Astron. Astrophys. 325, 1–8 (1997)

ASTRONOMY
AND
ASTROPHYSICS

A modification of the baryonic dark matter model

Attila Mészáros

Department of Astronomy, Charles University, 150 00 Prague 5, Švédská 8, Czech Republic

Received 7 August 1996 / Accepted 18 March 1997

Abstract. The conventional baryonic dark matter model of galaxy formation uses the concept of entropy fluctuations. The key idea of this picture is based on the assumption that already before the recombination on the scales characterized by masses not bigger than $\sim 10^9 M_\odot$ the entropic fluctuations were non-linear, and they were formed by non-gravitational processes. In this paper the basis of a modification of the baryonic dark matter picture is presented. This modification uses - instead of the entropic fluctuations - the concept of adiabatic perturbations existing only during and after the recombination. It reproduces the key idea of conventional picture: roughly for solar masses bold anisotropies should occur immediately during the recombination. Nevertheless, contrary to the conventinonal picture, this occurs without any non-gravitational phenomenons. All this follows from the so called "Bessel-Macdonald" instability, which is predicted to exist by a careful mathematical analysis of the key equation of adiabatic perturbations.

Key words: cosmic microwave background – dark matter

1. Introduction

The standard theory of baryonic dark matter (hereafter BDM) picture of the galaxy formation assumes that at the redshift z the order of the fluctuation of baryonic mass is roughly given by $\sim (1+z)^{-1}(10^{15} M_\odot/M)^{1/2}$, where $M_\odot$ is the mass of Sun, and M is the mass of fluctuation (Peebles 1993, p.662). This means that at $z \simeq 1000$ the fluctuations with masses not bigger than $\sim 10^9 M_\odot$ should be non-linear (i.e. they should be of order unity or bigger), and the first stars should be formed even during the recombination (Peebles 1993, p.663). There is no definite theory for the origin of such a behavior of baryonic fluctuations (Peebles 1993, p.663); only it is widely accepted that they formed before the recombination. This formation should happen due to some cosmic magnetic or other - more or less exotic non-gravitational - fields (see the whole Sect. 25 of Peebles' book (Peebles 1993) for details and further references; the behavior of such non-linear perturbations before and during the recombination is also discussed by Hogan (1993)). Nevertheless, it seems to be doubtless that before the recombination any adiabatic perturbations with masses smaller than $\sim 10^{13} M_\odot$ were destroyed due to the Silk damping (Börner 1993). In addition, the growth of any adiabatic perturbations in the pre-recombination era - "surviving the Silk-damping" - is also forbidden (Mészáros 1974). Hence, naturally, the BDM picture assumes that the perturbations before and during the recombination are dominated by entropy fluctuations.

The purpose of this paper is to modify the key ideas of BDM model. In this modified version, first, *no* non-linear perturbations will be assumed to occur *before* the recombination, second, the *adiabatic* perturbations will be used, and, third, these perturbations *will not* arise by exotic procedures. To do this a careful mathematical analysis of the basic equation of the linear adiabatic perturbations will be enough. Nevertheless, in this version the key idea of BDM picture - namely the assumption that at $z \sim 1000$ the perturbations with masses below $\sim 10^9 M_\odot$ are highly non-linear - will again be accepted. It will even be obtained as a natural consequence of mathematical results.

The paper is organized as follows. In Sect. 2. the key mathematical calculations of linear adiabatic perturbations are summarized. (This part of article does not contain any new result, and is presented for the purpose of further considerations.) Sect. 3. gives new mathematical solutions of the key equation of linear adiabatic perturbations, which are discussed both from the mathematical and from the cosmological point of view in Sect. 4. Finally, Sect. 5. summarizes the key ideas and results of paper.

2. The key equation of adiabatic perturbations

The standard theory of small (linear) adiabatic perturbations - during and after the recombination for the subhorizon scales - is based on the equation (cf. Peebles 1980, Eq. (10.2))

$$\ddot{\delta} + 2H\dot{\delta} - (b^2 a^{-2}\Delta + 4\pi G\rho)\delta = 0, \qquad (1)$$

where the infinitesimally small dimensionless δ is the density contrast ($|\delta| \ll 1$), $H > 0$ is the Hubble parameter, $b \geq 0$ is the sound velocity, $a > 0$ is the expansion function ($H = (\dot{a}/a)$), $\rho > 0$ is the density of unperturbed Friedmannian model, G is the gravitational constant, a dot denotes partial derivative with respect to time, and Δ is the Laplacian. δ is a function of time t and of dimensionless spatial comoving coordinates q_1, q_2, q_3; H, b, a, ρ are depending on time only. These functions are continuous, and any their derivatives used in this paper are also continuous.

From the mathematical point of view Eq. (1) is a standard second order linear hyperbolic partial differential equation, if $b > 0$; when $b = 0$, then Eq. (1) is a simple ordinary linear second order differential equation with respect to independent variable t.

Eq. (1) for $b = 0$ and $6\pi G\rho t^2 = 1$ (spatially flat Friedmannian model) gives the solution $\delta \sim t^{2/3} \sim (1 + z)^{-1}$. In the hyperbolic Friedmannian model the growth is even less rapid (Weinberg 1972, Sect. 15.9). In this paper this solutions these solutions will not be discussed, and therefore in what follows we will always suppose that $b > 0$.

For $b \neq 0$ only few solutions are known, which are based on the Fourier decomposition of δ, and which also give less rapid than $\delta \sim (1 + z)^{-1}$ growths (Weinberg 1972, Sect. 15.9). The use of modern gauge invariant methods (e.g. Bardeen 1980; Ellis & Bruni 1989) do not give any new results here, because this procedure is identical to the standard one for the scales smaller than the horizon at the matter dominated era.

In essence, in this paper it will be shown that Eq. (1) has also further - as far as it is known - fully new solutions leading to extremely fast growths on small scales. In order to do this a short recapitulation of the only solutions of Eq. (1) with $b > 0$ will be given in this section.

