PART A

THEORETICAL APPROACHES

NUMERICAL RELATIVITY ON A TRANSPUTER ARRAY

Ray d'Inverno

Faculty of Mathematical Studies, University of Southampton, Southampton, UK

Abstract. The area of numerical relativity is briefly reviewed and its status in general relativity is considered. The 3+1 and 2+2 approaches to the initial value problem in general relativity are described and compared. A 2+2 approach based on null cones emanating from a central timelike geodesic, together with an implementation on a transputer array is discussed.

1 EXACT SOLUTIONS OF EINSTEIN'S FIELD EQUATIONS

In 1915 Einstein proposed his field equations for the gravitational field

$$G_{ab} = \kappa T_{ab} \tag{1}$$

where G_{ab} is the Einstein tensor for the gravitational field with metric g_{ab}, κ is a coupling constant and T_{ab} is the energy-momentum tensor for any matter field present (d'Inverno (1992)). In the absence of matter fields, the equations reduce to the vacuum equations

$$G_{ab} = 0. \tag{2}$$

Einstein always considered the vacuum equations as being more fundamental in character. They may be viewed as second order non-linear partial differential equations for the metric potentials g_{ab}. The non-linearity means that the equations are difficult to solve, indeed Einstein originally thought that it would not be possible to solve them exactly. For example, they do not satisfy a superposition principle, and so complicated configurations cannot be analysed in terms of simpler constituent ones. It came as something of a surprise when Schwarzschild discovered an exact solution in 1916.

In the ensuing decades there were relatively few exact solutions discovered. However, the invariant techniques of the Petrov classification, optical scalars and Killing vectors led, in the 1960's, to the discovery of numerous exact solutions. It is difficult to count the number of known exact solutions, because many depend on parameters, or on solutions of subsidiary ordinary or partial differential equations. However, the number of authors involved in the discovery of exact solutions is certainly well into four figures. The area of exact solutions was for a long time a confused one until, in 1980, considerable progress was made in attempting to put known solutions into some sort of systematic framework through the publication of the exact solutions book (Kramer et al (1980)).

There are two major problems associated with this field, one practical and the other theoretical. The first concerns the problems associated with the horrendous algebraic calculations involved in work with exact solutions, especially when carried out by hand. The second involves the well-known equivalence problem: given two geometries g_1 and g_2, are they in fact different or is there a coordinate transformation which (locally) transforms one into the other? Significant advances with both these problems were made with the advent of computer algebra systems, some of which were specifically designed for the metric calculations involved in general relativity. Perhaps the best known and most used system in general relativity is the system Sheep (Frick (1977), d'Inverno and Frick (1982)). The power of a system like Sheep is that calculations can be undertaken which would have taken lifetimes to complete by hand. Moreover, the results are error-free.

The theoretical advance came with the discovery of the Karlhede algorithm for classifying a geometry (Karlhede (1980)). This is essentially achieved by introducing a frame in which the Riemann tensor and its covariant derivatives take on canonical forms. This classification can be undertaken, essentially automatically, in the extension of Sheep called Classi. Then, given two geometries, if the classifications are different then so are the geometries; if they are the same them they are candidates for identification. A search is then made for a coordinate transformation which can map one geometry into the other. This last procedure reduces, in general, to solving four algebraic equations, a process which is not algorithmic but which is often manageable in practise. The Karlhede classification program has led to the establishment of the computer database project, a joint research effort aimed at classifying and documenting all known exact solutions (Åman et al (1985)). The first aspiration of the project is to put all the solutions of the exact solutions book into the database. At present, several hundred solutions have been classified. The ultimate hope is that the database will be freely accessible to the scientific community and continuously kept up to date. Then any newly discovered solution can be checked out against the data base which can be updated if the solution is genuinely new.

The database project holds out the prospect of placing the field of exact solutions onto a much more coherent basis. Unfortunately, although this large number of exact solutions exist, very few would appear to be physically realistic or even approximately so. As is well known, partial differential equations admit large classes of solutions, many of which are pathological in nature. One usually has to apply boundary conditions or initial conditions to pick out the solutions which are of physical interest.

