ALTERRNATIVES TO FINITE DIFFERENCE METHODS
IN NUMERICAL RELATIVITY

Pablo Laguna
Department of Astronomy & Astrophysics, and
Center for Gravitational Physics & Geometry
Penn State University
University Park, PA 16802, USA

October 19, 1995

Abstract

Finite difference techniques have played a dominant role in Numeri-
cal Relativity. This situation will likely prevail; for instance, the current
Grand Challenge effort in the United States to simulate, by the end of the
century, black hole collisions is entirely based on finite difference codes.
Furthermore, the power of finite difference techniques has been recently
enhanced with the implementation of adaptive mesh refinements. In spite
of the finite difference success, there has been a significant number of nu-
merical studies in gravitation in which finite difference methods are either
not used, or applied in combination with other techniques. These lectures
review four alternative approaches to Numerical Relativity from those of
finite difference. The first lecture addresses solutions to the initial data
problem in General Relativity using multiquadrics and finite elements
methods. The second lecture reviews particle-mesh and smoothed particle
hydrodynamics methods used in conjunction with finite differences to
solve the Einstein/Hydro field equations.

1 Introduction

Finite difference (FD) methods are the most popular choice in computational
physics. In Numerical Relativity, the situation can be considered extreme; FD
methods currently enjoy an almost exclusive monopoly. We have recently seen,
however, signs that in the future this could change, perhaps not in a radical way,
but to the point in which it will be possible to choose from a family of numerical
techniques a method appropriate to the problem under consideration.
In general, numerical methods for solving partial differential equations can be divided into two categories: (A) Methods based on direct approximations of the derivatives in the differential equation and (B) methods that approximate the solution of the continuum differential equation by linear combinations of trial functions. FD methods belong to class A. This difference in methods has a direct impact on the construction of the computational grid. For FD methods the grid points are located along constant coordinate directions (Fig. 1a), so a natural approximation of derivatives can be constructed; on the other hand, methods in class B that approximate the solution with trial functions do not impose such a restriction (Fig. 1b) since the approximate derivatives are obtained after substitution of the approximate solution. One can be tempted to jump

Figure 1: Schematic representation of computational meshes for (a) FD methods and (b) methods based on approximating the solution by linear combinations of trial functions.

2
to the conclusion that, because of the restriction of grid points along constant coordinate directions, FD methods are not flexible enough to handle complicated physical domains or systems with vast ranges of characteristic scales. Two ways to remedy these difficulties are to find coordinate systems that conform to the boundaries and adapt to the variations in scales, or to use adaptive mesh refinements techniques (AMR). For obvious reasons, AMR constitutes the natural choice in complex three-dimensional problems.

The main goal of these lectures is to review those methods in class B that have been used in Numerical Relativity. The common property that these methods share is their intrinsic adaptive nature. Because of "space-time" restrictions, the two most important omissions in these lectures are spectral methods and Regge calculus. Spectral methods have been used in the study of gravitational collapse [1, 2] and for the construction of black hole collision initial data [3]. They have the desirable advantage of achieving high precision because of its fast convergence. In addition, spectral methods are well known for being capable of correctly handling any type of boundary conditions, as well as treating discontinuities without the use of artificial viscosity. One major drawback of spectral methods arises when applied to evolution equations; the Courant condition in this case is more restrictive than in FD methods.

The other important method that will not be discussed here is Regge Calculus. Regge calculus represents an interesting approach to Numerical Relativity since by construction it is a discrete version of General Relativity [4]. Under this formalism, dynamical variables are represented by finite distances, as oppose to the metric fields; that is, the space-time continuum is approximated by flat simplices with the curvature concentrated on the vertices. Unfortunately, the applications of Regge calculus performed so far [5, 6, 7] have not been able to show whether this technique is a competitive approach to Numerical Relativity in terms of accuracy for a given computational effort.

The lectures are organized as follows. A brief introduction to the 3+1 formalism is given in section 2. In section 3, the initial data problem in General Relativity is reviewed, and two numerical techniques for the construction of initial data are presented: Multiquadrics in section 3.1 and finite elements in section 3.2. Two particle techniques for the evolution of matter fields in curved spacetimes are presented in section 4: Particle-Mesh in section 4.1 and Smoothed Particle Hydrodynamics in section 4.2. Conclusions are given in section 5.

2 The 3+1 Formalism

The standard approach in Numerical Relativity to solving Einstein's equations is based on the 3+1 or ADM (Arnowitt-Deser-Misner [8]) formalism. Under this approach, space-time is viewed as the time history of the geometry of a space-like hypersurface (geometrodynamics) [9]. The starting point is a foliation of 3-surfaces, level surfaces of a scalar function that completely cover the 4-
dimensional space-time. Under this construction, the 4-metric, \((4)g_{\mu\nu}\), can be written in terms of the 3-metric, \(g_{ij}\), of the space-like hypersurfaces as \((4)g_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu\), where \(n_\mu\) denotes the unit normal to the hypersurfaces in the foliation.\(^1\) Equivalently, one can rewrite the 4-metric as

\[
    ds^2 = -\alpha^2 dt^2 + g_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt),
\]

(1)

where \(\alpha\), the lapse function, represents the freedom of choosing arbitrarily time coordinates, and \(\beta^i\), the shift vector, contains the freedom of relabeling coordinate points in the space-like hypersurfaces; that is, the lapse function and the shift vector represent kinematical variables. The unit normal \(n_\mu\) in terms of these variables reads \(n_\mu = (1, \beta^i)/\alpha\).