It is well-known that the necessary condition to have an unambiguous solution of Eq. (1) is to have be defined the initial values on a Cauchy surface (Morse & Feshbach 1953; Sect. 6.2). In what follows we will choose this Cauchy surface as the three-dimensional space at time $t = t_o > 0$. This means that the initial values are $\delta(t_o, q_1, q_2, q_3)$ and $(\partial\delta(t, q_1, q_2, q_3)/\partial t)|_{t=t_o}$. In other words, necessary condition to have an unambiguous δ for any $t \geq t_o$ is that the value of δ and its first time derivative be defined for a given time instant t_o.

Concerning the three spatial coordinates there are in essence two cases.

At the first case one has $-\infty < q_1 < \infty, -\infty < q_2 < \infty, -\infty < q_3 < \infty$. In that case we search for δ for $t \geq t_o$ for the whole three-dimensional space. In this case, of course, the initial data $\delta(t_o, q_1, q_2, q_3)$ and $(\partial\delta(t, q_1, q_2, q_3)/\partial t)|_{t=t_o}$ are also defined for any real q_1, q_2, q_3. If this is the case, then there is no complications concerning the existence of solution: There exists an unambiguous δ for $t \geq t_o$ and for any real q_1, q_2, q_3 (see again Morse & Feshbach 1953; Sect. 6.2.); the definition of initial values is also the sufficient condition to have unambiguous solution. It is essential to note that no further conditions for δ are required here for the existence of an unambiguous solution; of course, to find this solution technically may be highly complicated. (Trivially, here - from the physical point - one must still require $|\delta| \ll 1$.) Any further conditions on δ are innecessary for the existence of the unambiguous solution; these additional conditions may be given, but at that case Eq. (1) is overestimated. In other words, once there are further conditions on δ, then there can also be an unambiguous solution (technically these conditions may even help to find such solution), but then there is a loss of generality. For different additional conditions one may obtain fully different solutions even at the case, when the initial values for $t = t_o$ are the same. For example, be additionally assumed that $\int_{-\infty}^{\infty} dq_1 \int_{-\infty}^{\infty} dq_2 \int_{-\infty}^{\infty} dq_3 |\delta|$ is finite for any $t \geq t_o$. (Trivially, the necessary but not sufficient condition to have this finite integral is $\lim_{q_1\to\pm\infty} \delta = \lim_{q_2\to\pm\infty} \delta = \lim_{q_3\to\pm\infty} \delta = 0$; roughly speaking "$\delta$ must be zero at spatial infinity".) At this special case one may decompose δ into the three-dimensional Fourier-integral (cf. Bracewell 1978; p.9). In other words, trivially, to solve Eq. (1) for the whole three-dimensional space via the decomposition into Fourier-integral cannot be done at any case, even when the initial values are well defined and hence the existence of unambiguous solution is also ensured.

Even more complicated is the situation, when the spatial coordinates are restricted. For example, let the initial values $\delta(t_o, q_1, q_2, q_3)$ and $(\partial\delta(t, q_1, q_2, q_3)/\partial t)|_{t=t_o}$ be defined for a finite comoving cube defined by $-L \leq q_1 < L, -L \leq q_2 < L, -L \leq q_3 < L$ $(0 < L)$. These initial values are not enough to have unambiguous solution in the whole interior of this cube for $t > t_o$; this unambiguous solution may exist only at a limited part of cube, which part should be calculated via the characteristics (Courant & Hilbert 1968). To have unambiguous solution at the whole cube for any $t > t_o$ one has in essence two choices: Either to define the initial values not only in the cube alone but for any real q_1, q_2, q_3 and thus to have again the case discussed above, or to define some additional conditions on δ for $q_1 = \pm L, q_2 = \pm L, q_2 = \pm L$ ("to fix the boundary conditions"). In Sect. 6.3. of Morse & Feshbach (1953) it is explained that the unambiguous solution exists at the case, when on the boundary there is a linear relation between δ and its first time derivative for any time $t \geq t_o$. Of course, for different boundary conditions one may obtain quite different solutions even at the case, when the initial values are the same. For example, assume that $\delta(t, \pm L, q_2, q_3) = \delta(t, q_1, \pm L, q_3) = \delta(t, q_1, q_2, \pm L) = 0$ for any $t \geq t_o$. (There are "zero boundary conditions" on the six sides of comoving cube.) At that case, one may successfully use the three-dimensional Fourier-series to find the existing unambiguous solution. Nevertheless, these "zero boundary conditions" are not necessary from the mathematical point of view; there can be quite different ones, too. (Different types of boundary conditions are discussed in detail in Morse & Feshbach (1953); Sect. 6.3.) In addition, of course, the form of boundary may also be quite different (cube, cylinder, sphere, etc.).

The only known and in cosmology used solution of Eq. (1) for $b \neq 0$ is the solution, when the solution is given for a finite comoving cube with zero boundary conditions (Weinberg 1972; Sect. 15.9; Peebles 1980; Sect. 26; further discussions are also in several parts of monographs Peebles 1993; Börner 1993). From the previous considerations of this section it follows that this solution is a highly special solution, which surely does not represent the all possible solutions of Eq. (1). This is from the mathematical point of view a triviality. A restriction to the solutions given by Fourier-series is given in fact by heuristic physical arguments only. For example, it is generally believed that this restriction is from the physical point of view allowed (see pages 8-11 of Bracewell (1978) for illustration). Simply it is ad hoc widely assumed there is *no loss of generality*, once one restricts oneself to the solution of Eq. (1) in a comoving cube with zero boundary conditions.

It is doubtless, that one must take such conclusion carefully. Both the choice of cube as the spatial boundary and the choice of zero boundary conditions are highly artificial, and it is a triviality that to search for other physically reasonable solutions of Eq. (1) is highly required. This is the aim for further parts of this paper.

3. New solutions of Eq. (1)

Obviously, searching for further solutions of Eq. (1), which can have physical meaning, one must restrict oneself to the case when only a finite part of the three-dimensional space is considered. One cannot take $-\infty < q_1 < \infty, -\infty < q_2 < \infty, -\infty < q_3 < \infty$, because Eq. (1) - as the key equation of linear adiabatic perturbations - may be valid only for the subhorizon scales; i.e. for a limited part of three-dimensional space. Hence, some boundary and the some conditions on it must be defined in order to have an unambiguous solution. In what follows, this boundary will be a surface of a comoving sphere. The boundary conditions are kept still to be free, and will be defined later in order to keep the generality as far as possible.

