

Primordial black hole formation

Ilia Musco* †

SISSA, Trieste, Italy & CMA, Oslo, Norway

E-mail: ilia.musco@cma.uio.no

This work is a review of our main results obtained from general relativistic numerical computations of primordial black-hole formation during the radiation-dominated era of the universe. As initial conditions we consider supra-horizon-scale perturbations of a type which could have come from inflation, with only a growing component and no decaying component. We use a spherically symmetric Lagrangian code and study both super-critical perturbations, which go on to produce black holes, and sub-critical perturbations, for which the overdensity eventually disperses into the background medium. For super-critical perturbations, we confirm the results of previous work but noticing that the threshold amplitude for a perturbation to lead to black-hole formation is substantially reduced when one considers initial perturbations with a length scale sufficiently larger than the cosmological horizon, according to initial cosmological perturbations coming from inflation. For sub-critical cases, where an initial collapse is followed by a subsequent re-expansion, strong compressions and rarefactions are seen for perturbation amplitudes near to the threshold. In order to study perturbations with amplitudes extremely close to the supposed critical limit, we have introduced into the code an adaptive mesh refinement scheme. This allows us to see that scaling-law behaviour continues down to the smallest black hole masses that we are able to follow and we see no evidence of shock production such as has been reported in some previous studies and which led there to a breaking of the scaling-law behaviour at small black-hole masses. We attribute this difference to the different initial conditions used.

Black Holes in General Relativity and String Theory

24 - 30 August, 2008

Veli Losinj, Croatia

*Speaker.

†I am grateful to Prof. Lorian Bonora and the organization of the meeting for supporting my participation.

1. Introduction

It is possible that a population of black holes may have been formed during the radiative era of the early universe by means of gravitational collapse of cosmological perturbations. This was first proposed by Zel'dovich & Novikov (1969) [1] and then by Hawking (1971) [2]. The idea of these primordial black holes (PBHs) has subsequently attracted continuing attention, even if they have sometimes been regarded as uninteresting on the grounds that perturbations with amplitudes sufficiently large to form them are likely to have been extremely rare. Following the initial proposals, the possible formation process was then investigated further by Carr [3, 4] using an ingenious simplified model in which he considered an overdense spherical region, described as part of a closed Friedmann-Robertson-Walker (FRW) universe, surrounded by a spatially-flat FRW expanding background. Results from this simplified model suggested that PBH formation would occur if the perturbation amplitude δ , defined as the relative mass excess inside the overdense region measured when it had the same scale as the cosmological horizon, is greater than a threshold value of about $1/3$. A more precise determination of this threshold value, usually referred to as δ_c , requires a numerical computation and this was first done by Nadezhin, Novikov & Polnarev (1978) [5] and by Bicknell & Henriksen (1979) [6]. In the 1980s, attention shifted mainly to the cosmological consequences that a population of PBHs would have, and this was extensively studied in various scenarios (see, for example, the recent review by Carr [7]). In 1999 Niemeyer & Jedamzik [8] returned to the issue of making numerical calculations of the PBH formation process, and they were subsequently followed in this by other groups [9, 10] including ourselves [11, 12, 13]. All of these studies confirmed the overall picture of PBH formation, showing that the precise value of δ_c depends on the initial perturbation profile.

An issue with all of these calculations concerns the specification of initial conditions. In particular Niemeyer & Jedamzik in 1999 used non-linear horizon-scale perturbations as their initial conditions [8], finding that $\delta_c \simeq 0.7$ (in the range of $0.67 - 0.71$) for the three types of perturbation profile that they studied. In the same year, Shibata and Sasaki [9] presented an alternative formalism for studying PBH formation focusing on metric perturbations, rather than density perturbations, as had also been done previously [5, 14]. They pointed out that the initial conditions used in [8] were specified initially within a nonlinear regime of perturbations of the energy density and velocity field, and were therefore “inevitably contaminated by an unrealistic decaying mode” component that would diverge for $t \rightarrow 0$. In a recent paper [15] this analysis has been converted in terms of the perturbation amplitude δ , showing that the Shibata and Sasaki results, corresponding to a wide choice of perturbation shapes, are consistent with δ_c in the range $0.3 \lesssim \delta_c \lesssim 0.5$. Bearing all this in mind, we ourselves then made calculations in 2005 [11] specifying the initial conditions within the linear regime of energy density and velocity field, giving only growing mode solutions at the horizon crossing time. During the subsequent evolution, in fact, the decaying component of the perturbation had time to become negligible before the perturbation passed inside the cosmological horizon (at the “horizon-crossing” time) and started its non-linear evolution, making the treatment of the eventual black-hole formation a better representation of what would happen for perturbations coming from inflation. The simulations in [11] gave values of δ_c in the range $0.43 - 0.47$ (instead of $0.67 - 0.71$) for the same types of perturbation profile used in [8], explaining so the disagreement pointed out in [15].

All of these studies confirmed the overall picture of PBH formation, showing that the precise value of δ_c depends on the initial perturbation profile. However, the best thing to do is clearly to specify from the beginning perturbations containing only a growing component. We addressed this in 2007 [12], where the initial conditions were again specified as linear supra-horizon-scale perturbations of the fluid parameters, but now using an asymptotic quasi-homogeneous solution [16], which gives a consistent pure growing behaviour. When the perturbation length-scale is much larger than the cosmological horizon, the perturbation in the curvature can be non-linear (as necessary in order to subsequently produce a black hole) while the perturbations of the hydrodynamical quantities are still in the linear regime. In this situation, the coupled system of Einstein and hydrodynamical equations can be solved analytically, and one finds that the small perturbations in energy density, velocity, etc are generated uniquely by the curvature profile, which is then the only quantity that needs to be specified. In [12] the numerical scheme developed in [11] was modified so as to use initial data coming from this quasi-homogeneous solution, and the values of δ_c found in [11] were confirmed. The crucial point of this argument is that the main perturbations of interest in this cosmological context are ones coming from inflation which then re-enter the cosmological horizon at later times.