The metric \(g_{ij}\) only describes the internal geometry of the hypersurfaces. The embedding of these hypersurfaces in the 4-dimensional space-time is characterized by the extrinsic curvature \(K_{\mu\nu}\), defined as \(K_{\mu\nu} = -\nabla_\mu n_\nu\), where \(\nabla_\mu\) is the covariant derivative induced by the 3-metric \(g_{\mu\nu}\). The intrinsic metric and the extrinsic curvature pair \((g_{ij}, K_{ij})\) represents the dynamical quantities in Einstein’s theory. The evolution equations for \((g_{ij}, K_{ij})\) are obtained from the definition of the extrinsic curvature and the (space, space) components of Einstein’s equations, \(G_{\mu\nu} = 8\pi T_{\mu\nu}\). These equations read

\[
    \partial_t g_{ij} = -2\alpha K_{ij} + \nabla_i \beta_j + \nabla_j \beta_i, \quad (2)
\]

\[
    \partial_t K_{ij} = -\nabla_i \nabla_j \alpha + \beta^k \nabla_k K_{ij} + K_{ik} \nabla_j \beta^k + K_{kj} \nabla_i \beta^k + \alpha [R_{ij} + K K_{ij} - 2 K_{ik} K^k_j - 8\pi \{S_{ij} + g_{ij}(\rho_H - S)/2\}], \quad (3)
\]

where \(R_{ij}\) is the 3-Ricci tensor, \(R\) is its trace, and \(K\) is the trace of \(K_{ij}\). The source terms in (3) are obtained from the stress-energy tensor, \(T_{\mu\nu} = S_{\mu\nu} + 2J_{(\mu} n_{\nu)} + \rho_H n_\mu n_\nu\). These sources are explicitly given by

\[
    \rho_H = n^\mu n_\nu T_{\mu\nu}, \quad (4)
\]

\[
    J_\alpha = g_\alpha^\mu n^\nu T_{\mu\nu}, \quad (5)
\]

\[
    S_\beta = g_\beta^\mu n^\nu T_{\mu\nu}. \quad (6)
\]

Finally, the (time, time) and (time, space) components of Einstein’s equations yield the

\[
    \text{Hamiltonian Constraint:} \quad R + K^2 - K_{ij} K^{ij} = 16\pi \rho_H, \quad (7)
\]

\[
    \text{Momentum Constraint:} \quad \nabla_j (K^{ij} - g^{ij} K) = 8\pi J^i. \quad (8)
\]

\(^1\)Four-dimensional indices will be denoted with Greek letters and 3-dimensional indices with Latin letters; when needed, 4-dimensional tensors will be labeled with a superscript (4). Units are such that \(G = c = 1\).
function and shift vector. These equations represent conditions on the choices of initial data for \((g_{ij}, K_{ij})\). The Bianchi identities ensure that these constraints are preserved during the evolution.

In general terms, a 3+1 algorithm to obtain numerical solutions to Einstein’s equations splits naturally into two stages:

1. Construction of initial data that satisfies the Hamiltonian and momentum constraints, namely the initial data problem.

2. Evolution of the intrinsic metric and extrinsic curvature \((g_{ij}, K_{ij})\) using equations (2) and (3), respectively.

The evolution stage requires, in addition, specification of the lapse function and shift vector. Even though, the geometry of a space-time does not depend on the foliation, in numerical relativity, it is of fundamental importance to prescribe lapse functions and shift vectors that generate foliations covering as much as possible the future development of the initial data. In particular, for the evolution of black holes, construction of a foliation must be such that singularities are avoided.

Modifications to the above initial-data+dynamics scheme arise, for instance, when a particular structure of the metric, e.g. diagonal, is imposed. In these cases, evolution equations of the metric components are “traded” with equations for the shift vector [10]. Another possibility appears when the evolution is constrained or partially constrained [11, 12]. For such situations, the Hamiltonian constraint and/or the momentum constraint are used instead of some of the evolution equations.

3 The Initial Data Problem

Generally speaking, to obtain initial data, one faces the problem of finding out which four pieces of information within the twelve components of the initial data, \((g_{ij}, K_{ij})\), are to be solved from the four constraint equations (7) and (8). In some particular instances [11, 13], it is clear which metric or extrinsic curvature components need to be solved from the constraint equations; however, for a general situation, such as finding initial data for collision of black holes, this choice is unclear. The most popular and general formalism to construct initial data in numerical relativity is called the York procedure [14]. This procedure uses conformal transformations and trace-free, divergence-free tensor decompositions to isolate the true degrees of freedom in the construction of initial data.

The starting point is a conformal transformation of the 3-metric of the initial hypersurface

\[
g_{ij} = \phi^4 \tilde{g}_{ij},
\]  

(9)
where conformal quantities are denoted with hats. The next step involves a decomposition of the extrinsic curvature into its trace and trace-free parts,

$$K_{ij} = A_{ij} + \frac{1}{3} \delta_{ij} K,$$

(10)

followed by a conformal transformation of the trace-free part of the extrinsic curvature,

$$A_{ij} = \phi^{-2} \hat{A}_{ij}.$$

(11)

Finally, $\hat{A}_{ij}$ is split into its transverse and longitudinal parts,

$$\hat{A}^{ij} = \hat{A}^{ij}_t + (\hat{I} \hat{W})^{ij},$$

(12)

where by construction

$$\nabla_j \hat{A}^{ij}_t = 0,$$

(13)

and

$$(I \hat{W})^{ij} = \nabla^i W^j + \nabla^j W^i - \frac{2}{3} \delta^{ij} \nabla_k W^k.$$  

(14)

With the above transformations and decompositions, the Hamiltonian and momentum constraints take the form

$$8\Delta \phi = \hat{R}\phi - (\hat{A}_{ij} \hat{A}^{ij})\phi^{-7} + \frac{2}{3} K^2 \phi^{-5} - 16\pi \hat{\rho}_H \phi^{-3},$$

(15)

$$(\hat{\Delta}_i W)^i = \frac{2}{3} \phi^{-7} \nabla^i K + 8\pi J^i,$$

(16)

respectively, where $(\hat{\Delta}_i W)^i \equiv \nabla_j (\hat{I} \hat{W})^{ij}$ and $\hat{\Delta} \equiv \nabla^i \nabla_i$. The transformation rules for the sources are $\rho_H = \hat{\rho}_H \phi^{-8}$ and $J^i = \hat{J}^i \phi^{-10}$.