3.1. Decomposition of $\delta(t, q, \vartheta, \varphi)$ into spherical harmonics

Let the function $\delta(t, q, \vartheta, \varphi)$ be decomposed as follows (q, ϑ, φ are the usual spherical coordinates; $q_1 = q \sin\vartheta \cos\varphi,\ q_2 = q \sin\vartheta \sin\varphi,\ q_3 = q\cos\vartheta;\ 0 \le \vartheta \le \pi;\ 0 \le \varphi \le 2\pi;\ q \ge 0$)

$$\delta = \sum_{n=0}^{\infty}\sum_{m=0}^{n} A_{nm}(q,t) P_n^m(\cos\vartheta)\cos m\varphi + \sum_{n=1}^{\infty}\sum_{m=1}^{n} B_{nm}(q,t) P_n^m(\cos\vartheta)\sin m\varphi, \tag{2}$$

where n and m are integer non-negative numbers, $A_{nm}(q,t)$ and $B_{nm}(q,t)$ are the unknown functions, and the Legendre polynomials fulfil the equations

$$(\sin\vartheta)^{-1}\cos\vartheta(\partial P_n^m/\partial\vartheta) + (\partial^2 P_n^m/\partial\vartheta^2) + n(n+1)P_n^m - (\sin\vartheta)^{-2} m^2 P_n^m = 0. \tag{3}$$

Eq. (2) is obviously a standard decomposition of $\delta(t, q, \vartheta, \varphi)$ into spherical harmonics (modes). Substituting (2) into Eq. (1) one obtains for any n and m

$$\ddot{A}_{nm} + 2H\dot{A}_{nm} - b^2 a^{-2}(A_{nm}'' + 2q^{-1}A_{nm}' - q^{-2}n(n+1)A_{nm}) - 4\pi G\rho A_{nm} = 0, \tag{4}$$

where a comma denotes derivative with respect to q. The identical equation for B_{nm} need not be written down. (In the whole Sect. 3 everything what is said about A_{nm} holds also for B_{nm}; of course, except for the fact that B_{nm} are not defined for $n = 0$ and $m = 0$. This triviality will not be raised later.)

Using the decomposition (2) we in essence simplify Eq. (1). Eq. (1) is a hyperbolic partial differential equation with three spatial coordinates. This equation is in fact substituted with a set of partial differential equations of hyperbolic type (Eq. (4)) with one spatial coordinate.

The solution of Eq. (4) is searched for $0 \le q \le q_o$, where $0 < q_o < \infty$, and for $t \ge t_o > 0$. The finite q_o defines the boundary of space, inside of which the solutions of Eqs. (4) are searched. To have unambiguous $A_{nm}(t,q)$ one must define some boundary conditions on $q = q_o$ together with the initial values $A_{nm}(t_o, q)$ and $(\partial A_{nm}(t,q)/\partial t)|_{t=t_o}$. Note still that, trivially, because q is not defined for negative values, one has to fix some conditions also at $q = 0$; all this follows from the use of spherical coordinates.

3.2. Sturm-Liouville type decomposition of $A_{nm}(q,t)$

Be given the following decomposition

$$A_{nm}(q,t) = \sum_{s=1}^{\infty} R_{nm\lambda_s}(t) Q_{nm\lambda_s}(q), \tag{5}$$

where λ_s for any $s = 1, 2, 3, ...$ are real constants (no summation over n, m). For a given s we have an unambiguously defined λ_s; for different s we obtain different λ_s. Therefore - instead of the notation $R_{nm1}, R_{nm2}, R_{nm3}, ...$ - the notation $R_{nm\lambda_1}, R_{nm\lambda_2}, R_{nm\lambda_3}, ...$ is used; for functions $Q_{nm\lambda_s}$ this is also done. The functions $R_{nm\lambda_s}(t)$ are unknown ones, but the constants λ_s and the functions $Q_{nm\lambda_s}$ are assumed to be known.

The function $Q_{nm\lambda_s}(q)$ is defined as the solution of the following ordinary differential equation

$$q^2 Q_{nm\lambda_s}'' + 2qQ_{nm\lambda_s}' - n(n+1)Q_{nm\lambda_s} + q^2\lambda_s Q_{nm\lambda_s} = 0. \tag{6}$$

This equation is a standard Sturm-Liouville eigenvalue problem (cf. Morse & Feshbach 1953; Sect. 6.3.), and it has non-zero $Q_{nm\lambda_s}$ solution only for some λ_s. One may write

$$Q_{nm\lambda}(q) = q^n \sum_{k=0}^{\infty} f_k q^{2k}, \tag{7}$$

where the dimensionless real constants f_k are defined by the following recurrence formulae

$$f_0 = 1,\ f_k = -\lambda_s f_{(k-1)}(2k)^{-1}(2k + 2n + 1)^{-1},\ k \ge 1. \tag{8}$$

(Clearly, without loss of generality, one may take $f_0 = 1$. The solution with $f_0 = 0$ is identically vanishing, which need not be considered here; for $f_0 \neq 0$ the choice $f_0 = 1$ may be done, because the function $Q_{nm\lambda_s}$ is defined up to a non-zero constant; see Eq. (5).)

The function $Q_{nm\lambda_s}(q)$ is known for any n, m, λ_s and for any $0 \le q \le q_o$. Then the decomposition (5) is always possible and is unambiguous (Morse & Feshbach 1953, Sect. 6.3). Note that the function $Q_{nm\lambda_s}(q)$ is defined by Eqs. (7-8) for $q \ge 0$, but Eqs. (4) and (6) only for $q > 0$. This is not a problem, because $\lim_{q\to 0}(2q^{-1}Q_{nm\lambda_s}' - q^{-2}n(n+1)Q_{nm\lambda_s})$ is finite.