For understanding the cosmological impact of PBHs, it is necessary to determine the probabilities of different types of perturbation profile (see Hidalgo & Polnarev [17], and references therein) and the corresponding values of δ_c . However it is also necessary to have a consistent understanding of the collapse process and of the dependence of the PBH mass on the perturbation amplitude. An important contribution to this was made by Niemeyer & Jedamzik [8, 18] (reproduced by us in [11]) who showed also that the masses of PBHs produced in the radiative era by perturbations with a given profile type, follow the scaling-law behaviour of critical collapse, first discovered for idealized conditions by Choptuik [19], i.e. the masses of the black holes produced follow a power law $M_{BH} \propto (\delta - \delta_c)^\gamma$ if δ is close enough to δ_c . If one considers a perfect fluid with an equation of state of the form $p = we$ (where p is the pressure and e is the energy density), then the value of the critical exponent γ depends on the value of w [20]. Niemeyer & Jedamzik found a scaling law with $\gamma \simeq 0.36$ for the models which they studied, agreeing with the value of γ already found for idealized critical collapse in a radiation fluid (having $w = 1/3$) [21]. However, their code was only able to handle cases with a fairly small range of $\delta - \delta_c$ (and hence of PBH masses), leaving open the question of whether the scaling law would continue down to very low values of M_{BH} . In 2002 this issue was taken up by Hawke & Stewart [10] who used a sophisticated purpose-built code which was able to handle a much wider range of values of $\delta - \delta_c$ and deal with the strong shocks that they found arising for cases near to the critical limit ($\delta \rightarrow \delta_c$). They too found a scaling law over a certain range of PBH masses (spanning a factor of ~ 100) but with deviations away from this for higher and lower masses. In particular, for the perturbations having amplitudes closest to the critical limit, they found that the relationship levelled off at a minimum mass of $\sim 10^{-3.5}$ of the cosmological horizon mass at their initial time. However in [10] non linear initial conditions had again been considered, including a decaying component, and this could be the cause of these results. For this reason, in a third paper by us ([13] submitted), we went much deeper into the nearly-critical regime, using an updated version of our previous computer code implementing the scheme with an adaptive mesh refinement, but still using pure growing solutions as initial conditions. In the present report, we give a review of our main results which have been described more

extensively in [11, 12, 13], focusing in particular on our more recent results regarding behaviour in the nearly-critical regime and possible breaking of the scaling law (for which we, in fact, find no evidence).

2. Mathematical formulation of the problem

We use two different formulations of the general relativistic hydrodynamic equations: one for setting the initial conditions and the other for studying the black hole formation. Throughout, we are assuming spherical symmetry and that the medium can be treated as a perfect fluid; we use a Lagrangian formulation of the equations with a radial coordinate r which is co-moving with the matter.

2.1 The Misner-Sharp formalism

For the initial conditions, it is convenient to use a diagonal form of the metric, with the time coordinate t reducing to the standard FRW time in the case of a homogeneous medium with no perturbations. (This sort of time coordinate is therefore often referred to as “cosmic time”). We write this metric in the form given by Misner & Sharp (MS) [22], whose approach we follow here in writing the GR hydrodynamic equations:

$$ds^2 = -a^2 dt^2 + b^2 dr^2 + R^2 (d\theta^2 + \sin^2 \theta d\varphi^2), \quad (2.1)$$

with the coefficients a , b and R being functions of r and t and R playing the role of an Eulerian radial coordinate. Using the definitions

$$D_t \equiv \frac{1}{a} \left(\frac{\partial}{\partial t} \right), \quad (2.2)$$

$$D_r \equiv \frac{1}{b} \left(\frac{\partial}{\partial r} \right), \quad (2.3)$$

one defines the quantities

$$U \equiv D_t R, \quad (2.4)$$

and

$$\Gamma \equiv D_r R, \quad (2.5)$$

where U is the radial component of four-velocity in the “Eulerian” frame and Γ is a generalized Lorentz factor. The metric coefficient b can then be written as

$$b \equiv \frac{1}{\Gamma} \frac{\partial R}{\partial r}, \quad (2.6)$$

For a radiative medium where rest-mass makes a negligible contribution to the energy density, the GR hydrodynamic equations can then be written in the following form (where we use the notation that e is the energy density, p is the pressure, ρ is the compression factor, and M is the mass contained inside radius R):

$$D_t U = - \left[\frac{\Gamma}{(e+p)} D_r p + \frac{M}{R^2} + 4\pi R p \right], \quad (2.7)$$

$$D_t \rho = -\frac{\rho}{\Gamma R^2} D_r (R^2 U), \quad (2.8)$$

$$D_t e = \frac{e+p}{\rho} D_t \rho, \quad (2.9)$$

$$D_t M = -4\pi R^2 p U, \quad (2.10)$$

$$D_r a = -\frac{a}{e+p} D_r p, \quad (2.11)$$

$$D_r M = 4\pi R^2 \Gamma e, \quad (2.12)$$

plus a constraint equation

$$\Gamma^2 = 1 + U^2 - \frac{2M}{R}. \quad (2.13)$$

We also have the equation of state for a perfect fluid:

$$p = we \quad (2.14)$$

with $w = 1/3$ for radiation fluid, used in the results presented later.

2.2 Description of the initial conditions

The initial conditions are set by considering a perturbation of the otherwise uniform medium whose length-scale R_0 is much larger than the cosmological horizon $R_H \equiv H^{-1}$. Under these circumstances, the perturbations in e and U can be extremely small while still giving a large amplitude perturbation of the metric (as is necessary if a black hole is eventually to be formed) and the above system of equations can then be solved analytically to first order in the small parameter $\varepsilon \equiv (R_H/R_0)^2 \ll 1$, that is equivalent to having $R_0 \equiv NR_H$ with $N \gg 1$. The result, called the “quasi-homogeneous solution” [12], gives formulae for the perturbations of all of the metric and hydrodynamical quantities in terms only of the curvature perturbation. When $\varepsilon \ll 1$, the curvature perturbation is time-independent [23] and in our scenario we find it convenient to represent it in terms of a quantity $K(r)$ similar to the constant K appearing in the standard form of the FRW metric but now depending on r .

The system of equations (2.4) - (2.13) can be re-written as:

$$\dot{R} = aU \quad (2.15)$$

$$\frac{\dot{b}}{b} = a \frac{U'}{R'} \quad (2.16)$$

$$\frac{a'}{a} = -\frac{w}{1+w} \frac{e'}{e} \quad (2.17)$$

$$\dot{M} = -4\pi w e R^2 \dot{R} \quad (2.18)$$

$$M' = 4\pi e R^2 R' \quad (2.19)$$

$$\frac{R'^2}{b^2} = 1 + U^2 - \frac{2M}{R} \quad (2.20)$$

where the dot and dash denote differentiation with respect to t and r respectively. We use the equation of state (2.14) to express the pressure as a function of the energy density, equation (2.5) gives an expression for Γ , and ρ can be calculated from equation (2.8).

The background solution is a spatially flat FRW universe described by $K = 0$. The corresponding background value of the energy density (indicated with the suffix “b”) is calculated from the Friedmann equation

$$\left(\frac{\dot{s}}{s}\right)^2 = H_b^2 = \frac{8\pi}{3}e_b. \quad (2.21)$$

The quantity e_b is related to the scale factor $s(t)$ by the continuity equation (2.9) used for the unperturbed case

$$\frac{\dot{e}_b}{e_b} = -3(1+w)\frac{\dot{s}}{s}, \quad (2.22)$$

while the background values of the other quantities are obtained from the FRW metric and the set of equations (2.15) - (2.20):

$$a_b = 1 \quad (2.23)$$

$$b_b = s(t) \quad (2.24)$$

$$R_b = s(t)r \quad (2.25)$$

$$M_b = \frac{4}{3}\pi e_b R_b^3 \quad (2.26)$$

$$U_b = H_b R_b = \dot{s}(t)r \quad (2.27)$$

One can relate the scale factor to the comoving length scale of the perturbation r_0

$$s(t) = \frac{NR_H}{r_0}, \quad (2.28)$$

and define the comoving coordinate as

$$r = \frac{R_b}{s(t)} = \frac{r_0}{NR_H}R_b, \quad (2.29)$$

where the value of R_b is obtained by integrating equation (2.6) with $\Gamma = 1$ and r_0 is determined by the particular expression used for $K(r)$ [12].

