Hyperbolic Methods for Einstein's Equations

Oscar A. Reula
Universidad Nacional de Córdoba, FaMAF
Ciudad Universitaria
5000 Córdoba, Argentina
reula@fis.uncor.edu
http://surubi.fis.uncor.edu/reula

Abstract

I review evolutionary aspects of general relativity, in particular those related to the hyperbolic character of the field equations and to the applications or consequences that this property entails. I look at several approaches to obtaining symmetric hyperbolic systems of equations out of Einstein's equations by either removing some gauge freedoms from them, or by considering certain linear combinations of a subset of them.

Article Amendments

Contents

1   Introduction

General Relativity (GR) is a peculiar physical theory. GR is about geometries, and not fields, in a given space–time. That is, the theory's solutions are not metric tensors –or possible other matter tensors– but rather the equivalence class of these tensors under arbitrary smooth relabeling of points in space–time. This peculiarity makes the task of analyzing the dynamics of the theory difficult. One is used to evolving tensor fields, in fact tensor components, in a given coordinate system; while here the extra freedom of the theory makes these components non–unique. The values of some of these components can be given arbitrarily. Only certain relations between them are invariant –and so have a physical meaning. In particular, some components can be made arbitrarily large and rough, while the geometry is, for instance, flat. Thus, it is often hard to see, from just comparing tensor components, whether two solutions, that is two geometries, are close to each other during evolution. To overcome this problem, several proposals have been made to fix the evolution in a unique way and at the same time obtain well behaved solutions. In general, these proposals provide for equation systems equivalent, in a sense to be discussed at length later, to Einstein's equations which are hyperbolic, that is, whose evolution is continuous as a function of the initial data. This property is vital for many applications, ranging from Newtonian approximations to numerical simulations. The aim of this work is to review these proposals, paying special attention to the applications where they have proven fruitful.

1.1   Background and History

Hyperbolicity is a, basically algebraic, condition on the coefficients of a system of partial differential equations which grants that the Cauchy problem for the systems satisfying them is well posed; that is, if appropriate data for that system, in an appropriate hypersurface, are given, then a unique solution can be found in a neighborhood of that hypersurface, and that solution depends continuously, with respect to an appropriate norm, on the values of initial data given. Hyperbolicity naturally captures what one would expects to hold for most fundamental physical systems, since, besides the unique and continuous dependence on the initial data, it implies finite propagation velocities. In § 2 a rather complete description of some hyperbolicity conditions, for there are several variants of them, is given, along with references to modern literature on the topic. Here, and basically in order to provide a more definite idea of the sort of conditions involved, we introduce one of these notions of hyperbolicity, that of a symmetric hyperbolic system. This is the case which most often appears in physical problems:

Definition 1 *Let be a first order system of evolution equations,*

$$u_t = A^a(u)D_a u + B(u)u,$$

where $u = u(x,t)$ indicates a "vector" function of dimension s in $\Re^{n+1}$, u_t its time derivative, $A^a(u)$, a $s \times s$ matrix valued vector, and $B(u)$ a $s \times s$

matrix valued vector, whose components depend smoothly on u, and D_a a partial derivative operator on $\Re^n$. The system is called **symmetric hyperbolic** at a solution u_0 if there exists a neighborhood of u_0 and a smooth, positive definite, symmetric matrix $H(u)$ on it such that:

$$H(u)A^a(u) - A^{a\star}(u)H(u) = 0.$$

This condition ensures that the Cauchy problem is well posed, namely that there exists a time interval $[0, T)$ and a constant $C(T)$ such that if initial data for u is given at $t = 0$, $u(x, 0) = f(x)$, with $f(x)$ close enough to $u_0(x, 0)$ in a certain norm, then for the same norm we have:

$$||u(\cdot, t)|| \leq C||f(\cdot)||$$

A simple example of a symmetric hyperbolic system is the wave equation, see Example I 2.2

Well posedness of the Cauchy problem for Einstein equations was established in the early fifties by Choquet-Bruhat, [13], when the theory of evolutionary partial differential equations matured and general proofs for quasi-linear hyperbolic systems became available. With further refinement of the general theory, it was possible in the early seventies to improve on the result by lowering the minimal differentiability required in the proof, thus allowing more general initial data sets, [42], see also [18] and [16, 27, 28]. These works used the harmonic gauge introduced by Lanczos in the twenties. For a discussion on the harmonic gauge see § 3.1

In the seventies, with a new tool (the Weighted Sobolev Spaces) it became possible to enlarge the development region to contain asymptotically boosted slices relative to the initial slice, if the initial data were in a specific Weighted Sobolev space, [24, 17]. This result had two interesting by-products:

- *a)* It established that the evolution equations preserved the asymptotic decay of the initial data. That is, if they initially were in some weighted space (for some specific weight factor) then, in any other time slice given in the development asserted by the theorem –including boosted ones–, the induced data on it was also in the same weighted space.

- *b)* It established a relation between the size of the initial data in any bounded region and a lower bound for the time of existence of the solution, see [37] for a simple description.

Even more, it was possible to use as the function spaces for the initial data the same weighed spaces in which the constraint equations were solved, thus, for the first time obtaining a global (in space) control on Einstein's equations.

All the power of the techniques used in the above mentioned body of work was not enough to get a result which most people suspected would hold: the existence of complete, asymptotically flat space-times for generic, although small, initial data. An idea which made a breakthrough on this problem was to regularize Einstein's equations in terms of a conformally rescaled metric on the

corresponding conformally compactified space-time. Future and Null infinity are then at a finite distance, so standard, local in time, existence theorems can be used. Pursuing these ideas, Friedrich, [30, 29], was able to craft the full, and therefore non-linear, Einstein equations for the conformal fields into a regular symmetric hyperbolic system, and so to create a formidable tool to study these problems. Earlier results using regularized equations for the conformal metric included the local linear stability of null infinity[1], [38].

The variables used range from frames to Weyl tensor frame components. The regularization of the full equations is complicated and requires appending to the original Einstein equations the Bianchi identities as a new and independent set of variables. *For the first time, the harmonic gauge was not used.* This was also the first time a symmetric system was obtained which was not the mere and standard translation of quasi–linear second order wave equations into a first order system. This tool made it possible to show that, given any smooth and small enough initial data on a hypersurface reaching future null infinity at some cross section, a future development existed along which null infinity, to the future of such a hypersurface and up to future infinity, was included. A limitation of this technique is that initial data are not regular enough at space-like infinity to make the estimates work there, and so a complete, asymptotically flat space-time cannot be obtained, nor even a piece of null infinity starting from generic, although small, initial data on a space-like hypersurface; nevertheless see [25].

