ kpc, the dark-matter density is of the order of 0.5 GeV cm^{-3}. This value is in agreement with the results obtained by other authors [43]. Substitution of formula (163) into expression (162), gives the size of the giant halo of our galaxy:

$$R_g = R_G \left(\frac{5\%}{P} \frac{M_G}{3 \times 10^{44} g}\right)^{1/(3-\alpha)} , \quad R_G \approx 200 \text{ kpc} . \tag{164}$$

Hence, the size of the halo of our galaxy is of the order of 200 kpc.

It should be stressed that the distribution law of the dark-matter density in an NGS given by expression (123) is well supported by the rotation curves of other spiral galaxies (Fig. 15). On the other hand, determination of the size of the giant halo gives values similar to those in formula (164). but with a larger scatter which is associated mainly with the difference between the masses of such spiral galaxies [44].

The moment of creation z of galaxies is determined similarly. If we take into account an increase in the scales because of the Hubble expansion from the moment z to the present epoch, and if we assume that the scale of a given galactic system at the moment z of its creation is $R_0 = R_g$, we find that in the present epoch this size is

$$R = R_g (1+z) .$$

On the other hand, the size R can easily be expressed in terms of the baryonic mass M_G of a galaxy:

$$M_G = \frac{4\pi}{3} \rho_b R^3 , \quad R = \left(\frac{3}{4\pi} \frac{M_G}{\rho_b}\right)^{1/3} ,$$

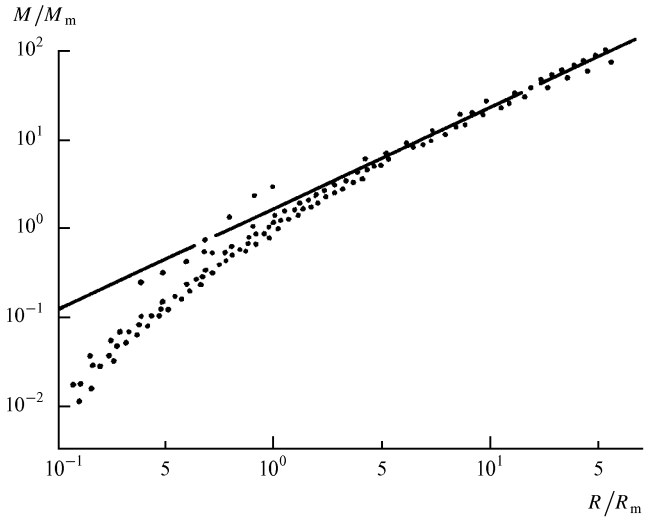

Figure 15. Distribution of the dynamic mass in the vicinity of spiral galaxies [39]. The continuous line is the theoretical dependence.

where $\rho_b \approx 3 \times 10^{-31}$ g cm^{-3} is the average density of the baryonic matter during the present epoch. Hence, it follows that

$$z = \frac{R}{R_g} - 1 .$$

For our galaxy, this relationship leads to $z \approx 9$.

The existence of a giant halo of galaxies whose size is given by formula (164) and in which the dark-matter distribution is described by expression (123) is typical only of the objects with a fully formed NGS. This condition is far from being satisfied in every case. For example, the structure of an NGS is disturbed by collisions of galaxies (it should be stressed that such galaxy collisions represent primarily the collisions of their NGSs), in the presence of a strong gradient of the gravitational field, and also in the case of objects which are not fully formed.

The last situation may be observed during the present epoch, first, because of the scatter of the moments of creation of galaxies and, second, because the time to form an NGS depends strongly on the structure of the initial fluctuations: in the case of perturbations characterised by a large value of ε, i.e. those which are strongly elongated or planar, the time needed for the formation of an NGS is considerably longer than the time before the appearance of the initial fluctuations with small values of ε (Sections 4 and 6). A criterion of completion of the formation of an NGS may be, for example, the shape of a galaxy: it is most likely that an NGS has not yet been formed in the case of irregular galaxies. Therefore, the ideas on the structure of dark-matter formations put forward in Ref. [45] on the basis of an analysis of several irregular galaxies cannot be regarded as sufficiently well-grounded.

Observational data confirming the existence of a giant dark-matter halo have started to appear recently [46]. Moreover, some support for the existence of a giant halo follows from the observations of gamma-ray bursts. This is discussed below.

11.2 Giant halo of neutron stars (model of the origin of gamma-ray bursts)

It is well-known that in the course of formation of a galaxy the initial baryonic matter — containing hydrogen, helium, and small amounts of light elements (Section 1) — becomes enriched with heavy elements that are created by nuclear reactions in stars. It is very important to stress that considerable enrichment occurs already at the very early stage of a protogalaxy [47]. It is usually assumed that this is the result of an explosion of a large number of supernovas.