Remark that for the special case, when $n = 0$, the solution of Eq. (6) need not be written down as a series, because from Eq. (6) one immediately obtains

$$Q_{00\lambda_s} = q^{-1}(-\lambda_s)^{-1/2}\sinh((-\lambda_s)^{1/2}q). \tag{9}$$

Using (8) it is easy to show the identity of solutions (9) and (7-8).

The concrete values of λ_s depends on the boundary conditions at $q = 0$ and at $q = q_o$. The boundary conditions at $q = 0$ are already given by Eqs. (7-8) as $\lim_{q\to 0}(q^{-n}Q_{nm\lambda_s}(q)) = 1$ and $\lim_{q\to 0}(q^{-n}Q_{nm\lambda_s}(q))' = 0$. On the other hand, the boundary conditions at $q = q_o$ are fully arbitrary yet from the mathematical point of view. In any case these must be linear relations of form (Morse & Feshbach 1953, Sect. 6.3.)

$$\lim_{q\to q_o} Q'_{nm\lambda_s}(q) = Y_{nm}Q_{nm\lambda_s}(q_o), \tag{10}$$

where Y_{nm} is a real constant (no summation over n and m). If $Y_{nm} = 0$, then there is no condition on $Q_{nm\lambda_s}(q_o)$, and we have the boundary condition $\lim_{q\to q_o} Q'_{nm\lambda_s}(q) = 0$. It may also be $Y_{nm} = \infty$; at that case the boundary condition is $Q_{nm\lambda_s}(q_o) = 0$, but $\lim_{q\to q_o} Q'_{nm\lambda_s}(q)$ is arbitrary. In fact, this is the case, when $A_{nm}(q_o, t) = 0$ is the boundary condition ("zero boundary condition").

Eq. (10) will define the values of λ_s; $s = 1, 2, 3, ...$ For the purpose of this paper it has a cardinal importance to clarify, when negative values of λ_s will occur. Fortunately, no detailed discussion of this mathematical problem is needed, because this problem is already well solved. The key results of the standard Sturm-Liouville problem may be summarized as follows (Morse & Feshbach, Sect. 6.3.). One has infinite number of eigenvalues, i.e. $s \to \infty$, and for $s \geq 1$ one has $\lambda_{(s+1)} > \lambda_s$; $\lim_{s\to\infty}\lambda_s = \infty$. In addition, the smallest value λ_1 is negative, if in Eq. (10) $\infty > Y_{nm} > 0$. This means that there will occur one and only one negative λ_1; in addition, not at any case, but only for $\infty > Y_{nm} > 0$.

Note that the case, when the "zero boundary condition" is required (i.e. when one has formally $Y_{nm} = \infty$ and the boundary condition is $Q_{nm\lambda_s}(q_o) = 0$), does not need any detailed discussion. From Eq. (8) it is obvious that for $\lambda_s < 0$ all constants f_k are positive; from the recurrence formula it follows that there is no alternation of sign there. Hence, for $q_o > 0$ the relation $Q_{nm\lambda_s}(q_o) = 0$ cannot occur for negative λ_1; for $0 < \lambda_1$ it must be $Q_{nm\lambda_s}(q_o) > 0$. There is no negative λ_1 for the case of "zero boundary condition". Clearly, for $\lambda_1 < 0$ both Q_{nms} and its derivative must be positive; neither of them can be zero or negative. Hence, trivially, the necessary condition to have negative λ_1 is in Eq. (10) $\infty > Y_{nm} > 0$. This condition is also sufficient, as the detailed analysis of Sect. 6.3. of Morse & Feshbach (1953) shows.

3.3. Solutions for $R_{nm\lambda_s}$

Substituting (5) into Eq. (4) the following ordinary differential equations are arising for any λ_s

$$\ddot{R}_{nm\lambda_s} + 2H\dot{R}_{nm\lambda_s} + (b^2a^{-2}\lambda_s - 4\pi G\rho)R_{nm\lambda_s} = 0. \tag{11}$$

Note that the same equation is already known and solved for $a \sim t^{2/3}$ and for $\lambda_s > 0$ (Weinberg 1972; Sect. 15.9.). In what follows this Weinberg's solution is in essence generalized and also the case with negative λ_s will be studied.

First, Eq. (11) is solved for the spatially flat case. Here one has $a(t) = a_0(t/t_0)^{2/3}$, where a_0 is the value of expansion function at time $t_0 > 0$. Let $T = (t/t_0)$ be the "dimensionless" time. (This trick is usual also for the computer simulations; Börner 1993, Sect. 12.) The choice of time instant t_0 is free; usually it defines the time of recombination, and this is done also here. For the post-recombination era the sound velocity is expected to be given by $b(T) = b_0(a_0/a) = b_0T^{-2/3}$, where b_0 is the sound velocity at time $T = 1$ (Weinberg 1972, Sect. 15.8). Therefore, here Eq. (11) takes the form

$$(d^2R_{nm\lambda_s}/dT^2) + 4(3T)^{-1}(dR_{nm\lambda_s}/dT) + (W\lambda_sT^{-8/3} - 2/(3T^2)^{-1})R_{nm\lambda_s} = 0, \tag{12}$$

where $W = b_0^2t_0^2a_0^{-2}$ is a dimensionless positive constant. For $\lambda_s > 0$ the solution of Eq. (12) is presented by Weinberg (1972; Sect. 15.9); and takes the form

$$R_{nm\lambda_s}(T) = T^{-1/6}R^{(1)}_{0nm\lambda_s}J_{(5/2)}(3(W\lambda_s)^{1/2}T^{-1/3}) + T^{-1/6}R^{(2)}_{0nm\lambda_s}J_{(-5/2)}(3(W\lambda_s)^{1/2}T^{-1/3}), \tag{13}$$

where $R^{(1)}_{0nm\lambda_s}$ and $R^{(2)}_{0nm\lambda_s}$ are dimensionless constants defined by initial data; $J_{(\pm 5/2)}(3(W\lambda_s)^{1/2}T^{-1/3})$ are the Bessel functions of first kind. On the other hand, for $\lambda_s < 0$ the solution of Eq. (12) is not written down yet. This case is also solvable, and one has

$$R_{nm\lambda_s}(T) = T^{-1/6}R^{(1)}_{0nm\lambda_s}I_{(5/2)}(3(-W\lambda_s)^{1/2}T^{-1/3}) + T^{-1/6}R^{(2)}_{0nm\lambda_s}K_{(5/2)}(3(-W\lambda_s)^{1/2}T^{-1/3}), \tag{14}$$

where $I_{(5/2)}(3(-W\lambda_s)^{1/2}T^{-1/3})$ and $K_{(5/2)}(3(-W\lambda_s)^{1/2}T^{-1/3})$ are the modified Bessel functions; the second one is also called as Macdonald function (Watson 1922).