A different path was followed by Christodoulou and Klainerman, [23]. They also made use of the detailed structure of Einstein equations, but in physical space-time, to show global existence. The special structure of Einstein equations allows the use of other energy estimates beyond the traditional one. The estimates are *boosted energies*, [48, 49, 47], and are crucial to establishing this global result. Christodoulou and Klainerman did not use a conformal compactification of space-time. They were able to obtain complete asymptotically flat space-times, i.e. asymptotically including space-like regions, out of rather generic initial data in a Cauchy surface reaching that infinity. Christodoulou and Klainerman found that at null infinity the differential structure does not seems to be C^∞. That is, they claimed there are smooth initial data sets whose development is not smooth at null infinity and that only finite differentiability remains. Christodoulou and Klainerman also did not use a harmonic gauge condition. Rather, their strategy was to use the equations for tensorial quantities built out of higher order derivatives of the metric. After obtaining estimates for these tensorial quantities, estimates for the metric and its first derivatives were obtained from elliptic theory, and the maximal slicing condition.

As expected, both methods used detailed properties of the Einstein equations to assert global existence of small data solutions. It is believed that both

---

[1]In this work, the linearized Einstein equations in terms of the conformal metric in a neighborhood of $\mathcal{J}$ were regularized and made hyperbolic. Thus local existence results applied to these equations showed that perturbations with smooth initial data at a hypersurface reaching $\mathcal{J}$ stayed bounded during evolution, and so the asymptotic structure was preserved to the future of that hypersurface.

methods use the same properties, but with different techniques. That is, the property that allows the conformal Einstein equations to be regular should be the same property that allows boosted energies to be estimated. In fact both make estimates in terms of the Bel-Robinson tensor. However, to my knowledge, this has not yet been fully explored.

1.2 Main Subject and Plan of the Review

In the nineties, researchers have realized that other or more general forms of gauge conditions are needed to address some pressing problems, in particular, the numerical studies of regions of high curvature –like black hole formation. One should also mention the study of slow solutions, that is, solutions which weakly depend on the value of the velocity of light –in particular those which have a well behaved Newtonian limit. For the first problem, the issue has been whether the coordinate system chosen, that is, whether the gauge condition, stays well behaved during evolution and does not cause unphysical singularities, singularities which, for the numerical problem, are as bad as real ones. For the second problem, the issue has been whether the gauge is well behaved and does not prevent uniform continuity of slow solutions as the speed of light goes to infinity.

In both cases it is clear that one can choose bad gauges which would make the problem intractable and which nevertheless are genuine smooth coordinate conditions in the whole region of interest.

The rest of this review describes the efforts triggered by the need to deal with the above mentioned problems. An attempt to compare these efforts in all respects appears to be a hopeless task. The variables they use are different. Thus we shall concentrate on revealing the points these theories hold in common and discuss the key properties of each. These properties can vary a lot from one system to another. The properties can be seen as virtues or defects of the system according to the uses to which people put them. For instance, a gauge condition can be incorporated into the systems as:

- *i)* a solution to a hyperbolic equation, and so be incorporated as a part of the hyperbolic system;

- *ii)* a solution to an elliptic equation, given a mixed hyperbolic-elliptic system;

- *iii)* a solution to a parabolic equation, given a mixed hyperbolic-parabolic system;

- *iv)* or it can be given as a fixed function in space-time, chosen by a rule of thumb by looking at the initial data.

Each one of these possibilities can be implemented in a mathematically rigorous manner in most of the general schema proposed. And probably each of them would be of relevance in some specific implementation. The point, in other words, is that it is very difficult to go beyond a superficial or general description

of the methods before some particular, extended, and fruitful use legitimizes the job.

In § 2 I give a short summary of hyperbolic theory, describing the case which is completely understood, namely the constant coefficient case, and mention what of it can be extended to the case of interest here, namely quasi-linear systems. For physicists, this discussion should be a complement to Geroch's lecture notes on symmetric hyperbolic systems [36]. This discussion follows very closely chapter II of [50], see also chapter IV of [39].

In § 3 I first describe the general problem of adapting the theory of hyperbolic systems to general relativity and the gauge issue. A companion to this section is, besides Geroch's lectures, the paper of Friedrich in *Hyperbolic Reductions* [32]. In particular we discuss the standard approach, that is, the harmonic gauge.

In § 4 I present the more recent approaches, and divide them into four classes, according to the type of variables used.

In § 5 I compare the different implementations that have been made of the approaches introduced in the previous sections, and discuss the impact these approaches have had on the problems where they have been applied.

In § 6 I consider the role the constraint equations play in these new systems. In the harmonic gauge, constraint equations become evolution equations. The consistency of the gauge is all that is needed to ensure equivalence between Einstein's equations and the harmonic system. In the new systems, one is not incorporating the constraints, and so one should make sure that, if the constraint equations are satisfied initially, they are satisfied during evolution. I claim that, in the initial value formulation, this follows from uniqueness of solutions to the equations and from the fact that the modified evolution vector fields proposed are tangent to the constraint sub-manifold. I also mention the difficulties that appear when considering an initial-boundary value problem.

2   The Theory of Linear Constant Coefficients Evolution Equations and Generalizations to Quasi–linear Systems

In this section, I summarize the main results of the theory of first order evolutionary partial differential equation systems. I do this by first developing the theory of linear constant coefficients evolution equation systems in $\Re^n$, that is, equations of the type:

$$u_t = P(D)u,$$

where $u = u(x,t)$ indicates a "vector" valued function of dimension s in $\Re^{n+1}$, u_t its time derivative, $P(D)$, a $s \times s$ matrix whose components smoothly depend on: $D_\nu := \frac{\partial^{|\nu|}}{\partial x_1^{\nu_1}...x_n^{\nu_n}}$. For most of the results, no particular form for the dependence of P on D is needed, as long as it is continuous. But for simplicity one can think of P as given by:

$$P(D) := \sum_{|\nu| \leq m} A^\nu D_\nu.$$

We shall focus on the Cauchy (or initial value) Problem for the above system, namely under what conditions it is true that given the value of u at $t = 0$, $f(x)$, say, there exists a unique solution, $u(t,x)$, to the above system with $u(0,x) := f(x)$. Later we shall mention a related problem which is important on most numerical schemes used in relativity, namely the initial-boundary value problem, where one also prescribes some data on time-like boundaries.

What follows is a short account of chapter II of [50], see also chapter IV of [39]. After this, I indicate what aspects of the theory generalize to quasi–linear systems, and under which further assumptions this is so. I also give some indications of the relation of this theory to the stability issues of numerical simulations. This section can be skipped by those not interested in the mathematical theory itself or those who already know it.

2.1   Existence and Uniqueness of Smooth Solutions

Let M_0 be the space of functions of the form:

$$f(x) := \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} e^{i\omega \cdot x} \phi(\omega) d\omega, \quad \phi(\omega) \in C_0^\infty(\Re^n).$$

Since the integral exists at each point, it can be differentiated inside the integral sign where it gives another compactly supported integrand, thus these functions are smooth ($C^\infty(\Re^n)$). What's more, since the Paley–Wiener theorem holds, they are analytic.

