

Article

Resolution of Initial Value Problems of Ordinary Differential Equations Systems

Josep Vicent Arnau i Córdoba ¹ and Màrius Josep Fullana i Alfonso ^{2,*} 

¹ Departament de Matemàtica Aplicada, Universitat de València, Avinguda Vicent Andrés Estellés, s/n, 46100 València, Spain; arnau@uv.es

² Institut Universitari de Matemàtica Multidisciplinària, Universitat Politècnica de València, Camí de Vera, s/n, 46022 València, Spain

* Correspondence: mfullana@mat.upv.es

Abstract: In this work, we present some techniques applicable to Initial Value Problems when solving a System of Ordinary Differential Equations (ODE). Such techniques should be used when applying adaptive step-size numerical methods. In our case, a Runge-Kutta-Fehlberg algorithm (RK45) has been employed, but the procedure presented here can also be applied to other adaptive methods, such as N-body problems, as AP3M or similar ones. By doing so, catastrophic cancellations were eliminated. A mathematical optimization was carried out by introducing the objective function in the ODE System (ODES). Resizing of local errors was also utilised in order to address the problem. This resize implies the use of certain variables to adjust the integration step while the other variables are used as parameters to determine the coefficients of the ODE system. This resize was executed by using the asymptotic solution of this system. The change of variables is necessary to guarantee the stability of the integration. Therefore, the linearization of the ODES is possible and can be used as a powerful control test. All these tools are applied to a physical problem. The example we present here is the effective numerical resolution of Lemaitre-Tolman-Bondi space-time solutions of Einstein Equations.



Citation: Arnau i Córdoba, J.V.; Fullana i Alfonso, M.J. Resolution of Initial Value Problems of Ordinary Differential Equations Systems.

Mathematics **2022**, *10*, 593.

<https://doi.org/10.3390/math10040593>

Academic Editor: Ioannis K. Argyros

Received: 18 December 2021

Accepted: 1 February 2022

Published: 14 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: numerical; initial value problem; ordinary differential equations systems; exact solutions of einstein equations

1. Introduction

In mathematics, a dynamical system is a system in which a function describes the time-dependence of a point in a geometrical space. Numerous examples can be found in the literature, applied to a wide variety of fields in science and technology. In order to make a prediction about the system's future behaviour, an analytical solution of such equations or their integration over time through computer simulations must be carried out. When modelling such a dynamical system to be able to investigate certain aspects of its behaviour, we end up generating a system of ordinary differential equations (ODES), partial differential equations (PDE), and algebraic equations (AE), whose analogous solution is unknown to us in the vast majority of cases. We need to resort to numerical methods for ODES which allow us to evaluate approximate solutions with a certain tolerance after fixing the indeterminate constants of the system, either by initial values (IVP or initial value problem) or boundary conditions (boundary Cauchy problem). In any case, we will find a large number of numerical methods at our disposal that can be divided into two large groups:

(i) The predictor-corrector type (also called Adams), either implicit or explicit. These are usually multistep methods which we can use when we want to increase the order of the method (that is, their precision), and usually require uniform pitch spacing $x_{i+1} = x_i + h$ (prior discretization of the interval $[x_0, x_{N_h}]$). These are more effective methods in the sense that they usually require fewer evaluations from the ODES, but it is clear that if the value in $x_0 + h$ depends on the previous values of k , x_{k-1}, \dots, x_0 , it will be necessary to generate

or impose more initial conditions, and it is not very clear whether we can consider the integration of the ODES as an IVP.

(ii) Runge–Kutta (RK)-type [1] are also explicit or implicit, but they are one-step and manage to increase their order by calculating in more intermediate points of the interval $[x_i, x_i + h]$. In this way, we can say that the value in $x_i + h$ depends only on the value in x_i , so we are talking about the IVP in each subinterval. This feature of the RK methods allows us to use a variable h (different in each step) and adapt them to the particular ODES to achieve the same estimated error order in each step. See the Ref. [2].

Consequently, our first choice was between an Adams or an RK. We opted for the RKF45 or explicit Runge-Kutta-Felberg45 [3], which is one of the most widely used among the adaptive step RKF45. However, as we intended to also address more realistic simulations with density defects compensated by a wall at a certain distance, we implemented the Cash–Karp variant [4], which allowed us to choose between an RKF45 of order $O(h^5)$ or one of variable order. In such a case, rapid changes in the gradient of some functions must be expected, and must therefore be taken into account so as to be prepared for such an eventuality.

Once our dynamical system is modelled, we must proceed to establish the objective function that models the particular aspect that we want to simulate for investigating its evolution and be in a position to predict future values (meteorology, evolution of the pandemic, volcanology, behaviour of a structure in the face of earthquakes, and so forth.). The fields of application are multiple, and the accuracy of the predictions depends on the accuracy of the model and the rigour of the tests required for the applied integration method.

Nowadays, there are applications of all kinds that include sophisticated *solve*, *odeint*, or similar commands, such as MAPLE [5], MATLAB [6], Mathematica [7] or Python [8], among others. These methods require writing our system in a certain way and assigning values to certain parameters that configure the method to be applied. After a while, it prints results and/or graphs which should be interpreted without clearly knowing what has been done, why the method has changed in certain cases, and so forth, and it is therefore difficult to make modifications to the integrating algorithm. At most, we can change our dynamic system or try to vary its parameters in a series of tests and errors that we hope will provide us with a more or less stable response. Then, we calculate means, variances, and other statistical properties which are published. Of course they make our work easier, but we must always keep in mind that *those applications or commands are quite generic*, and they have not been written for our specific problem. They will be very sophisticated, but it is very difficult for them to be optimal to give us control of the process. What is very fast is usually not very reliable. If we aim to take control over the evolution of our numerical simulation, some tools of Numerical Analysis must be taken into account whenever we face the resolution of a Dynamical System. Such tools are described in Section 2. When solving Dynamical Systems, the convergence and stability analysis of the problem must also be considered; see, for instance [9,10]. In the example we present here, such study was also undertaken.

In this paper, we solve the ODE (Initial Value Problem) associated with the gravitational effect of a pressureless, spherically symmetric perturbation on free photons described within the General Relativity Theory (GRT) and follow null geodesics. These null geodesics describe the Cosmic Microwave Background (CMB) Radiation. We numerically compute the anisotropies on CMB caused by a great inhomogeneity in the universe [11–18]. A detailed explanation of the main physical results and where they were published are presented in Section 8. Some important definitions related with our problem are presented in Appendix A.

Examples of modern physical problems where analytic solutions of differential equations are not possible, such as the problem we present in the present paper, is found in quantum systems. Many body quantum systems appear and have to be solved by applying advanced algorithms [19–21]. The ODEs are the Schroedinger and Dirac equations [20,22,23]. In such kinds of problems, the consideration of the finite size of nucleons

and nuclei needs the use of advantageous numerical techniques [20,21,23]. Nowadays, there exists an increasing interest in dealing with numerical solutions of wave functions for N-body problems using the field of machine learning methods [24–27].

Other kinds of problems which also deal with the development of numerical integration schemes for physical systems are based on the theory of variations. The integration methods, known as discrete variational integrators, which treat the discretization of the Hamilton's principle, can be found in the said theory. It is worth noting that the preservation of the geometric features of the dynamical system makes these methods appropriate for both conservative and nearly dissipative systems [28–31]. Systems that present highly oscillatory behaviour [28] are one of the most difficult challenges to solve ODEs. Here, standard numerical methods require a huge number of (time) steps to trace the oscillations, while geometric integrators are advantageous tools [32]. Geometric integrators face highly qualified simulations for long-term integration processes avoiding spurious effects [32–37]. It is interesting to derive numerical integration schemes that respect the geometric characteristics of such systems [38,39].

In a totally different field such as image processing, the Ref. [40] deals with the blind deconvolution of an image looking for the convolution kernel using fractional powers of the Laplacian (see e.g., [41,42] and references therein), which turns the method into a dynamic assimilable regularization system to our rescaled variables, although ODEs are replaced by successive FFTs (Fast Fourier Transform). Likewise, the Ref. [43] deals with the restoration of images by means of non-linear Gaussian approximations on the Luminance of the image.

The example we present here is the effective numerical resolution of Lemaître-Tolman-Bondi (LTB) space-time solutions of Einstein Equations (EE); see [44]. These techniques can also be used in other LTB applications, so knowing them is of great interest. In the near future, we plan to explain similar techniques that we have applied to N-body algorithms in the cosmological treatment of large-scale structure, such as in the Hydra project (see [45,46]). All in all, we also intend to share some of the important features of the algorithms we have designed in relativistic positioning systems (see [47]), which we are currently working on. All of these and similar techniques are powerful tools for other astrophysics, physics, and engineering issues in general. We find it appropriate to present them so that they may be used, where appropriate, in the improvement of numerical methods for solving systems of equations both in our field and in science and technology in general.

The paper is organized in the following sections. Section 2 describes some previous considerations about the techniques we apply. Section 3 introduces the main features of the physical problem that is solved with our numerical and computational techniques. Section 4 gives details of the EE solutions used in the cosmological problem solved. In Section 5, the numerical algorithm CS is described. Section 6 deals with other applications of the CS algorithm. Section 7 introduces the numerical resolution of the null geodesic equations that describe the CMB anisotropies in a LTB universe. In this section, our particular application of the RKF algorithm is also described. Some words about the convergence and stability of the algorithms are also presented in Section 7. Section 8 presents the main results and a brief discussion of them, in conjunction with other papers of the current literature. Section 9 has a summary of our paper and the main conclusions. This section includes: (i) a short summary of the paper's novelties and its scientific contribution, (ii) the conclusions extracted by the present study, and (iii) future extensions and outlook. At the end of the paper, we included two appendices which try to add more information about some of the aspects we consider relevant in our work.

2. Numerical Analysis Tools

In this section, we specify certain tools of Numerical Analysis that must be considered to solve a Dynamical System. These tools allow us to take control of the evolution of our numerical simulation. Sometimes, this process uses more CPU time, and in some cases,

also RAM memory. Nevertheless, following it, we have a guarantee that the solution will fit the problem to be solved.

The aspects that should be considered the most in these integrations are:

- (1) Working with computers that use finite precision mathematics, the first question to keep in mind is that in computational models, you have to try to avoid, whenever possible, the *subtraction* of functions that can have values of the same order. Such a subtraction produces a smaller result than that of the operands and it can lead to the loss of the most significant digits of the given value and produce the named *catastrophic cancellation*; see [48] (chapter 1). This operation is especially disastrous in numerical derivation processes as well as in recursive algorithms which are quite fashionable for their beauty and the simplicity of their implementation, as checked in the Ref. [48] (problem 2, chapter 1). It is also dangerous when a derivative (total or partial) is expected to change signs (detection of extremals in the corresponding function) because if it changes signs, it means that it has had to reach zero and lost all its significant digits. Therefore, we must always look for alternative formulas to those that appear in books and manuals (see Appendix B for the IEEE standards) for floating point computation, for intrinsic functions of programming languages, and for numerical derivatives.
- (2) The next step is to try to base our model on another with an analytical solution that is the limit of ours when a certain parameter tends to zero, for example (asymptotic model). If the asymptotic model has an analytical solution that provides us with certain AE to obtain values of some variables *near those of our model*, such values eventually serve to *rescale or renormalize* the variables of our model. This technique is very common in the N-body processes of physics (renormalization of sources), and even in the numerical solution of hyperbolic ODE and PDE which requires reconstruction of the axes at each step to avoid distortions if the coefficients vary with time.
- (3) At this point we need to verify that the functions we are going to use are evaluable with a precision greater than the tolerance we impose (standard IEEE). This analysis shows to what extent it is feasible to solve our model and which problems require modifications of the integrating algorithm.
- (4) A very important tool in simulations is to make variable changes, such as passing an affine parameter λ to the t variable that focuses more on knowing how much we have advanced and how much we still have to. We can also rely on the easy asymptotic solution to be calculated in order to define new variables that are quotients of the same variable in each system. These ratios have a very small range of variation, and we reduce the impact of noise (local error) on the objective function. Sometimes it is even possible to decouple some ODEs, as in the case discussed in the present paper.
- (5) We enter the phase of incorporating our *objective function* into the dynamical system that we have modelled through its analytical derivation (if it is an AE) so that it can evolve with the global model. If the values that this function take are very small compared to those obtained from the integration from other ODEs, we have to rescale the estimated values of the relative errors of each ODE *in each step* to avoid situations where the big ones dominate the small ones.
- (6) The decisive step is to modify the integrating method so that it only takes into account the relative error of the principal variables for the purpose of re-evaluating the integration step.
- (7) Extremal detection is also very important in order to modify the step. Usually, some of these extremals (zeros of the derivatives) indicate the areas of greatest effect and should not be ignored. In our RKF45ng algorithm, we indicate a way to achieve this, which has given us very good results in the past, although it is probably very improvable.
- (8) Finally, any modification we introduce must be intensively tested. The most accurate test is to get the algorithm of our nonlinear dynamical system to obtain the values of the asymptotic system when we make the corresponding parameter zero

(background test), at least within the precision of the machine used for the defined variables. If the final algorithm passes this test, we can be sure that it will pass any other reasonable test.