Apart from changes in the chemical composition and the heating of the primordial gas, supernovas should also generate relic neutron stars [48]. These neutron stars suffer hardly any collisions. Consequently, their dynamics during the stage of formation of the galactic structure is identical with the dynamics of nondissipative matter. Specifically, neutron stars occupy a region of size R_g and their density ρ_N is distributed in accordance with the law

$$\rho_N = \rho_{N_0} \left(\frac{r}{R_g}\right)^{-\alpha}, \quad \rho_{N_0} = \frac{3-\alpha}{4\pi R_g^3} N , \tag{165}$$

where N is the total number of relic neutron stars.

The existence of such a halo is pointed out in Ref. [42] and it is suggested there that relic neutron stars are the sources of gamma-ray bursts. This makes it possible to explain the main statistical properties of these bursts, which are their highly spherical symmetry and a considerable concentration at the centre of our galaxy, which is known as the $\log N - \log S$ curve. The observational data of Ref. [49] are compared in Fig. 16a with the results of a calculation of the parameters of gamma-ray bursts based on the distribution given by expression (165). The agreement in the region of nearby bursts can be proved if account is taken of the creation and expansion of neutron stars after the explosions of supernovas in the galaxy at a later stage (dashed curve) [42].

The weak asymmetries in the distribution of gamma-ray bursts are very important. They are due to two factors. First, the Sun is not at the centre of our galaxy but is shifted relative to this centre in the galactic plane to a distance $r_\odot = 8$ kpc. Second, there should be an asymmetry due to the interaction between the giant halos of our galaxy and the M34 (Andromeda) galaxy located at a distance of 600 kpc and moving in the direction towards us at the velocity of 200 km s^{-1}. In view of the sizes of the giant halos of the two galaxies, their peripheral structure should be distorted because of the interaction.

An analysis of recent data obtained by the COMPTON Observatory [50] shows that the BATSE (Burst and Transient Source Experiment) sensitivity is sufficient to detect gamma-ray bursts at distances up to 150–200 kpc. The observational data reveal both dipole and weak local asymmetries, demonstrating clearly the interaction with the giant halo of the Andromeda galaxy, in satisfactory agreement with the giant halo model (Fig. 16b) [51].

It therefore follows that the model of a giant halo of relic neutron stars, which is based on the theory of a giant dark matter halo, can describe quite satisfactorily not only the spherical symmetry and the spatial distribution of the gamma-ray bursts, but also their weak asymmetry. The model is also in satisfactory agreement with recent observational data.

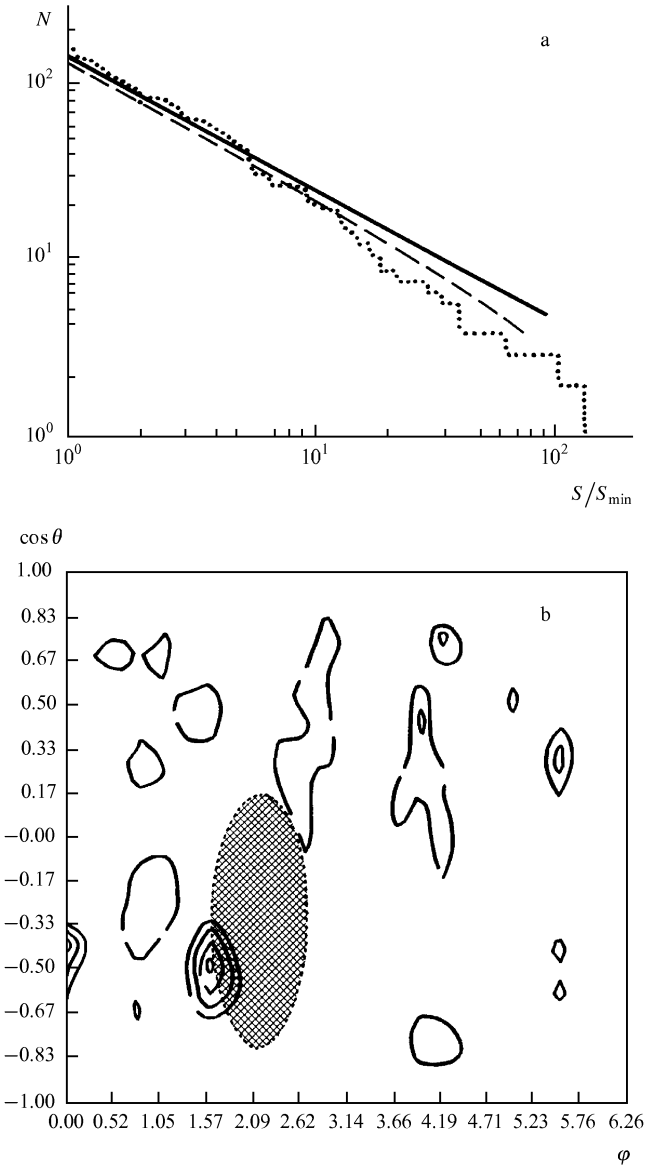

Figure 16. (a) Curve representing the $\log N - \log S$ dependence, plotted in accordance with Eqn (165) on the basis of the 'standard candle' assumption. The dots represent the observational data [49]. (b) Contours of fluctuations of the number of sources of gamma-ray bursts corresponding to 1σ, 2σ, and 3σ. The shaded area is the giant halo of the dark matter of Andromeda [50].