To this space also belongs the function,

$$u(x,t) := \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} e^{i\omega \cdot x} e^{P(i\omega)t} \phi(\omega) d\omega, \quad \phi(\omega)$$

This function is smooth not only along space directions, but also along time directions. In fact, it is straightforward to check that the function satisfies the evolution equation above for the initial condition $f(x)$. Thus we see that initial data in M_0 always produces solutions which at each constant time slice are in M_0. In fact, defining a M_0 solutions as:

Definition 2 *A function $u(x,t)$ is a M_0 solution if:*

   i) $u(\cdot, t) \in M_0$ *for all $t \geq 0$;*

  ii) *its Fourier transform, $\hat{u}(\omega, t)$, is continuous, and vanishes for $|\omega| > k$ where k is some constant independent of t;*

 iii) *u is a classical solution; that is u_t exists and u satisfies the equation at each point (x, t).*

a direct application of the uniqueness of the Fourier representation for smooth functions shows:

Lemma 1 *Given a constant coefficients linear evolution equation, for each initial data in M_0 there exists a unique M_0 solution and it is given by the above formula.*

Thus we see that there are plenty of smooth solutions, whatever the system is. But it was realized by Hadamard, [40], that there were not enough solutions, since the space M_0 is not closed. Furthermore, in general there are no topologies on the space of initial data, and of solutions for which solutions depend continuously on initial data. Lack of continuity of solutions with respect to their initial data would not only imply lack of predictability from the physical standpoint, for all data are subject to measurement errors, but also lack of realistic possibilities of numerically computing solutions, due to truncation errors. Thus it is important to characterize the set of equations for which continuity holds. There are several possibilities for the choice of the topologies for the spaces of initial data and of solutions. Here we restrict consideration to those which have been more prolific with respect to results and generalizations to non–linear, non–constant coefficient equations systems.

Definition 3 *A system of partial differential equations is called well posed if there exists a norm (usually a Sobolev norm) and two constants, k, α, such that for all initial data in M_0 and all positive times,*

$$||u(t, \cdot)|| \leq ke^{\alpha t}||f(\cdot)||.$$

Remarks:

- It is possible to define weaker conditions for well posedness in which the norm for the initial data is different (weaker) than the norm for the solution. This is unsatisfactory for equations in which there is no preferred time direction. Besides, in general it produces results which are not robust under lower order term (in differentiation) perturbations of the equations, and so do not generalize to variable coefficients or non–linear equations.

- Defining continuity through norms is a limitation and rules out certain more general types of hyperbolicity conditions, in particular those due to Leray and Ohya, [54, 55], see more below.

- Well posedness and linearity imply that we can extend by continuity the set of solutions to all those which are generated by initial data on the completion generated by M_0 on the given norm.

Theorem 1 *A system is L^2 well posed if and only if there exists constants k and α such that for all positive times,*

$$|e^{P(i\omega)t}| \leq ke^{\alpha t} \quad \forall \omega \in \Re^n,$$

where the above norm is the usual operator norm on matrices.

   If a system is well posed for the L^2 norm, [recall that the L^2 norm of a function is the square root of the integral of its square], then it is well posed for any other Sobolev norm, (as follows from the above theorem), since the constants are independent of ω. The above theorem reduces the problem of well posedness to an algebraic one which we further refine in the following theorem:

Theorem 2 *[**Kreiss** [51]] The following conditions are equivalent:*

i) *The system is L^2 well posed.*

ii) *There exist constants k, and α, and a positive definite Hermitian form $H(\omega)$ such that:*

$$k^{-1}I \leq H(\omega) \leq kI \quad and \quad H(\omega)P(i\omega)+P^\star(i\omega)H(\omega) \leq 2\alpha H(\omega) \ \forall \omega \in \Re^n.$$

   This result is central to the theory. The proof that *ii)* implies *i)* is simple and follows directly from the inequality:

$$\frac{d}{dt}(\hat{u}, H(\omega)\hat{u}) = (\hat{u}, H(\omega)P(i\omega) + P^\star(i\omega)H(\omega)\hat{u}) \leq 2\alpha(\hat{u}, H(\omega)\hat{u}),$$

that is, from the construction of an *energy norm*. We see that for any well posed problems this special energy norm can be constructed, so one can always attempt to approach the problem by trying to find, usually with the help of the physics behind the problem, the correct energy norm. Condition *ii)* is usually referred to as the *semiboundedness of the operator $P(D)$ with respect to the norm H* (induced on functions in $\Re^n$ by Fourier Transform).

2.2   First Order Systems

We shall now restrict consideration to systems which have at most one space derivative, i.e. systems of the form,

$$u_t = A^a D_a u + Bu \tag{1}$$

Using the above theorem, it is easy to see that if a system $P(D)$ is well posed, then so is the system $P(D) + B$, where B is any constant matrix. For the particular case at hand, this means that we can further restrict attention, without loss of generality, to the *the principal part* of the operator, namely

$$P_1(D) := A^a D_a u,$$

where A^a is a $s \times s$ matrix valued vector in $\Re^n$. In this case we can improve on the above condition by showing that well posedness implies no growth of the solution, that is that we can choose $\alpha = 0$ above.

Theorem 3 *A first order system is well posed if and only if there exist, a constant k, and a positive definite Hermitian form $H(\omega)$ such that:*

$$k^{-1}I \leq H(\omega) \leq kI \quad and \quad H(\omega)A^a\omega_a - A^{a\star}\omega_a H(\omega) = 0 \quad \forall\omega_a \ \ with \ \ |\omega_a| = 1.$$

If A^a satisfies the above condition for some $H(\omega)$, then we say that $P(D)$ is *strongly hyperbolic*, which, as we see, is equivalent for first order equation systems to well posedness. If the operator H does not depend on ω_a, a case that appears in most physical problems, then we say the system is *symmetric hyperbolic*. Indeed, if H does not depend on ω_a, then there is a base in which it just becomes the identity matrix. (One can diagonalize it and re–scale the base.) Then the above condition in the new base just means that $A^a\omega_a$ –with the upper matrix index lowered– is symmetric for any ω_a, and so each component of A^a is symmetric. Even in the general (strongly hyperbolic) case, one can find a base (ω_a dependent) in which $P(i\omega)$ can be diagonalized, basically because it is symmetric with respect to the (ω_a dependent) scalar product induced by $H(\omega)$. In this diagonal version, it is easy to see that the well posedness requires all eigenvalues of $iA^a\omega_a$ to be purely imaginary. Thus we see that an equivalent characterization for well posedness of first order systems is that their principal part (i.e. $iA^a\omega_a$) has purely imaginary eigenvalues, and that it can be diagonalized by an invertible, ω_a-dependent, transformation. The classical example of a symmetric hyperbolic system is the wave equation.

For simplicity we consider the wave equation in $1 + 1$ dimensions. Choosing Cartesian coordinates we have,

$$\phi_{tt} - \phi_{xx} = 0,$$

and so defining the "vector" $u := (\phi, \phi_t, \phi_x)$ we have the following first order system:

$$u_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} u_x + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} u$$

There are several other notions of hyperbolicity that appear in the literature:

- A first order system is called *weakly hyperbolic* if the eigenvalues of the principal part are purely imaginary. This condition, clearly weaker than strong hyperbolicity, is not enough to assert well posedness in the sense I have defined here, and so I do not discuss it further.