3. The Physical Problem to Solve

In the present paper, we show the numerical algorithm built to solve the equations that compute the gravitational effect of a pressureless spherically symmetric perturbation on free photons radiation. Those photons are described in the General Relativity Theory (GRT) and follow null geodesics. Moreover, they account for Cosmic Background Radiation (CBR). For some important definitions related with our problem, see Appendix A.

Now, the applied cosmological theory is presented. A spherically symmetric pressureless perturbation may be described in Cosmology using the following exact solutions of Einstein Equations:

- (a) Friedmann-Lemaître-Robertson-Walker solution (FLRW) which represents a homogeneous and isotropic space-time (with constant energy density ρ_B). This is the background Universe.
- (b) Lemaître-Tolman-Bondi solution (LTB) which accounts for a Universe with a unique great perturbation (an excess or a defect of energy density ρ over the mean) which is a pressureless spherically symmetric solution and which asymptotically tends to a FLRW metric (background). This is a parametric solution and requires numerical methods to determine the necessary functions in every space-time point. See ([44,49]).

If an observer O is placed in an arbitrary point of LTB, photons reaching from different directions will be affected in a different way by the perturbation. Given a direction ψ at t_0 , we call $T(\psi)$ to the photon temperature received by O evolving in LTB and $T_B(\psi)$ that is one received by the same observer but evolving in the asymptotic FLRW. We are interested in computing; see [50]

$$\delta_B = \frac{T(\psi) - T_B(\psi)}{T_B(\psi)}.$$

This temperature contrast is so far named the gravitational anisotropy produced by the perturbation, and it is supposed to have values of the order of 10^{-5} . Notice that the spherical symmetry and the absence of external fields allows us to ensure that all planes with $\varphi = constant$ are equivalent, and for this reason only, the case $\varphi = 0$ will be studied. This fact can be seen in Figure 1.

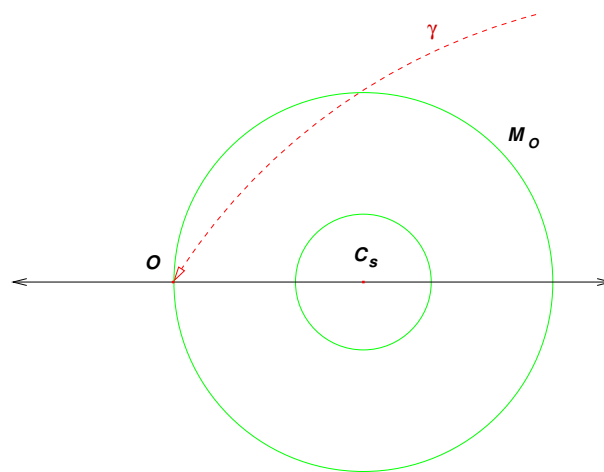


Figure 1. LTB plane with $t = constant$ and $\varphi = 0$. The slices M_O , observer, and another arbitrary slice are plotted, as well as the symmetry centre C_s . The discontinuous line γ represents the null geodesic of a photon arriving at the observer, O , at time t_0 .

The procedure to solve the problem is as follows:

1. Build a suitable algorithm (we will name it CS) to numerically solve the parametric LTB metric with sufficient precision.
2. Generate a second numerical algorithm (we will call it RKF45ng) that allows the integration of the null geodesics equations (or paths of free photons) in LTB with appropriate initial conditions.
3. Compute the gravitational anisotropy.
4. Design suitable controls and tests to check the consistency of the obtained values. In order to do that, the following data are considered:
 - 4.1. If the LTB perturbation, while photons are crossing it, stays with small density contrasts $\rho - \rho_B \leq 0.01\rho_B$, it is admissible to make a first-order analytical approximation, both in the Einstein Equations and in the null geodesics. Then one can integrate this linear analytical approximation. A **linear** code has been built for such a case. This code supplies a method to compare with the case when $\rho \neq \rho_B$.
 - 4.2. If one integrates the null geodesics of LTB generated with a null perturbation ($\rho = \rho_B$), one should obtain a FLRW Universe and, therefore, null values of the anisotropy. This is a good test to detect systematic errors and loss of precision. Moreover, it will be the main tool to guarantee that the RKFng is stable. (Background test).

That means our LTB null geodesics integration algorithm is **nonlinear** in the sense that no analytical approximations of the equations to integrate are done, although it should also be useful for the first studied linear case. Such a requirement should originate in additional problems in the error control and in the numerical stability of the system to integrate in these two assumptions: **(1)** When LTB is very close to the FLRW limit (points far away from the symmetry centre); and **(2)** When the null geodesic is very near to the symmetry centre. The main goal of our research is the detection and solution of such kinds of problems.

4. Solutions of EE Used

4.1. LTB Solution of EE

The spherically symmetric conditions imposed to the LTB universe, as well as the use of comoving coordinates and null pression, make most of the components of Einstein’s tensor null. Thus, their evolution equations are reduced to:

$$\mathcal{G}_{11} = 0; \quad \mathcal{G}_{14} = 0; \quad \mathcal{G}_{44} = -8\pi G\rho(M, t).$$

The $\mathcal{G}_{44} = -8\pi G\rho(M, t)$ equation becomes in a relation between M and $R(M, t)$ through $\rho(M, t)$. This relation may be expressed through one of the two following equivalent equations:

$$M = 4\pi \int_0^R x^2 \rho(x, t) dx, \quad \rho = \left(4\pi R^2 \frac{\partial R}{\partial M}\right)^{-1}, \tag{1}$$

and the second one allows us to obtain $\partial R/\partial M$; then, this partial derivative can be obtained with the same relative precision as that obtained in R and ρ .

The equations $\mathcal{G}_{11} = 0$ and $\mathcal{G}_{14} = 0$ are analytically integrable and lead to the parametric solution of LTB with both arbitrary functions $f(M)$ and $t_1(M)$ to be determined with the initial conditions. This solution is:

$$f > 0 \quad \text{non connected shell} \quad \left\{ \begin{array}{l} R = \left(\frac{GM}{f}\right) (\text{Ch } \eta - 1) \\ t + t_1(M) = \left(\frac{GM}{f^{3/2}}\right) (\text{Sh } \eta - \eta) \end{array} \right. \tag{2}$$

$$f < 0 \quad \text{connected shell} \quad \left\{ \begin{array}{l} R = \left(\frac{GM}{-f}\right)(1 - \cos \eta) \\ t + t_1(M) = \left(\frac{GM}{(-f)^{3/2}}\right)(\eta - \sin \eta) \end{array} \right. \quad (3)$$

$$f = 0 \quad \text{limiting shell} \quad R = \left(\frac{9GM}{2}\right)^{1/3} (t + t_1(M))^{2/3} \quad (4)$$

On the one hand, the spherically symmetric LTB metric stands for:

$$ds^2 = \frac{1}{\Phi^2} \left(\frac{\partial R}{\partial M}\right)^2 dM^2 + R^2(d\theta^2 + \sin^2 \theta d\varphi^2) - dt^2, \quad (5)$$

where:

$$\Phi^2 = 1 + f(M) = 1 + \left(\frac{\partial R}{\partial t}\right)^2 - \frac{2GM}{R}. \quad (6)$$

Function $f(M)$ is subjected to the restriction $f(M) > -1$ so that the metric has no problems. Furthermore, once $f(M)$ is known, $\partial R/\partial t$ can be computed. Then, this function establishes the variation profile of the radial coordinate with respect to t . On the other hand, $f(M)$ plays the role of the difference between kinetic and potential energies of the M shell. Its sign allows to distinguish the connected and non-connected shells of the gravitational field. As $f(M)$ does not depend on time, the connexion of the shell is determined by the initial conditions.

Once this equation has been solved, the next step is to compute the Christoffel symbols which appear in the null geodesics equations, and read:

$$\Gamma_{11}^4 = \frac{\partial R}{\partial M} \frac{\partial^2 R}{\partial t \partial M} \frac{1}{\Phi^2}; \quad \Gamma_{22}^4 = R \frac{\partial R}{\partial t}; \quad \Gamma_{24}^2 = \frac{1}{R} \frac{\partial R}{\partial M}; \quad (7)$$

$$\Gamma_{11}^1 = -\frac{1}{\Phi} \frac{\partial \Phi}{\partial M} + \frac{\partial^2 R}{\partial M^2} \frac{\partial M}{\partial R}; \quad \Gamma_{12}^2 = \frac{1}{R} \frac{\partial R}{\partial t}; \quad (8)$$

$$\Gamma_{14}^1 = \frac{\partial^2 R}{\partial t \partial M} \frac{\partial M}{\partial R}; \quad \Gamma_{22}^1 = -R\Phi^2 \frac{\partial M}{\partial R}. \quad (9)$$

Such functions will appear as coefficients of a coupled differential equations system when integrating the null geodesics. Therefore, maximum precision is required. As a matter of fact, the Γ_{11}^1 symbol is the only one which is the sum of two functions and can have precision problems in the case where it becomes very small.

4.2. FLRW Solution of EE

For a FLRW universe, one has:

$$ds^2 = \frac{a^2}{1 \pm r^2} dr^2 + a^2 r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) - dt^2, \quad (10)$$

and that is why if one takes:

$$\Phi^2 = 1 \pm r^2, \quad R = a(t)r, \quad M = 4\pi\rho_B R^3/3, \quad (11)$$

then FLRW becomes a LTB with constant $\rho = \rho_B$ for every t and r as a comoving radial coordinate (independent of t) and of $a(t)$ as the scale factor. The computations to complete, in this case, are evident. Note that these equations are AE, not ODE or PDE.

5. CS Algorithm

In order to complete the effective numerical computation of the parametric LTB solution, one has to perform the following steps:

5.1. Initial Functions

To determine the two arbitrary functions of the parametric solution, one has to proceed in the following way:

The function $t_1(M)$ usually becomes zero at t_D , although it could not be so.

To determine the function $f(M)$ through the density profile defined at t_D (see Equation (A1) at Appendix A), one has to proceed this way:

(i) **Computation of $R(M,t)$:** From Equation (1) and the density profile, one has

$$M = \frac{4\pi\rho R_v^3}{3} \left(\left[\frac{R}{R_v} \right]^3 + 3\varepsilon \left[\frac{R}{R_v} - \text{Arctg} \left(\frac{R}{R_v} \right) \right] \right). \tag{12}$$

This equation has a unique positive solution of $x = R/R_v$ for each $M \geq 0$. Once M is known, it can be written in the following way:

$$x^3 + 3\varepsilon(x - \text{Arctg } x) - d = 0, \quad d = \frac{3M}{4\pi\rho R_v^3}$$

and it can be applied to any method of roots computation as the solution is unique and simple. As in the FLRW, one has $M_B = 4\pi\rho_B R_B^3/3$, and $x_0 = R_B/R_v$ can be taken as an initial value for the Newton method. It still has to be determined whether the error tolerance is admissible for R . Therefore, a method to obtain the radial coordinate $R(M, t_D)$ associated to each shell is achieved with a prefixed relative error bound. That is, the profile $R(M, t_D)$ is obtained.