11.3 Pair correlation functions

In the strong correlation range ($\xi \gg 1$) the main contributions to the pair correlation function comes from binary systems (Section 10). Their contribution not only dominates the parameter β, but it rises most rapidly with reduction in the distance r.

It follows from expression (155) that the steady-state pair correlation functions of galaxies, clusters, superclusters, etc. should increase with reduction in the distance in accordance with $r^{-\alpha}$, where $\alpha \approx 1.8$ is the universal scaling parameter governed by the laws of three-dimensional nonlinear compression of self-trapped nondissipative dark matter. It is well known that this dependence is in full agreement with the observed pair correlation functions of galaxies and clusters [52]. The examples shown in Fig. 17 demonstrate a good agreement between the theory and observations.

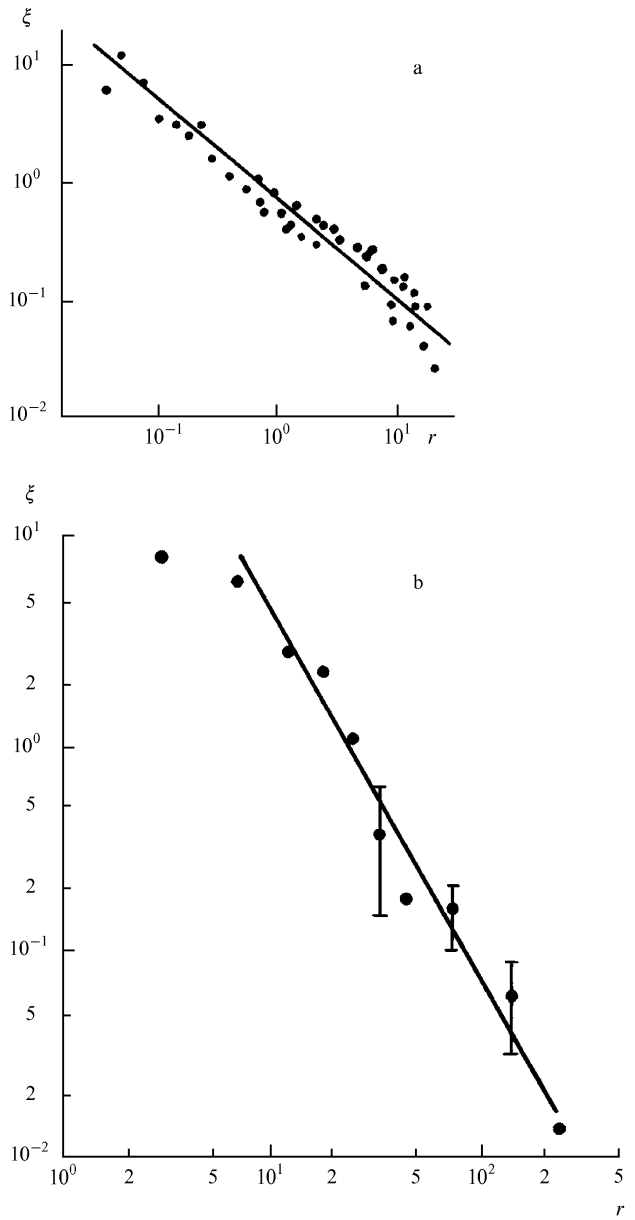

Figure 17. Correlation functions of galaxies (a) and galaxy clusters (b).

In recent years the same dependence of the pair correlation function has been reported also for other objects such as groups of galaxies, quasars, and various clusters. We can therefore conclude that the nonlinear compression theory [26–30] presented above is confirmed by the observational data. Moreover, since the law described by expres-sion (123) is derived relying strongly on the low thermal velocities of the dark-matter particles, we may conclude that the observations provide indirect support for the conclusion that the dark matter is cold.