The initial perturbations are defined as:

$$R = R_b(1 + \varepsilon\tilde{R}) \quad (2.30)$$

$$U = H_b R(1 + \varepsilon\tilde{U}) \quad (2.31)$$

$$b = \frac{R'}{\sqrt{1 - K(r)r^2}}(1 + \varepsilon\tilde{b}) \quad (2.32)$$

$$a = 1 + \varepsilon\tilde{a} \quad (2.33)$$

$$e = e_b(1 + \varepsilon\tilde{e}) \quad (2.34)$$

$$M = \frac{4}{3}\pi e_b R^3(1 + \varepsilon\tilde{M}) \quad (2.35)$$

where here, and in the following, the tilde denotes a perturbation quantity. Considering ε as a small parameter allows one to keep only first order terms with respect to ε in the hydrodynamical equations (2.15) - (2.20). Inserting the perturbations just defined into this set of equations, we get

the explicit form of the so called quasi-homogeneous solution, expressing all of the tilde-quantities just as functions of $K(r)$:

$$\tilde{\epsilon} = \Phi \frac{1}{3r^2} [r^3 K(r)]' r_0^2 \quad (2.36)$$

$$\tilde{a} = -\Phi \frac{\gamma}{1+\gamma} \frac{1}{3r^2} [r^3 K(r)]' r_0^2 \quad (2.37)$$

$$\tilde{U} = \frac{1}{2} [\Phi - 1] K(r) r_0^2 \quad (2.38)$$

$$\tilde{M} = \Phi K(r) r_0^2, \quad (2.39)$$

$$\tilde{R} = -I_1 \frac{1}{3r^2} [r^3 K(r)]' r_0^2 + I_2 \frac{K(r)}{2} r_0^2, \quad (2.40)$$

$$\tilde{b} = I_1 r \left[\frac{1}{3r^2} (r^3 K(r))' \right]' r_0^2. \quad (2.41)$$

with

$$\Phi = \frac{3(1+w)}{5+3w} \quad (2.42)$$

$$I_1 = \frac{3w}{(1+3w)(5+3w)} \quad (2.43)$$

$$I_2 = -\frac{2}{(1+3w)(5+3w)} \quad (2.44)$$

being three parameters that depend only on the value of w in the equation of state ¹. The time variation of the perturbation quantities is given only by $\varepsilon(t)$:

$$\varepsilon(t) \propto \left(\frac{t}{t_0} \right)^{\frac{2(1+3w)}{3(1+w)}}, \quad (2.45)$$

which has the same behaviour as the standard solution for a growing mode in cosmological perturbation theory (see for example [24, 25]).

2.3 General properties of the curvature profile

From the definition of b given by (2.32), we obtain an obvious mathematical requirement

$$1 - K(r)r^2 > 0 \Rightarrow K(r) < \frac{1}{r^2}. \quad (2.46)$$

This corresponds to the physical condition that a perturbed spherical region of co-moving radius r should not be causally disconnected from the rest of the Universe. This condition was formulated in [7] and recently analyzed in [26], where the perturbation was described by a region with $K = 1$, embedded in a region with $K = 0$.

Another important requirement is given by the condition that the total excess of mass in the initial configuration is equal to zero

$$\int_0^\infty 4\pi r^2 \tilde{\epsilon}(r) dr = 0. \quad (2.47)$$

¹As explained in [12], the quasi-homogeneous solution holds also for multiple fluids with $w = w(t)$; in that case, there is not a simple analytical expression for Φ, I_1 and I_2 and the tilde-quantities depend on both r and t .

Using (2.36), this is equivalent to

$$\lim_{r \rightarrow \infty} r^3 K(r) = 0. \quad (2.48)$$

Condition (2.47) follows from conservation of total mass-energy and is related to the fact that we consider only perturbations within a spatially flat Universe. This requirement is related to the causality principle (see [12] for further explanation).

An important parameter characterising the curvature profile is the value of r_0 that specifies the comoving length-scale of the overdense region in the configuration. From (2.36) we have

$$\tilde{\epsilon}(r_0) = 0 \Rightarrow K(r_0) + \frac{r_0}{3} K'(r_0) = 0. \quad (2.49)$$

and this allows one to determine the value of r_0 to be used in the quasi-homogeneous solution, when the explicit form of $K(r)$ has been chosen.

To characterize the amplitude of the perturbation, we use the integrated quantity δ , that measures the relative mass excess within the overdense region, as frequently done in the literature. In the quasi-homogeneous solution, this is given by

$$\delta(t) \equiv \left(\frac{4}{3} \pi r_0^3 \right)^{-1} \int_0^{r_0} 4\pi \frac{e^{-e_b}}{e_b} r^2 dr = \Phi \epsilon(t) K(r_0) r_0^2 \quad (2.50)$$

which also has the familiar linear growth with cosmic time. For discussing the scaling law, we need to have a measure of the perturbation amplitude defined consistently for different cases, and for this we use δ as given by expression (2.50) with ϵ set equal to 1. This gives very similar values to those actually calculated at horizon crossing, with non-linear evolution taken into account, but it can be calculated with greater precision since it is given analytically.

In the present work, we make the choice of using a Gaussian profile for $K(r)$, normalized to be equal to 1 at the centre:

$$K(r) = \exp\left(-\frac{r^2}{2\Delta^2}\right). \quad (2.51)$$

Substituting this into (2.36), one obtains

$$\tilde{\epsilon} = \frac{2}{3} r_0^2 \left[1 - \frac{r^2}{3\Delta^2} \right] \exp\left(-\frac{r^2}{2\Delta^2}\right). \quad (2.52)$$

Since $r = r_0$ at the outer edge of the overdensity (where $\tilde{\epsilon} \rightarrow 0$), one then has $r_0^2 = 3\Delta^2$ and so (2.52) can be rewritten as

$$\tilde{\epsilon} = 2\Delta^2 \left[1 - \left(\frac{r}{r_0}\right)^2 \right] \exp\left(-\frac{3}{2} \left(\frac{r}{r_0}\right)^2\right), \quad (2.53)$$

or

$$\tilde{\epsilon} = 2\Delta^2 \left[1 - \left(\frac{R_b}{R_0}\right)^2 \right] \exp\left(-\frac{3}{2} \left(\frac{R_b}{R_0}\right)^2\right), \quad (2.54)$$

in terms of the background radial coordinate $R_b = s(t)r$, where $s(t)$ is the cosmological scale factor. This expression is that for the well-known mexican-hat profile, as used in previous calculations by Niemeyer & Jedamzik [8] and ourselves [11, 12]. Inserting the expression for r_0 into (2.50) gives

$$\delta = 2\Delta^2 \exp\left(-\frac{3}{2}\right) \quad (2.55)$$

when $\varepsilon = 1$ and $w = 1/3$, and so the perturbation amplitude can be characterized alternatively by Δ . In [11] and [12] the δ_c for this profile was found to be $\simeq 0.45$, which corresponds to $\Delta_c \simeq 1$. For further details about setting initial conditions using the quasi-homogeneous solution, the reader is referred to [12].