- A first order system is called *strictly hyperbolic* if all eigenvalues of the principal part are purely imaginary and distinct. If the eigenvalues are distinct, then the eigenspaces invariant under the action of $P_1(i\omega)$ are one dimensional and the system can be diagonalized. Thus this class is contained in the strongly hyperbolic one. Due to the degeneracies of systems induced by symmetries in more than one dimension, physical problems are seldom strictly hyperbolic. Sometimes this definition is used to mean that the eigenvectors belonging to the different invariant spaces generated by the symmetries, which must be also invariant under $iA^a\omega_a$, have distinct eigenvalues. With suitable conditions on the symmetries, this implies the full diagonalization of the principal part, and so equivalence with strong hyperbolicity.

- There is a slightly different notion of hyperbolicity due to Leray see [54] and [55]. There a system is called *strictly hyperbolic*, or just hyperbolic, if it satisfies certain conditions which amount to having the Cauchy problem well posed in the sense we have used. A system is called *non–strictly hyperbolic* if it satisfies conditions implying well posedness of the Cauchy Problem, but where the continuity notion is not given by a norm, but rather through Gevrey classes of functions. In particular these spaces are subspaces of smooth, C^∞ functions, and so the data must be also smooth. I doubt very much can be done with them in terms of studying the stability of numerical methods, so we shall not concentrate on them.

2.3   Generalization to Variable Coefficient and Non-linear Systems

We shall consider in what follows a first order system of the form:

$$u_t = A^a(x, t, u)\nabla_a u + B(x, t, u)u,$$

where the vector valued matrix A^a, and the matrix B, are assumed to be smooth[2] functions of all its arguments.

Systems of this type are called *quasi-linear* because the derivative appears linearly. This property allows one to use most of the machinery for constant coefficient equations to prove well posedness, thus the local existence is well understood, via linearization techniques. There are few global results, and in general they depend on more refined knowledge of the equation systems for which they apply.

The behavior of solutions to quasi–linear equations is not yet fully understood. Most of the solutions develop singularities in a finite time for most initial data, even if they are in M_0. This is the case for convective systems, or more generally for genuinely non-linear systems, see [46, 57], for the definition and main results, a class which includes systems like perfect fluids and relativistic

---

[2]Actually, for the machinery of proving well posedness, finite differentiability to some higher order is needed.

dissipative fluids –for they contain as part of the system the perfect fluid equations. This is also the case for general relativity, where singularity theorems (see [41, 60]) tell us about the development of singularities, although of a different type. Thus the concept of well posedness has to be modified to account for the fact that solutions only last for a finite time and this time depends on the initial data. Basically, the most we can pretend to show in the above generality is the same type of well posedness one requires from an ordinary system of equations. Which is quite a lot! The non-linear aspect of the equations implies also that it is not possible to generalize their solutions to be distributions. The minimum differentiability needed to make sense of an equation depends on the particular equation. Furthermore, there are cases (e.g. convection) in which, for some function spaces of low differentiability, the equation makes sense and some solutions exist, but they are not unique[3].

Definition 4 *Let $u_0(t, x)$, $t \in [0, T_0)$, $T_0 < +\infty$ be a smooth solution of a quasi–linear evolution system. We shall say the system is well posed at the solution u_0 with respect to a norm $|| \ ||$ if given any $\delta > 0$ there exists $\varepsilon > 0$ such that for any smooth initial data $f(x)$ such that $||f - f_0|| < \varepsilon$, with $f_0(x) := u_0(0, x)$ there exists a smooth solution $u(t, x)$ defined in a strip $0 \le t < T$, with $|u(t, \cdot) - u_0(t, \cdot)| < \delta$, $|T - T_0| < \delta$.*

In order not to worry about the possibility that the smoothness of the solutions be too stringent a requirement, one can smooth out the equation using a one parameter family of mollifiers, and require that the relation $\delta(\varepsilon)$ be independent of that parameter family.

To obtain results about well posedness, we just have to slightly modify the concepts of hyperbolicity already discussed in the constant coefficient case. Since in the constant coefficient case the matrices did not depend on the points of space-time, nor on the solution itself, we had only two cases. In one case, the norm H did not depend on ω_a, and so in some base the matrix A^a was symmetric. In the other case, the norm H did depend on ω_a, and we had a general strongly hyperbolic system. In the latter case, it can be seen that $H(\omega)$ is piece–wise continuous and so integrable, which is, in that case, all that is needed to proceed with the proof. In the general case with which we are now dealing, H would in general depend not only on ω_a, but also on the point of space-time and on the solution, $H = H(\omega_a, t, x, u)$.

This difference has caused terminology to be not uniform in the literature, so I have taken advantage of this and establish terms in the way I consider best suited for the topic.

Certain authors call some systems symmetric hyperbolic and others symmetrizable. They call symmetric hyperbolic only those systems where the symmetrizer does not depend on the unknown variables nor on the space-time variables (or at most depends only on the base space variables $H := H(t, x)$); they call the other systems symmetrizable. This is a rather arbitrary distinction,

---

[3]Extra conditions have to be imposed, like entropy growth across shocks, to obtain uniqueness.

since the methods of proof used are valid for both with no essential difference. Thus, if H does not depend on ω_a but depends smoothly on all other variables, $H := H(t, x, u)$, then we shall still say the system is **symmetric hyperbolic**. In this case the non-singular transformation which symmetrizes $A^a(t, x, u)$ is smooth in all its variables.

The existence and smoothness proof is based, as in the constant coefficient case, on energy norm estimates, but now supplemented by Sobolev inequalities. Since the norm is built out of H and it does not depend on ω_a, no passage to Fourier space is needed.

If H does also depend on ω_a, and is smooth on all variables, $H := H(t, x, u, \omega_a)$, we shall say the system is **strongly hyperbolic**. The existence and smoothness proof now requires the construction of a pseudo-differential norm out of H, and so pseudo-differential calculus is needed, which implies that H has to be smooth in all its entries, in particular in ω_a.

We shall not discuss weak hyperbolic systems, for they are generically unstable under perturbations, nor shall we discuss strictly hyperbolic systems, i.e. systems with strictly different eigenvalues of $A^a \omega_a$, for they seldom appear in physical processes in more than one dimension.

With this concept of well posedness we have the following theorem [See for instance [58] pg. 123]:

Theorem 4 *Let $r > \frac{n}{2} + 1$, then a strongly hyperbolic system is well posed with respect to the Sobolev norm $|| \ ||_r$. The solution is in $C([0, T), H^r)$, the time of existence depends only on $||f||_r$.*

Remarks:

- In the generic case the value of r cannot be reduced from the above value, but of course it can for certain special types of systems. In general relativity, a slight improvement, $(r > \frac{n}{2})$, is obtained from the fact that the matrix A^a only depends on a subset of variables (the metric).

- The time of existence is the same for all r. That is, when a solution loses differentiability, it loses all of it at the same time. This is reflected in the following result: [58]. If a solution exists until T_0 and there $||u(t, \cdot)||_{C^1_\star}$ is bounded, (where $C^1_\star$ is a Zygmund space, see Appendix A in [58]), then the solution can be extended further. Again in general relativity a slight improvement can be obtained, see [42], and [18].