In the range of high values of the correlation function, not only the scaling law of expression (155) is obeyed, but also the amplitude of the function depends on the average distance between the objects. Fig. 18 gives the results of observations taken from Ref. [53]. In the range of distances $20h^{-1} < d_i < 80h^{-1}$ Mpc the amplitude of the correlation function A_i increases in accordance with the law

$$A_i = \xi_i\,(1\,\mathrm{Mpc}) = (0.4d_i)^{1.8}\ . \qquad (166)$$

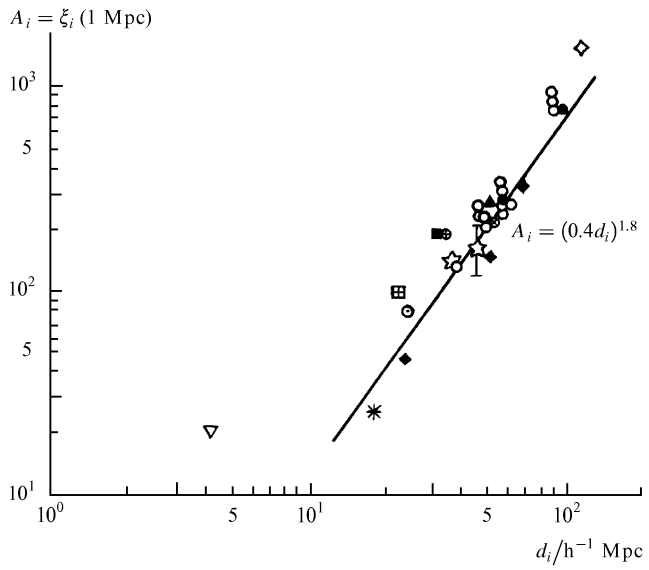

Figure 18. Dependence of the amplitude of the pair correlation function on the average distance between objects [53].

As pointed out above, objects with these dimensions have not yet completed (or have done so quite recently) the nonlinear stage of the evolution process, i.e. the red shift at the moments of their formation is small and lies somewhere in the interval $0 < z_i < 0.5$. In this case the scatter of z_i is unimportant and the dependence of A_i on d_i, which follows from theoretical formulas (155) and (156), is in agreement with the observed dependence (166).

It is evident from Fig. 18 that galaxies do not obey the law (166): the amplitude of their correlation function is considerably greater. The same effect, though to a lesser degree, is observed also for quasars. However, galaxies appear much earlier than the other objects for which data are plotted in Fig. 18. If in expression (156) we substitute the average red shift at the moment of formation of galaxies $z_{\mathrm{g}} \approx 5-8$ and the corresponding value for quasars $z_{\mathrm{q}} \approx 1.5-2$, we find that the results are in agreement with the observations. It should be noted however that the constant 0.4 in the experimental law (166) is approximately twice that obtained from expressions (155) and (156) on the assumption that $z_i < 0.5$.

We have discussed so far only the main contribution to the correlation function in the range of its high values ($\xi \gg 1$), where expression (155) obtained above is valid. This contribution is related solely to binary systems.

The next, in terms of the parameter β, contribution to ξ depends on the distance as $r^{3-2\alpha}$. The contribution is small if the correlation function is calculated by averaging over a sufficiently large region of space. However, it may become comparable with the main contribution if we consider, for example, only the correlation of objects of one type in close vicinity to their large cluster. In this case the dependence $r^{-\alpha}$ should change to $r^{3-2\alpha}$ as ξ is reduced.

It is interesting to note also that a similar change in the slope is indeed observed for the correlation functions of galaxies inside the clusters [54]. However, a reliable comparison requires both an increase in the amount of observational data and a further development of the theory proposed here. In particular, in the calculation of the correlation function in the range of values $\xi \sim 1$ we

need to take account accurately of the initial correlations and the distribution of objects in accordance with their formation times.

It is important to stress that the pair correlation function manifests strikingly the specific properties of NGSs not only because they represent the strongest singularities in respect of the density compression, but also because they are spherically symmetric. The contribution of other large-scale structural singularities, such as filaments or pancakes (Section 7), to the pair correlation function is much smaller because of the averaging over directions.

There is therefore a considerable interest in the study of higher correlations. For example, a triple correlation function depends on two vectors: $\boldsymbol{R}_1 = \boldsymbol{r}_1 - \boldsymbol{r}_2$ and $\boldsymbol{R}_2 = \boldsymbol{r}_2 - \boldsymbol{r}_3$. The presence of filament or pancake structures should manifest itself by a singularity of the triple correlation function $\xi^{(3)}$ when the dependence on the angle θ between the vectors $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ is considered: the function $\xi^{(3)}$ diverges in the limit $\theta \to 0$.

Similar singularities are naturally expected also for higher correlations. One would hope that the progress in the gathering of observational data and their analysis will make it possible to identify in future such singularities of the higher correlation functions of the observational data.

11.4 Centre of a nondissipative gravitational singularity

In addition to a giant halo, a fully formed NGS has also a sharply defined centre, which is a density singularity described by expression (123). The presence of this singularity is important in dark-matter diagnostics [55]. In fact, although the cross section of the interaction of the dark-matter particles with one another is extremely small, irrespective of the nature of dark matter, the radiation flux created by this interaction is

$$F \propto \int \sigma \rho^2 \, \mathrm{d}^3 r \,, \tag{167}$$

where σ is the particle interaction cross section. Substitution of expression (123) for the density into the above relationship shows that the above integral diverges at $r = 0$, i.e. at the centre of an NGS.