(ii) **Computation of $f(M)$:** By taking $z = (Rf)/(GM)$ in Equations (2) and (3), one has:

$$z > 0; \quad \left(\frac{GM}{R^3} \right)^{1/2} t_D = \frac{\sqrt{z^2 + 2z} - \text{ArgCh}(z + 1)}{z^{3/2}}, \tag{13}$$

$$z < 0; \quad \left(\frac{GM}{R^3} \right)^{1/2} t_D = \frac{\text{Arccos}(z + 1) - \sqrt{-z(z + 2)}}{(-z)^{3/2}}. \tag{14}$$

If $z \neq 0$, one of these equations has a solution (see Figure 2), from which z is obtained through a roots calculus and, afterwards, $f(M)$ with the same precision when numerically solved for an arbitrary M and its corresponding $R(M, t_D)$. If no equations have a solution, that means that in the shell, the following is accomplished:

$$R = \left(\frac{9GM}{2} \right)^{1/3} (t_D)^{2/3}$$

and then, $f(M) = 0$ (limiting shell).

5.2. The Other Functions at t_D

In order to determine the rest of the necessary functions, one should proceed this way:

1. As ρ is known at t_D (initial profile), in the Equation (1), one can directly compute the value

$$\frac{\partial R}{\partial M} = \frac{1}{4\pi\rho R^2}$$

with the same relative error bound as R .

2. Furthermore, once $f(M)$ and $R(M, t_D)$ are known, one can also obtain

$$\left| \frac{\partial R}{\partial t} \right| = \sqrt{f(M) + \frac{2GM}{R}}$$

with the same relative error bound. So as to know the sign of this derivative, one needs to know whether, at a later time, $R(M, t)$ increases or decreases for the same

shell. The advantage of this procedure is that the derivative value is precise and no significant digits are lost (see Appendix B).

3. The other functions appearing in the Christoffel symbols such as $\partial^2 R / \partial t \partial M$, $\partial^2 R / \partial M^2$ and $\partial \Phi / \partial M$ need derivation numerical techniques (a $t = \text{constant}$) of the known functions $\partial R / \partial t$, $\partial R / \partial M$ and $\Phi = 1 + f(M)$, although one cannot now guarantee a fixed relative error bound, as the *catastrophic cancellation* phenomenon can appear through the loss of significant digits with the subtraction of similar numbers. The logarithmic derivative is crucial here.

We tried to choose ΔM so that the possible maximum between the truncation and the rounding errors was the smaller one (see Appendix B).

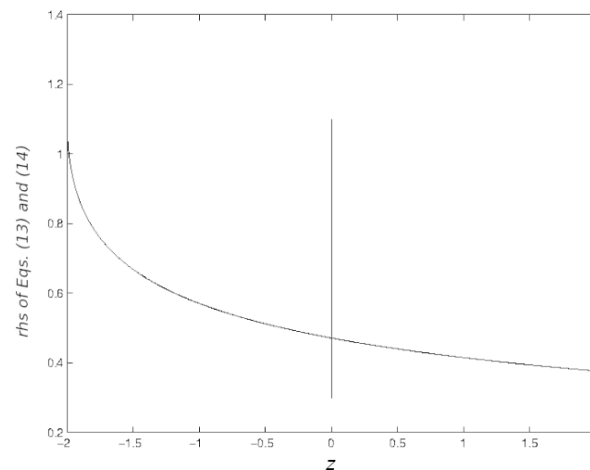


Figure 2. Plot of the right-hand side of Equations (13) and (14) for $z \in [-2, 2]$. The plot of negative values of z corresponds to Equation (14) and the plot of positive values of z corresponds to Equation (13).

5.3. Functions for $t > t_D$

As $f(M)$ is now known, given a value of M , one can directly choose among the Equations (2)–(4) to determine which is suitable for this shell. By substituting t_D by the desired t and, with the same change $z = (Rf)/(GM)$, one has:

$$f > 0; \quad \frac{f^{3/2}}{GM} t = \sqrt{z^2 + 2z} - \text{ArgCh}(z + 1), \tag{15}$$

$$-1 < f < 0; \quad \frac{(-f)^{3/2}}{GM} t = \text{Arccos}(z + 1) - \sqrt{-z(z + 2)}, \tag{16}$$

$$f = 0; \quad R = \left(\frac{9GM}{2}\right)^{1/3} (t)^{2/3}.$$

Let us now compute the approximated root for z (if necessary) and let us obtain the corresponding $R(M, t)$. From now on, easy computations allow us to obtain the rest of the necessary functions at each point (M, t) .

For example, for known $R(M, t)$ and $f(M)$, Equation (6) gives the value of $|\partial R / \partial t|$, and a numerical derivation at a constant t value gives $\partial R / \partial M$. Then Equation (1) gives the value of $\rho(M, t)$. The rest of the procedure is as in the former case.

Once all the commented functions are obtained, **one can compute the Christoffel symbols at any shell M and at any time t .**

6. Additional Utility of the CS Algorithm

A variation of this algorithm allows to obtain, with different FLRW background universes, the ϵ and R_v parameters at t_D , and this lets us “simulate” the present profiles of

certain realistic structures, such as the Boötes Void, the Great Attractor, the Virgo Cluster, and so forth. To do so, one proceeds as follows:

1. Choose an $\Omega \in [0.2, 1[$ and, together with the critical density value $\rho_c(t_0)$, determine $\rho_B(t_0) = \Omega\rho_c(t_0)$ for the asymptotic FLRW.
2. With the core “radius” $R_v(t_0)$, choose $R_{v1}(t_D) \approx a(t_D)R_v(t_0)/a(t_0)$ (application of the scale factor). Additionally choose a value of the contrast for t_D ε_1 and compute the $\rho_B(t_D)$ value of the FLRW.
3. Simplify the CS algorithm to obtain only the $R(M, t)$ and $\rho(M, t)$ values. Apply it only at t_D and t_0 for a predetermined sample of shells M_i and obtain $R(M_i, t_0)$ and $\rho(M_i, t_0)$.
4. Compute $\rho_0 = \rho(0, t_0)$ by extrapolation and $\widehat{\varepsilon}(t_0) = (\rho_0 - \rho_B)/\rho_B$ (maximum present density contrast).
5. Additionally compute \widehat{R}_v or the present core radius by interpolation, first searching the value of the core shell, M_v , such that $\rho(M_v, t_0) - \rho_B = \rho_B\widehat{\varepsilon}(t_0)/2$ and the corresponding value of $R(M_v, t_0)$.
6. In this way, a function of two variables has been defined:

$$[\widehat{\varepsilon}(t_0), \widehat{R}_v] = \text{PARAM}(\varepsilon_1, R_{v1})$$

evaluable for each pair of initial data.

7. Now the numerical secant method for systems of equations allows to perform iterations, starting with two initial values (ε_1, R_{v1}) y (ε_2, R_{v2}) , until it is able to find the desired present maximum contrast and core radius (in the case where it converges, of course).

This process is able to simulate perturbations with positive contrast (overdensities), with the same present linear contrast (as shown in Figure 3) and for universes with different values of Ω . Initial parameters of each overdensity are shown in Table 1b. For these cases, high precision is not required in the PARAM algorithm as the present contrast is obtained from observational low-precision data. The initial R_v are all of the order of $180/51 \text{ h}^{-1}$ Mpc, and that is the reason why they are not written in Table 1b.

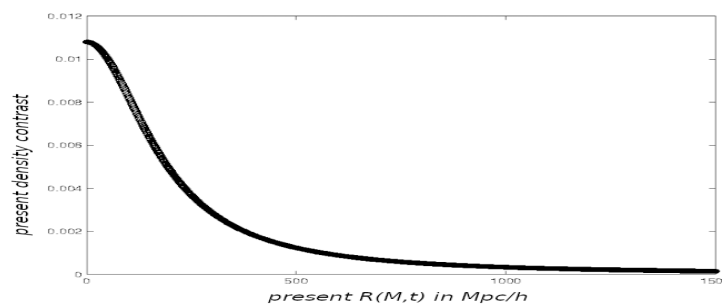


Figure 3. Present linear density contrast as a function of R , for the perturbations O_1 – O_7 , with $R_v(t_0) = 180 \text{ h}^{-1} \text{ Mpc}$ and $\varepsilon(t_0) = 0.0108$. The initial values at t_D that generate each overdensity appear in Table 1b.

So as to generate initial values that lead to nonlinear negative present contrasts (underdensities), a negative present profile was chosen (see Figure 4) and the data found are written in Table 1a. The relation $R_{v1} \approx 22.5/51 \text{ h}^{-1} \text{ Mpc}$ is accomplished here as well, but is not included in Table 1a, as it is the same one for all underdensities.

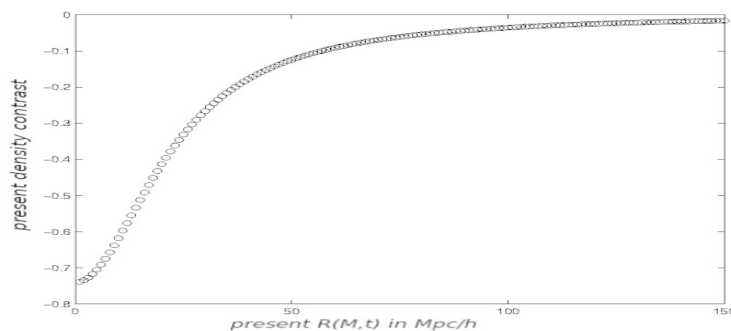


Figure 4. Present nonlinear negative density contrast as a function of R , for the perturbations $U_1 - U_{10}$, with $R_v(t_0) = 22.5 \text{ h}^{-1} \text{ Mpc}$ and $\epsilon(t_0) = -0.74$. The initial values at t_D that generate each underdensity appear in Table 1a.

Table 1. Initial parameters, for universes with different Ω , with the same present energy density contrast, δ .

| (a) Nonlinear Underdensities | | |
|------------------------------|----------|------------------------|
| Case | Ω | $\epsilon(t_D)$ |
| U_1 | 0.2 | -1.10×10^{-1} |
| U_2 | 0.3 | -8.42×10^{-2} |
| U_3 | 0.4 | -7.00×10^{-2} |
| U_4 | 0.5 | -6.09×10^{-2} |
| U_5 | 0.6 | -5.45×10^{-2} |
| U_6 | 0.7 | -4.97×10^{-2} |
| U_7 | 0.8 | -4.60×10^{-2} |
| U_8 | 0.9 | -4.30×10^{-2} |
| U_9 | 0.95 | -4.18×10^{-2} |
| U_{10} | 0.999 | -4.06×10^{-2} |
| (b) Overdensities | | |
| Case | Ω | $\epsilon(t_D)$ |
| O_1 | 0.2 | 5.985×10^{-4} |
| O_2 | 0.3 | 4.496×10^{-4} |
| O_3 | 0.4 | 3.700×10^{-4} |
| O_4 | 0.5 | 3.200×10^{-4} |
| O_5 | 0.6 | 2.870×10^{-4} |
| O_6 | 0.7 | 2.550×10^{-4} |
| O_7 | 0.8 | 2.400×10^{-4} |

The data of Table 1b have been used to generate overdensities. The effects of these overdensities in the CMB anisotropies have been used to compare the results of our nonlinear algorithm with the linear algorithm stated before. Meanwhile, the underdensities created with the data of Table 1a have been used to test the nonlinear results.