2.4 The Hernandez-Misner formalism

The MS approach using cosmic time slicing is convenient for setting initial conditions but has a well-known drawback for calculations of black hole formation in that singularities are formed rather quickly when using it and then further, potentially observable, evolution cannot be followed unless an excision procedure is used. Various slicing conditions can be used to avoid this difficulty but for calculations in spherical symmetry it is particularly convenient to use null slicing. In our work we have used the ‘‘observer time’’ null-slicing formulation of Hernandez & Misner [27] where the time coordinate is taken as the time at which an outgoing radial light ray emanating from an event reaches a distant observer. (In the original formulation, this observer was placed at future null infinity but for calculations in an expanding cosmological background we use an FRW fundamental observer sufficiently far from the perturbed region so as to be unaffected by the perturbation.) We use the MS approach for setting up initial data and for evolving it so as to produce data on a null slice. This is then used as input for our observer-time code with which we follow the black hole formation.

For the observer-time calculation, the metric (2.1) is re-written as

$$ds^2 = -f^2 du^2 - 2fb dr du + R^2 (d\theta^2 + \sin^2 \theta d\varphi^2), \quad (2.56)$$

where u is the observer time and f is the new lapse function. The operators equivalent to (2.2) and (2.3) are now

$$D_t \equiv \frac{1}{f} \left(\frac{\partial}{\partial u} \right), \quad (2.57)$$

$$D_k \equiv \frac{1}{b} \left(\frac{\partial}{\partial r} \right), \quad (2.58)$$

where D_k is the radial derivative in the null slice and the corresponding derivative in the Misner-Sharp space-like slice is given by

$$D_r = D_k - D_t. \quad (2.59)$$

The hydrodynamic equations can then be formulated in an analogous way to what was done in cosmic time and we refer the reader to our discussion in [11].

3. The method of calculation

As in our previous work, the present calculations have been made with an explicit Lagrangian hydrodynamics code based on that of Miller & Motta (1989) [28] but with the grid organized in a way similar to that in the code of Miller & Rezzolla (1995) [29] which was designed for calculations in an expanding cosmological background. The code has a long history and has been carefully tested in its various forms. Full details of the methods used have been given in the above papers, in [11, 12], and in other papers cited there.

The code uses a staggered grid and the co-moving coordinate is a mass-type coordinate which we will refer to as μ in the discussion below. The basic grid has logarithmic spacing, allowing it to reach out to very large radii while giving finer resolution at small radii. The initial data for our present calculations was derived from the quasi-homogeneous solution and was specified on a space-like slice (at constant cosmic time) with $\varepsilon = 10^{-2}$, giving $R_0 = 10R_H$. The outer edge of the grid was placed at $90R_H$, which is far enough away so that there is no causal contact between it and the perturbed region during the time of the calculations. The initial data was then evolved using the MS equations (2.7-2.13), so as to generate a second set of initial data on a null slice (at constant observer time). To do this, an outgoing spherical light pulse was traced out from the centre, and parameter values were noted as it passed each grid zone. It is not necessary to continue to evolve the solution significantly behind the wave front in this part of the calculation, since that region (within which singularities might form) does not have causal contact with the wave front. The null-slice initial data, constructed in this way, was then evolved using the observer-time equations [11]. From here on, we will refer to the observer-time coordinate as t rather than u .

In [13], our aim was to study PBH formation for much smaller values of $(\delta - \delta_c)$ than was possible with our previous version of the code ([11, 12]). The particular problem which needed to be overcome is that as the PBH is forming, a semi-void appears around it and this becomes progressively deeper as δ gets closer to δ_c . For a Lagrangian code with a co-moving grid, this means that the Eulerian widths of zones in the semi-void region become very large, producing truncation errors that crash the code. The situation is made worse by the fact that a highly-relativistic wind blows matter across this region and this needs good resolution in order to be well-represented. As mentioned above, the basic code was already using a logarithmically-spaced grid in order to be able to simultaneously treat both the small scales on which the PBH was forming and also the large scales necessary for representing the surrounding expanding universe up to scales much larger than the cosmological horizon. For dealing with the problem of the semi-void region, we introduced an adaptive mesh refinement scheme (AMR) on top of the existing logarithmic grid. This has allowed us to follow PBH formation for values of $(\delta - \delta_c)$ up to eight orders of magnitude smaller than we could do before.

Our AMR uses a fully-threaded-tree algorithm, making its refinement on the basis of the Eulerian width of the zone ΔR (which then, in turn, impacts on the truncation error). The outline of the procedure is as follows. Every 20 time-steps, a check is made to see whether the fractional zone spacing $\Delta R/R$ is greater than 10% for any zone. If it is, that zone is subdivided introducing a new grid point at its centre (in terms of the co-moving radial coordinate μ) and quantities at the new grid-point and in the two newly-created zones to either side of it are calculated by means of cubic interpolation. Only one zone is subdivided at a time, allowing the solution to relax before performing any subsequent sub-division. Doing this, we have not found it necessary to introduce any artificial viscosity into the code for suppressing numerical noise, an important point for interpreting the appearance or absence of shocks. Also every 20 time steps (but offset by 10 time steps), we carry out a check to see whether $\Delta R/R$ has become less than 4% across any zone resulting from a previous subdivision. If so, then it is merged with the zone from which it was previously separated if that zone has the same $\Delta\mu$. Again, only one merging is carried out at a time to allow relaxation. Other grid-management routines are used so as to maintain a good overall grid structure.

We have successfully used the scheme with more than thirty levels of refinement and its per-

formance can be judged from the results which we present in the next section. All of the features shown have been fully resolved with the AMR.

4. Description of the results

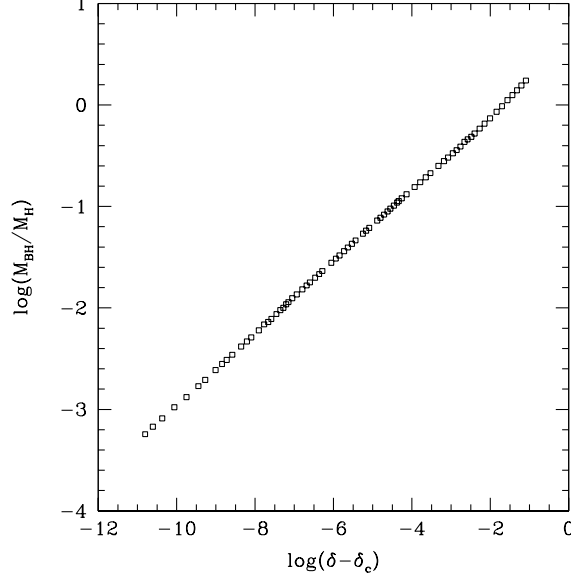


Figure 1: Scaling behaviour for M_{BH} as function of $(\delta - \delta_c)^\gamma$ calculated for a radiative perfect fluid. For $M_{BH} \lesssim M_H$, the points are well-fitted by a scaling law with $\gamma = 0.357$ and $K = 4.02$.