Apart from black hole, cosmological and plane wave solutions, the likelihood is that the remaining solutions are indeed pathological in nature. More importantly, we do not possess exact solutions corresponding to or approximating to important physical scenarios such as a 2-body system, an n-body system, a radiative source, the interior of a rotating object undergoing gravitational collapse, and so on. Yet these are precisely the objects that are of interest to us, especially on an astrophysical scale. This is where, I believe, numerical relativity comes in.

2 NUMERICAL RELATIVITY

Numerical Relativity consists of solving Einstein's equations numerically on a computer. The standard scenario is to specify the 3-metric ${}^{(3)}g$ — the intrinsic geometry — of some spacelike slice ($t = t_1 =$ constant, say), and use the field equations to compute the 3-metric at some future time ($t = t_2 > t_1$). The significance of being able to do this is that we can thereby model physically interesting scenarios. Indeed, given the freedom to vary the initial configuration, we can consider the resulting numerical simulations as being in the arena of experimental relativity. This significance will become more pronounced when the long awaited detection of gravitational waves is at last reported and we move into the era of gravitational astronomy. The need will then arise of finding theoretical justifications for actual observations, and this need will likely push numerical relativity into the forefront of general relativity.

There are, in essence, three distinct approaches to numerical relativity: the 3+1 approach, the 2+2 approach and the Regge calculus. These proceedings are largely concerned with the first two approaches, and so we shall overview them briefly in turn.

3 THE 3+1 APPROACH

The basis of this approach is to decompose 4-dimensional space-time into families of 3-dimensional spacelike hypersurfaces and 1-dimensional timelike lines (see, for example, article of York in Smarr (1979)). In more geometrical language, space-time is decomposed into a spacelike foliation and a (transvecting) timelike fibration (figure 1). We can introduce a constructive procedure for generating the decomposition if we start off with a 4-dimensional manifold possessing no metrical or affine structure on it — a so-called bare manifold — and prescribe on it a vector field which transvects some 3-dimensional submanifold i.e. the vector field nowhere lies in the submanifold (figure 2). We then use the vector field to propagate the initial submanifold or hypersurface into a family of hypersurfaces (technically by Lie dragging). The standard initial value problem (or IVP for short), sometimes called the Cauchy IVP, consists of specifying a positive definite 3-metric on the initial hypersurface Σ_0 and then using the vacuum field equations to determine the 3-geometries on successive hypersurfaces Σ_t, say.

There is an analogous initial value problem when matter fields are present.

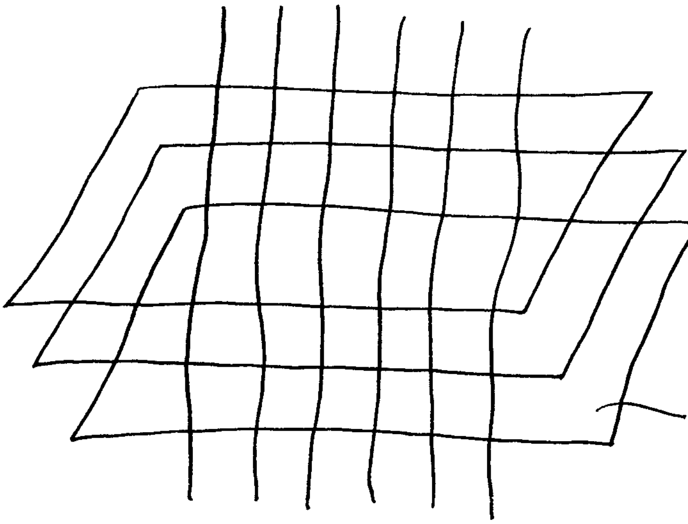

Figure 1. Spacelike foliation and timelike fibration.

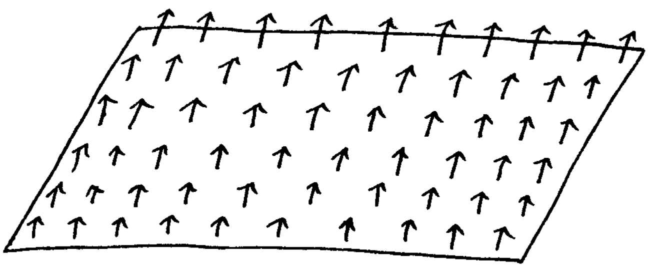

Figure 2. 3-dimensional submanifold and transvecting vector field.