2.4 Hyperbolicity and Numerical Simulations

Knowing that the field equations for GR can be cast in a symmetric hyperbolic form, we can now ask how this fact can be of help for numerical calculations, (besides the extremely important fact that the problem would be well posed and so tractable!). There are at least two reasons why one should use variables in which the system is hyperbolic when performing numerical simulations. The first reason is that having a strongly hyperbolic system allows for standard

constructions of numerical codes which are stable. If the system is symmetric, hyperbolic codes with better properties can be constructed. In particular, schema for symmetric hyperbolic systems can be constructed with numerical dissipation of lower order than what is needed for generic strongly hyperbolic ones (see chapter VI of [39]). In variables where the system becomes diagonal, one can also use methods which take advantage of that structure.

The second reason for using a hyperbolic formulation for numerical analysis is that well posedness of the system gives bounds on the growth of the solution and its derivatives, as long as the solution is smooth[4]. This property, when used in conjuction with stable algorithms, implies that one can bound the errors made on the simulation. That is, one knows not only that the error goes to zero as some power of the step size, but also the proportionality factor of that power law. In simulating phenomena whose observation would require hundreds of millions of dollars, a tight control of the accuracy reached should be required. Nevertheless it should be noted that raw hyperbolicity estimates alone usually give exponential bounds with very large growth coefficients, and that they are not of much value for numerical work.

Many of the systems we shall analyze can be cast as flux conservative equations with sources. This is a direct consequence of the facts that the principal part of the equations depends only on the metric variable, and that the equation for the time derivative of it does not contain derivatives of any of the dynamical variables. This property is important when using codes with variable grid spacing, even more if one considers that there are many standard codes for fluids –which are truly flux conserving– with adaptive grid schema.

It has to be said that flux conservation is important when dealing with systems that develop shock waves, that is in convective or more precisely in genuinely non–linear systems (for a definition of this term and many of the results, see [46, 57]). One should be cautious about any expectation of improvement by using flux conservative properties in general relativity, since here the shocks would probably not develop –in particular the systems are not genuinely non-linear. Rather, when singularities appear, they would be much worse than mere discontinuities of some of the dynamical fields.[5] Due to bad gauge choices, discontinuities resembling shocks have been observed in numerical simulations, see [5]. Perhaps, instead of trying to devise an algorithm which allows one to go through these discontinuities, one should concentrate on finding better gauges, where it could even happen that the system cannot be put in flux conserving form. Thus, it is not clear whether flux conservative forms are relevant for vacuum general relativity.[6]

---

[4]That is, sufficiently smooth for the Sobolev energy, with minimum differentiability possible to close the bounds, to exist.

[5]One could of course put a discontinuity on the derivatives of the metric on initial data. The discontinuity would propagate along the characteristics. In general this is not considered a shock, for it is not generated by the dynamics and does not propagate along a different characteristic than the neighboring continuous regions. These are called contact discontinuities.

[6]If matter sources are included, like fluids, then one might need to consider flux conservative schema for the whole system of equations. But probably it would be much better to use relativistic dissipative fluids –whose global existence for small data has been proven recently

Going Further

For the reader wishing to delve further into this precious theory of hyperbolic systems, while keeping a physicist's approach, I recommend [36]. For those wishing to see more of the machinery at work, I recommend the book of Kreiss and Lorentz [50]. Finally, for those who really want to get the latest on the technical aspects and the modern approach to the problem, I recommend Taylor's book, [58]. Considerations about numerical analysis and algorithms can be found in [39]. In particular, that book contains general stable algorithms for strongly and symmetric hyperbolic systems and numerical error bounds in terms of analytic bounds of the exact solutions applicable to non–linear systems.

---

[52]–to dispense of shocks altogether when considering weak data.

3 The Problem of Hyperbolicity in General Relativity

What follows are descriptions of the problem of hyperbolicity in general relativity and of the main approaches that have been proposed to deal with the problem.

In studying Einstein's field equations we are faced with a problem (see [23, 36, 32]): While the theory of partial differential equations has developed as a theory for tensor components in a given coordinate system, or at best for tensor fields in a given metric space, Einstein's equations acquire their full meaning –and the characteristics which distinguish them from all other physical theories– when they are viewed as equations for geometries, that is, for equivalent classes of metric tensors. The object of the theory is not a metric tensor, but the whole equivalence class to which it belongs –all other metrics related to the first one by a smooth diffeomorphism. This fact is contained in the equations, for they are invariant under those diffeomorphisms. To clarify the concept, and see the problem, let us assume we have a solution to Einstein equations in a given region of a manifold. Take a space–like hypersurface across it, Σ_0, and a small "lens shaped" region which can be foliated by smooth space-like surfaces Σ_t starting at Σ_0. If Einstein's equations were hyperbolic for the metric tensor, then uniqueness of the solution (the metric tensor) in the lens shaped region would follow from the standard theory once proper initial data at Σ_0 is given. But we know that if we apply a diffeomorphism to the original metric tensor solution, which is different from the identity only in a region inside the lens shaped one but which does not intersect the initial Σ_0 slice, the resulting metric would also satisfy Einstein's equations, thus contradicting uniqueness, and so the possibility that the system be hyperbolic.

Since, as shown in § 2, hyperbolicity is equivalent to the existence of norms which are bounded under evolution, we see that for Einstein's equations there cannot be such norms on the space of metric tensors. Norms are not only important for well posedness, but also for other related issues which often appear in general relativity, in particular when one tries to see whether some approximation schema is indeed an approximation. Examples of this appear in very unrelated cases, for instance, in numerical algorithms and post-Newtonian approximations. Thus, a method is needed to find relevant norms on metric tensors, that is, to break the diffeomorphism invariance. The norms thus obtained are not natural, and so by themselves do not imply any physical closeness of metrics in numerical values. They have to be considered in their topologically equivalent class. Physically relevant notions of closeness can still be obtained by building, out of the metric tensor and its derivatives, diffeomorphism invariant quantities and making the comparisons with then.

Can we avoid this detour into tensors and make a theory of diffeomorphism invariant objects? It is not clear whether this can be done. Some attempts in this direction have been made by trying to build norms which have some partial diffeomorphism invariance. Here the norms are made out of scalars built out

of curvature tensor components of the metric, in particular see [23, 32]. But
I think a fully geometrical theory needs other types of mathematics than the
theory of partial differential equations, a theory which might be emerging from
parallel questions in quantum gravity.

3.1 The Standard Approach, or the 4-D Covariant Approach

The standard approach to overcome this problem has been to "fix the gauge",
that is, by imposing some extra condition on the metric coordinate components
which would select one and only one representative from each equivalent class
of Einstein's geometries. With a clever choice of gauge fixing, commonly the
so-called harmonic gauge, Einstein's equations can be "reduced" to a hyperbolic
system by removing from them the parts which the gauge condition would make
vanish. This reduced system is equivalent to the full Einstein equations if one
can prove that, possibly after imposing further initial data set constraints, the
solutions of the reduced system satisfy the gauge conditions previously imposed
to the system. Alternatively, one can think of this method as fixing some com-
ponents of the tensor obtained by taking the difference between two connections,
the one associated with the metric tensor at which we are looking and the one
associated with some other arbitrary background metric [26, 41].