For the numerical methods discussed in the following sections, we restrict our attention to the construction of initial data for the collision of black holes. In this case, the matter sources vanish; furthermore, it is usually assumed that initially the extrinsic curvature vanishes (maximal embedding) and the hypersurface is conformally flat. Under these conditions, the constraint equations (15) and (16) take the form

$$8\Delta \phi = - (\hat{A}_{ij} \hat{A}^{ij})\phi^{-7},$$

(17)

$$\nabla_j \hat{A}^{ij} = 0.$$  

(18)

An immediate consequence of $K = 0$ is that the constraint equations decouple. One can then show [16, 15] that a solution to equation (18) for $N$ black holes, each with linear momentum $P^o$ and angular momentum $S^o$, is given by

$$A_{ij} = \frac{3}{2} \sum_{\alpha=1}^N \left\{ \frac{1}{r^3_{\alpha}} \left[ P_{\alpha j} n_{\alpha}^o + P_{\alpha k} n_{\alpha}^o + (n_{\alpha j} - n_{\alpha}^o n_{\alpha}^o) P_{\alpha k} n_{\alpha}^o \right] \\
+ \frac{1}{r^3_{\alpha}} \left[ \epsilon_{kij} S_{\alpha}^i n_{\alpha}^o n_{\alpha}^j + \epsilon_{kij} \hat{S}_{\alpha}^i n_{\alpha}^o n_{\alpha}^j \right] \right\},$$

(19)
where \( n^\alpha_i \) is the unit normal at the throat of the \( \alpha \)-th black hole, \( r_\alpha \) the distance to the center of the \( \alpha \)-th hole, and \( n_{ij} \) is the flat-metric.

In order to solve the Hamiltonian constraint (17), we need to specify boundary conditions at the throats of the black holes and at the outer boundaries. The condition on the conformal factor \( \phi \) at large distance from the holes (asymptotic flatness) is the Robin condition [17]:

\[
\frac{\partial \phi}{\partial r} = \frac{1 - \phi}{r}.
\]  

(20)

On the other hand, inversion symmetry requires that at the throat of the holes [18] the conformal factor satisfies

\[
n^\alpha_i \nabla_i \phi + \frac{\phi}{2r_\alpha} = 0.
\]  

(21)

3.1 Multiquadrics

The multiquadrics (MQ) approximation scheme\(^2\) is based on the principle that a continuous function, \( f(x) \), may be written as

\[
f(x) = \sum_{i=1}^{N} a_i g_i(x - x_i),
\]  

(22)

where

\[
g_i(x - x_i) = [d_i^2(x - x_i) + s^2]^{1/2}
\]  

(23)

are \( N \) continuously differentiable and integrable basis functions. The parameter \( s^2 \) is an input parameter that controls the shape of the basis functions, \( x_i \) are \( N \) specified data points, and

\[
d_i^2(x - x_i) = (z - z_i)^2 + (y - y_i)^2 + (z - z_i)^2.
\]  

(24)

If \( N \) values of \( f(x) \) are known at \( \{x_j\}_{j=1..N} \), the standard procedure is then to solve for the coefficients \( a_i \) from the linear set of equations that one obtains after substituting \( f(x_j) \) into equation (22). Once the coefficients \( a_i \) are found, equation (22) provides an approximation of \( f(x) \) at any location inside the computational domain. Experiments show that the MQ interpolator (22) performs particularly well when the function to be approximated is steep [20]. Tarwater [21] found that the root-mean-square error in MQ is a function of \( s^2 \). The error decreases when \( s^2 \) is increased, until it reaches a minimum; beyond that value, the error grows with the increase of \( s^2 \). Kansa [22] showed that the accuracy of MQ gets considerably enhanced if \( s^2 \) is made basis dependent. Having a "local"

\(^2\)Results presented in this section are based on work by Mark Dubal [19].
The $s^2$ parameter yields a diverse family of basis function shapes, allowing a wider range of functions that can be approximated. A convenient choice is

$$s^2_i = s^2_{\text{min}} \left( \frac{s^2_{\text{max}}}{s^2_{\text{min}}} \right)^{(i-1)/(N-1)}.$$  \hspace{1cm} (25)

with $s_{\text{max}}/s_{\text{min}}$ input parameters. Experiments show that the ordering for labeling the data points, and consequently the $s^2_i$ parameters, has no significant impact on the accuracy of the solutions.

So far, we have only addressed the issue of approximating functions via MQ. The application of MQ to solve boundary value problems such as the initial value problem in General Relativity proceeds as follows [19, 22]: In Cartesian coordinates, one uses MQ given by (22) with an appended constant; that is,

$$f(x, y, z) = a_1 + \sum_{i=2}^{N} a_i \tilde{g}_i(x - x_i),$$  \hspace{1cm} (26)

where $\tilde{g}_i = g_i - g_1$, and

$$g_i = [(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 + s^2_i]^{1/2},$$  \hspace{1cm} (27)

with $s^2_i$ given by (25). Then the first and second derivatives of $f$ are given by

$$\frac{\partial f}{\partial x} = \sum_{i=2}^{N} a_i \frac{\partial \tilde{g}_i}{\partial x}$$

$$= \sum_{i=2}^{N} a_i \left[ \frac{(x - x_i)}{g_i} - \frac{(x - x_1)}{g_1} \right]$$  \hspace{1cm} (28)

$$\frac{\partial^2 f}{\partial x^2} = \sum_{i=2}^{N} a_i \frac{\partial^2 \tilde{g}_i}{\partial x^2}$$

$$= \sum_{i=2}^{N} a_i \left[ \frac{1}{g_i} - \frac{1}{g_1} - \frac{(x - x_i)^2}{g_i^3} + \frac{(x - x_1)^2}{g_1^3} \right].$$  \hspace{1cm} (29)

respectively. Other derivatives can be constructed in a similar fashion.

The MQ procedure for solving linear boundary value problems starts by scattering $N$ grid points anywhere throughout the computational domain, including the boundaries. Here is where the power of MQ arises: one can accurately represent boundary conditions in irregular domains and set a higher density of grid points in regions where the solution is known to vary more rapidly. Once the grid points are placed, derivatives in the equations and boundary conditions are approximated using the expressions (26), (28), (29) and so on. This substitution leads to a set of $N$ linear algebraic equations, $S \cdot a = b$, for each grid point.
from which the coefficients \( a = \{a_i\}_{i=1,...,N} \) can be obtained. When this procedure is applied to the Hamiltonian constraint, linearization is required since this equation is quasi-linear.

Perhaps, the main reason why MQ techniques have not reach the popularity of other methods has its roots on the numerical difficulties [19] that one faces in solving the MQ matrix system, \( S \cdot a = b \). The MQ matrix \( S \) is a full and highly ill-conditioned \( N \times N \) matrix [23]. This should not be surprising since the method couples each grid point with the remaining \( N - 1 \) points. The end result is that standard methods, such as LU decomposition, are not appropriate for solving the system. Dubal [19] has shown that the MQ matrix system can be successfully solved using a single-value decomposition (SVD) [24]. The SVD method consists of a \( S = U \Lambda V^T \) decomposition, where \( U \) and \( V \) are orthogonal \( N \times N \) matrices, and \( \Lambda \) is a \( N \times N \) diagonal matrix containing the singular values of \( S \). The inverse of \( S \) can be then obtained from \( S^{-1} = V \cdot \text{diag}(1/\Lambda) \cdot U^T \).