Second, Eq. (11) will be solved in the case of open Friedmannian model, when the expansion is in its asymptotic phase (see, e.g., Weinberg 1972). Here one has $a(t) = ct$, $H = t^{-1}$, $b = b_0(a_0/a) = b_0T^{-1}$, $\rho = 0$, where c is the velocity of light. Then Eq. (11) takes the form

$$(d^2R_{nm\lambda_s}/dT^2) + 2T^{-1}(dR_{nm\lambda_s}/dT) + W\lambda_sT^{-3}R_{nm\lambda_s} = 0. \tag{15}$$

The solution is given either by (for $\lambda_s > 0$)

$$R_{nm\lambda_s}(T) = T^{-1/2}R^{(1)}_{0nm\lambda_s}J_1(2(W\lambda_s)^{1/2}T^{-1/2}) + T^{-1/2}R^{(2)}_{0nm\lambda_s}J_{(-1)}(2(W\lambda_s)^{1/2}T^{-1/2}), \tag{16}$$

or by (for $\lambda_s < 0$)

$$R_{nm\lambda_s}(T) = T^{-1/2}R^{(1)}_{0nm\lambda_s}I_1(2(-W\lambda_s)^{1/2}T^{-1/2}) + T^{-1/2}R^{(2)}_{0nm\lambda_s}K_1(2(-W\lambda_s)^{1/2}T^{-1/2}). \tag{17}$$

Weinberg (1972; Sect. 15.9) also notes that an other time dependence of the sound velocity is also possible, because it is not excluded that $b = b_0(a_0/a)^{1/2}$ occurs at least during and shortly after the recombination. Therefore, as a third case, the spatially flat Friedmannian model with this "alternative" sound velocity is considered. Here, instead of Eq. (12), one obtains ($b = b_0 T^{-1/3}$)

$$(d^2 R_{nm\lambda_s}/dT^2) + 4(3T)^{-1}(dR_{nm\lambda_s}/dT) + \Big(W\lambda_s - \frac{2}{3}\Big)T^{-2}R_{nm\lambda_s} = 0. \tag{18}$$

The solution of this equation is clearly given by

$$R_{nm\lambda_s} = R^{(1)}_{0nm\lambda_s}T^{\alpha_1} + R^{(2)}_{0nm\lambda_s}T^{\alpha_2},$$
$$\alpha_{1,2} = -\frac{1}{6} \pm \frac{1}{6}(25 - 36W\lambda_s)^{1/2}. \tag{19}$$

As the fourth case, the open Friedmannian model in its asymptotic case with the "alternative" sound velocity is considered. Here $H = t^{-1}, \rho = 0, b = b_0T^{-1/2}, a = ct$ holds, and Eq. (11) takes the form exactly identical to Eq. (15).

Note that the constants $R^{(1)}_{0nm\lambda_s}$, $R^{(2)}_{0nm\lambda_s}$ in Eqs. (14), (16), (17) and (19) should again be determined by initial data similarly to the case (13).

4. Remarks

A. Consider, first, the solution of Eq. (11) given by Eq. (19). For $W\lambda_s > (25/36)$ there is no growth of the given mode; there exist only sound waves, which is not new (Weinberg 1972; Sect. 15.9.). For $\lambda_s < 0$ the mode is growing faster than $\sim T^{2/3}$. Because λ_s may be arbitrary (see Eq. (10)), this growth may also be arbitrarily fast.

B. Discuss, second, the solution (14) of Eq. (12). (The solution (13) for $\lambda_s > 0$ is discussed by Weinberg (1972, Sect. 15.9).) Let the value of T be expected to run from $T = 1$ to $T = 10^3$ (i.e. from $z = 1000$ to $z = 10$). From the theory of Bessel functions it is known that the Macdonald function $K_\nu(x)$, where ν is arbitrary real and x is dimensionless, is for $x \gg 1$ nearly identical to $\exp(-x)$ (Watson 1922). Consider, first, the case when $W\lambda_s = -1$. Then $x = 3(-W\lambda_s)^{1/2}T^{-1/3}$ runs from 3 to 0.3. Thus here the identity

$$K_{(5/2)}(3(-W\lambda_s)^{1/2}T^{-1/3}) \approx \exp(-3(-W\lambda_s)^{1/2}T^{-1/3}) \tag{20}$$

is practically not fulfilled. (The precise values of some Macdonald functions are tabulated by Watson (1922).) Consider, second, the case when $W\lambda_s = -100$. Then x is decreasing from 30 to 3, and thus between $T = 1$ and $T = 1000$ Eq. (20) is in essence already fulfilled. For increasing T the function $\exp(-const.T^{-1/3})$ is increasing (if $const. > 0$, which is the case here). To illustrate the magnification of mode consider two time instants T_2 and T_1, where $T_1 \geq 1,\ T_2 = \omega \times T_1;\ \omega > 1$. Then the ratio

$$y(\omega, T_1, (-W\lambda_s)) = \frac{R_{nm\lambda_s}(T_2)}{R_{nm\lambda_s}(T_1)} = \omega^{-1/6}\exp(3(-W\lambda_s)^{1/2}T_1^{-1/3}(1 - \omega^{-1/3})) \tag{21}$$

defines the relative magnification between T_1 and T_2. This ratio may be a huge number for the large $-\lambda_s$. For example, consider the values $T_1 = 1$, $\omega = 10^3$, $-W\lambda_s = 100$. Then Eq. (21) gives $y(10^3, 1, 10^2) = 10^{-1/2}\exp(30 \times 0.9) \simeq 10^{11}$. Roughly speaking, the Macdonald function in Eq. (14) defines an extremely fast growth for $-\lambda_s \gg 1$. This effect may be called as "Bessel-Macdonald instability" after the key function. Note still that the second mode in Eq. (14) defined by function $I_{5/2}(x)$ may be discussed similarly. But here the function $I_\nu(x)$ may be approximated for $x \gg 1$ by $\exp x$; hence, quite similarly, one obtains that for large $-\lambda_s$ there is an extremely fast reduction instead of magnification.