In this section we discuss the numerical results which we have obtained for PBH formation in the radiative era, using the AMR algorithm to go as close to $\delta = \delta_c$ as we can. The main result is shown in figure 1 where the log of the black-hole mass M_{BH} normalized with M_H (the mass inside the cosmological horizon at the moment of horizon crossing) is plotted against the log of $(\delta - \delta_c)$. Note that the cosmological horizon mass continues to increase after the horizon-crossing time, and so it is not inconsistent to have $M_{BH} > M_H$.

We find that for $M_{BH} \lesssim M_H$ (which corresponds to $(\delta - \delta_c) \lesssim 2 \cdot 10^{-2}$), the points for different cases rather accurately follow a scaling law

$$\frac{M_{BH}}{M_H} = K(\delta - \delta_c)^\gamma \quad (4.1)$$

with $\gamma = 0.357$ – close to the expected value for a perfect fluid of radiative particles [21]. For larger values of the masses, the curve steepens slightly, as seen also in the past [8, 10, 11] but there is no sign of divergence from the scaling law at lower masses, even though we have now covered a range of masses in the scaling-law region of more than one thousand. The most extreme case shown has $(\delta - \delta_c) \sim 10^{-11}$ and gives a black hole mass of $\sim 5 \cdot 10^{-4} M_H$. These results are in contrast with those seen by Hawke & Stewart [10], where the scaling-law behaviour persisted for a range of masses of ~ 100 but terminated at low masses with the curve flattening out at a minimum value of M_{BH}/M_H . However, we want to stress that great care needs to be taken in

making a comparison (see [13] for an extended discussion). Their explanation for the minimum mass seen in their calculations was related to the occurrence of strong shocks in the lower-mass cases. With our initial conditions, we do not see these shocks (hence the continuation of the scaling law) but if we impose more general non-linear initial conditions within the cosmological horizon scale, then we *do* often see shocks which are consistent with those reported in [10] (although our code is not equipped to handle the strong shock conditions which they saw and so we can see only the beginnings of this behaviour). We therefore attribute the difference in results to the different initial conditions used.

In our simulations, we start with supra-horizon-scale perturbations derived from the quasi-homogeneous solution, which are still well within the linear regime as far as the hydrodynamical variables are concerned (although they are non-linear in the curvature, as mentioned previously). These perturbations, with only a growing component, have a very special relationship between the density profile and the velocity field. If we imposed as our initial conditions, at the same supra-horizon-scale, a perturbation of density but not of velocity (as we did in [11]), then the perturbation would reconfigure as it evolved, going back to this special relationship between the density profile and the velocity field before horizon crossing. (Comparing with a similar perturbation which is linear in *all* quantities, including the curvature, one can think there of it as being a superposition of growing and decaying modes: as it evolves, the decaying mode dies away, leaving only the growing mode.) When this special type of perturbation becomes non-linear in the fluid quantities as well as in the curvature, it still holds together and does not produce shocks in collapses producing black holes, as a general perturbation set up at late times can do. Shibata & Sasaki [9] who used a similar type of initial perturbation to ours, but formulated in a completely different way, also did not mention seeing shock formation in their collapses producing black holes.

We next present some further results from our simulations of PBH formation near to the critical limit which give further insight into the nature of the process. The critical case ($\delta = \delta_c$) separates cases which collapse to give a black hole from ones which bounce and merge back into the ambient medium. In figure 2 we show some plots for a nearly-critical case and a subcritical one. The top two frames show results from a run which was too close to critical for us to tell whether it would produce a black hole or not by the time that it terminated. Here one can see an important characteristic feature of nearly-critical collapse. The quantity $2M/R$ is a key diagnostic of the behaviour, going to 1 at horizons, both the cosmological horizon and the apparent horizon of the forming black hole. In the top left-hand frame, $2M/R$ is plotted as a function of R/R_H at different time levels, where R_H is the horizon scale at the moment of horizon crossing. The dashed curve represents the initial conditions used by the observer-time code. This plot is showing just the inner part of the solution where the collapse occurs: the outer part where $2M/R$ rises to 1 again at the cosmological horizon is at far larger radii. As time proceeds, the maximum of $2M/R$ initially decreases to around 0.5, moving outwards in radius, but then it moves towards the centre, maintaining an almost constant value but with a very slow decrease towards an eventual value of ~ 0.48 . This marks a “critical surface” that separates cases giving collapse to a black hole from ones which do not [30]. Super-critical and sub-critical cases eventually deviate away from this. The right-hand frame shows the corresponding behaviour of the radial four velocity U (note that these profiles look rather different when viewed in observer-time from the familiar picture in cosmic-time). Initially the velocity is positive everywhere with the general cosmological expansion just being slowed somewhat in the