It follows that the main radiation flux, which can be used in dark-matter diagnostics, is associated with the centre of an NGS. It is at this centre that we can expect the greatest effect. The radiation flux is determined by truncation of the law governing the increase in the NGS density described by expression (123). The actual truncation of the divergence of the density described by this expression depends primarily on the stage in the hierarchy in which a given NGS has formed.

For example, if an NGS has grown from the maxima of scale $k = k_{\max}$, the truncation region is extremely small: it is determined by the decaying mode in the initial spectrum of the fluctuations (Appendix). However, if an NGS has not yet formed, the truncation radius of the singularity is determined by the current moment of time and decreases with time.

The assumption that an NGS in our galaxy has already developed is used in Ref. [54] to calculate the expected radiation flux from the centre of this singularity. It is assumed that dark matter is in the form of weakly interacting supersymmetric particles. It is shown that the strongest truncation of the NGS in our galaxy can come from a massive black hole. A comparison of the predicted radiation flux with observations sets limits on the mass of supersymmetric particles.

The existence of a singularity is important also in the dynamics of the protogalactic baryonic gas. As pointed out above, the baryonic matter density is only a few percent of the dark matter density. Therefore, during the initial stage, the baryonic gas moves in a given field of an NGS and drops to the bottom of a potential well.

The density of the baryonic gas near a singularity then increases strongly and the motion of this gas is always of potential nature. Consequently, a massive protostar forms near the centre. Nuclear processes in such objects are known to be very fast [56] and they lead to the formation of a giant black hole. An estimate of the mass of the resultant black hole can be found in Ref. [28].

One might mention here also other predictions which should be checked in detail by comparison with observational data. These predictions include, first, a giant halo of established clusters with the dark matter distribution in accordance with the law described by expression (123). The existence of such a halo should give rise to rotation curves of galaxies or other objects trapped in the gravitational field of the halo. The size of the halo is estimated to be 2–5 Mpc. The second prediction is the appearance of pancake, filament, and node structures, and of the structure of the biggest large-scale objects which have not reached the stage of the initial singularities (these are objects of the Great Attractor type [40]).

We should mention also the qualitative differences between the regions in space which are inside and outside the giant halos of galaxies. In view of the Hubble expansion the dark matter density and, consequently, the average density of the baryonic matter are very different in these regions. This difference is likely to affect the formation of condensed objects, for example, small galaxies. It can be investigated in particular by considering the example of the near cosmic space of our galaxy and its closest neighbours.† In principle, it should be possible also to observe directly the interaction of planets with dark matter [57].‡

It should be stressed that the analytic theory of the large-scale structure of matter presented in this review is as yet far from complete. Among the many problems that need to be discussed further, we would like to mention the following.

(1) It is necessary to generalise the solution of the dynamic problem of the formation of an NGS to the case of an arbitrary value of the parameter ε, to include in greater detail the influence of the anisotropy of the initial data, and also to discuss fully the process of three-dimensional mixing. A more effective use of numerical methods is needed to solve these problems.

(2) It is desirable to study the process of collisional relaxation. It should be stressed that this means collisions of galactic systems and primarily the collisions of giant dark-matter halos. The influence of the baryonic matter and, in particular, of the baryonic gas, can then be very significant: the baryonic matter determines dissipation and this must affect the overall dynamics of the process.

†This was pointed out by G V Chibisov.

‡A numerical error was made in estimating a constant in Ref. [57]. In spite of the underestimation of this constant, the effect in question should remain within the limits of observability.

(3) It is essential to extend the statistical theory in order to take account of both the initial correlations and the correlations of higher orders. Moreover, it is necessary to separate the singularities of the correlation functions of the third and fourth orders when these singularities correspond to dynamic filament or pancake structures (Section 9). It is obviously equally important to identify the same types of singularities in the higher correlation functions also in the observational data.

There are also other interesting problems. However, it is obviously most important to find a way for correct matching of this anaytical theory to the direct numerical simulation methods. This approach could make the greatest contribution to further developments of the theory.

Acknowledgements. We are grateful to V L Ginzburg for valuable discussions and his continuous interest, and to D A Kirzhnits, M I Zel'nikov, and Yu V Medvedev for helpful discussions, and to V A Sirota for the help in preparation of this paper for press.

The work was financed by Grant No. 93-02-17065 from the Russian Fund for Fundamental Research.