A convenient way to describe this scheme is by introducing a background
metric, $\tilde{g}^{ab}$, thus the gauge is not a coordinate condition, but rather a condition
which links the physical metric with the background one. In this approach,
see [41], the basic variable is a densitized symmetric tensor, $\Phi^{ab} := g\delta g^{ab}$,
where g is the metric determinant with respect to the background one, [7] and
$\delta g^{ab} := g^{ab} - \tilde{g}^{ab}$.

In these variables Einstein's equations become,

$$\frac{1}{2}g^{cd}\tilde{\nabla}_c\tilde{\nabla}_d\Phi^{ab} - g^{c(a}\tilde{\nabla}_c\Psi^{b)} + \frac{1}{2}g^{ab}\tilde{\nabla}_c\Psi^c \tag{2}$$

$$+(terms\ in\ \tilde{\nabla}_c\delta g^{de}\ and\ \delta g^{de}) \tag{3}$$

$$= 8\pi gT^{ab} - g\tilde{G}^{ab}, \tag{4}$$

where $\tilde{\nabla}_c$ is the covariant derivative associated with $\tilde{g}^{ab}$, $\tilde{G}^{ab}$ its Einstein tensor,
and $\Psi^a := \tilde{\nabla}_b\Phi^{ab} = \tilde{\nabla}_b\sqrt{\frac{g}{\tilde{g}}}g^{ab}$.

In that gauge, Einstein's equations are "reduced" to a hyperbolic system by
removing from them all terms containing Ψ^a, for this quantity is assumed to
vanish in this gauge. By doing this one gets a set of coupled wave equations, one
for each metric component. Thus by prescribing at an initial (space-like) hyper-
surface values for Φ^{ab} and of its normal derivatives one gets unique solutions to
the reduced system. When are such solutions solutions to Einstein's equations?

---

[7]If ε_{abcd} is the Levi-Civita tensor corresponding to the physical metric, g^{ab} and $\tilde{\varepsilon}_{abcd}$ the
one corresponding to background metric, $\tilde{g}^{ab}$, then $\varepsilon_{abcd} = \sqrt{g}\tilde{\varepsilon}_{abcd}$.

That is, under what conditions does Ψ^a vanish everywhere? It turns out that the Bianchi identity grants that when Φ^{ab} satisfies the reduced equations, then Ψ^a satisfies a linear homogeneous second order hyperbolic equation. Standard uniqueness results for such systems implies that if initially Φ^{ab} is chosen so that Ψ^a and its normal derivative vanish at the initial surface, then they vanish everywhere on the domain of dependence of that surface. Thus the question is now posed on the initial data, that is, on whether it is possible to choose appropriate initial data for the reduced system, $(\Phi^{ab}, \dot{\Phi}^{ab})$, in such a way that $(\Psi^a, \dot{\Psi}^a)$ vanish initially. It turns out, using that the reduced equations are satisfied at the initial surface, that one can indeed express Ψ^a and its normal derivative at the initial surface, in terms of Φ^{ab} and its derivatives (both normal and tangential to the initial surface). Thus one finds there are plenty of initial data sets for which solutions to the reduced system coincide with solutions to the full Einstein system. Are they all possible solutions to Einstein's equations, or are we loosing some of them by imposing this scheme? The answer to the first part of the question is affirmative (subject so some asymptotic and smoothness conditions), for one can prove that given "any" solution to the Einstein equations, there exists a diffeomorphism which makes it satisfy the above harmonic gauge condition.

It is important to realize that it is not necessary to set Ψ^a to zero to render the Einstein equations hyperbolic; it just suffices to set it equal to some given vector field on the manifold, or any given vector function of the space-time points and on the metric, but not its derivatives. So there are actually many ways to hyperbolize Einsteins's equations via the above scheme. We shall call all of them **harmonic gauge conditions**, and reserve the name **full harmonic condition** to the one where $\Psi^a \equiv 0$.

An important advantage of this method is that some gauge conditions, like the full harmonic gauge, are four-dimensional covariant –although a background metric is fixed– a condition which can be very useful for some considerations.

One drawback of this method, at least in the simplest version of the harmonic gauge, i.e. the full harmonic gauge, was recognized early, [14]. The drawback is the fact that this gauge condition can be imposed only locally, and generically breaks down in a finite evolution time. A related problem has been discussed recently in [5] in the context of the hyperbolizations of the ADM variables with the harmonic gauge along the temporal direction. The above disadvantage can be considered just a manifestation of another: the lack of ductility of the method, that is the fact that one has been able do very little besides imposing the full harmonic gauge condition, and that for each new harmonic gauge condition one would like to use, a whole study of the properties of the reduced equations would have to be undertaken. Although there are many other gauge conditions besides the harmonic one, the issue of the possibility of their global validity, or the search for other properties of potential use, do not seem to have been considered. For a detailed discussion of this topic, see [31, 32], and also [41].

One can summarize the situation by noticing that in this setting one needs to prescribe a four vector as a harmonic gauge condition. Since the theory keeps its four dimensional covariance, then it is hard to choose any other vector but

zero, that is the full harmonic gauge. Since recently there have been no advances in this area, I do not elaborate on it.

3.2 The Modification of the Field Equations Outside the Constraint Sub-manifold, or the 3+1 Decomposition Point of View

Another approach to deal with the diffeomorphism freedom of Einstein's equations is by first removing the diffeomorphism invariance. This is done by prescribing the time foliation along evolution, that is, by prescribing a lapse-shift pair along evolution. This removes the diffeomorphism invariance up to three-dimensional diffeomorphisms at the initial surface. Sometimes four dimensional covariance is also broken by splitting Einstein's equations, and possibly other supplementary equations, with respect to that foliation, and then recombining the split pieces in a suitable way. The resulting equations are equivalent to the original ones –for it is just a linear combination of them– so they have the same solutions. Notice that here we are doing more than just a $3+1$ decomposition, since in general one is recombining space-space components of the Einstein tensor with time-time, and time-space components in a non-covariant way, and taking as equations this combination, or even transforming the equations to first order in derivatives of the variables by defining new variables and equations and modifying them.

After this procedure is done, one obtains a system which is symmetric hyperbolic for most choices of *given* lapse-shift functions, once they are suitably re-scaled. Subsequent arguments go very much on adding equations for the lapse-shift vector in order to make the whole system well posed, and presumably useful for some application. It is instructive to think of these modifications of the evolution equations from the point of view of the initial value formulation. There one starts by solving the constraint equations, the time-time and time-space components of the Einstein tensor, at the initial surface. With the initial data thus obtained, one finds the solution to the evolution equations which are taken to be the space-space components of the Einstein tensor. Since that evolution preserves the constraints, (The vector field generating the flow in phase–space is tangent to the constraint sub-manifold.), one can forget about the constraint equations and think of the evolution equations as providing an evolution for the whole phase–space. In this sense, the modification one is making affects the evolution vector outside the constraint sub-manifold, leaving the vector intact at it. Uniqueness of solutions, which follows from the well posedness of the system, then implies that the solutions stay on the sub-manifold. Nevertheless, and we shall return to this point, as shown in [33], there is no guarantee that the sub-manifold of constraint solutions is stable with respect to the evolution vector field as extended on the whole phase–space. This is an important point for numerical simulations.