C. Solution (17) shows that the Bessel-Macdonald instability occurs also in the case of open Friedmannian model, because there is no essential difference between Eqs. (14) and (17). The small substitutions (instead of (-1/6) one has to use (-1/2), instead of $K_{(5/2)}$ K_1, etc...) do not change the general behavior of instability. In addition, the Bessel-Macdonald instability occurs also for the alternative sound velocity, if the universe is open (see the end of previous section).

D. It is obvious that the necessary condition for the Bessel-Macdonald instability to occur is the occurence of large negative λ_1. This may occur, indeed. The necessary and sufficient condidition to have negative λ_s is $\infty > Y_{nm} > 0$ in Eq. (10), as it was said at the end of subsection 2.2. For $\lambda_1 < 0$ the concrete value $-\lambda_1$ may be an arbitrarily large number. To illustrate this consider the function $A_{00}(t, q)$ and hence the function $Q_{00\lambda_s}$ (see Eqs. (5) and (9)). Assume that in the boundary condition (10) one has $\infty > Y_{00} > 0$. Hence from Eqs. (9-10) it follows

$$1 + Y_{00}q_o = (-\lambda_s)^{1/2}q_o\coth((-\lambda_s)^{1/2}q_o), \quad q_o > 0. \tag{22}$$

This algebraic trancendent equation has one negative solution (λ_1), and infinite positive solutions (λ_2, λ_3,...), which should be solved numerically. Nevertheless, even without a long discussion it is obvious that for larger and larger positive Y_{00} one clearly obtains larger and larger $-\lambda_1$. For $Y_{00}q_o \gg 1$ one immediately has $Y_{00} \simeq (-\lambda_1)^{1/2}$. For the remaining values one clearly obtains the equation ($s \geq 2$)

$$1 + Y_{00}q_o = \lambda_s^{1/2}q_o\cot(\lambda_s^{1/2}q_o), \tag{23}$$

which is solvable numerically, and one has

$$(s-2)\pi < \lambda_s^{1/2} < (s-2)\pi + \frac{\pi}{2}. \tag{24}$$

All this is a standard matter in the theory of Sturm-Liouville problems (Morse & Feshbach 1953, Sect. 6.3), and no further discussion is necessary here. For our purpose it is cardinal that $-\lambda_1$ may be *arbitrary large* due to suitable boundary condition (10).

E. One may think that such a choice of initial condition is unphysical, and cannot occur. It can be shown that this is *not* the case. Consider again the special case when the perturbation is spherically symmetric; i.e. $\delta = \delta(q, t) = A_{00}(t, q)$. Consider this function for $0 \leq q \leq q_o < \infty$, and there is some boundary condition (10) for $q = q_o$. This does not mean that there is no further perturbation for $q > q_o$. But, because of the spherical symmetry, any existence of such perturbations at $q > q_o$ is *unimportant* for the region $q \leq q_o$; as it is well-known, one may simply *neglect* the existence of any matter at $q > q_o$, once there is a spherical symmetry (Weinberg 1972; Sect. 11.7.). Then the parameter Y_{00} in Eq. (10) (see also Eq. (22)) may be arbitrarily large at $q = q_0$; simultaneously, $A_{00}(t_0, q_0)$ may be arbitrarily small. In addition, there is also a freedom in the choice of q_o itself. This means that once we have a spherical symmetry, and for an arbitrarily small thin of spherical shell at the comoving radial distance $q = q_o$ we have $[(\partial\delta(q,t)/\partial q)|_{q=q_o}]/\delta(q_o,t) = Y_{00} \gg 1$, then the Bessel-Macdonald instability will occur. Around $q \simeq q_o$ bold non-linearities will arise. Note that similar conclusion - from ad hoc initial conditions - was already obtained via heuristic considerations (Mészáros 1991, Mészáros 1995). Note here that in essence this Remark E. is the key part of this article; it is shown here that the extremely fast growth of Bessel-Macdonald instability is not only mathematically correct, but it is also physically well *reasonable*.

F. In the general case the significance of parameters λ_s is from Eqs. (5-6) obvious; $\mid \lambda_s \mid^{-1}$ characterizes the comoving size of the first and second order derivatives of A_{nm} (and, hence, also of δ) with respect to the spatial coordinate q. This is especially clear for the special case of a spherical symmetry, i.e. when only the term A_{00} is non-zero in Eq. (2). Then from Eq. (9) and from Remarks D and E it is obvious that for $\lambda_1 < 0$ the value $(-\lambda_1)^{-1/2}$ defines the characteristic comoving length of first order spatial derivative. Roughly speaking, for increasing $\mid \lambda_s \mid$ the spatial derivatives of δ are also increasing; one may obtain huge δ also from very small initial data, if the spatial gradients alone are large enough.