References

1. Zwicky F *Helv. Phys. Acta* **6** 110 (1993)
2. Walker T P, Steigman G, Schramm D N, Olive K A, Kang H S *Astrophys. J.* **76** 51 (1991)
3. Peebles P J E *Nature (London)* **321** 27 (1986)
4. White S D M, Davis M, Efstathiou G, Frank C S *Nature (London)* **330** 451 (1987)
5. Smoot G F, Bennett C L, Kogut A, et al. *Astrophys. J. Lett.* **396** L1 (1992)
6. Shandarin S F, Doroshkevich A G, Zel'dovich Ya B *Usp. Fiz. Nauk* **139** 83 (1983) [*Sov. Phys. Usp.* **26** 46 (1983)]
7. Peebles P J E *Astrophys. J. Lett.* **263** L1 (1982)
8. Primack J R, Blumental G R *Formation and Evolution of Galaxies and Large Structures in the universe* (Ed. J T V Tran) (Dordrecht, Netherlands: Reidel, 1983) p. 163
9. Bond J R, Efstathiou G *Astrophys. J. Lett. Ed.* **285** 45 (1984)
10. Zel'dovich Y B *Adv. Astron. Astrophys.* **3** 241 (1965)
11. Harrison E R *Rev. Mod. Phys.* **39** 862 (1967)
12. Linde A D *Fizika Elementarnykh Chastits i Inflyatsionnaya Kosmologiya* (Physics of Elementary Particles and Inflation Cosmology) (Moscow: Nauka, 1990)
13. Davis M, Efstathiou G, Frenk C S, White S D M *Astrophys. J.* **292** 371 (1985)
14. Bertschinger E, Gelb J M *Comput. Phys.* **5** 164 (1991)
15. Park C, Gott J R *Mon. Not. R. Astron. Soc.* **249** 288 (1991)
16. Cen R *Astrophys. J. Suppl.* **78** 341 (1992)
17. Ostriker J P, in *Annual Review of Astronomy and Astrophysics* (New York: Pergamon Press, 1992) p. 689
18. Myers S T, Readhead A C S, Lawrence C R *Astrophys. J.* **405** 8 (1993)
19. Primack J R, Holtzman J, Klypin A, Caldwell D O (submitted to *Phys. Rev. Lett.*)
20. Chandrasekhar S, Munch G *Astrophys. J.* **115** 103 (1952)
21. Peebles P J *The Large-Scale Structure of the universe* (Princeton, NJ: Princeton University Press, 1980)
22. Sirovich L, Smith L, Yakhot V *Phys. Rev. Lett.* **72** 344 (1994)
23. Zel'dovich Ya B *Astrofizika* **6** 319 (1970)
24. Arnol'd V I *Tr. Semin. im. I G Petrovskogo* **8** 21 (1982)
25. Arnold V I, Shandarin S F, Zeldovich Ya B *Geophys. Astrophys. Fluid Dyn.* **20** 111 (1982)
26. Gurevich A V, Zybin K P *Zh. Eksp. Teor. Fiz.* **94** (1) 3 (1988) [*Sov. Phys. JETP* **67** 1 (1988)]
27. Gurevich A V, Zybin K P *Zh. Eksp. Teor. Fiz.* **94** (10) 5 (1988) [*Sov. Phys. JETP* **67** 1957 (1988)]
28. Gurevich A V, Zybin K P *Zh. Eksp. Teor. Fiz.* **97** 20 (1990) [*Sov. Phys. JETP* **70** 10 (1990)]
29. Gurevich A V, Zybin K P, Medvedev Yu V *Zh. Eksp. Teor. Fiz.* **104** 3369 (1993) [*J. Exp. Theor. Phys.* **77** 593 (1993)]
30. Gurevich A V, Zel'nikov M I, Zybin K P *Zh. Eksp. Teor. Fiz.* (in press)
31. Durrer R *Astron. Astrophys.* **208** 1 (1989)
32. Landau L D, Lifshitz E M *Fluid Mechanics* 2nd edition (Oxford: Pergamon Press, 1987)
33. Hockney R W, Eastwood J W *Computer Simulation Using Particles* (New York: McGraw-Hill, 1981)
34. O'Neil T *Phys. Fluids* **8** 2285 (1965)
35. Lifshitz E M, Pitaevskii L P *Physical Kinetics* (Oxford: Pergamon Press, 1981)
36. Gurevich A V, Zybin K P, Medvedev Yu V, Preprint No. 35 (Moscow: Lebedev Physics Institute, Russian Academy of Sciences, 1994)
37. Bahcall N A, Cen R *Astrophys. J. Lett.* **407** L49 (1993)
38. Bardeen J M, Bond J R, Kaiser N, Szalay A S *Astrophys. J.* **304** 15 (1986)
39. Burstein D, Rubin V C *Astrophys. J.* **297** 423 (1985)
40. Dressler A *Sci. Am.* **257** (3) 38 (1987)
41. Gurevich A V, Zybin K P, Paper presented at the International Astronomical Union Symposium on the Interstellar Disk–Halo Connection in Galaxies, Leiden, Netherlands, 1990
42. Gurevich A V, Beskin V S, Zybin K P, Ptitsyn M O *Zh. Eksp. Teor. Fiz.* **103** 1873 (1993) [*J. Exp. Theor. Phys.* **76** 925 (1993)]
43. Flores R A *Phys. Lett. B* **215** 73 (1988)
44. Gurevich A V, Zybin K P, Sirota V A *Phys. Lett. A* (in press)
45. Flores R A, Primack J R *Astrophys. J. Lett.* **427** L1 (1994)
46. Zaritsky D, White S D M *Astrophys. J.* **435** 599 (1994)
47. Marochnik L S, Suchkov A A *Galaktika* (The galaxy) (Moscow: Nauka, 1984)
48. Imshennik V S, Nadezhin D I, in *Sovremennye Problemy Fiziki i Evolyutsii Zvezd* (Current Problems in Physics and Evolution of Stars) (Ed. A G Masevich) (Moscow: Nauka, 1989) p. 37
49. Meegan C A, Fishman G J, Wilson R B, et al. *Nature (London)* **355** 143 (1992)
50. Bhat P N, Fishman G J, Meegan C A, et al. *Nature (London)* **359** 217 (1992)
51. Gurevich A V, Zharkov G F, Zybin K P, Ptitsyn M O *Phys. Lett. A* **192** 305 (1994)
52. Bahcall N A, Soneira R M *Astrophys. J.* **270** 20 (1983)
53. Bahcall N A, West M J *Astrophys. J.* **392** 419 (1992)
54. Einasto M *Month. Not. R. Astron. Soc.* **258** 571 (1992)
55. Berezinsky V S, Gurevich A V, Zybin K P *Phys. Lett. B* **294** 221 (1992)
56. Bond J R, Arnett W D, Carr B J *Astrophys. J.* **280** 825 (1984)
57. Braginsky V B, Gurevich A V, Zybin K P *Phys. Lett. A* **171** 275 (1992)
58. Moutarde F, Alimi J M, Bouchet F R, Pellat R *Astrophys. J.* **441** 10 (1995)