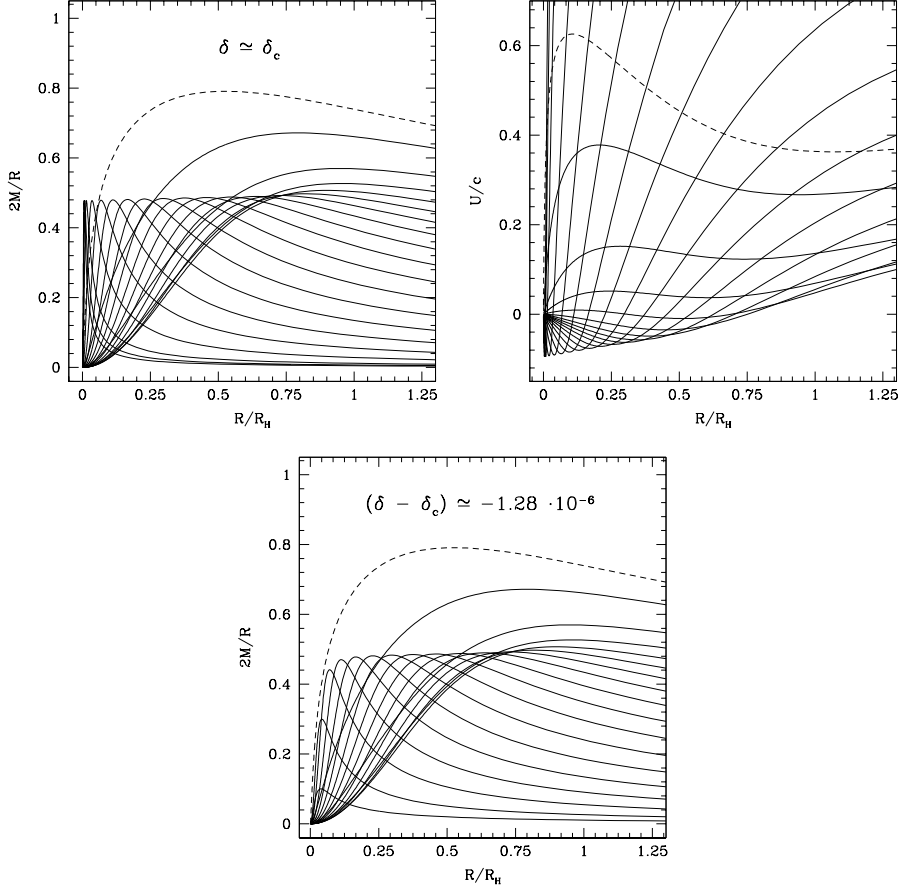


Figure 2: The top two frames show the behaviour of $2M/R$ and the radial four-velocity U for a nearly-critical case ($\delta \simeq \delta_c$), plotted against R/R_H at different time levels, with the dashed curve representing the initial conditions used by the observer-time code. R_H is the horizon scale at the moment of horizon crossing. The bottom frame shows the equivalent plot of $2M/R$ for a sub-critical case with $(\delta - \delta_c) \simeq -1.28 \cdot 10^{-6}$ which does not produce a black hole.

region of the perturbation. As time goes on, the expansion is progressively slowed down both because the universe as a whole is decelerating and because the effect of the perturbation becomes progressively greater in the central regions. After the dashed curve marking the initial time, the maximum in the velocity profile becomes progressively less pronounced and eventually the expansion reverses into a collapse in the central regions. As time proceeds further, the maximum infall velocity increases but the size of the collapsing region shrinks, eventually tending towards zero. Outside the collapsing region, one sees increasing positive velocities, representing a wind which takes matter away from the central condensation. We will analyse this in more detail when we consider a super-critical collapse which produces a black hole. During the “equilibrium phase”, when the maximum of $2M/R$ (which from now on we will refer to as $(2M/R)_{peak}$) remains roughly constant, the location of this maximum corresponds to that where the velocity goes to zero, separating the collapsing region from that of the outgoing wind. At the location of $(2M/R)_{peak}$, there is a continuing close balance between gravitational forces and pressure gradients which is a key feature

of critical collapse. As the collapsing region shrinks, a balance is kept all the way down to zero size in the special critical case $\delta = \delta_c$. For sub-critical cases, with $\delta < \delta_c$, pressure eventually wins over gravity: the remaining material in the collapsing region bounces, with a consequent decrease of $(2M/R)_{peak}$ (see the bottom frame of figure 2), and disperses into the surrounding medium. For nearly-critical cases, this is a surprisingly violent process, as we will discuss later, at the end of this section. For super-critical cases, with $\delta > \delta_c$, gravity eventually wins over pressure, $(2M/R)_{peak}$ then grows towards 1 and a black hole is formed. The rest of the present section contains a detailed discussion of some of the features of this, as seen in the simulations.

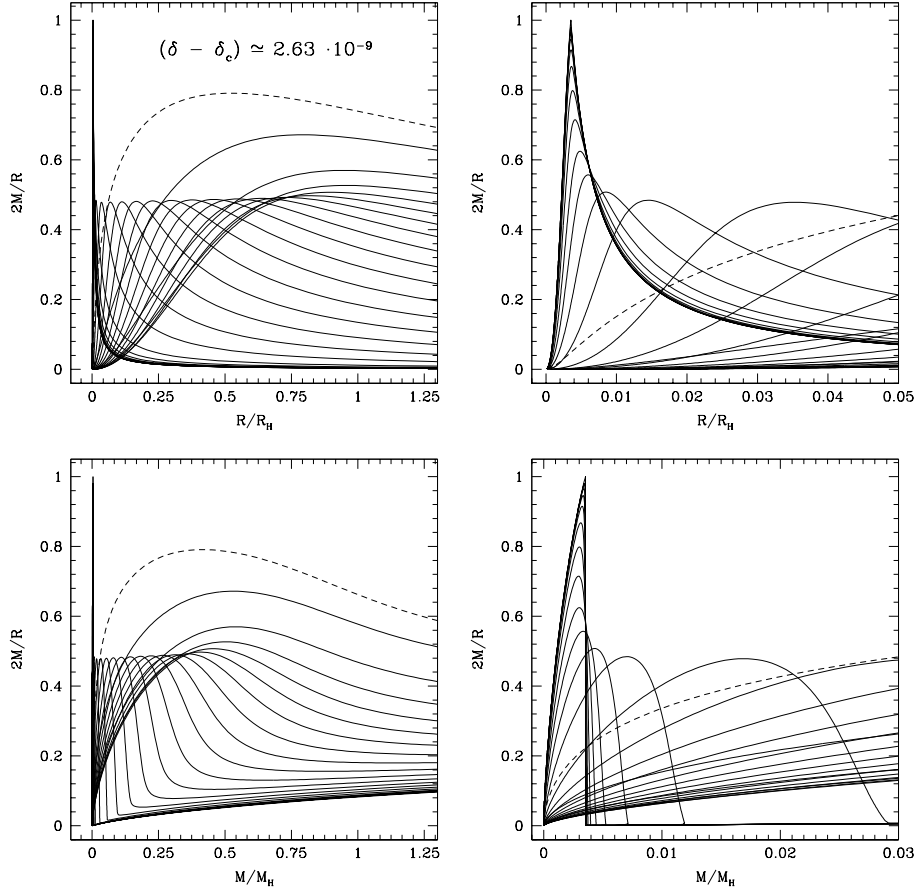


Figure 3: These plots, for a representative but fairly extreme case where a black hole is formed, show different views of the behaviour of $2M/R$ as a function of R/R_H or M/M_H at different time levels with the dashed curve representing the initial conditions used by the observer-time code. The perturbation has $(\delta - \delta_c) \simeq 2.63 \cdot 10^{-9}$ and the collapse gives rise to a black hole with a mass $M_{BH} \simeq 3.54 \cdot 10^{-3} M_H$. In the top two frames, $2M/R$ is plotted against R/R_H with the right-hand frame being an enlargement of the inner parts of the left-hand one. The bottom two frames show the same data plotted as a function of M/M_H .