Another point that needs to be stressed is that the condition number increases with \( s^2 \). This behavior of the condition number with \( s^2 \) has a direct impact on the limitations of MQ for approximating flat functions. Representing flat surfaces requires flat basis functions. From (23), one sees that to produce flat basis functions, large \( s^2 \) values are needed. The net effect is then an increase on the condition number of the system, yielding large values of \( a_i \)'s with alternating signs.

The problem of representing flat functions using MQ plays a fundamental role in the procedure required for the construction of initial data in black hole collisions. One of the conditions used to derive the Hamiltonian constraint in the form given by (17) was that of asymptotic flatness. When solving (17), MQ yields results plagued with inaccuracies since the conformal factor \( \phi \) is required to be flat (\( \phi \rightarrow 1 \)) far away from the holes. A solution to this problem [25] is to make the following substitution

\[
\phi = 1 + \left[ \left( \frac{c_1}{r_1} \right)^3 + \left( \frac{c_2}{r_2} \right)^3 + \left( \frac{c_1 c_2}{r_1^2 r_2^2} \right)^{3/2} \right] \psi, \quad (30)
\]

where \( c_{1,2} \) are the radii of the black holes and \( r_{1,2} \) the distances to the positions of the black holes. With this substitution, equation (17) becomes more cumbersome (see equation (17) in [25]); however, the solution function \( \psi \) in this case has the non-flat asymptotic behavior \( \psi \propto r^2 \) that leads to accurate results.

In Dubal et al. [25], construction of Misner type data [26] was obtained using 2042 particles or grid points; errors of < 1% were obtained when compared with the analytic solution. A direct comparison of MQ with FD methods using Cartesian and Ćadež coordinates was performed by Cook et al. [27]. The FD approach using Ćadež curvilinear coordinates yielded, in spite of the presence of coordinate singularities, the most accurate solutions for a given amount of computational effort. This is understandable since the topology of Ćadež coordinates is such that constant-coordinate surfaces coincide with the throats of
the black holes. It is unlikely that such coordinate systems can be found during the evolution of black hole space-times. On the other hand, MQ and Cartesian FD with a uniform grid produced comparable accuracies, but, as mentioned before, MQ was computationally more expensive.

### 3.2 Finite Elements

Another interesting method that has been used for the construction of initial data in General Relativity is finite elements (FE). FE methods belong to the class known as weighted residual methods (WRM) [29]. WRM assume that there exists an analytic approximation of the solution of the form

$$f(\vec{x}) = \sum_{i=1}^{N} a_i \phi_i(\vec{x}),$$

where $a_i$ are unknown coefficients and $\phi_i(\vec{x})$ known continuously differentiable and integrable basis functions. Let us consider a function $f$ satisfying the differential equation

$$L(\hat{f}) = 0.$$  

The residual of this equation is obtained by substituting the approximate solution (31) into (32); that is,

$$\hat{L}(f) = R.$$  

By considering a sufficiently large number of basis functions, it is possible to obtain coefficients $a_i$ such that the residual is small throughout the computational domain. The goal is then to find the coefficients $a_i$ from the system of equations that arise when the following condition is imposed:

$$\int_{\Omega_k} Rd^3 \vec{x} = 0.$$  

Different choices of weighting functions $\Omega_k$ yield different methods, in particular the FE method.

The Galerkin FE method is based on the following three properties:

1. The weighting functions $\Omega_k$ are selected from the basis functions $\phi_i$, i.e., $\Omega_k = \phi_k$.

2. The coefficients $a_i$ are directly the values of the unknown solution at a given set of grid points, i.e., $a_i = f_i$.

3. The basis functions $\phi_i$ are low-order piecewise polynomials with compact support.

---

\(^3\)Results presented in this section are based on work by Arup Mukherjee and Douglas Arnold [28].
For example, the one-dimensional linear basis functions read

\[
\phi_i = \begin{cases} 
0 & (z < x_{i-1}) \\
\frac{x - x_{i-1}}{x_i - x_{i-1}} & (x_{i-1} \leq z \leq x_i) \\
1 & (z = x_i) \\
\frac{x_{i+1} - x}{x_{i+1} - x_i} & (x_i \leq z \leq x_{i+1}) \\
0 & (z > x_{i+1}).
\end{cases}
\] (35)

The linear approximation within two adjacent elements, \((x_{i-1} \leq z \leq x_i)\) and \((x_i \leq z \leq x_{i+1})\), is then given by

\[
f = \begin{cases} 
f_{i-1} \phi_{i-1} + f_i \phi_i & (x_{i-1} \leq z \leq x_i) \\
f_i \phi_i + f_{i+1} \phi_{i+1} & (x_i \leq z \leq x_{i+1}).
\end{cases}
\] (36)

The general structure of the FE method involves:

1. Construction of a mesh from geometrical elements (e.g. line segments in 1D, triangles in 2D, and tetrahedra in 3D).

2. Piecewise interpolation within the finite elements from information at the nodes, i.e.

\[
f(\bar{x}) = \sum_{\text{FE nodes}} f_i \phi_i(\bar{x}).
\] (37)

3. Derivation of algebraic equations for the node values \(f_i\) from a WRM condition (34).

4. Solution of the system of algebraic equations.

One important feature of FE methods is that the mesh or triangulation is not required to be regular or even uniform; elements can have different sizes and orientations across the computational domain. As a consequence, with FE methods it is not only possible to use computational domains with irregular boundaries, but also study systems in which the solution has steep, as well as shallow, gradients. These are precisely the numerical difficulties that one faces in constructing initial data for the collision of black holes, namely irregular computational domain (boundaries at the hole throats) and steep gradients near the holes and asymptotic flatness far away from the holes.