Appendix: Influence of a decaying mode on the formation of a nondissipative gravitational singularity

In considering the nonlinear theory of a growing perturbation mode we have ignored completely a decaying mode and we have thus narrowed down the class of initial conditions. In fact, arbitrary initial conditions are set by four scalar functions $\delta_i(\boldsymbol{x})$ and $v_i(\boldsymbol{x})$, whereas in the case of a growing mode they are set by just one function. However, since separation into growing and decaying modes is possible only during the linear stage, it is necessary to consider the influence of all possible perturbations that we have ignored on the solution described by expression (123).

The velocity of particles in a decaying mode is

$$v = -\dot{D}_2(t)\,\frac{a(t)}{E}\left\{\dot{D}_1(i)\int \mathrm{d}^3x'\,\delta_i(\boldsymbol{x}')\,\frac{\boldsymbol{x}'-\boldsymbol{x}}{|\boldsymbol{x}'-\boldsymbol{x}|^3} - D_1(i)\frac{v_i^{\mathrm{d}}}{a(t)}\right\} + v_i^{\mathrm{r}}\,\frac{a_i}{a(t)}\,. \tag{A1}$$

Let us first consider the irrotational motion, i.e. let us assume that $v_i^{\mathrm{r}} = 0$. We shall be interested in the solution given by expression (A1) in a region close to the effective density maximum considered above. Then, in general, the velocity v can be expanded as a Taylor series. If only the nonlinear terms are retained, the result is

$$u_\alpha = U_{\alpha,\beta}\,x_\beta\,,$$

where the velocity u_α and the coordinate x_α are made dimensionless in accordance with formula (46).

The description of the velocity by expression (A1) is valid as long as $\delta < 1$. If $\delta \geqslant 1$, we have to consider a contracting solution, described by relationships (51). A characteristic tensor $U_{\alpha,\beta}$ at the moment when $\delta = 1$, will be denoted by ϵ. It depends on the rate of expansion of the Universe and on the fluctuation amplitude δ_i. If $\Omega = 1$, then

$$\epsilon \approx \delta_i^{3/2} \ll 1\,. \tag{A2}$$

An analysis of the solution described by expression (A1) demonstrates that, without limiting the generality of the discussion, we need to consider only those initial values of the field of velocities of the decaying mode which are characterised by $\operatorname{div} v = 0$. Then, on reduction of the matrix $U_{\alpha,\beta}$ to the diagonal form (bearing in mind that $\operatorname{curl} v = 0$), we obtain

$$u_x = \frac{\epsilon}{3}\,x\,, \quad u_y = -\frac{2}{3}\,\epsilon y\,, \quad u_z = \frac{\epsilon}{3}\,z\,. \tag{A3}$$