4   Recent Approaches to the Problem

4.1   The ADM representation

In the ADM formulation of Einstein's equations, the fields involved are detached from the under-laying space-time and brought into an abstract three dimensional manifold product a segment of the real line. To obtain the relation between these abstract fields and the metric tensor defining a solution of Einstein's equations, we start by pretending we have such a solution, that is a space-time (M, g_{ab}). In this space-time we look at a **Cauchy surface**, Σ_0, that is, an everywhere space-like hypersurface such that any in-extendible time-like piece-wise smooth curve pierces it once and only once, and a **time flow**, that is, a smooth time-like vector field, t^a. From the definition of the Cauchy surface, t^a is never tangent to it and every point of M falls in an integral curve of t^a. Thus, in assuming the existence of Σ_0 we are restricting attention to manifolds of the type $S \times \Re$. To make this structure more apparent we define a function t by setting to zero the parameter defining the integral curves of t^a at Σ_0, that is, the value of t at $p \in M$ is defined as the value of the parameter the integral curve of t^a takes at p if defined in such a way that at Σ_0 is takes the value 0. We shall call the surfaces of constant t by Σ_t. Thus $t^a \nabla_a t = 1$, notice that nevertheless, in general, they are not space-like. When they are space-like, we shall call t a **time function**. In this case we also say we have a **space-like foliation** of space-time (M, g^{ab}). We shall assume that this is the case, but one must take into account that when we are solving for a space-time we do not know for how long this would continue to hold. Using this structure we can split tensor into "space" and "time" parts with respect to the surfaces Σ_t. If n^a is the normal to them, that is, $n^a := -g^{ab} \frac{\nabla_b t}{\sqrt{g^{cd}\nabla_c t \nabla_d t}}$ then $h^{ab} := g^{ab} - n^a n^b$ is the induced metric on each Σ_t. We also define the lapse shift pair, (N, N^a), as the "time-like" and "space-like" parts of t^a with respect to Σ_t, that is, $t^a := Nn^a + N^a$.

Given the foliation, and the triad (h^{ab}, N, N^a) we can reconstruct the metric as $g^{ab} = h^{ab} - N^{-2}(t^a - N^a)(t^b - N'^b)$. Given another foliation, (Σ_0, t'^a), but the same triad, we get another metric tensor, the relation between both is a diffeomorphism which leaves invariant Σ_0 and is generated by the integral curves of $t^a - t'^a$. Alternatively we can choose another pair (N, N^a) and so construct another metric tensor, the relation between both metric thus obtained is also a diffeomorphism. Since one is interested in geometries, that is in metric tensors up to diffeomorphisms, both metric obtained are equivalent, but one needs to pick up an specific one in order to write down, (and solve), Einstein's equations.

Using the above splitting vacuum Einstein's equations become two **evolution equations:**

$$\dot{h}_{ab} = \qquad 2Nh^{-1/2}(P_{ab} - \frac{1}{2}h_{ab}P) + 2D_{(a}N_{b)},$$

$$\dot{P}^{ab} = - \quad Nh^{1/2}(R^{ab}(h) - \frac{1}{2}R(h)h^{ab})h^{1/2}(D^a D^b N - h^{ab}D^c D_c N)$$

$$+ \quad \frac{1}{2}Nh^{-1/2}h^{ab}(P_{cd}P^{cd} - \frac{1}{2}P^2) - 2Nh^{-1/2}(P^{ac}P^b{}_c - \frac{1}{2}PP^{ab})$$
$$+ \quad D_c(N^cP^{ab}) - 2P^{c(a}D_cN^{b)},$$

where the dot means a Lie derivative with respect to t^a, $P^{ab} := \sqrt{h}(K^{ab} - Kh^{ab})$ and K^{ab} is the **extrinsic curvature** of Σ_t with respect to the four-geometry, that is, $K^{ab} := h^{ac}\nabla_c n^b$. D_a is the covariant derivative associated with h^{ab} on each Σ_t, and R^{ab} its Ricci tensor.

And two **constraint equations:**

$$hR(h) - P^{ab}P_{ab} + \frac{1}{2}P^2 = 0,$$
$$D_aP^{ab} = 0.$$

Note that there are two "dynamical" variables, h_{ab}, and P^{ab}, while the lapse–shift pair (N, N^a), although necessary to determine the evolution, is undetermined by the equations. Note also that they do not enter into the constraint equations, for as said above a change on the lapse–shift pair leave the fields at initial surface unchanged.

It is important step back now and see that these equations can be thought as "living" in a structure completely detached from space–time. To see this, identify all points laying in the same integral curve of t^a, thus the equivalence class is a three dimensional manifold S, homeomorphic to any Σ_t. On it, for each t we can induce space-contravariant tensors, such as $h_{ab}(t)$, $P_{ab}(t)$, $N_a(t)$, and scalars, as $N(t)$. As long as the surfaces are space–like, the induced metric is negative definite and we can invert it, thus we can perform all kinds of contractions and write the above equations as dynamical equations on the parameter t on fields on the same manifold, S. This is of course the setting in which one sets most of the schemes to solve the equations, and it is hard to keep control, even awareness, that the surfaces defining the foliation can become null or nearly so. Einstein's equations "feel" that effect since they are causal, and this is abruptly fed back via the development of singularities on the solutions . They do not have any thing to do with real singularities of space–time, but rather with foliations becoming null.

The first application of this idea to Einstein's equations appears to have been [19]. There a hyperbolic system consisting of wave equations for the time derivative of the variable hh^{ab} is obtained when the shift is taken to vanish and the lapse is chosen so as to impose the time component of the harmonic gauge. The shift is set to zero, but, as stated in the paper, this is an unnecessary condition. Thus, it is clear that one gains in flexibility compared with the standard method above. This seems to correspond with the fact that in one of the equations forming the system, the time derivative of the evolution equation for the momentum, the momentum constraint has been suitably added, thus modifying the evolution flow outside the constraint sub-manifold. As stated in the paper, they could not use the other constraint, the Hamiltonian one, to modify that equation.

The condition for that system to be (symmetric) hyperbolic is that the term below should not have second derivatives of P^{ab}.

$$(h^{ab}\Delta - D^a D^b)(\frac{\dot{N}}{N} - \frac{N}{2}P),$$

where h^{ab} is the induced three metric on a hypersurface, D_a is the covariant derivative at that hypersurface compatible with h_{ab}, $\Delta := h^{ab}D_a D_b$, N is the lapse function, and $P := h_{ab}P^{ab}$, with P^{ab} the momentum field conjugated to h_{ab}.