For most of the three-dimensional systems of interest in Numerical Relativity, the system of algebraic equations for the nodal functions \(f_i\) is huge. A powerful technique to solve such systems is multigrid (for a review see [24]). The use of multigrid methods introduces additional numerical difficulties. An algorithm is needed for the bisection of elements. The bisection procedure needs to be such that the outcome of refinements are conforming meshes. Furthermore, error estimators are required, so a decision can be made regarding which elements need to be subdivided and for how many times.
Mukherjee and Arnold [28] have developed a 3D FE/Multigrid code to solve the Hamiltonian constraint for multiple black holes systems. They have been able to obtain 3D Misner data solutions using ~ 1500 nodes with errors < 2% in computational domains with the outer boundary at 100M (M the mass of the holes). The code exhibits second-order convergence and is capable of finding solutions to the Hamiltonian constraint using up to 30,000 nodes in ~ 100 seconds in DECalpha workstations.

4 Matter Evolution in Curved Space-times

Few attempts have been made to solve the evolution equations for the dynamics of the geometry of space-like hypersurfaces using methods other than FD. Mann [30] used a mixed FE method (FD in time + FE in space) for the relativistic spherical collapse. In the following two sections, I will discuss two particle methods for the evolution of fluid matter fields in curved space-times. These methods may be used to solve the Einstein-hydro equations in which a hybrid approach is followed: the geometry \((\alpha, \beta, g_{ij}, K_{ij})\) is computed using FD; on the other hand, the matter fields are evolved following a particle approach. An important ingredient in these hybrid or mixed approaches is the communication between the geometry and matter sectors; that is, special care must be paid to interpolations between the mesh, where the metric variables \((\alpha, \beta, g_{ij}, K_{ij})\) are defined, and the particles that carry matter attributes.

4.1 Particle-Mesh

The dynamical evolution of a collisionless gas of particles in General Relativity has interesting application in cosmology and astrophysics [31]. In general terms, the problem consists of solving the collisionless Boltzmann equation in General Relativity.

The starting point is the construction of a stress-energy tensor consisting of the sum of the stress-energy tensors for each particle

\[
T^{\mu\nu} = \sum_A m_A n_A U_A^\mu U_A^\nu ,
\]

where \(m\) is the rest mass of the particles, \(n\) is the number density in the comoving frame and \(U^\mu\) is the 4-velocity of each particle with \(U^\mu U_\mu = -1\). The subscript \((A)\) labels particle categories. For collisionless systems, the equations governing the dynamics of the matter fields reduce to

\[
U^\mu \nabla_\mu U^\nu = 0 ,
\]

the geodesic equation of motion for the particles. From the computational
point of view, it is more convenient to rewrite the geodesic equation (39) as

\[
\delta_t z^i = \frac{\partial^{(4)} U_i}{U^4}, \tag{40}
\]

\[
\delta_t U_i = -\frac{U_\alpha U_\beta \partial^{(4)}_\beta g^{\alpha \beta}}{2U^4}, \tag{41}
\]

where \( \delta_t \equiv \partial_t + V^i \partial_i \), \( z^i \) is the coordinate position of each particle, and \( V^\mu = U^\mu / U^4 \) is the "transport" velocity of the particles as measured by the observers at rest with respect to the coordinate grid points. The value of \( U^4 \) is obtained from the normalization condition of the 4-velocity. It is important to point out that one evolves the spatial components of the covariant 4-velocity. This choice simplifies the evolution equations considerably. Given then the components of the metric \( g_{\alpha \beta} \) at any instant of time, one can solve the system of equations (40) and (41) to compute new positions and 4-velocities of the particles.

The collisionless matter sources for the ADM equations are obtained by substituting the expression (38) for \( T_{\mu \nu} \) into equations (4) through (6); this yields

\[
\rho_M = \sum_{mn} mn \gamma^2, \tag{42}
\]

\[
J_i = -\sum_{mn} mn U_i \gamma, \tag{43}
\]

\[
S_{ij} = \sum_{mn} mn U_i U_j, \tag{44}
\]

where \( \gamma \) is given by \( \gamma = -n_\mu U^\mu \).

In Newtonian gravity, particles in collisionless systems have pair-wise interactions via a \( 1/r \) potential field. On the other hand, the situation in General Relativity is considerably more complicated as particles propagate along geodesics through a curved space-time that the particles themselves distort. As mentioned before, in the hybrid approach considered here, the space-time field variables are computed using FD, so particles are pushed around a grid using forces derived from a discrete potential field. It is then natural that standard particle-mesh (PM) methods [32] are easily adapted to solving the geodesic equations of motion (40) and (41).

A direct consequence of this hybrid approach for solving the Einstein-hydro equations is that one requires the source terms (42) through (44) to be defined on a lattice of grid points and the metric "forces" at the particle positions. A procedure to reconstruct grid data from the unstructured particle data and vice versa must be designed. One possibility is to construct mesh quantities from particle quantities with

\[
f(\vec{z}_j) = \sum_{i=1}^{N_p} f(\vec{p}_i) W(|\vec{z}_j - \vec{p}_i|), \tag{45}
\]
where \( \vec{x} \) and \( \vec{p} \) denote the cell center and particle positions, respectively. \( W(|\vec{x} - \vec{p}|) \) is a weighting function that smoothly interpolate the particle data to the grid. For example, the density source (42) defined in the grid can be obtained from particle information as

\[
\rho_H(\vec{x}_j) = \sum_{i=1}^{N_p} m_i \gamma_i^2 W(|\vec{x}_j - \vec{p}_i|).
\]

(46)

Similar procedures may be used for the other sources (43) and (44). For \( W(|\vec{x} - \vec{p}|) \) one uses the Cloud-in-Cell (CIC) formula which assigns a density function that is continuous as a particle moves across a cell boundary. That is,

\[
W(\vec{x}) = w(x)w(y)w(z),
\]

(47)

with

\[
u(x) = 1 - |x - p_x|/L \quad \text{for} \quad |x - p_x| < L,
\]

(48)

and zero otherwise. Similar expression for \( w(y) \) and \( w(z) \) hold.

![Diagram of two-dimensional schematic set up of Particle-Mesh algorithms. Particles have the same shape and size as grid cells but otherwise free to move everywhere in the computational domain. The dark shaded area represents the contribution of the particle to the grid cell.](image)

Figure 2: Two-dimensional schematic set up of Particle-Mesh algorithms. Particles have the same shape and size as grid cells but otherwise free to move everywhere in the computational domain. The dark shaded area represents the contribution of the particle to the grid cell.