G. The concrete value of the dimensionless constant $W = b_0^2 t_0^2 a_0^{-2}$ may straightforwardly be calculated. For example, consider the spatially flat case. Here one has $a(t) = U\eta^2,\ ct = (U/3)\eta^3$, where U is a constant having the dimension of length, and η is the conformal time. There is a freedom in the choice of U and η; it is only necessary that $U\eta^3$ be not changing. Thus, without loss of generality, one may take $U = a_0$. Then $\eta_0 = (3ct_0/a_0)^{1/3}$ defines the conformal time at t_0. But $a(t_0) = a_0$, and hence $\eta_0 = 1$. Therefore, in this choice $W = b_0^2(3c^2)^{-1}$, and the comoving distance $\eta_0 = q = 1$ defines the distance of the particle horizon at $t = t_0$. The sound velocity b_0 must be defined independently by physical considerations. As it is clear from Weinberg (1972), during the recombination b_0^2 quickly decreases from $\simeq (c^2/3)$ to $(5kT/(3m))$, where k is the Boltzmann constant, $T \simeq 4000$ K is the temperature during the recombination, and m is the mass of hydrogen atom. Hence, one may expect that W quickly decreases from $\simeq 10^{-1}$ to 10^{-9}. Because, for the occurence of Bessel-Macdonald instability one needs at least $W|\lambda_s| \sim 10$ (Remark B), the maximum of still allowed comoving distance $|\lambda_s|^{-1/2}$ should quickly decrease from $\sim 10^{-1}$ to $\sim 10^{-5}$. The relevant maximum mass, for which the Bessel-Macdonald instability still occurs, is then the $\sim (10^{-3} - 10^{-15})$ fraction of the mass that is included inside the horizon at $z \sim 1000$. Thus, the maximum of mass of Bessel-Macdonald instability is drastically decreasing from $\sim 10^{12-15} M_{\odot}$ to $\sim (1 - 10^3) M_{\odot}$ (Weinberg 1972; Sect. 15.8.). We see that the ad hoc qualitative assumption of BDM picture - namely that during the recombination there are bold non-linearities for masses smaller than $10^9 M_{\odot}$ (see Introduction) - is quite naturally reproduced. In this modification the upper threshold seems to be even smaller and directly the stellar masses are preferred, because - fully heuristically - the value $\sim (1 - 10^3) M_{\odot}$ seems to be more reasonable. Nevertheless, the concrete quantitative value of this upper threshold is highly scattered, and we can only claim that it exists; very roughly it reproduces the value of conventional BDM picture, because this maximum may in principle be from the whole interval $\sim (1 - 10^{15}) M_{\odot}$.

H. The theory of small perturbations in a Friedmannian universe was first presented by Lifshitz (1946). The Fourier decomposition was introduced by Bonnor (1957). Since that time, as far as it is known, exclusively only this method is used for the cases with $b \neq 0$. From the mathematical point of view, in this procedure one proceeds as follows. First, one defines the Fourier decomposition of δ, and, second, one substitutes the hyperbolic partial differential equation with three spatial coordinates (Eq. (1)) by a set of ordinary differential equations having the same form as Eq. (11), in which the terms $b^2 a^{-2} const.$ and $4\pi G\rho$ will always have the opposite signs for any mode (see, e.g., Weinberg, 1972; Sect. 15.9.; Peebles, 1980; Sect. 26). To have this procedure one must use "zero boundary conditions", as it was already noted in Sect. 2. Then one obtains a special concrete solution of Eq. (1). To search for other solutions for a given finite part of three-dimensional space one may either use different boundary (instead of cube e.g. a sphere), or to consider "non-zero boundary conditions" for the cube, or even both. In this paper this last possibility was done; there were taken a finite comoving sphere with radius q_o, and on it there were "non-zero boundary conditions" (Eq. (10) with $\infty > Y_{nm} > 0$). In fact, there are two quite different arguments for this choice: 1. One may try to search for solutions also inside a comoving cube with "non-zero boundary conditions". In fact, from the mathematical point of view, this may easily be done; it is enough to generalize trivially the considerations of Morse & Feshbach (1953) (Sect. 6.3.) done for the two-dimensional space for three spatial dimensions. In this case, not on all six sides of cube, but only on five sides are "zero boundary conditions", and one again obtains the fast growth due to the Bessel-Macdonald instability. But the author means that this mathematically correct solution is so artificial that hardly can have any physical significance. 2. One may also consider the interior of a comoving sphere with zero boundary conditions. But in this case the Bessel-Macdonald instability will not occur, because at that case λ_1 cannot be negative (see subsection 2.2.). All this means that the occurence of Bessel-Macdonald instability is not caused by the choice of spherical spatial region, but by the choice of "non-zero initial conditions" (see Eq. (10)). In fact, the occurence of this instability may be expected for any type of closed comoving region; the choice of comoving sphere was simply given by the fact that then mathematical solutions may also have well defined physical significance (Remark E). On the other hand, oppositely, one may claim that for any type of spatial region, once on its boundary there are "zero boundary conditions", no Bessel-Macdonald instability will occur.

I. The key counter-argument against the galaxy formation theory based on adiabatic perturbations in a baryon dominated universe is the proclaim that the perturbations cannot grow faster than $\sim (1+z)^{-1}$. Then this model needs $\sim 10^{-3}$ intrinsic anisotropies of cosmic microwave background radiation, which are clearly not present (for details see, for example, Börner 1993; Sect. 11.2.2.). This counter-argument does not hold further, because from any small initial perturbations non-linearities can occur, as we have seen. This means that the rejection of adiabatic perturbations in the baryon dominated universe is premature (in Peebles (1993); p.655; this model is mentioned as "only of historical interest").

J. Some theoretical studies trying to obtain fast growth of adiabatic perturbations were already done. These efforts can be separated into two classes. The first class contains the considerations based on Eq. (1) (Mészáros 1991, Mészáros 1993a, Mészáros 1995). (This paper belongs here, too. As far as it is know, no such efforts were done by other authors yet). From these papers it follows that under some artificial choice of initial values one may be in accordance with observations. Nevertheless, only this paper avoids the special choice of initial values. The second class contains the efforts, when Eq. (1) is substituted by a different equation (Mészáros 1993b; Peebles 1980; Sect. 18). It is known that Eq. (1) is an approximation usable only for $|\delta| \ll 1$, and $|v/(Hd)| \ll 1$, where v is the velocity of perturbations and d is the characteristic spatial size of these perturbations (Peebles 1980; Eq. (10.1)). One may think that for great spatial gradients used in this paper the condition $|v/(Hd)| \ll 1$ need not be fulfilled (because d should be small), and hence the Bessel-Macdonald instability should not occur. This is not the case. Assume that only the condition $|\delta| \ll 1$ occurs. Then Eq. (1) should be substituted in cartesian coordinates by (Peebles 1980; Eq. (18.1))