In writing down the relationships in expression (A3) we are assuming that the gradient of the field of velocities near the point $r = 0$ of the density maximum is directed along the y axis. We shall now consider how the selection at the moment $t = t^*$ $(\delta = 1)$ of the initial velocity given by expression (A3) affects the process of compression of a bunch when this process is due to the growing mode. Since we are interested in the vicinity of $r = 0$, at the compression stage we can expand the solution given by expression (51) near this specific point:

$$v = -\frac{2}{3}\,\frac{\boldsymbol{r}}{1-\tau}\,, \quad \rho = \frac{2}{3(1-\tau)^2}\,. \tag{A4}$$

Here, the quantity r is normalised to a and τ is normalised to t_{c}.

Subtraction of formulas in expression (A4) from the complete hydrodynamic system of equations (28) and designation of the differences by δv, $\delta\rho$, $\delta\psi$, yields

$$\begin{aligned} &\frac{\partial}{\partial\tau}\,\delta v - \frac{2}{3(1-\tau)}\,(\boldsymbol{r}\cdot\nabla)\,\delta v + \frac{2}{3(1-\tau)}\,\delta v + \nabla\delta\psi + (\delta v\cdot\nabla)\,\delta v = 0\,,\\ &\frac{\partial}{\partial\tau}\,\delta\rho + \frac{2}{3(1-\tau)^2}\,(\nabla\cdot\delta v) - \frac{2}{3(1-\tau)}\,\nabla\cdot(\boldsymbol{r}\delta\rho) + \nabla\cdot(\delta\rho\,\delta v) = 0\,. \end{aligned} \tag{A5}$$

The initial conditions for the system of equations (A5) are specified in expression (A3). We shall seek the solution of the system of equations (A5) in the form

$$\delta v_k = r_k h_k \eta^4\,, \quad \delta\rho = q(\eta)\,\eta^4\,, \quad \nabla\delta\psi = \frac{1}{3}\,\boldsymbol{r}\,\delta\rho\,, \tag{A6}$$

where $\eta = (1-\tau)^{-1/3}$, and there is no summation over k.

Substitution of the relationships in expression (A6) into the system of equations (A5) gives

$$\begin{aligned} &\frac{1}{3}\,\frac{\partial}{\partial\eta}\,h_k + h_k^2 + \frac{1}{3}\,q\eta^{-4} = 0\,,\\ &\frac{1}{3}\,\frac{\partial}{\partial\eta}\,q - \frac{2q}{3\eta} + q\sum_k h_k + \frac{2}{3}\,\eta^2\sum_k h_k = 0\,. \end{aligned} \tag{A7}$$

Since, according to expression (A2), we have $\epsilon \ll 1$, it is sufficient to consider just the linear solution of the system of equations (A7). Dropping the nonlinear terms from the system of equations (A7) and substituting the initial conditions given in expression (A3), we obtain

$$\begin{aligned} &\delta\rho = 0\,, \quad \delta v_k = g_k\,r_k(1-\tau)^{-4/3}\,,\\ &g_1 = \frac{\epsilon}{3}\,, \quad g_2 = -\frac{2}{3\epsilon}\,, \quad g_3 = \frac{\epsilon}{3}\,. \end{aligned} \tag{A8}$$

It therefore follows from expression (A8) that in the linear approximation the density does not rise, i.e. $q = 0$, but the velocity rises faster than the velocity of the main stream described by expression (A4) and near a singularity it is much higher than the latter.

We shall therefore consider the nonlinear solution of the system of equations (A5) on the assumption that $n \to \infty$. It follows asymptotically from the first equation of this system (A7) that

$$\frac{1}{3}\,\frac{\partial h_k}{\partial\eta} + h_k^2 = 0\,. \tag{A9}$$

The above equation described the kinematic motion of particles. Its solution is

$$h_k(\eta) = \frac{h_k(0)}{1 + 3h_k(0)(\eta - 1)}\,. \tag{A10}$$

It is important that along the coordinate y the perturbations grow faster than along the other two coordinates. If $\eta = 1 + \epsilon/2$, it follows from the initial conditions given by expression (A3) that a singularity appears in the solution given by expression (A10). This singularity [see Eqn (A9)] is the result of a nonlinear kinematic reversal. It is one of the possible Lagrangian singularities [25]. This means that the central singularity of expression (123) discussed by us spreads out over a certain small region around $r = 0$ because of kinematic mixing. Estimates of the energy show however that the characteristic size δr of the spreading region is extremely small:

$$\delta r \approx \epsilon^2 = \delta_i^3\,. \tag{A11}$$

Consequently, under real physical conditions this process is of little importance.

We shall conclude by noting that inclusion of low rotational (vortex) velocities v_i^{r} leads to a similar small region of smoothing out of the centre of an NGS.