The next step is to introduce some formalism which is adapted to the 3+1 decomposition. The most elegant method is to work with a 4-dimensional formalism, that is one which is manifestly covariant, coupled with projection operators to accomplish the decomposition and Lie derivatives to accomplish the propagation. In this overview, we shall simply use adapted coordinates in which the hypersurfaces have equation $x^0 = t =$ constant, and possess intrinsic coordinates x^α (where latin indices run from 0 to 3 and greek from 1 to 3). Then Lichnerowicz has shown that the vacuum field equations are equivalent to six evolution equations (d'Inverno (1992))

$$R_{\alpha\beta} = 0 \tag{3}$$

and four constraint equations

$$G^0{}_a = 0. \tag{4}$$

If these constraint equations hold everywhere on Σ_0 and the evolution equations hold everywhere, then the contracted Bianchi identities reveal that the constraint equations are automatically satisfied everywhere. A typical numerical scheme consists of taking the dynamical variables to be the 12 variables consisting of $g_{\alpha\beta}$, the 3-metric, and $K_{\alpha\beta}$, its extrinsic curvature, which is defined as (apart from an unimportant numerical factor)

$$K_{\alpha\beta} = \dot{g}_{\alpha\beta} \tag{5}$$

where a dot denotes differentiation with respect to time. The evolution equations then reduce to the first order propagation equations

$$\dot{g}_{\alpha\beta} = \text{known} \tag{6}$$

$$\dot{K}_{\alpha\beta} = \text{known} \tag{7}$$

where the right hand sides are known functions of $g_{\alpha\beta}$, $K_{\alpha\beta}$ and spatial derivatives. The initial data consists of prescribing $g_{\alpha\beta}$ and $K_{\alpha\beta}$ on Σ_0, and use of the propagation equations means that both these quantitites are known on the next neighbouring hypersurface. By taking the time derivatives of equations (6) and (7) we can repeat the process on the next neighbouring hypersurface. Proceeding in this way, we obtain an iterative procedure for generating a solution forward in time. In a numerical regime the derivatives are obtained by a finite difference procedure. There are many variants to this approach, but this short description should serve to illustrate the essential characteristics of the standard approach.

There are two main problems associated with this approach. First of all, the initial data is not freely specifiable but must satisfy the constraints initially. These can be decomposed into the Hamiltonian constraint

$$^{(3)}R - K^{\alpha\beta}K_{\alpha\beta} + (K^{\alpha}{}_{\alpha})^2 = 0 \tag{8}$$

and the momentum constraint

$$^{(3)}\nabla_{\alpha}(K^{\alpha}{}_{\beta} - K^{\gamma}{}_{\gamma}\delta^{\alpha}{}_{\beta}) = 0. \tag{9}$$

This problem can be resolved by extracting a conformal factor from the 3-geometry and investigating the resulting elliptic partial differential equations. This reveals that the gravitational field possesses two true degrees of freedom, namely, in Hamiltonian language, two coordinates (associated with the $g_{\alpha\beta}$) and two momenta (associated with $K_{\alpha\beta}$). The second problem relates to the extent of the development. Although there are existence theorems which say that a solution can be generated for some finite time to the future of the initial slice, they do not indicate how far this may be. Moreover, the approach fails if the foliation goes null. Yet null foliations are important in there own right as we shall next see.

4 THE 2+2 APPROACH

The basis of this approach is to decompose space-time into two families of space-like 2-surfaces. We can view this as a constructive procedure in which an initial 2-dimensional submanifold S_0 is chosen in a bare manifold, together with two vector fields v_1 and v_2 which transvect the submanifold everywhere (figure 3).

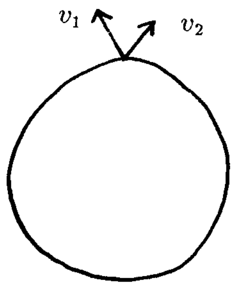

Figure 3. 2-dimensional submanifold and two transvecting vector fields.