7. Resolution of the Null Geodesics EDOS

From Equation (5) and the fact that $\varphi = \text{constant}$ for spherical symmetry reasons, along any null geodesic, one has: $ds^2 = 0 = g_{11}dM^2 + g_{22}d\theta^2 - dt^2$, where

$$g_{11} = \frac{1}{\Phi^2} \left(\frac{\partial R}{\partial M} \right)^2, \quad g_{22} = R^2.$$

On the other hand, if one defines λ as the affine parameter of the geodesic and the tangent components in a point of the geodesic as (K^1, K^2, K^3, K^4) , the following relations stand:

$$K^1 = \frac{dM}{d\lambda}, \quad K^2 = \frac{d\theta}{d\lambda}, \quad K^3 = 0, \quad K^4 = \frac{dt}{d\lambda}.$$

Finally, the free photons evolution in GRT must accomplish the following condition $dK^\alpha/d\lambda = -\Gamma_{\beta\gamma}^\alpha K^\beta K^\gamma$. In LTB, that means the resolution of the following system:

$$\text{System A: } \begin{cases} K^1 = dM/d\lambda, & K^2 = d\theta/d\lambda, & K^4 = dt/d\lambda \\ dK^1/d\lambda = -2\Gamma_{14}^1 K^1 K^4 - \Gamma_{11}^1 K^1 K^1 - \Gamma_{22}^1 K^2 K^2 \\ dK^2/d\lambda = -2K^2(\Gamma_{24}^2 K^4 + \Gamma_{12}^2 K^1) \\ dK^4/d\lambda = -\Gamma_{11}^4 K^1 K^1 - \Gamma_{22}^4 K^2 K^2 \end{cases} \quad (17)$$

Control Eq.: $(K^4)^2 = g_{11}(K^1)^2 + g_{22}(K^2)^2$
 Objective: $\delta_B = a(t_0)K_0^4/a(t_D)K_D^4 - 1$

where the Christoffel symbols we must compute with the CS algorithm appear. The independent variable is the affine parameter λ and the dependent variables are $M, \theta, t, K^1, K^2, K^4$.

These equations form a system of six coupled differential first-order equations that have to be integrated with the suitable initial conditions. As when one gives conditions in t_D , one has many problems to ensure that the photon arrives at the observer at t_0 , and the best procedure is to do it in the inverse way. It is supposed that the observer is looking in the direction that forms an angle Ψ_0 with the OC_s axis, and then the initial conditions should be:

$$\lambda_0 = 0, \quad K_0^1 = -\frac{\cos \Psi_0}{\sqrt{g_{11}(M_O, t_0)}}, \quad K_0^2 = -\frac{\sin \Psi_0}{\sqrt{g_{22}(M_O, t_0)}}, \quad K_0^4 = -1,$$

where the 0 subscript stands for the initial value while M_O is the M coordinate of the observer shell.

Let us also suppose that CMB is homogeneous in the hypersurface $t = t_D$ and that the temperature of all the photons that start from that hypersurface is constant. In these conditions, it is well-known that the gravitational anisotropy to be computed verifies:

$$\delta_B = \frac{a(t_0)K_0^4}{a(t_D)K_D^4} - 1, \quad (18)$$

where $a(t)$ is the scale factor.

Theoretically, one only has to compute K_D^4 by integrating (backwards in time) the system of equations. In any way, it can be supposed that problems with the stability of the system will appear in the case where you need to obtain δ_B of the order of 10^{-5} with Equation (18), which should be the precision required in K_D^4 to obtain at least seven correct decimals in the quotient and to be able to guarantee that δ_B has at least three correct significant digits. If what one tries to do is to obtain quasi null anisotropy (FLRW limit), the precision requirements should be even greater.

As it can be noticed, the problem is not only the election of an appropriate integration method. One will also need for **the integration step to also adapt** to a precision that one elects, as for great values of M , LTB tends to FLRW. A very small step is only valid to increase the global round error and the CPU cost. Moreover, in these conditions, the gravitational anisotropy will change very little. (b) For small values of M , the density contrast increases in one direction and decreases in the opposite one, so one has to be very precise in the chosen step, as in this case, the greater changes of the gravitational anisotropy are present.

We have chosen a Runge-Kutta-Fehlberg of order 4(5) with a variable step requiring, at first, a local truncation relative error of the order of 10^{-10} in every variable. Of course, this forces us to demand that the roots x and z of the equations solved with the CS algorithm also have the same relative precision, although the functions determined through numerical derivation may not have it.

7.1. Numerical Instability of the Initial System

When system A was tried to be integrated for $\epsilon = 0$ (FLRW test) with the VAX (From the IFIC, of València) computer, it was found that, in some directions Ψ_0 , this integration stopped without reaching t_D because it found negative values of M , despite the quadruple precision required to minimize the rounding errors. This absurd result indicated uncontrolled oscillations in the system coefficients (numerical instability). When $\Gamma_{\beta\gamma}^\alpha$ implied values where plotted **amplified oscillations** were detected in the Γ_{11}^1 , as it was predictable for, it was the only term with sum/subtraction of functions.

7.2. First Numerically Stable System

Problem 1. In order to avoid the problems with the Γ_{11}^1 term, it must not appear in the system to integrate.

Solution 1: Substitute the equation where the Γ_{11}^1 term appears by the control equation. At the same time, the independent variable is also changed considering now the t variable instead of the affine parameter λ . In this way, a greater control is achieved on the time that the integration must be stopped.

Once these changes are made, the new system to solve is the following one:

$$\text{System B: } \begin{cases} K^1/K^4 = dM/dt, & K^2/K^4 = d\theta/dt \\ 1 = g_{11}(dM/dt)^2 + g_{22}(d\theta/dt)^2 \\ dK^2/dt = -2(K^2/K^4)(\Gamma_{24}^2K^4 + \Gamma_{12}^2K^1) \\ dK^4/dt = (-\Gamma_{11}^4K^1K^1 - \Gamma_{22}^4K^2K^2)/K^4 \end{cases} \quad (19)$$

Control Eq.: $K^4 dK^1/dt = -2\Gamma_{14}^1K^1K^4 - \Gamma_{11}^1K^1K^1 - \Gamma_{22}^1K^2K^2$

Objective: $\delta_B = a(t_0)K_0^4/a(t_D)K_D^4 - 1$

where the $d\lambda = dt/K^4$ equation is unnecessary because the value of this parameter is not interesting and the equation eliminated is now used as the control equation with the purpose of controlling the Γ_{11}^1 oscillations (by comparing the obtained value when clearing the unknown in this equation, with the value obtained in the CS algorithm).

The integration of System B with $\epsilon = 0$ (FLRW test) allows to arrive at t_D if $|\Psi_0| \geq 0.1$ grades. Then, the numerical instability of the system has disappeared, and the diagnosis of Problem 1 has been confirmed. Nevertheless, the final values of δ_B are of the order of 10^{-5} , depending on the initial value of Ψ_0 , although they should be close to 0. This fact indicates that, either (1) there are systematic errors in the integration or (2) the step fit of the integration is not suitable. As a matter of fact, both cases appear when the observation angle Ψ_0 is small.

In the case that Ψ_0 is small, the trajectory of the photon which comes from t_D is each time nearer to C_s , and therefore travels through shells with decreasing M until it starts to go far away from C_s . This indicates that dM/dt changes its sign at some point of the photon trajectory defining a minimum at $M < M_O$. This minimum should be detected by the integration method or else systematic errors might appear. Moreover, the objective function does not help in the adjustment of the integration step, and this is dangerous as if significant digits should be lost, then:

7.3. Inclusion of the Objective Function

In the previous section, (Section 7.2), some troubles have been presented. They should be summarized this way:

Problem 2. (2a) The possible change in the sign of dM/dt must be detected.
 (2b) The objective function must be included in the system of equations.

Solution 2.

(2a) The second line of system B suggests that dM/dt could be cleared as a function of $d\theta/dt$ or vice versa, but then the sign will not be well-defined. This equation is of the type $1 = \cos^2 \alpha + \sin^2 \alpha$ as both g_{11} and g_{22} are positive, and therefore, an angle α must exist, such as one that can define these derivatives through trigonometric functions.

As a matter of fact, let $(K^1, K^2, 0, 0)$ be the spatial projection of the tangent vector to the null geodesic and $(\sqrt{g_{11}}(M, t), 0, 0, 0)$ be the radial direction which joins C_s with the photon. The cosine of the angle formed by these two vectors is given by the quotient of the dot product of both vectors and their norms, that is:

$$\cos(\alpha) = \frac{K^1 \sqrt{g_{11}}}{\sqrt{K^1 K^1 g_{11} + K^2 K^2 g_{22}}},$$

which when using the first and second equations of system B, the following is obtained:

$$\frac{dM}{dt} = \frac{\cos \alpha}{\sqrt{g_{11}}}, \quad \frac{d\theta}{dt} = \frac{\sin \alpha}{\sqrt{g_{22}}}$$

and allows for rewriting of the other equations in terms of α . Now the sign of dM/dt is perfectly defined, and this derivative changes its sign when $\alpha = \pi/2 \pm n\pi$ with n as an integer.

On the other hand, the initial value of α at t_0 is, evidently, Ψ_0 (see Figure 5), and the following can be written:

$$\frac{d\alpha}{dt} = \sin \alpha \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos \alpha - \frac{\Gamma_{12}^2}{\sqrt{g_{11}}} \right].$$

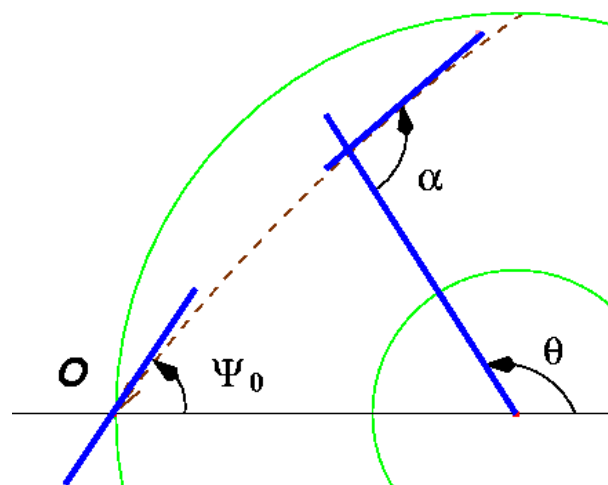


Figure 5. Plane of the LTB with $t = constant$ and $\varphi = 0$. The angle $\alpha(t)$ and the observation direction Ψ_0 are plotted.

Now the integration from t_0 to t_D is done in two parts when $\cos \alpha$ changes its sign, both of them with constant signs. In this way, it is definite that no “jumps” in the extremal

point are produced because this is in a great step-interval of integration. This modifies the RKF45ng in the following way (if $|\Psi_0| < \pi/2$): (i) While $|\alpha| \leq \pi/2$ both in t and in $t - h$, the RKF45ng controls the change of the step. (ii) The first time that a change of sign between t and $t - h$ appears, the increase of the step h is forbidden until indicated, the values in $t - h$ are not accepted, and the step is divided by 10. The integration follows with such a restriction and every time that $|\alpha| > \pi/2$ is overcome, the new value is not accepted and a division by 10 is made in h . (iii) When $|h| < tol_h$ or $\pi/2 - |\alpha| < tol_\alpha$, the control of the step is returned to the RKF45ng so that the integration is completed by increasing or decreasing the step.

As our goal is to be near the extremal point with a desired precision and it is also known that the extremal point is unique (if it exists), no more changes are needed, except by the obvious ones that must be done when the integration is from t_D to t_0 in the double integration test proposed further on.