For this discussion of super-critical collapse, we will focus on a particular representative case which is typical of those fairly close to the critical limit. This case has $(\delta - \delta_c) \simeq 2.63 \cdot 10^{-9}$ and forms a black hole with mass $M_{BH} \simeq 3.54 \cdot 10^{-3} M_H$. Figure 3 shows various views of the $2M/R$ profiles. In the top two frames, $2M/R$ is plotted against R/R_H with the right-hand frame being an enlargement of the inner parts of the left-hand one. The top left-hand frame is the counterpart

of the plots which we have shown for the critical and sub-critical cases and one can see the “intermediate state” with its almost-constant value of $(2M/R)_{peak}$ and the eventual precipitous rise of this towards 1 as the black hole forms. (Note that in observer time, the black hole only fully forms asymptotically, when the time as measured by a distant observer tends to infinity, but we terminated the calculation when the central value of the lapse f had fallen below 10^{-10} .) The top right-hand frame is an enlargement of the central part, showing how the solution diverges away from the intermediate state and enters the final black-hole collapse phase. The bottom two frames show the same data plotted as a function of M/M_H . This is useful for showing in a clear way that when the collapsing region reaches the intermediate state, it then evolves by losing material to the outside while maintaining almost the same compactness until the final precipitous collapse. Again, the right-hand frame is an enlargement of the central part of the left-hand one, showing more detail of the departure from the intermediate state.

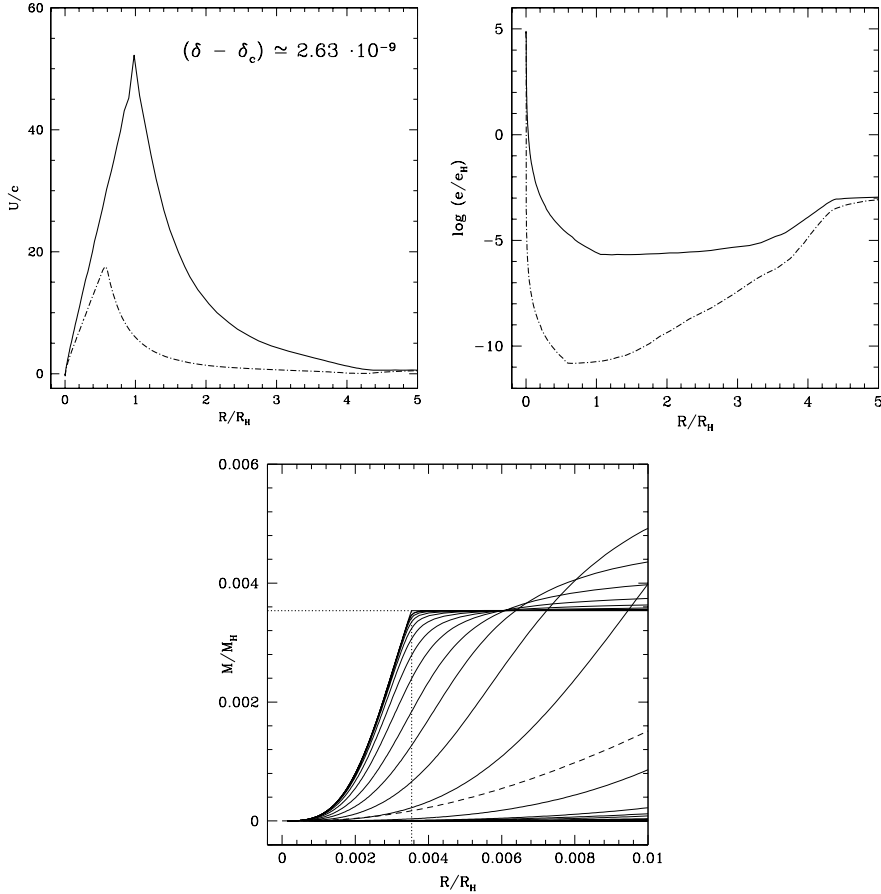


Figure 4: This figure shows further details for the same case as in figure 3. The top two frames show the profiles of radial four velocity U and energy density e at two key moments: the time when the wind away from the central region reaches its maximum strength (solid line) and the moment when the void starts to refill from the outside (dot-dashed line). Note that the black hole is at a very small scale on the left-hand side of these plots. The bottom frame shows the mass profiles at different time levels, with the dashed curve representing the initial conditions used by the observer-time code and the dotted lines marking the mass and radius of the eventual black hole.

As material is shed from the central condensation when it is in the intermediate state, this forms a strong relativistic wind (as mentioned earlier) which excavates a deep semi-void around it. This is shown in the top two frames of figure 4 where the profiles of radial four velocity and energy-density are plotted. Two times are shown. The first (marked by the solid curves) is when the wind reaches its maximum strength, here with the maximum of U/c being around 50 (for more extreme cases, we see values of more than 100). By this stage, a substantial semi-void has already formed around the central condensation and this then proceeds to deepen further as more material is blown outwards by the wind which is being accelerated by the steep pressure gradient at its inner edge. The second time shown (marked with the dot-dashed curves) is when the outward velocity has dropped to zero at the outer edge of the void and it then starts to refill. At this stage, the AMR scheme used for our simulations has reached 23 levels of refinement (a factor of almost 10^7). When we follow the subsequent refilling of these voids, we find that it proceeds very gently and takes place on a timescale which is long compared to the dynamical timescale of the final collapse producing the black hole but short in cosmological terms. We then expect that further accretion onto the central black hole would proceed in a standard way. The bottom frame of figure 4 shows the profiles for the mass at successive times, with the dashed curve again marking the initial data for the observer-time code. Initially, when all of the medium is expanding, $M(R)$ is, of course, decreasing at all values of R but later, when the collapse starts, $M(R)$ increases for the values of R inside the collapsing region, giving asymptotically the final mass of the black hole, evaluated where $(2M/R)_{peak}$ tends to 1 and the lapse f tends to 0 (this location is marked by the dotted lines). The effect of the void can be seen in the very flat profile of $M(R)$ outside this point (the last curve shown corresponds to the same time as the dot-dashed curves in the top two frames). Paper [13] also contains some discussion about the self-similar behaviour that we have observed in these simulations, reinforcing our feeling that the scaling law is likely to be preserved all the way down to very small values of $(\delta - \delta_c)$.

In the rest of this section we describe further the situation for subcritical perturbations. When δ is considerably smaller than δ_c , the perturbation initially grows but then subsides back into the surrounding medium in an uneventful way. However, for perturbations with δ sufficiently close to δ_c , some very interesting behaviour is seen and we present here some results from a representative case of this (more details can be found in [11]). Our calculations for subcritical perturbations use only the Misner-Sharp code and the run presented has $\delta - \delta_c = -3 \times 10^{-3}$.