An inverse process for reconstructing grid quantities, such as metric functions and their gradients, at the particle locations is also needed. One can use the
same CIC interpolation function \( W(|\vec{x} - \vec{p}|) \) for this assignment procedure; that is, a grid-based function \( f(\vec{x}) \) is assigned at a particle location \( \vec{p}_j \) by means of

\[
f(\vec{p}_j) = \sum_{i=1}^{8} f(\vec{x}_i) W(|\vec{x}_i - \vec{p}_j|), \tag{49}
\]

which is basically the weighted average of the function at the eight corners of the grid cell in which the particle is located (see Fig. 2). However, this computation is expensive when performed over all needed quantities, so one uses instead a Nearest-Grid-Point scheme where

\[
w(x) = 1 \quad \text{for} \quad |x - p_x| \leq \frac{L}{2}. \tag{50}
\]

This speeds the code up by approximately a factor of two, but it generates, however, phase errors of order half the grid spacing when particles approach the cell boundaries.

The first and most extensive PM numerical program for solving Einstein equations in the presence of collisionless matter has been carried out by Shapiro and Teukolsky [31, 33]. They work includes relativistic stellar dynamics [34], collapse of dense star clusters to supermassive black holes [35], axisymmetric star clusters [36], and more recently the formation of naked singularities [37]. Recently, a PM + FD approach has been undertaken by Anninos and Laguna [38] to study the collapse of large scales structures at horizon scales.

### 4.2 Smoothed Particle Hydrodynamics in Curved Space

The variables that completely characterize a general relativistic fluid are: the unit 4-velocity \( U^\mu \) of an observer comoving with the fluid, the baryon rest mass density \( \rho \), the specific internal energy \( \epsilon \), and the isotropic pressure \( P \). Other auxiliary quantities are the energy density \( \rho \epsilon \) and the total inertial–carrying energy density \( \rho \omega \) with \( \omega = 1 + \epsilon + P/\rho \) the relativistic enthalpy. Notice that in general \( U^\mu \neq n^\mu \); that is, the fluid 4-velocity is not necessarily orthogonal to the hypersurfaces of the foliation. The inner product of these two unit timelike vectors yields the relativistic gamma or boost factor \( \gamma \) between Eulerian observers \( (n^\mu) \) and Lagrangian observers \( (U^\mu) \):

\[
\gamma = -n_\mu U^\mu = \alpha U^t = (1 + g^{ij} U_i U_j)^{1/2}, \tag{51}
\]

where \( n_\mu = (-\alpha, 0) \).

The goal is to rewrite the general relativistic hydrodynamic equations in a form resembling their Newtonian counterparts [39]. One starts by defining a 3-dimensional transport velocity,

\[
V^t = \frac{\alpha}{\gamma} U^t. \tag{52}
\]
This is the velocity that will be used to move the particles during the simulation. The equations determining the evolution of the fluid are the following conservation laws:

\begin{align}
\text{Baryon Conservation:} & \quad (4) \nabla_\mu (\rho U^\mu) = 0, \\
\text{Energy-Momentum Conservation:} & \quad (4) \nabla_\mu T^{\mu\nu} = 0.
\end{align}

where \((4) \nabla_\mu\) is the covariant derivative with respect to the 4-metric \((4) g_{\mu\nu}\). In the case of perfect fluids, the stress-energy tensor reads

\[ T_{\mu\nu} = \rho \omega U_\mu U_\nu + P g_{\mu\nu}. \]

Using equation (52), the equation of baryon conservation can be written as

\[ (4) \nabla_\mu (\rho \gamma V^\mu / \alpha) = 0. \]

We now define the red-shifted energy density \(D \equiv \rho \gamma\). Since in particle methods one usually adopts a Lagrangian description, we rewrite the above expression in terms of the Lagrangian derivative

\[ \delta_t \equiv V^\mu \delta_\mu = \partial_t + V^i \partial_i. \]

Equation (56) can then be rewritten as

\[ \delta_t D + D(\nabla_i V^i + \partial_i \ln \sqrt{g}) = 0. \]

The time derivative of the metric determinant can be obtained from the evolution equation for the spatial metric (2); that is,

\[ \partial_t \ln \sqrt{g} = -\alpha K + \nabla_i \beta^i. \]

Substitution of (59) into (58) yields

\[ \delta_t D + D[\nabla_i (V^i + \beta^i) - \alpha K] = 0. \]

Using the stress-energy tensor (55), the energy conservation equation for a perfect fluid is given by:

\[ U^\nu [\nabla_\mu (\rho \omega U^\mu U_\nu) + \partial_\nu P] = 0, \]

which can be rewritten with the help of the baryon conservation equation as

\[ \rho U^\mu \partial_\mu \epsilon + P \nabla_\mu (4) U^{\mu} = 0. \]

Here again, one can use equation (52) to obtain from (62)

\[ \delta_t \epsilon + \left( \frac{\gamma P}{D} \right) [\nabla_i (V^i + \beta^i) - \alpha K - \delta_t \ln \gamma] = 0. \]
Finally, from the divergence of the stress-energy tensor (55), one obtains
\[
\nabla^{(4)}_{\mu} (\rho \omega U^\mu U_\nu) + \partial_\nu P = 0.
\] (64)

Using the baryon conservation equation and examining only the spatial components, the momentum equation can be re-expressed as follows:
\[
\partial_t S_i + \left( \frac{\alpha}{2 \omega \gamma} \right) S_\mu S_\nu \delta^{(4)}_{\mu\nu} g^\rho_\sigma + \frac{\alpha}{D} \partial_\mu P = 0,
\] (65)

where \( S_i \equiv \omega \gamma V_i / \alpha \) is a momentum variable.