(2b) With the aim that the integration algorithm also controls the error of the objective function, the auxiliary function is defined as follows:

$$\delta(t) = \frac{a(t_0)K_0^4}{a(t)K^4(t)} - 1,$$

that verifies the relation $\delta_B = \delta(t_D)$. The total derivative with respect to the time of this equation,

$$\frac{d\delta}{dt} = (1 + \delta) \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos^2 \alpha + \Gamma_{24}^2 - \frac{1}{a} \frac{da}{dt} \right],$$

gives a new differential equation for δ with initial value $\delta(t_0) = 0$ and that leads to the following system:

$$\text{System C: } \begin{cases} \frac{dM}{dt} = \frac{\cos \alpha}{\sqrt{g_{11}}} \\ \frac{d\delta}{dt} = (\delta + 1) \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos^2 \alpha + \Gamma_{24}^2 - \frac{1}{a} \frac{da}{dt} \right] \\ \frac{d\alpha}{dt} = \sin \alpha \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos \alpha - \frac{\Gamma_{12}^2}{\sqrt{g_{11}}} \right] \end{cases}$$

with the variables α , δ and M decoupled from the other ones, although the integration of system C with $\varepsilon = 0$ does not substantially improve the results from System B.

Our conclusion is that the true problem with this test is in the different value ranges that the variables and their total derivatives with respect to t have. Thus one has that: (i) M has a very great range of positive values that depends on the observation direction Ψ_0 , but dM/dt has a constant sign in some directions and changes in other ones. (ii) α always has values in $]0, \pi]$ (if the observation direction is positive) and has monotonous behaviour when travelling along the null geodesic. (iii) δ achieves a maximum (As M achieves a minimum) whether $|\Psi_0| \bmod \pi$ is small but, in any case, its variation range is very small.

7.4. Rescaling of the Variables. RKF45ng Algorithm

Therefore, **rescaling the truncation error in each variable** or the integration method is necessary, which will adjust the step according to the truncation error of the dominant variable.

The question here is that we are building a general algorithm, and there is no “perfect rescaling” that allows us to consider all the possible combinations of LTB, linear or not, with success guarantees. Moreover, the scaling of errors will depend on the observation direction, and this cannot be achieved with the classical method of defining a “relative weights” vector for the variables.

The only way to make all the integration variables have a similar truncation error is to make a change in the variables M and α to another y_1 and y_2 , that represents small changes in their values and their derivatives values. Such a change of variable must be considered

as variable rescaling, at every integration step, that helps the RKF45ng to control the local truncation error according to the physical conditions of the problem.

Given the Ψ_0 direction and the null geodesic of the LTB γ that arrives to O in this direction, let γ_B be a fictitious null geodesic of FLRW that arrives to O in the same direction. (a) Associate point of both geodesics with the same coordinate t . The point $(M, \alpha, t) \in \gamma$ is made corresponding with the point $(M_B, \alpha_B, t) \in \gamma_B$. (b) Use M_B and α_B to define the integration variables, $y_1 = M_B/M, y_2 = \alpha_B/\alpha$. Now, the variables y_1, y_2, δ will be the ones that control the step and the two first variables have the range reduced in the neighbourhood of 1.

In terms of the variables δ, y_1 and y_2 , the system to solve is formed by the following equations:

$$\text{Final System: } \left\{ \begin{array}{l} \frac{d\delta}{dt} = (1 + \delta) \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos^2 \alpha + \Gamma_{24}^2 - \frac{1}{a} \frac{da}{dt} \right], \\ \frac{dy_1}{dt} = \frac{1}{M} \left(\frac{dM_B}{dt} - y_1 \frac{dM}{dt} \right), \\ \frac{dy_2}{dt} = \frac{1}{\alpha} \left(\frac{d\alpha_B}{dt} - y_2 \frac{d\alpha}{dt} \right), \\ \frac{d\alpha}{dt} = \sin \alpha \left[(\Gamma_{14}^1 - \Gamma_{24}^2) \cos \alpha - \frac{\Gamma_{12}^2}{\sqrt{g_{11}}} \right], \\ \frac{d\alpha_B}{dt} = \sin \alpha_B \left(\frac{\sqrt{1+r^2}}{ar} \right), \\ \frac{dM_B}{dt} = -3M_B \left(\frac{d\alpha_B}{dt} \right) \text{ctg } \alpha_B, \end{array} \right.$$

with $M_B = 4\pi\rho_B a^3 r^3 / 3$, being $M_B = M, \alpha_B = \alpha$ at t_0 . Now the two last equations allow us to obtain some acceptable values of M_B and α_B , based in the FLRW background, to rescale the variables M and α at y_1 and y_2 .

If the order of the variables chosen is: $\delta, y_1, y_2, \alpha, \alpha_B, M_B$, the rescaling vector can be defined as $scale = (fac_\delta, 1, 1, 0, 0, 0)$ that multiplies the truncation error of each variable. This definition implies that the RKF45ng **will not consider** the variables α, α_B, M_B to adjust the step. For y_1 and y_2 will simply consider their truncation error, while for δ it will consider its truncation error multiplied times fac_δ . Such an election first allows the integration with $fac_\delta = 1$ and, if the results are not satisfactory with the new suitable tests, fac_δ may progressively be increased and force the RKF45ng to make smaller steps.

More concisely, define: (a) A $\Delta_i = (y_{n+1})_i - (y_{n+1}^*)_i$ for each integration variable (see [4]). (b) Now the vector $reesc_i = \Delta_i * scale_i$. (c) Finally, $\Delta = \|reesc\|_1$ (the maximum modulus component). Then the RKF45ng is now prepared so that it only considers the truncation errors of the first three variables to adjust the step.

Theoretically, there is no limit for fac_δ . Thus, with $fac_\delta = 10^5$, the condition $\Delta \approx 10^{-10}$ reads $10^5 * \Delta_1 \leq 10^{-10}$. The local truncation error in δ is required to be of the order of 10^{-15} ! and the relative error in the final step of the order of 10^{-12} in 1000 steps. This must be avoided as the greater fac_δ is, the greater the CPU cost is, more integration steps are done, and the global rounding error increases.

The variables y_1, y_2 should oscillate around the unity, so that the $fac_\delta = 1$ and the condition of absolute error of local truncation is $\approx 10^{-10}$, and leads to us asking whether 9 or 10 are correct decimal digits in y_1, y_2 and δ . If δ reaches values of the order of 10^{-5} , at least two or three correct digits are expected in 1000 integration steps. Such precision is acceptable to determine the order of the gravitational anisotropy, and it makes no sense to demand more precision given the CPU cost and the problem conditions.

7.5. Convergence and Stability. FLRW Test for RKF45ng with $\epsilon = 0$

Once the final system is integrated with $\epsilon = 0, tol_r = 10^{-10}, fac_\delta = 1$, with local truncation of the order of tol_r and quadruple precision in the VAX computer of the IFIC:

(a) The RKF45ng arrives at t_D if $|\Psi_0| \geq 0.001$ grades in a number of steps between 800 and 1200, according the direction. (b) The $|\delta_B|$ values obtained are in the interval $[10^{-26}, 10^{-24}] \leq tol r^2 \times 10^{-4}$. As some of the computed functions by derivation should have a relative error greater than the local truncation error, it must be concluded that **the final system is numerically stable** for the FLRW limit.

7.6. Convergence and Stability. New Tests for the RKF45ng When $\epsilon \neq 0$

Once the stability of the system for the FLRW limit has been checked, the guarantee that the final system is working properly for great values of M is achieved, and therefore, one can be confident that the process of integration does not “contaminate” the δ variable at t_D . A check of the final results obtained for LTB is lacking. In order to do this, the following test was designed.

7.6.1. Double-Integration Test

Let us call *rkf45ng* the algorithm of integration of the Final System of the null geodesics. If *valini* stands for the initial set of values and *valfin* for the final set of values, with the notation $[valfin] = rkf45ng(valini, \pm 1)$, where the + sign indicates integration from t_D to t_0 and vice versa. If:

1. $[valfin] = rkf45ng(valini, -1)$,
2. $[valini1] = rkf45ng(valfin, +1)$,

the difference $|valini - valini1|$ allows to “estimate” the recovered digits of each initial variable.

7.6.2. Direction Test

Integrate the final system with Ψ_0 , where $-\Psi_0$ and $\Psi_0 \pm 2\pi$ allows us to detect the possible disarrangements in the rescaling vector when the modulus of α or its sign are changed.

7.6.3. Extrapolation Test

The final plots of δ_B as a function of Ψ_0 analysis allows us to determine the smoothness of this curve and the appropriate working of the extrapolation to $\Psi_0 = 0$.

7.6.4. Linear Test

If some of the inhomogeneities that are in a linear regime at present are generated, one can obtain δ_B with the linear code and with the *rkf45ng*, and compare their results.

For the case of $\epsilon \neq 0$, $tolr = 10^{-10}$, $fac_\delta = 1$, with local truncation of the order of rescaled *tolr* and quadruple precision in the VAX of the IFIC.

For the linear test, the following steps are given:

1. With both codes, linear and *rkf45ng*, the present temperature contrast δ_B is computed for values of $\Psi_0 = 1, 2, \dots, 180$. In $\Psi_0 = 0$ the result is obtained by extrapolation. This way, the values of δ_B are known in 181 evenly spaced points.
2. By taking into account that $T(t_0)/T_D = (1 + \delta_B)/(1 + Z_D)$, the $T_0(\Psi_0)$ temperatures in each direction and their mean are computed:

$$\langle T_0 \rangle = \frac{1}{2} \int_0^\pi T_0(\Psi_0) \sin(\Psi_0) d\Psi_0.$$

3. Now the new contrast is computed:

$$\delta_T(\Psi_0) = \frac{T(\Psi_0) - \langle T_0 \rangle}{\langle T_0 \rangle}.$$

As a consequence of this definition, the monopole of δ_T is null.

4. The dipolar contribution to δ_T is related with the Observer–Symmetry Centre relative velocity (Doppler Effect). The other contributions to δ_T are named δ_R or **nondipolar anisotropy contribution**, and are computed by subtracting the dipolar part from δ_T .
5. Given any inhomogeneity O_3 – O_7 of Table 1b, δ_T plots of the linear solution and of the *rkf45ng* solution are **indistinguishable**.
6. For the case O_4 , the plots $\delta_R(\Psi_0)$ are shown in Figure 6a. The continuous plot corresponds to the linear computations, while the discontinuous plot corresponds to the non-linear ones (*rkf45ng* algorithm). The figure shows that both results are very similar. This means there is very simultaneous effectiveness in both schemes. For the other inhomogeneities of Table 1b, we also obtained very similar results for both algorithms (linear and non-linear). Figure 6b plots the non-linear computations of $\delta_R(\Psi_0)$ for the cases O_3 – O_7 of Table 1b. In Figure 6b, we can appreciate that the CMB anisotropy grows as the values of Ω decrease. Therefore, these kinds of CMB anisotropies, $\delta_R(\Psi_0)$ values, are greater as the model of universe is more open.

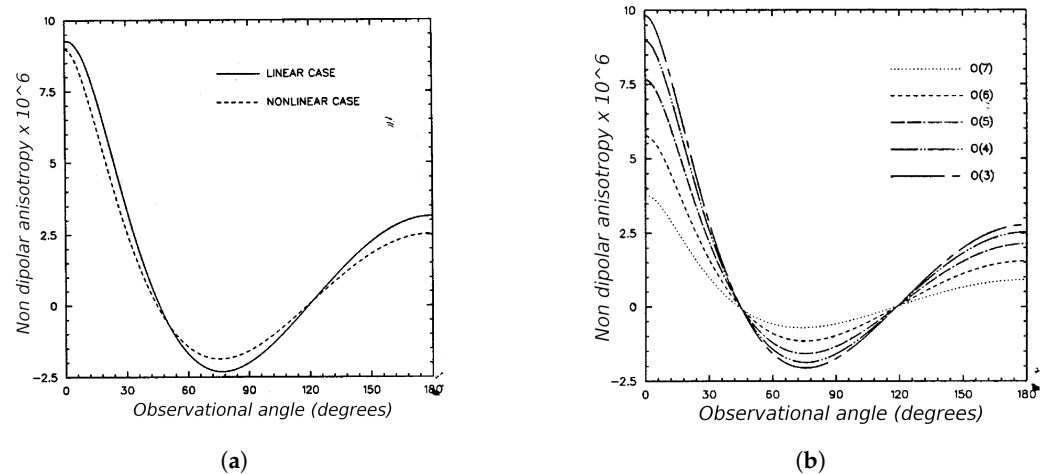


Figure 6. Plot of $\delta_R \times 10^6$ as a function of the observational angle, Ψ_0 , in degrees. (a) Inhomogeneity O_4 . The continuous plot is the result of the linear computation. The discontinuous plot is the result of *rkf45ng* computation. (b) Anisotropies of overdensities in universes with different Ω values.