$$\ddot{\delta} + 2H\dot{\delta} = b^2 a^{-2}\Delta\delta + 4\pi G\rho\delta + a^{-2}\partial^2(v^\alpha v^\beta)/(\partial q^\alpha \partial q^\beta), \quad (25)$$

where $\alpha,\ \beta = 1, 2, 3$ and v^α are the components of the velocity of perturbations. (Once Eq. (25) is taken, then mathematically one needs further three equations to have a system of four partial differential equations for four unknowns δ, v^α. This system is unsolvable analytically. The details of these mathematical questions are in Sect. 9 of Peebles (1980).) The effect of the last "velocity term" on the right-hand-side of Eq. (25) is estimated in Sect. 18 of Peebles (1980). (The term $b^2 a^{-2}\Delta\delta$ is "the pressure term"; $4\pi G\rho\delta$ is "the gravitational term"; the last one is the "velocity term".) It is shown that the velocity term should even increase the growth, because the velocity term and the gravitational term should have the same sign. From this fact one may conclude at once that the velocity term cannot destroy the Bessel-Macdonald instability. This instability is in essence caused by the same sign of pressure term and gravitational term (Remark H). But, to avoid the Bessel-Macdonald instability, one would necessarily need an opposite sign for the velocity term compared with the remaining two ones. But, because the velocity term and the gravitational term should have the same sign, this cannot occur; the velocity term should even increase of the huge growth. Substitution of Eq. (1) by Eq. (25) hardly can change the conclusions of this paper.

K. It is quite remarkable that the Bessel-Macdonald instability leads to non-linearities during the recombination on the scales of solar masses; exactly this was introduced by the conventional BDM model based on entropy fluctuations (see Introduction). But in the modification outlined in this paper no non-gravitational phenomenons are needed, and, in addition, adiabatic perturbations - not necessarily existing before the recombination - were also enough. Any small initial perturbations on small scale can at once be non-linear due to Bessel-Macdonald instability, once the Mészáros effect (Mészáros 1974) does not work. It is well-known that Eq. (1) is wrong for the perturbations before the recombination, and, in addition, due to the Silk damping + the Mészáros' effect (Peebles 1971; p.220; Mészáros 1974; Peebles 1980; Sect. 11), it is natural to assume that no adiabatic perturbations existed before the recombination. Thus, it is highly problematic to assume the existence of any adiabatic perturbations in the pre-recombination era, and, actually, they are not needed here. (Entropy fluctuations may exist, but the use of them seems to be unnecessary.) Thus, it is reasonable to assume that the Bessel-Macdonald instability occurs either during or immediately after the recombination from extremely small (say, of order $\sim 10^{-5}$) perturbations surely existing during the recombination. For the sake of precisity it must be added that the claims of "the non-existence of adiabatic perturbations before the recombination" means that such perturbations were not bigger than $\sim 10^{-5}$; see Börner (1993), Sect. 11.2.2. It is sure that some departures from the exact homogeneity and isotropy below this limit surely existed at the pre-recombination era. For example, the temperature of the microwave background radiation should decrease as $\sim (1+z)^{-1}$, but the temperature of plasma as $\sim (1+z)^{-2}$, and the mixture of them cannot be in an exact statistical equilibrium necessary for the perfect Friedmannian model; for details see Peebles (1971), p.230. Hence, such extremely small perturbations cannot be excluded even before the recombination. It is natural to expect that even from these small perturbations during and immediately after the recommendation the first objects arose. (The details of these phenomenons should be studied in near future.)

5. Conclusion

The main mathematical results of this paper show that new solutions of Eq. (1) allow an extremely fast growth of δ during and after the recombination. The initial $|\delta(t_0, q, \vartheta, \varphi)|$ may be arbitrarily small; it is enough that the spatial gradients themselves be large enough (concretely, in Eq. (10) the parameter Y_{nm} should be a large positive number). It is also noted that the Fourier method with "zero boundary condition" used earlier does not give these solutions; in fact, the "zero boundary conditions" exclude the possibility of the so called Bessel-Macdonald instability. It is shown that, once there are ”non-zero boundary conditions”, this instability may occur. The key conclusion of this paper follows from the fact that for a spherical comoving region this effect may have well defined physical significance (Remark E.).

These results should have an essential cosmological importance, because it can lead to the ”resurrection” of the simplest and most natural galaxy formation theory based on the adiabatic perturbations in an Universe dominated by baryonic dark matter. The development of the details of this modified version of BDM picture - mainly the investigations of physical aspects - is the subject of further studies.

Acknowledgements. The author would like to thank the valuable remarks of an anonymous referee.

References

Bardeen J.M., 1980, Phys. Rev. D22, 1882
Bonnor W.B., 1957, MNRAS 117, 104
Börner G., 1993, The Early Universe. Springer, Berlin
Bracewell R.N., 1978, The Fourier Transform and its Application, McGraw-Hill, Auckland
Courant R., Hilbert D., 1968, Methoden der mathematischen Physik II, Springer, Berlin
Ellis G.F.R., Bruni M., 1989, Phys. Rev. D40, 1804
Hogan C.J., 1993, ApJ 415, L63
Lifshitz E.M., 1946, J.Phys.(Moscow) 10, 116
Mészáros A., 1991, MNRAS 253, 619
Mészáros A., 1993a, Astro. Spa. Sci 207, 5
Mészáros A., 1993b, A&A 278, 1
Mészáros A., 1995, Acta Phys. Hung. New Ser. 1, 107
Mészáros P., 1974, A&A 37, 225
Morse P.M., Feshbach H., 1953, Methods of Theoretical Physics. McGraw-Hill, New York
Peebles P.J.E., 1971, Physical Cosmology. Princeton Univ. Press, Princeton
Peebles P.J.E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton
Peebles P.J.E., 1993, Principles of Physical Cosmology. Princeton Univ. Press, Princeton
Watson G.N., 1922, Theory of Bessel Functions. Cambridge Univ. Press, Cambridge
Weinberg S., 1972, Relativity and Cosmology. J.Wiley, New York

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.