The two vector fields can then be used to drag the initial 2-surface out into two foliations of 3-surfaces. The character of these 3-surfaces will depend in turn on the character of the two vector fields. The most important cases are when at least one of the vector fields is taken to be null. For example, if the two vector fields are null, then they give rise to a double-null foliation (indicated schematically in figure 4).

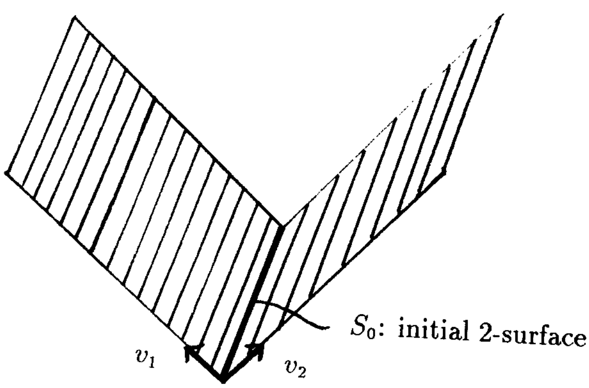

Figure 4. Double-null foliation.

Or if one is null and the other is timelike, this gives rise to a null-timelike foliation (figure 5).

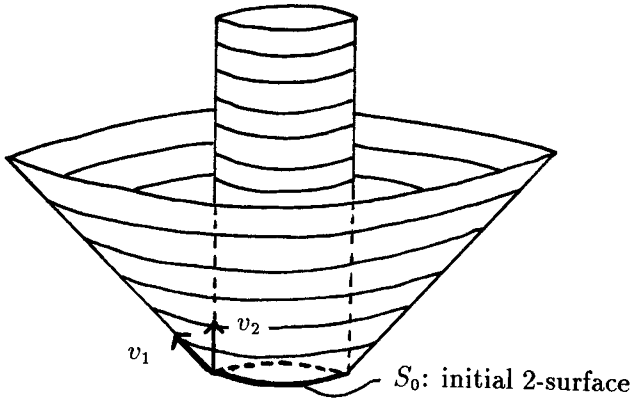

Figure 5. Null-timelike foliation.

The most elegant way of proceeding is to introduce a formalism which is manifestly covariant and which uses projection operators and Lie derivatives associated with the two vector fields. The resulting formalism is called the 2+2 formalism (d'Inverno and Smallwood (1980), d'Inverno (1984)). When the vector fields are of a particular geometric character, then this can be refined further into a 2+(1+1) formalism. Finally, one extracts a conformal factor from the spacelike 2-geometries to isolate the gravitational degrees of freedom.

The 2+2 formalism leads to a number of advantages. First of all, it identifies the two gravitational degrees of freedom as the conformal 2-geometry (d'Inverno and Stachel (1978)). Secondly, this data is unconstrained. Thirdly, the data satisfy ordinary differential equations along the vector fields. Most importantly, unlike the 3+1 approach, the formalism applies to situations where the foliations go null. Such IVPs are called null or characteristic IVPs (CIVPs for short). These are the natural vehicles for studying gravitational radiation problems (since gravitational radiation propagates along null geodesics), asymptotics of isolated systems (since I^+ and I^- are null hypersurfaces) and problems in cosmology (since we gain information about the universe along our past null cone).

The null or characteristic approach, however, suffers from one main drawback resulting from the fact that, in general, null hypersurfaces develop caustics. There are then two quite distinct ways of proceeding. One approach is to develop techniques for generating solutions through caustics (Corkhill and Stewart (1983)). The other is to restrict attention to caustic-free scenarios. This is possible by considering models which are close to spherical symmetry and where it can be proved rigorously that caustics will not occur. This assumption still allows the investigation of an important class of models including stellar collapse, oscillations and supernovae.