8. Results and Discussion

Next, some of the most important physical results of the numerical work presented in this paper are discussed, as well as an indication of where they have been published. The detailed numerical and computational techniques that were not presented in such papers are presented here. The value of the results presented here and their original features have been analysed throughout our present paper.

The first results of the physical application presented in this paper can be found in the Ref. [11]. This was the first time our method was presented. Linear tests and preliminary results with non-linear structures were demonstrated. The first conclusion was that, in linear structures, both our non-linear and linear algorithms produced graphically indistinguishable results. The results had to be expanded down to the quadrupole to appreciate the difference. As a side-effect, we found that our non-linear algorithm produced higher levels of precision in δ_B than expected. The reason for this was that even the octupole (with values between 10^{-7} and 10^{-9}) “appeared” to behave properly, although only $tolr = 10^{-10}$ was required as the local error bound in each step and between 800 and 1200 steps were performed in the cases treated there. A less robust code would have presented octupole noise contamination in all likelihood.

Concrete simulated experiments with the non-linear version (part) of the algorithm were performed on the clearly non-linear structures of the Great Attractor and the Virgo cluster [12]. These calculations referred to the computation of the quadrupole. This was

separated into two parts: (a) Q_D or relativistic quadrupole and (b) Q_R or the remainder. Comparisons with results of COBE satellite data were made, which revealed that $Q_R \approx 10^{-7}$ was too small to be detected with the technology of the epoch (nineties of the past century). At the same time, the robustness of our code in [11] was confirmed, which reinforced our results. The effect of a Great Attractor-like structure located at Z_0 was also studied between redshifts 2 and 30 [13]. If the density parameter Ω_0 was varied between 0.15 and 0.4, the conclusion was that the pure gravitational effect of these structures was sufficient to explain a significant part of the large-scale anisotropy of the Universe. Such a conclusion meant that it might be necessary to review the concept of large-scale angular anisotropic generation. The results presented in the Ref. [13] were published as a rapid Letter due to the importance of their findings.

More results were obtained after the previous papers. In the Ref. [14], a modification of the density profile at t_D was presented. This modification included two contrast values $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$ and two radii of the core $R_{v1} \gg R_{v2}$, that is, two superimposed disturbances with the same centre of symmetry. These values should be adjusted so that at a certain $R \gg R_{v1}$, in t_D , the density contrast would be null (compensated structures at a predetermined distance). To do this, the PARAM algorithm was modified in a non-trivial way. Many graphs and tables which helped clarify the analysis were generated, and important conclusions, for instance, the compatibility with observational data, were obtained. The most interesting findings were produced for very open universes and objects very far away from the observer. A deeper study of the structures presented in the Ref. [13] can be found in [15]. A more in-depth study of great underdensities was also performed. In the Ref. [16], the study of the Boötes void with a similar technique to that of paper [14] was carried out; that is, simulating a density vacuum surrounded by a “wall” with excess density which compensated the internal void.

Moreover, we have the following papers about CMB anisotropies in LTB universes: [51–56]. The reader may find comprehensive explanations, including important consequences of previous applications performed with our algorithm, in the book ?? and in the review paper of the Ref. [18]. The latter paper also outlines an overview of CMB and CMB anisotropies research so far.

At this stage, we would like to say some words about the results and the convergence and stability of our algorithms. They are written in the next three paragraphs.

With the example we were trying to solve, we have proven that with the expected signal of our objective function, the temperature contrast, or anisotropy of the CMB in an LTB universe, was persistently polluted by the integration method. If $\varepsilon = 0$, the density contrast was null and our objective function should be 0 for any value of t , in particular for $t = t_D$, since we would be integrating in a FLRW universe or asymptotic limit. It could be argued that the EDOS of the LTB is inadequate for the FLRW, but it was clear that at a great distance from the centre of the disturbance in both universes they almost coincided in any direction of observation and the integration method was contaminating the signal we were studying. Our integration method was one of the most widely used among Runge–Kutta ones—the Kasp–Carl version of the RKF45. However, since we did not need to resort to the coefficients which provide methods of order h^3 , h^2 and h^1 , we had actually used an RKF45. The convergence and theoretical stability of such a method is sufficiently analysed in the bibliography (see, for instance, [3,4]).

Our tools are now successively presented to show which specific problems improve, and to make a normalized analysis of the value of each tool, and with that aim, specific parameters of precision requirements are given. These were: (i) $tolr = 10^{-10}$ as the bound of the local truncation error and of the relative precision of the root calculation when it was performed, (ii) $\varepsilon = 0$ so that our objective function gave us the order of the polluting noise and (iii) calculation in quadruple precision to minimize the effect of rounding error (see Appendix B). With this we intended to estimate the impact of the noise in our objective function. We assumed that the initial system could not reach t_D in some directions because it obtained $M < 0$ in some layers, which was absurd. We rewrote our ODES to avoid

this problem. Thus, we managed to get t_D up to 0.1 degree observation angles in a first correction. It was only after applying all the mentioned tools and getting the final system, that we verified that our objective function was less than $tolr^2/100^2$. This was a resounding zero given the tolerance, $tolr$, required for each function and its derivatives. It must be noted that if the local truncation error was of the order of $tolr$, when performing about 1000 integration steps, the theory told us that the global truncation error could be of the order of $1000 \times tolr$, that is, it could become of the order of 10^{-7} . On the other hand, we checked that the overall noise was of an order less than 10^{-24} when working with quadruple precision, that is, about 38 decimal places. These findings ensure that our algorithms are robust, sufficiently converge, and are very stable.

Our algorithms have been used in a series of papers that are described above. They have been applied both in linear and non-linear structures. Some interesting results about stability and convergence can be summarized as follows: (i) The double integration test gives values in the rang $|valini - valini1| \in [10^{-8}, 10^{-7}]$ for δ_B . As the initial value of $\delta_B = 0$ and the final value is $\delta_B \approx 10^{-5}$, it is deduced that the final value should have at least 8–9 correct decimal digits, which is enough. As a matter of fact, it is highly possible that it has more correct digits due to the fact that the way to “force” the approximation to $\alpha = \pi/2$ is not equivalent for both integrations. (ii) The other initial variables give similar relative error results, i.e., $|valini - valini1|/|valini| \in [10^{-8}, 10^{-7}]$. (iii) All the published plots are smooth enough to allow extrapolation at $\Psi_0 = 0$.

There are recent examples of optimization of numerical methods similar to ours. For instance, the numerical resolution of the Dirac equations (see [19] where the DiracSolver is introduced. See also references therein). Another technique which improves numerical algorithms following a procedure that has some similarities with our work is that presented in the Ref. [38], in which the author extended methods of deriving variational integrators to provide solutions of systems having oscillatory behaviour. His techniques are applied to problems of discrete Lagrangian functions, discrete Euler-Lagrange equations and the definition of a weighted sum of “continuous intermediate Lagrangians” each of them evaluated at an intermediate time node. Additionally, the author introduced different numbers of intermediate points. He assessed the accuracy, convergence and computational time of the proposed technique by testing and comparing them with well-known standards. This work has points of connection with our research as the tests have similarities in the standards used (see also references therein in the Ref. [38]). There is a recent numerically research in Astrophysics which could be connected to ours too. in the Ref. [57] the authors numerically worked with the McVittie’s metric and the Mathisson–Papapetrou (MP) equations in the post-Newtonian approach [58]. They analytically and numerically investigated the orbits of spinning particles around black holes in the post-Newtonian limit and in the presence of cosmic expansion. On the one hand, we could connect the study with LTB and LFRW metrics with McVittie’s one. On the other hand, the spinning particle orbits could be compared with our numerical Relativistic Positioning Systems techniques used in the Refs. [47,59].

9. Summary and Conclusions

We have presented here the resolution of an ODES applied to a cosmological problem. We chose a RKF45 instead of an Adams. Here, we describe all the adaptations we have made to solve our physical problem. The *objective function* has been defined. The originality of our work is that the numerical tools used in our algorithms, which are particular to our problem, should be extended to other problems in the fields of physics and engineering as we remark in the text. The accuracy of the predictions depends on the accuracy of the model and the rigour of the tests. This is also discussed along the text. The numerical algorithms built to solve the physical problem have been described for the use of adaptive step-size numerical methods. In their application, some general properties have been used, and can be extrapolated to other problems. Catastrophic cancellations have been suppressed, and such kinds of suppression can be considered in similar problems. Additionally, the

introduction of the objective function helps the optimization of the resolution of the ODES. Moreover, local error resizing is applied, and its usefulness is explained in our paper. This resizing is performed by the use of the asymptotic solution of the ODES. This technique can also be applied to similar problems. The change of variables guarantees the stability of the solution. The linearization of the problem is used as a powerful control test. What we state here is that this kind of tool should solve a wide range of problems that are unstable or do not have a proper solution with the modern numerical tools. These results represent one of the main novelties of our paper and their scientific contribution.

Some conclusions can be extracted by our present study. The IVP techniques used in our research can be applied to solve ODES. Such techniques should be used when applying adaptive step-size numerical methods with variables in a wide range of values and that seek an almost residual objective function value. In order to consider these Dynamical Systems as a disturbance of the asymptotic system, it is necessary to find an asymptotic solution on which to rely. In this case, the Dynamical System is remodeled in order to study the evolution of this disturbance. This is the meaning of our change to the variables y_1 , y_2 and the rescaling of the errors. The RKF45 algorithm used and its adaptive methods can be applied to other similar problems, such as N-body problems, like AP3M, or similar. The way that catastrophic cancellations are eliminated should be used in problems where ODES have to be solved. Mathematical optimization was carried out by introducing the objective function in the ODES. We encourage the use of this procedure when an ODES has similar treatment. Local error resizing was also executed to solve the problem. This resize implies the use of certain variables to adjust the integration step, while the other variables are used as parameters to determine the coefficients of the ODES. This resize is made by using the asymptotic solution of this system. This approach has demonstrated to be very useful and should be applicable to other similar problems as ours. The change of variables is necessary to guarantee the stability of the integration. The linearization of the ODES, when possible, should work as a powerful control test. All the skills named in these paragraphs are powerful tactics to be applied to physical problems where ODES are needed.

We will now say a few words about future extensions of our research and an outlook. The case we have treated here is the effective numerical resolution of LTB space-time solutions of EE. These tools can also be applied to other LTB problems which make them relevant. In addition, our procedures may be used with other metrics, such as McVitte's one [57]. As a matter of fact, we have used them in N-body algorithms when studying the large-scale structure of the universe in Cosmology, as well as the CMB anisotropies, both in the framework of the Hydra project (see [45,46]) or similar (see [60]). The Hydra code is based on AP3M (Adaptive Particle-Particle, Particle-Mesh) and is adapted with our routines to the evolution of Dark Matter (with and without gas) in Cosmology. At present, some of our techniques are also being applied in Relativistic Positioning Systems in the Solar System (see [47,59]). We solved an ODES of 36 equations of which initial values were obtained from ephemeris and evolved with EE using an asymptotic system of Schwarzschild's metric. Our rescaling approach is crucial as the differences from satellites' and celestial objects' semi-axes in the solar system must be solved. In the near future, we want to present the progress of our skills and the new algorithms built so far in another paper. All of these and similar techniques are powerful tools for other astrophysics, physics and engineering issues in general. We hope all the procedures presented here and in other publications will be useful in the resolution of ODES, both in our specific field and in science and technology in general.