To close the system of equations and unknowns, we need to introduce an equation of state. For an ideal gas the equation of state is given by \( P = \epsilon \rho (\Gamma - 1) \), where \( \Gamma \) is the ideal gas adiabatic exponent. In terms of our variables this equation of state reads
\[
P = \epsilon D (\Gamma - 1) / \gamma.
\] (66)

In summary, the hydrodynamic equations for fluid flows in curved spacetimes read

• Baryon conservation:
\[
\partial_t D + D [\nabla_i (V^i + \beta^i) - \alpha K] = 0.
\] (67)

• Energy conservation:
\[
\partial_t \epsilon + \left( \frac{\gamma P}{D} \right) [\nabla_i (\epsilon V^i + \beta^i) - \alpha K - \partial_t \ln \gamma] = 0.
\] (68)

• Momentum conservation:
\[
\partial_t S_i + \left( \frac{\alpha}{2 \omega \gamma} \right) S_\mu S_\nu \delta^{(4)}_{\mu\nu} g^\rho_\sigma + \frac{\alpha}{D} \partial_\mu P = 0.
\] (69)

The Newtonian limit of these equations is obtained by: \( U^\mu \rightarrow (1, V^i) \), \( \gamma \rightarrow 1, \omega \rightarrow 1, D \rightarrow \rho, S_i \rightarrow U_i, \alpha \rightarrow 1, \beta^i \rightarrow 0 \) and \( \sqrt{g} \rightarrow \text{flat space 3-metric determinant} \). From the computational point of view, the three most important differences, besides factors of \( \alpha \) and \( \gamma \), between these general relativistic hydrodynamic equations and their Newtonian counterparts are:

1. The time derivative of the relativistic boost factor, \( \partial_t \ln \gamma \), appears in the right hand side of the energy conservation equation. Handling this term presents some difficulty for time-explicit integration. Usually, one introduces a modified version of "operator splitting." For each time step, one uses equation (68) without the \( \partial_t \ln \gamma \) to find a first approximation to the specific energy density \( \epsilon \). This value of \( \epsilon \) is then used, in conjunction with the velocity re-normalization condition \( (U^\mu U_\mu = -1) \), to obtain an updated value of \( \gamma \). A time derivative of \( \gamma \) can then be computed which is used to finally update the specific energy density \( \epsilon \).
2. Gravitational forces on the fluid are computed via the space-time metric (see terms with and $\delta^{(4)}_{ij} g^{jk}$ in the momentum equation). If the space-time metric is a known analytic function, this term has a straightforward implementation.

3. The velocity of the fluid, used for moving the particles, is obtained from the *momentum* $S_i = \omega \gamma V_i / \alpha$. That is, the velocity of the particles is derived from $V^i = \alpha g^{ij} S_j / \omega \gamma$

Smoothed particle hydrodynamics (SPH) is playing a fundamental role in astrophysics and cosmology because of its intrinsic adaptive nature and simple implementation (for a review see [40]). SPH discretization of the relativistic hydrodynamic equations (67-69) will be derived under the assumption that the metric field variables ($\alpha, \beta, g_{ij}, K_{ij}$) are computed in a grid using FD and interpolated to the location of the particles. Given a function $f(x)$, its mean smoothed value $\langle f(x) \rangle$ can be obtained from

$$
\langle f(x) \rangle = \int W(x, x'; h) f(x') \sqrt{g} \, d^3 x', \quad (70)
$$

where $W$ is the kernel, $h$ the smoothing length and $\sqrt{g} \, d^3 x'$ the volume element in the space-like hypersurface with intrinsic metric $g_{ij}$. One can show [41] that this smoothing procedure is second-order accurate in $h$. For the particular case of spherically symmetric kernels in flat space, the truncation error introduced by the smoothing is

$$
\langle f(x) \rangle = f(x) + \alpha \ h^2 \Delta f + O(h^3), \quad (71)
$$

where

$$
\alpha = \int W(|x - x'|; h) |x - x'|^2 h^3 \, d^3 x' \quad (72)
$$

is independent of $h$; in the one-dimensional case, $\alpha = 1/4$.

For consistency, a normalization condition on $W$ is required:

$$
\int W(x, x'; h) \sqrt{g} \, d^3 x' = 1. \quad (73)
$$

To satisfy the above condition, a convenient choice of the kernel is

$$
W(x, x'; h) = \xi(x) \Omega(v), \quad (74)
$$

where $v = |x - x'| / h$. $\Omega(v)$ is any of the standard spherical kernels, for instance, the spline kernel [42]

$$
\Omega(v) = \frac{1}{\pi h^3} \begin{cases} 
1 - \frac{3}{2} v^2 + \frac{3}{4} v^3 & (0 \leq v \leq 1) \\
\frac{1}{4} (2 - v)^3 & (1 \leq v \leq 2), \\
0 & (2 \leq v)
\end{cases} \quad (75)
$$
By construction, this kernel satisfies
\[ \int \Omega(v) \sqrt{\eta} \, d^3 x = 1, \tag{76} \]
where $\eta$ is the flat, spatial metric determinant. Thus, from (74) and (73),
\[ \xi^{-1}(x) = \int \Omega(v) \sqrt{g'} \, d^3 x' \tag{77} \]
is the normalization function of the kernel $\Omega(v)$ in curved space.

Using (70), the smooth approximation of gradients of scalar functions reads
\[
\langle \nabla f(x) \rangle = \int W(x, x'; h) \nabla' f(x') \sqrt{g'} \, d^3 x' \\
= \xi(x) \int \Omega(v) \nabla' f(x') \sqrt{g'} \, d^3 x'.
\]
Integration by parts yields,
\[
\langle \nabla f(x) \rangle = -\xi(x) \int f(x') \nabla' \Omega(v) \sqrt{g'} \, d^3 x' \\
= \xi(x) \int f(x') \nabla \Omega(v) \sqrt{g'} \, d^3 x'. \tag{78}
\]
In deriving (78), surface terms were ignored, and the assumption that the kernel $\Omega(v)$ is spherical, $\nabla' \Omega(x, x'; h) = -\nabla \Omega(x, x'; h)$, was used. Finally, equation (78) can be rewritten as
\[
\langle \nabla f(x) \rangle = \nabla \int f(x') W(x, x'; h) \sqrt{g'} \, d^3 x' \\
- \nabla \ln \xi(x) \int f(x') W(x, x'; h) \sqrt{g'} \, d^3 x';
\]
that is,
\[ \langle \nabla f(x) \rangle = \nabla \langle f(x) \rangle - \langle f(x) \rangle \nabla \ln \xi(x). \tag{79} \]

Similarly, the smooth approximation of the divergence of a vector reads
\[ \langle \nabla \cdot A(x) \rangle = \nabla \cdot \langle A(x) \rangle - \langle A(x) \rangle \cdot \nabla \ln \xi(x). \tag{80} \]