5 THE STATUS OF NUMERICAL RELATIVITY

In numerical relativity, we are concerned with dynamical situations, and hence it is conventional to refer to the dimension of a problem in terms of its spatial dimension. Some pioneering work was carried out in the mid sixties involving spherical collapse by May and White, and Hahn and Lindquist. However, the field really only came of age in the mid seventies with the work of Smarr on two black hole collisions (Smarr (1977)) There are now a large number of successful codes in existence including codes in spherical collapse, dust collapse, 2-dimensional black hole collision, 2-dimensional axisymmetric neutrino star bounce, Brill waves, Teukolsky waves, planar symmetry solutions, colliding gravitational waves, cylindrically symmetric solutions, accretion disks, shock waves, inflationary cosmolgy, n-body calculations, collapse of massless scalar fields, evolution of 3-dimensional wave packets and 3-dimensional relativistic hydrodynamics. Most of the fully 3-dimensional work undertaken to date has involved Newtonian models of one sort or another. We are just on the verge of 3-dimensional relativistic codes; indeed some of these codes are reported elsewhere in this volume. These codes will make enormous demands (by present standards) on computer time and memory.

There are a number of problems associated with numerical relativity. The main one relates to the role of the constraint equations. The finite difference version of Einstein's equations leads to an overdetermined system in which the constraints are either ignored (free evolution) or artificially imposed. In the latter case, one method involves imposing the constraints after finite intervals of times (chopped evolution) and another is to impose them at every stage of integration (fully constrained evolution). Unfortunately, each method has associated drawbacks. For example, computations with particular exact solutions have demonstrated that a free evolution drifts further away from the true solution as it evolves in time. Similar problems arise with chopped and fully constrained evolutions. Piran has indicated this schematically in figure 6, where the plane represents the subspace of solutions which satisfy the constraint equations.

Other problems relate to the finite difference approximation. Unfortunately, there are an infinite number of possible finite element difference schemes, each with its own solution, of which a large number will bear little resemblance to the exact solution of the original equations. This is because of instabilities which arise due to an incorrect discretization of space-time. Even if one is using a stable scheme, another major source of inaccuracy occurs in truncation errors. These latter errors stem from the fact that one is essentially approximating a function by a finite part of a Taylor series

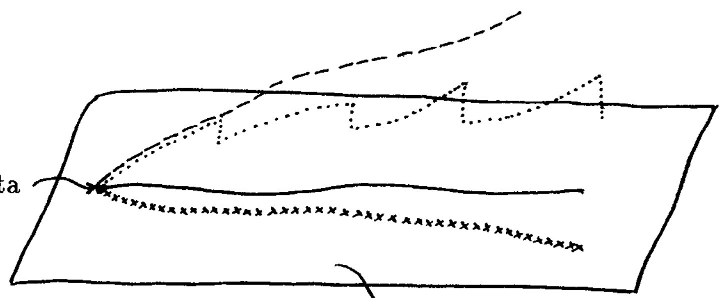

Figure 6. Actual trajectory of solution (solid line); free evolution (dashes); chopped evolution (dots); fully constrained (crosses).

expansion. Other difficulties involve applying appropriate coordinate (or gauge) conditions, coordinate singularities and the boundary conditions associated with the use of a finite numerical grid. Then there is the problem of representing and interpreting the solution: in a 3-dimensional code what quantities should be computed and how should they be displayed? Then, as we have indicated before, numerical relativity makes enormous demands on computer time and memory, which produces limits on what is attainable at any one time.

Another issue is that certain formalisms involve long and complicated algebraic computations leading to lengthy expressions which often require conversion into a particular coding format — a process which could well introduce errors. In 1986, Nakamura used the computer algebra system Reduce to generate such algebraic expressions, and then exploited Reduce's ability to convert algebraic expressions into their Fortran equivalent, prior to numerical computation (Nakamura (1986)). Here, computers are being used for both algebraic and numeric work. Nakamura proposes as a name for this combined area CAR — Computer Aided Relativity.

Historically, relativists were originally distrustful of results emanating from computer algebra systems, because they were not convinced that the results were reliable. It was only after very complicated calculations had been checked successfully against each other using algebra systems based on different machines employing different software and design philosophies, that confidence was eventually established in the tool. A similar problem would seem to apply to numerical relativity. The one thing that you can virtually guarantee about a numerical calculation is that it will produce a result; but is the result correct? The field is still a young one, perhaps only some fifteen years old, and it is a small, albeit growing, one. Again, it would appear that