Author Contributions: Conceptualization, J.V.A.i.C. and M.J.F.i.A.; formal analysis, J.V.A.i.C. and M.J.F.i.A.; funding acquisition, M.J.F.i.A.; investigation, J.V.A.i.C. and M.J.F.i.A.; methodology, J.V.A.i.C. and M.J.F.i.A.; project administration, M.J.F.i.A.; resources, M.J.F.i.A.; software, J.V.A.i.C. and M.J.F.i.A.; supervision, M.J.F.i.A.; validation, J.V.A.i.C. and M.J.F.i.A.; visualization, J.V.A.i.C.; writing—original draft, J.V.A.i.C. and M.J.F.i.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Spanish Ministerio de Ciencia, Innovación y Universidades and the Fondo Europeo de Desarrollo Regional, Projects PID2019-109753GB-C21 and PID2019-109753GB-C22, the Generalitat Valenciana Project AICO/2020/125 and the Universitat de València Special Action Project UV-INVAE19-1197312. The APC was funded by the Spanish Ministerio de Ciencia, Innovación y Universidades and the Fondo Europeo de Desarrollo Regional, Projects PID2019-109753GB-C21 and PID2019-109753GB-C22

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We get a lump in our throat and we get emotional every time we remember (and we can well assure you that they are many times) our dear friend and companion Diego Pacual Sáez Milán to whom we owe a lot of our knowledge, and we will be forever deeply grateful for his guidance and mentorship. It was an honour and privilege to know him and work alongside him, for much of our lives. That is why we want to take the opportunity that this space gives us to thank him once again so much. Diego left us four years ago, but his person and his memory will be in this universe as long as we are in it. With him we shared this and most of the work of our research, developed over so many years of patience, debates, discussions, hardships, laughs and joy; in other words, many moments of happiness to remember for as long as our breath and force be with us. There are no words to describe our gratitude, both for his human and scientific teachings. My friend and colleague Diego, our esteem is with you, accompanies us and pushes us through our work, both scientifically and of any sort. Thank you very much, dear colleague! This work was supported by the Spanish Ministerio de Ciencia, Innovación y Universidades and the Fondo Europeo de Desarrollo Regional, Projects PID2019-109753GB-C21 and PID2019-109753GB-C22, the Generalitat Valenciana Project AICO/2020/125 and the Universitat de València Special Action Project UV-INVAE19-1197312.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Some Important Definitions

- Comoving coordinates.** These coordinates assign constant spatial coordinate values to observers who measure the universe as isotropic. As such observers move along with the Hubble flow, they are called comoving observers. CMB is also perceived to be isotropic. It is useful, for our purposes, to assign comoving coordinates to every point of the universe because LFRW and LTB universes are in expansion. Moreover, due to the spherical symmetry of such universes we choose coordinates $P(M, \theta, \varphi, t)$, with the origin in the symmetry centre C_s of the perfect fluid which is their source. Furthermore, we take the coordinates $O(M_O, \pi, 0, t)$ (being O the observer), as the φ coordinate becomes useless in our study. Additionally, for symmetry reasons, it is sufficient to study the spatial half-plane that contains the OC_s axis (see Figure 1). The M coordinate is a *label* associated to every spatial spherical surface with centre in C_s , named **shell**, and is related to the mass-energy contained in the surface defined by the shell. The *radial* coordinate $R(M, t)$ of such a shell depends on the time and is related with M and $\rho(M, t)$ through the EE. In FLRW, we choose an arbitrary point as the origin of coordinates and an arbitrary direction as the axis (at constant t , all the axes are equivalent) and the coordinates $Q(r, \theta, \varphi, t)$ are also defined, where r is a radial coordinate, which is also comoving. The shells in this universe are a mere auxiliary definition which depends on the origin of coordinates.
- Decoupling time.** This is defined as the time t_D , where it can be considered that the dominant effect over the free photons is gravitational. This moment is as far as it is measured by the spectral lines *Redshift* Z_D . Decoupling time should be taken at $Z_D \approx 1100$. From this moment, photons freely propagate in the gravitational field of the Universe and its structures. As we are interested in computing the gravitational effect produced by nonlinear structures, we can take initial values at some suitable

$Z_D < 1100$. In our work, we take initial values at $Z_D = 50 \approx 10^3$ Mpc (1 Mpc $\approx 3.2 \cdot 10^6$ lightyear).

3. **Scale factor.** This is the time function $a = a(t)$ included in the FLRW metric. The distances among galaxies are proportional to the scale factor and the following relation stands $a(t) = a(t_0)/(1 + Z)$, where t_0 is the present time and Z is the redshift at time t .
4. **Critical density** This is the mass density value ρ_c such that, when it exceeds, the effect of gravity in the universe is great enough to stop its expansion in finite time. In our work, we used it to define different values of ρ_B as background FLRW universes for our model of LTB. This definition is made through the density parameter $\Omega = \rho_B/\rho_c$. Thus, values such that $\Omega < 1$ originate FLRW open universes, while $\Omega > 1$ values originate closed FLRW universes. The $\Omega = 1$ case corresponds to a flat FLRW.
5. **Density profile.** This is an initial condition to determine the LTB metric. At the initial time t_D we take, in this work, a density profile given by:

$$\rho(M, t_D) = \rho_B \left(1 + \frac{\varepsilon}{1 + \left(\frac{R}{R_v}\right)^2} \right), \tag{A1}$$

where ε gives the maximum density contrast $(\rho - \rho_B)/\rho_B$ and R_v indicates the radial coordinate at which this contrast is reduced to a half (*we call it core*). As t varies, $\rho(M, t)$ changes. The new profile in t may change its shape if the perturbation is nonlinear. Nevertheless, with linear perturbations it is acceptable that the shape of the profile remains the same.

Both R_v and ε besides Ω , are initial values, for the searched algorithm. $\Omega \in [0.2, 1[$ have been taken in the present work. With this initial profile, the spherical symmetry is guaranteed and the asymptotic FLRW limit with $\rho_B = cte$.

Appendix B. Numerical Aspects to Take into Account

Appendix B.1. IEEE Norm

The standard representation of numbers in floating point arithmetic (scientific calculus) is of the type $y = \pm(.d_1d_2 \dots d_t)_\beta \times \beta^e$, and is defined by four integer values. The base β , the precision or word length t , and the range of exponents $e_{min} \leq e \leq e_{max}$. The mantissa $m = \pm(.d_1d_2 \dots d_t)$ accomplishes $0 \leq m < 1$ and $0 \leq d_i < \beta$. In order to ensure unique representation, $d_1 \neq 0$, and, in this case, the representation is said to be normalized; if the base is also 2, the $d_1 = 1$, so it is not necessary to save it as the decimal point.

The word length t is chosen for each variable by assigning a type such as **float**, **double**, **quadruple**, and so forth, and here comes the concept of precision and rounding error for each variable. The IEEE standard specifies three basic formats, among others. In order not to extend ourselves too much, we limit ourselves to binary machines (see Table A1).

Table A1. Storage of variables in the IEEE standard in binary machines. Rounding units and range of variables according to the declared type.

| Type | Long | Mantissa | Expo | Rounding Unit | Rank |
|--------|----------|----------|---------|---|-----------------|
| Float | 32 bits | 24 bits | 8 bits | $2^{-24} \approx 5.96 \times 10^{-8}$ | $10^{\pm 38}$ |
| Double | 64 bits | 53 bits | 11 bits | $2^{-53} \approx 1.11 \times 10^{-16}$ | $10^{\pm 308}$ |
| Quad | 128 bits | 113 bits | 15 bits | $2^{-113} \approx 9.63 \times 10^{-35}$ | $10^{\pm 4932}$ |

The IEEE standard also specifies that if w is an elementary operation on the operands (x, y) , the result of $z = (x w y)$ must be given as if the operation $(x w y)$ was performed exactly and, when obtaining the result, it was rounded. This standard allows us to guarantee that the relative rounding error of the elemental operations is of the order of the values

given in column 5 of the table above. Naturally, the loss of significant digits has nothing to do with rounding, but with the negligence of the programmer.

Appendix B.2. Intrinsic Functions

We refer in this section to the functions that compilers incorporate as elementary functions \exp, \log, sqrt , all hyperbolic and trigonometric functions and their inverses, and so forth. Such functions are optimized not only in CPU time, but also choose the most suitable method to the value of the input variable to ensure that the error in the output is of the order of the rounding unit. Such functions should be used whenever possible in order to have an estimate of rounding.

Appendix B.3. Numerical Derivation

In numerical analysis, it is well-known that given $g(x) \in C^2$ and $\forall h$, Taylor’s development allows us to write

$$g(x + h) = g(x) + g'(x)h + \frac{1}{2}g''(\xi)h^2, \quad \xi \in]\min(x + h, x), \max(x + h, x)[,$$

so it is very common to calculate the numerical derivative of the function $g(x)$ using the Mean Value Theorem

$$g'(x) = \frac{g(x + h) - g(x)}{h}, \quad E(g') \leq \frac{1}{2} \max_{\xi} |g''(\xi)| |h| \tag{A2}$$

where $E(g')$ designates a bound of the truncation error for the derivative, error because we have truncated the Taylor expansion in the linear approximation.

With the relative error of g' we are already presented with the first problem because we should calculate $E(g') / |g'(x)|$ and that can give problems in an environment of the extremals of $g(x)$ where the derivative is close to 0. It is true that if the derivative vanishes it should be $g(x + h) = g(x)$, but in finite precision arithmetic finding, strict equality is a lottery and we cannot subject our algorithms to such chance. Additionally, $E(g')$ can lose significant digits if h is too small, making our attempt to control the error even more disastrous.

We refer to this problem when we say that *the truncation of the derivative is of order $\mathcal{O}(h)$ while rounding is of order $\mathcal{O}(1/h)$* . We are actually talking improperly and we should say that rounding is in the order of the rounding unit mentioned in the IEEE standard, divided by $|h|$. If h is small, we are losing precision in the derivative, as if our variable **double** became **float** in a single calculation when $h \approx 10^{-8}$. Due to this circumstance, it is essential **to implement ways to detect the extremals** by means of geometric or trigonometric formulas that allow us to avoid these problems.

While it does not seem feasible to avoid amplification by h of any errors in the evaluation of $g(x)$, we can protect ourselves against the possible *catastrophic cancellation* in $\Delta_g = g(x + h) - g(x)$ when both evaluations have the same sign since otherwise, there is no problem. The solution is so simple as to use the Taylor expansion of the function $\log(|g(x + h)|)$. Suppose $g(x + h), g(x) > 0$, and we will have:

$$\log(g(x + h)) = \log(g(x)) + \frac{g'(x)}{g(x)}h + \frac{1}{2} \left(\frac{g''(\nu)}{g(\nu)} - \left[\frac{g'(\nu)}{g(\nu)} \right]^2 \right) h^2, \tag{A3}$$

with the same conditions for ν that ξ had before. If now we are left with the linear approximation and we solve $g'(x)$, we are left with:

$$g'(x) = g(x) \frac{\log(g(x + h)) - \log(g(x))}{h} = \frac{g(x)}{h} \log\left(\frac{g(x + h)}{g(x)}\right). \tag{A4}$$

Now the quotient of the evaluations of g is an elementary operation and \log is an intrinsic function so all operations are within the scope of the rounding unit corresponding to the type of declared variable. The h can be very small and will only affect the value of g' , but since both values $g(x+h)$ and $g(x)$ will be very close, the quotient will be close to 1 and its logarithm tends to zero. It does not solve all the problems, but it helps to a high degree in our case to calculate the three partial derivatives with respect to M in the CS algorithm.