The next step is to obtain discrete representations of the smoothing procedures (70), (79) and (80). This is accomplished by introducing the number density distribution of particles
\[ \langle n(x) \rangle \equiv \sum_{a=1}^{N} \delta(x - x_a)/\sqrt{g}, \tag{81} \]
with \( \{ x_a \}_{a=1,...,N} \) the collection of \( N \)-points (particles) where the functions are known. With the help of (81), one can now evaluate the integrands in (70), (79) and (80) and obtain

\[
(f(x_a)) = \xi(x_a) \sum_{b=1}^{N} \frac{f(x_b)}{\langle n(x_b) \rangle} \Omega_{ab} 
\]

(82)

\[
(\nabla f(x_a)) = \xi(x_a) \sum_{b=1}^{N} \frac{f(x_b)}{\langle n(x_b) \rangle} \nabla x_b \Omega_{ab} 
\]

(83)

\[
(\nabla \cdot A(x_a)) = \xi(x_a) \sum_{b=1}^{N} \frac{A(x_b)}{\langle n(x_b) \rangle} \cdot \nabla x_b \Omega_{ab} 
\]

(84)

respectively, where \( \Omega_{ab} \equiv \Omega(x_a, x_b; h) \). It is important to point out that these discrete approximations contain two sources of truncation errors. One error arises from the smoothing procedure (71), and the other error is introduced when the integrals in the smoothing procedure are approximated by discrete sums.

Finally, the approximations (82-84) can be used to derive the SPH version of the relativistic hydrodynamic equations (67-69). These equations read [47]

- **Baryon conservation:**

\[
\delta_t D_a = - \sum_{b=1}^{N} m_b \xi_a (\bar{V}_b + \bar{\beta}_a) \cdot \nabla_a \Omega_{ab} + \alpha_a K_a D_a. 
\]

(85)

- **Energy conservation:**

\[
\delta_t e_a = - \left( \frac{\gamma_a P_a}{D_a} \right) \left[ \sum_{b=1}^{N} \left( \frac{m_b \xi_b}{D_b} \right) (\bar{V}_b + \bar{\beta}_a) \cdot \nabla_a \Omega_{ab} \right. 
\]

\[
- \left. \alpha_a K_a - \delta_t \ln \gamma_a \right]. 
\]

(86)

- **Momentum conservation:**

\[
\delta_t S_a (\bar{Z}_a) = - \left( \frac{\alpha_a}{2 \omega a \gamma_a} \right) [S_a \delta_i^{(4)} g^{\mu \nu}]_a 
\]

\[
- \alpha_a \sum_{b=1}^{N} m_b \xi_a \left( \frac{P_a}{D_a^2} + \frac{P_b}{D_b^2} \right) \nabla_a \Omega_{ab}, 
\]

(87)

where \( A_a \equiv \langle A(\bar{Z}_a) \rangle \), and \( \nabla_a \) denotes gradient \( \partial \) with respect to \( \bar{Z}_a \). In deriving (85-87), the number density distribution was rewritten as \( \langle n(x_a) \rangle = D_a / m_a \), where \( m_a \) represents the rest mass of the fluid particles.
As with $N$-body algorithms, SPH exhibits pair-wise particle interactions. The interactions, though, are only between particles (neighbors) inside the compact support of the kernel; after all, hydrodynamics involves only local or contact interactions. Without any information regarding the particles that are neighbors to each particle, SPH calculations require, in principle, computational work of $O(N^2)$, an $N-1$ search for each of the $N$ particles in the system. It is then crucial to find a computationally efficient method to search for the neighbors to each particle. Once these neighbors are known, updating a system of $N$ particles is a $O(N_{h} N)$ procedure, with $N_{h} << N$ and $N_{h}$ the characteristic number of neighbors of the particles. Thanks to the development of hierarchical tree methods, the problem of finding neighbors has been transformed from a $O(N^2)$ to $O(N \ln N)$ operation [43, 44], making possible these days simulations with up to $10^7$ particles in massively parallel supercomputers.

One drawback of SPH arises when dealing with systems requiring a higher resolution along particular directions, e.g. thin slab symmetries. This situations could lead to expensive and inefficient calculations since the interpolating functions (kernels) in SPH are usually spherically symmetric. There have been recently attempts to design a truly adaptive SPH in which the interpolant kernels are non-spherical, so they adapt and change shape to achieve higher resolution along particular directions, without increasing the overhead in the other directions [45, 46].

The first application of SPH in curved space addressed the problem of tidal disruptions of stars by supermassive black holes [48]. In this problem, the curvature of space-time is completely dominated by the black hole. One can then treat the system as fluid motion (star) in a fixed, curved background (black hole). At every step in the evolution, the metric variables, and consequently the forces on the fluid elements, are known analytic functions. Currently, there is an effort [38] to include an SPH approach in the simulation of dynamical space-time systems coupled to fluid flows.

Conclusions

The main goal of these lectures was to review some of the interesting, non-traditional approaches to numerical relativity. Four numerical techniques were presented, two for the construction of initial data of space-times containing black holes and two for evolving matter fields in curved space-times.

The MQ methods constitute a powerful tool to achieve high resolution at low cost (number of grid points); however, it has the serious drawback of requiring the solution of an algebraic system of equations in which the coefficient matrix is ill-conditioned. Perhaps this difficulties can be overcome by reformulating MQ including domain decomposition and blending techniques [49].

FE techniques have been recently the focus of considerable attention for solving the constraint equations in General Relativity. When combined with
multigrid solvers. FE algorithms clearly constitute a powerful alternative to FD methods because of their natural adaptive mesh nature. It is important to point out, however, that a key ingredient required when dealing with FE methods is the design of a suitable element bisection algorithm that preserves the quality of the mesh at all refinement levels.

Finally, PM and SPH methods were reviewed as tools for solving Einstein’s equations in the presence of collisionless matter or fluid flows. The basic idea of these hybrid methods is that the geometry is computed via FD and the matter using a particle approach. In this construction, careful attention must be paid at the interpolation between particle and mesh based quantities. An interesting generalization to these methods would be to construct a “pure” particle method in which the gravitational fields are also computed at the particle locations.

Acknowledgment

I thank Peter Anninos, Matt Choptuik, Mark Dubal, Richard Matzner, Arup Mukherjee and Philippos Papadopoulos for numerous discussions and helpful suggestions. Work supported in part by NSF grants Nos. PHY-9357219 (NY1), PHY-9309834 and PHY/ASC93-18152 (arpa supplemented).

References


23


24