In Figure 5, the fluid worldlines are plotted. Initially, the perturbation grows within the expanding fluid, following the quasi-homogeneous solution while the fluid perturbations are in the linear regime. The deceleration in the perturbed region is larger than that in the unperturbed region and its expansion lags progressively behind that of the outer matter until eventually it starts to re-contract shortly after horizon crossing. The contraction is not strong enough to produce a black hole and the fluid bounces out again, expanding until it encounters the surrounding matter which did not participate in the contraction. A compression wave forms where the two regions of fluid meet, while the density becomes very low at the centre of the perturbation. The compression wave proceeds out into the surrounding material but also some matter is sent back into the middle of the rarefaction where it undergoes a second bounce which is much more extreme than the first with a very abrupt change of velocity in the central regions. Whereas the outward moving compression

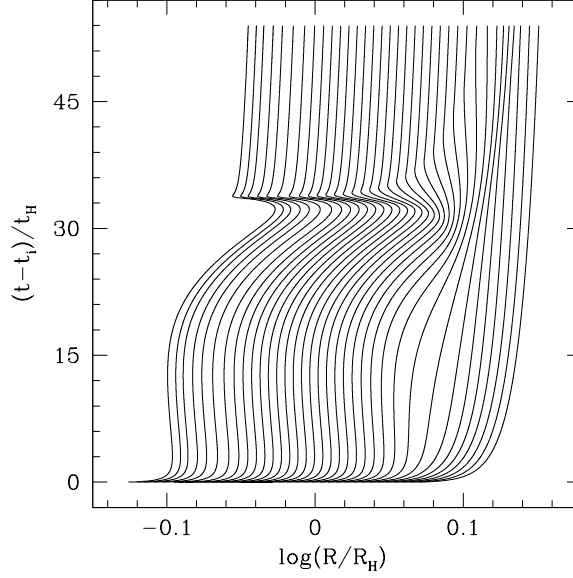


Figure 5: Worldlines for a Mexican-hat perturbation with $(\delta - \delta_c) = -3.0 \times 10^{-3}$. This plot shows alternating collapse and expansion of the perturbed region while the outer material continues to expand uniformly. The “cosmic” time is measured in units of the time at horizon crossing.

is damped geometrically as it proceeds to spherical surfaces with progressively larger areas, the inward-moving wave of material is geometrically amplified by the inverse process. The reason for the second collapse and bounce being more violent than the first is that while the first is a collapse of an overdensity which is resisted throughout by internal pressure, the second is essentially the collapse of a “shell” with near vacuum inside it and is hence close to free-fall until just before the bounce. The compression wave formed by the second bounce propagates out into the surrounding medium following the first one. Both proceed to damp geometrically and eventually the medium returns to a uniform state.

5. Conclusions

In this paper, we have presented results from our investigation of the extent to which primordial black hole formation in the radiative era of the early universe can be considered as a manifestation of the critical collapse phenomenon. For doing this we have made numerical simulations including both the collapse producing the black hole and the continuing expansion of the surrounding universe, focusing on the behaviour of initial perturbations of a type which could have come from inflation, having only a growing component and no decaying component. Implementing an AMR scheme within our Lagrangian code, we have been able to follow formation of black holes over a range of more than a thousand in mass. For the type of perturbation that we are studying, we find that scaling-law behaviour persists down to the smallest masses that we are able to follow, with no sign of a levelling-off such as had been reported in earlier work. We attribute this difference to the different type of initial conditions used; the present ones do not lead to the formation of shocks during collapses giving rise to black holes, which were a key feature of previous work. For

perturbations of the type coming from inflation, we have not seen any evidence to make us think that there would be a deviation away from the scaling law at lower masses until either there is a change in the matter model or the quantum regime is reached.

For slightly subcritical cases, we see surprisingly violent behaviour at the time of the second bounce which warrants further investigation.

Acknowledgements: I gratefully acknowledge first Prof. John Miller, Dr. Alexander Polnarev and Prof. Luciano Rezzolla, my collaborators on the papers of which this review is a synthesis. Then I would also like to gratefully acknowledge the following colleagues, with whom I have had helpful discussions during these years: (in alphabetical order) Carlo Baccigalupi, Marco Bruni, Bernard Carr, Carsten Gundlach, Ian Hawke, Karsten Jedamzik and Stefano Liberati.

References

- [1] Zel'dovich Ya.B. & Novikov I.D. 1966 *Astron.Zh.* **43** 758 [*Sov.Astron.* **10** 602 (1967)]
- [2] Hawking S.W. 1971 *MNRAS* **152** 75
- [3] Carr B.J. & Hawking S.W. 1974 *MNRAS* **168**, 399
- [4] Carr B.J. 1975 *Astrophys.J.* **201** 1
- [5] Nadezhin D.K., Novikov I.D. & Polnarev A. G. 1978 *Astron.Zh.* **55** 216 [*Sov.Astron.* **22(2)** 129 (1978)]
- [6] Bicknell G.V. & Henriksen R. N. 1979 *Astrophys.J.* **232** 670
- [7] Carr B.J. 2003 *Lect. Notes Phys.* **631** 301
- [8] Niemeyer J.C. & Jedamzik K. 1999 *Phys.Rev.D* **59** 124013
- [9] Shibata M. & Sasaki M. 1999 *Phys.Rev.D* **60** 084002
- [10] Hawke I. & Stewart J. M. 2002 *Class. Quantum Grav.* **19** 3687
- [11] Musco I., Miller J.C., Rezzolla L. 2005 *Class. Quantum Grav.*, **22**, 1405
- [12] Polnarev A.G. & Musco I. 2007 *Class. Quantum Grav.*, **24**, 1405
- [13] Musco I, Miller J.C. & Polnarev A.G. 2008 arXiv:0811.1452 [gr-qc]. To be published in *Class. and Quantum Grav.*
- [14] Ivanov P. 1998 *Phys.Rev.D* **57** 7145
- [15] Green A.M., Liddle A.R. Malik K.A., Sasaki M. 2004 *Phys.Rev.D*, **70**, 041502
- [16] Lifshits E.M. & Khalatnikov I.M. 1963 *Usp. Fiz. Nauk.* **80**, 391 [*Sov. Phys. Usp.* **6**, 496 (1964)]
- [17] Hidalgo C. & Polnarev A.G. 2008 arXiv: 0806.2752 [astro-ph]
- [18] Niemeyer J.C. & Jedamzik K. 1998 *Phys.Rev.Lett* **80** 5481
- [19] Choptuik M.W. 1993 *Phys. Rev. Lett.* **70** 9
- [20] Neilsen D.W. & Choptuik M.W. 2000 *Class. Quantum Grav.*, **17**, 761
- [21] Evans C.R. & Coleman J.S. 1994 *Phys. Rev. Lett.* **72** 1782
- [22] Misner C.W. & Sharp D.H. 1964 *Phys.Rev* **136** B571

- [23] Lyth D.H., Malik K.A., Sasaki M. 2005 *JCAP* **May** 004
- [24] Liddle A.R. & Lyth D.H. 2000 *Cosmological Inflation and Large-Scale Structure* Cambridge University Press
- [25] Padmanabhan T. 1993 *Structure Formation in the Universe* Cambridge University Press
- [26] Harada T. & Carr B.J. 2005 *Phys.Rev.D* **71** 104009
- [27] Hernandez W.C. & Misner C.W. 1966 *Astrophys.J.* **143** 452
- [28] Miller J.C. & Motta S. 1989 *Class. Quantum Grav.* **6** 185
- [29] Miller J.C. & Rezzolla L. 1995 *Phys.Rev.D* **51** 4017
- [30] Gundlach C. & Martín-García J.M. 2007 *Living Rev. Relativity* **5**
[<http://www.livingreviews.org/lrr-2007-5>]