References

- Perrin, C.L. *Numerical Recipes in Fortran 90: The Art of Scientific Computing, Volume 2 (3 CD-ROMs and Manual)*; Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., Eds.; Cambridge University Press: New York, NY, USA, 1997.
- Butcher, J.C. A history of runge-kutta methods. *Appl. Numer. Math.* **1996**, *20*, 247–260.
- Fehlberg, E. *Low-Order Classical Runge-Kutta Formulas with Stepsize Control and Their Application to Some Heat Transfer Problems*; National Aeronautics and Space Administration: Washington, DC, USA, 1969; Volume 315.
- Cash, J.R.; Karp, A.H. A variable order runge-kutta method for initial value problems with rapidly varying right-hand sides. *ACM Trans. Math. Softw. TOMS* **1990**, *16*, 201–222.
- Lynch, S. *Dynamical Systems with Applications Using MapleTM*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
- Lynch, S. *Dynamical Systems with Applications Using MATLAB®*; Springer: Berlin/Heidelberg, Germany, 2014.
- Lynch, S. *Dynamical Systems with Applications Using Mathematica®*; Springer: Berlin/Heidelberg, Germany, 2017.
- Lynch, S. *Dynamical Systems with Applications Using Python*; Springer: Berlin/Heidelberg, Germany, 2018.
- Zhu, Q. Stabilization of stochastic nonlinear delay systems with exogenous disturbances and the event-triggered feedback control. *IEEE Trans. Autom. Control* **2018**, *64*, 3764–3771.
- Wang, H.; Zhu, Q. Global stabilization of a class of stochastic nonlinear time-delay systems with siss inverse dynamics. *IEEE Trans. Autom. Control* **2020**, *65*, 4448–4455.
- Arnau, J.V.; Fullana, M.J.; Monreal, L.; Sáez, D. On the microwave background anisotropies produced by nonlinear voids. *Astrophys. J.* **1993**, *402*, 359–368.
- Sáez, D.; Arnau, J.V.; Fullana, M.J. The imprints of the great attractor and the virgo cluster on the microwave background. *Mon. Not. R. Astron. Soc.* **1993**, *263*, 681–686.
- Arnau, J.V.; Fullana, M.J.; Sáez, D. Great attractor-like structures and large-scale anisotropy. *Mon. Not. R. Astron. Soc.* **1994**, *268*, L17–L21.
- Fullana, M.J.; Sáez, D.; Arnau, J.V. On the microwave background anisotropy produced by great attractor-like structures. *Astrophys. J. Suppl. Ser.* **1994**, *94*, 1–16.
- Sáez, D.; Arnau, J.V.; Fullana, M.J. Effects of great attractor-like objects on the cosmic microwave background. *Astrophys. Lett. Commun.* **1995**, *32*, 75.
- Fullana, M.J.; Arnau, J.V.; Sáez, D. On the microwave background anisotropy produced by big voids in open universes. *Mon. Not. R. Astron. Soc.* **1996**, *280*, 1181–1189.
- Fullana i Alfonso, Màrius Josep. *Anisotropies de la Radiació de fons de Microones Produïdes per Inhomogeneïtats Cosmològiques No Lineals*; Servei de Publicacions de la Universitat de València: València, Spain, 1996.
- Fullana, M.J.; Arnau, J.V.; Sáez, D. Looking for the imprints of nonlinear structures on the cosmic microwave background. *Vistas Astron.* **1997**, *41*, 467–492.
- Tsoulos, I.G.; Kosmas, O.T.; Stavrou, V.N. Diracsolver: A tool for solving the dirac equation. *Comput. Phys. Commun.* **2019**, *236*, 237–243.
- Kosmas, T.S.; Lagaris, I.E. On the muon–nucleus integrals entering the neutrinoless $\mu \rightarrow e$ -conversion rates. *J. Phys. G Nucl. Part. Phys.* **2002**, *28*, 2907.
- Kitano, R.; Koike, M.; Okada, Y. Detailed calculation of lepton flavor violating muon-electron conversion rate for various nuclei. *Phys. Rev. D* **2002**, *66*, 096002, .
- Stoica, S.; Mirea, M. New calculations for phase space factors involved in double- β decay. *Phys. Rev. C* **2013**, *88*, 037303.
- Stoica, S.; Mirea, M.; Nițescu, O.; Nabi, Ja.; Ishfaq, M. New phase space calculations for β -decay half-lives. *Adv. High Energy Phys.* **2016**, <https://doi.org/10.1155/2016/8729893>.
- Carleo, G.; Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **2017**, *355*, 602–606, .
- Snyder, J.C.; Rupp, M.; Hansen, K.; Müller, Kl.; Burke, K. Finding density functionals with machine learning. *Phys. Rev. Lett.* **2012**, *108*, 253002.
- Arsenault, Lo.; Lopez-Bezanilla, A.; Lilienfeld, O.A.v.; Millis, A.J. Machine learning for many-body physics: The case of the anderson impurity model. *Phys. Rev. B* **2014**, *90*, 155136.
- Torlai, G.; Mazzola, G.; Carrasquilla, J.; Troyer, M.; Melko, R.; Carleo, G. Neural-network quantum state tomography. *Nat. Phys.* **2018**, *14*, 447–450.

28. Engquist, B.; Fokas, A.; Hairer, E.; Iserles, A. *Highly Oscillatory Problems*; Cambridge University Press: Cambridge, UK, 2009; p. 366.
29. Wendlandt, J.M.; Marsden, J.E. Mechanical integrators derived from a discrete variational principle. *Phys. D Nonlinear Phenom.* **1997**, *106*, 223–246.
30. Kane, C.; Marsden, J.E.; Ortiz, M. Symplectic-energy-momentum preserving variational integrators. *J. Math. Phys.* **1999**, *40*, 3353–337.
31. Marsden, J.E.; West, M. Discrete mechanics and variational integrators. *Acta Numer.* **2001**, *10*, 357–514.
32. Hairer, E.; Lubich, C.; Wanner, G. Geometric numerical integration illustrated by the störmer–verlet method. *Acta Numer.* **2003**, *12*, 399–450.
33. Leimkuhler, B.; Reich, S. *Simulating Hamiltonian Dynamics*; Cambridge University Press: Cambridge, UK, 2004; Number 14.
34. Ober-Blöbaum, S. Galerkin variational integrators and modified symplectic runge–kutta methods. *IMA J. Numer. Anal.* **2017**, *37*, 375–406.
35. Ober-Blöbaum, S.; Saake, N. Construction and analysis of higher order galerkin variational integrators. *Adv. Comput. Math.* **2015**, *41*, 955–986.
36. Campos, C.M.; Junge, O.; Ober-Blöbaum, S. Higher order variational time discretization of optimal control problems. *arXiv* **2012**, arXiv:1204.6171.
37. Kosmas, O.; Leyendecker, S. Variational integrators for orbital problems using frequency estimation. *Adv. Comput. Math.* **2019**, *45*, 1–21.
38. Kosmas, O. Energy minimization scheme for split potential systems using exponential variational integrators. *Appl. Mech.* **2021**, *2*, 431–441. <https://doi.org/10.3390/applmech2030024>.
39. Kosmas, O.; Leyendecker, S. Family of higher order exponential variational integrators for split potential systems. *J. Phys. Conf. Ser.* **2015**, *574*, 012002.
40. Romero, P.D.; Candela, V.F. Blind deconvolution models regularized by fractional powers of the laplacian. *J. Math. Imaging Vis.* **2008**, *32*, 181–191.
41. Abbas, M.I.; Ragusa, M.A. On the hybrid fractional differential equations with fractional proportional derivatives of a function with respect to a certain function. *Symmetry* **2021**, *13*, 264.
42. Treeby, B.E.; Cox, B.T. Modeling power law absorption and dispersion in viscoelastic solids using a split-field and the fractional laplacian. *J. Acoust. Soc. Am.* **2014**, *136*, 1499–1510.
43. Candela, V.; Falcó, A.; Romero, P.D. A general framework for a class of non-linear approximations with applications to image restoration. *J. Comput. Appl. Math.* **2018**, *330*, 982–994.
44. Lemaître, G. Expansion of the universe, the expanding universe. *Mon. Not. R. Astron. Soc.* **1931**, *91*, 490–501.
45. Fullana, M.J.; Arnau, J.V.; Thacker, R.J.; Couchman, H.M.P.; Sáez, D. Estimating small angular scale cosmic microwave background anisotropy with high-resolution n-body simulations: weak lensing. *Astrophys. J.* **2010**, *712*, 367.
46. Fullana, M.J.; Arnau, J.V.; Thacker, R.J.; Couchman, H.M.P.; Sáez, D. On the estimation and detection of the rees–sciama effect. *Mon. Not. R. Astron. Soc.* **2016**, *464*, 3784–3795.
47. Neus Puchades Colmenero, José Vicente Arnau Córdoba, and Màrius Josep Fullana i Alfonso. Relativistic positioning: Including the influence of the gravitational action of the sun and the moon and the earth’s oblateness on galileo satellites. *Astrophys. Space Sci.* **2021**, *366*, 1–19.
48. Amat, S.; Aràndiga, F.; Arnau, J.V.; Donat Beneito, R.M.; Mulet Mestre, P.; Peris Sancho, R. *Aproximació Numèrica*; Servei de Publicacions de la Universitat de València: València, Spain, 2002.
49. Bondi, H. Spherically symmetrical models in general relativity. *Mon. Not. R. Astron. Soc.* **1947**, *107*, 410–425.
50. Bardeen, J.M. Gauge-invariant cosmological perturbations. *Phys. Rev. D* **1980**, *22*, 1882–1905.
51. Saez, D.; Arnau, J.V. On the Tolman Bondi solution of Einstein’s Equations. Numerical Applications. In *Recent Developments in Gravitation*; World Scientific Publishing Co Pte Ltd.: Singapore, 1990; p. 415.
52. Fullana, M.J.; Saez, D.P.; Monreal, L.; Arnau, J.V. The great attractor and the anisotropy of the microwave background. In *Recent Developments in Gravitation*; World Scientific Publishing Co Pte Ltd.: Singapore, 1992; p. 131. <https://doi.org/10.1142/9789814537070>.
53. Fullana, M.J.; Arnau, J.V. The Boötes void and the Cosmic Microwave Background. In *Inhomogeneous Cosmological Models*; World Scientific Publishing Co Pte Ltd.: Singapore, 1995; p. 125.
54. Fullana, M.J.; Saez, D.; Arnau, J.V. The imprints of great voids on the cosmic microwave background. In *Relativistic Astrophysics and Cosmology*; World Scientific Publishing Co Pte Ltd.: Singapore, 1997; p. 121.
55. Saez, D.; Arnau, J.V. Accurate simulations of the microwave sky at small angular scales. In *Microwave Background Anisotropies, Proceedings of the XVIth Moriond Astrophysics Meeting, Les Arcs, Savoie, France, 16–23 March 1996*; Editions Frontières: Singapore, 1997; p. 245.
56. Sáez, D.; Arnau, J.V. Learning from observations of the microwave background at small angular scales. *Astrophys. J.* **1997**, *476*, 1.
57. Antoniou, I.; Papadopoulos, D.; Perivolaropoulos, L. Spinning particle orbits around a black hole in an expanding background. *Class. Quantum Gravity* **2019**, *36*, 085002. <https://doi.org/10.1088/1361-6382/ab0fc1>.
58. Myron, M. Neue mechanik materieller systemes. *Acta Phys. Polon* **1937**, *6*, 163–2900.

-
59. Fullana Alfonso, M.J.; Saez Milán, D.P.; Arnau, J.V.; Colmenero, N.P. Some improvements on relativistic positioning systems. *Appl. Math. Nonlinear Sci.* **2018**, *3*, 161–166.
 60. Puchades, N.; Fullana, M.J.; Arnau, J.V.; Saez, D. On the rees–sciama effect: maps and statistics. *Mon. Not. R. Astron. Soc.* **2006**, *370*, 1849–1858.