The simplest condition to guarantee this is:

$$\dot{N} = \frac{N^2}{2}P,$$

which in view of the definition of P^{ab}, which implies $\dot{h} = hNP$, has as a solution,

$$N = (\frac{h}{e})^{1/2},$$

where $\frac{h}{e}$ is the determinant of the metric h^{ab} with respect to a constant in time background metric e^{ab}. This is precisely the harmonic condition for the time component of Ψ^b in notation introduced in the paper's introduction. That is, $\Psi^b n_b \equiv 0$, where n_b it the normal to the foliation. If the determinant of e_{ab} is not taken to be constant in time, then one gets,

$$\frac{\dot{N}}{N} - \frac{N}{2}P = \frac{\dot{e}}{e},$$

and so the system remains hyperbolic. Thus, we see that, up to the determinant of the metric, the lapse can be prescribed freely. This freedom is very important because it gives ductility to the approach, since this function can be specified according to the needs of applications. We shall call this a *generalized harmonic time gauge.*

Although in the introduction of [19] there is a remark dedicated to numerical relativists about the possible importance of having a stable system, the paper did not spark interest until recently, when applications required these results to proceed. In recent years, a number of papers have appeared which further elaborate on this system, [2, 4, 1]. In particular I would like to mention [4], where the authors look at the system in detail, writing it as a first order system, and introduce all variables which are needed for that. In these recent papers, the generalized harmonic time gauge has been included, as well as arbitrarily prescribed shift vectors. If one attempts attempts to write down the system as a first order one, that is, to give new names to the derivatives of the basic fields until bringing the system to the form of equation 1, the resulting system is rather big, it has fifty four variables, without counting the lapse-shift pair. We shall see that there are first order hyperbolic systems with half that number of variables.

Two similar results are of interest: In [9] a system is introduced with basically the same properties, but of lower order, that is, only first derivatives of the basic variables are taken as new independent variables in making the system first order. In this paper, it is realized that the same trick of modifying the evolution equations using the constraints can be done by modifying, instead of the second time derivative of the momentum, the extra equation which appears when making the ADM equations a first order system, that is the equation which fixes the time evolution of the space derivatives of the metric, or alternatively the time evolution of the Christoffel symbols. When this equation is suitably modified by adding a term proportional to the momentum constraint, and when the harmonic gauge in the generalized sense used above is imposed, a symmetric hyperbolic system results. In [11] the generalized harmonic time gauge is included, as well as arbitrarily prescribed shift vectors. For the latest on this approach see, [10]. I shall comment more on this in next section, § 4.

In [34] a similar system is presented. In this case, the focus is on establishing some rigorous results in the Newtonian limit. So a conformal rescaling of the metric is employed using the lapse function as conformal factor. The immediate consequence of this transformation is to eliminate from the evolution equation for P^{ab} the term with second space derivatives of the lapse function, precisely the term giving rise to one of the terms in equation 4.1. The end consequence is that the conformal metric is flatter to higher order. With this re–scaling, and using the same type of modification of the evolution equation for the space derivatives of the metric that the above two approaches use, a symmetric hyperbolic system is found, for arbitrary shift and lapse. [8] This freedom of the lapse and shift was used to cancel several divergent terms of the energy integrals in the Newtonian limit by imposing an elliptic gauge condition on the shift, which also determined uniquely the lapse. This resulted in a mixed symmetric hyperbolic–elliptic system of equations. In [35] an attempt is made to explore what other possibilities there are of making symmetric hyperbolic systems for general relativity with arbitrarily prescribed lapse–shift pairs. A set of parametrized changes of field variables and of linear combinations of equations are made, and it is shown that there exists at least a one parameter family of symmetric hyperbolic systems. In these systems generalized harmonic time gauge is replaced by:

$$N := (\frac{h}{e})^\delta \quad \delta > 0.$$

So, the dependence of the lapse on the determinant of the metric can be modified, but never suppressed. Since the changes in the parameter imply changes in the dynamical variables, while the factor proportional to the momentum added to the evolution equation for the connection is unique, and so fixed, it is not clear whether this can be of help for improving numerical algorithms. We shall see this type of dependence arise in one of the developments of one of the above mentioned approaches.

---

[8]Note that the lapse used in this work already had the generalized harmonic time gauge built into it.

In [32] a similar system, in the sense of using variables from the $3 + 1$ decomposition, is obtained by imposing also the same generalized harmonic time gauge. This system, as is the original system of [19], is of higher order because it includes the electric and magnetic parts of the Weyl tensor in the $3 + 1$ decomposition. As such, it contains more variables (fifty) than the two discussed above (thirty).

4.2   The Frame Representation

Another way of dealing with Einstein equations is through the frame representation. There, instead of using the metric tensor as the basic building block of the theory, a set of frame fields is used. At first this representation seems to be even less economical than the metric tensor representation of geometries, since, on top of the diffeomorphism freedom, one has the freedom to choose the frame vector fields. In short, one has sixteen variables instead of ten. But the antisymmetry of the connection coefficients compared with the symmetry of the Christoffel symbols levels considerably the difference, ending, after adding the second fundamental form and others, with twenty-eight variables, instead of the thirty of the conventional system. But this count is not entirely correct. As mentioned above, in order to close the system of equations one has to add the evolution equations for the electric and magnetic part of the Weyl tensor, thus ending with a total of thirty four variables. Actually one can close the system with the twenty four variables at the expense of making the equations into second order wave equations, (See for instance [59].), so effectively adding more variables when re–expressing it as a first order system.

The more important application of the frame representation has been the conformal system obtained by Friedrich, [30, 29], (see § 1), where in a fixed gauge he got a symmetric hyperbolic system which allowed him to study global solutions. Later, using similar techniques and spinors, he found a symmetric hyperbolic system with the remarkable property that lapse and shift appear in an undifferentiated form, allowing for greater freedom in relating them to the geometry without hampering hyperbolicity [31]. In [32] he introduces new symmetric systems for frame components where one can arbitrarily prescribe the gauge functions, which in this case does not only include the equivalent to the lapse–shift pair, but also a three by three matrix fixing the rotation of the frame. In this case, these gauge functions enter up to first derivatives. This compares very favorably with the ADM representation schema where the lapse-shift entered up to second order derivatives. Friedrich also finds a symmetric hyperbolic system with the generalized harmonic time condition. Contrary to the systems in the ADM formalism, where the issue is rather trivial, these systems do not seem to allow for a writing in flux conservative form. We do not consider that a serious drawback. The structure of Einstein's equations is very different than those of fluids, where the unavoidable presence of shocks makes it important to write them that way. Indeed the reason fluids have shocks can be attributed to their genuinely non-linear character, [46], a property not shared by Einstein's theory. (More about this in the next section, § 4.)

4.3 Ashtekar's Representation

The example of the one dimensional wave equation 2.2 can be slightly improved by the following construction: We define a two dimensional vector $u := (u^1, u^2) := (\phi, \phi_t - \alpha\phi_x)$. We then have $u^1_t = \alpha u^1_x + u^2$, while, $u^2_t = -\alpha u^2_x + (1 - \alpha^2)\phi_{xx}$. Thus we have a diagonal –and so symmetric hyperbolic– system if $\alpha = \pm 1$, namely,

$$u_t = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix} u_x + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} u,$$

The two possible values α can take correspond to the two characteristic directions the wave equation defines. This trick can not be extended two more dimensions, basically because the space derivative of ϕ is a vector, and so can not be properly mixed with its time derivative. But in other dimensions one can implement similar schema if the fields are not scalars but appropriate tensors.

Einstein's equations in Ashtekar's' variables [7, 8] is one beautiful example of this, since they have the remarkable property of naturally constituting a first order evolution system. Because of this reason it is also a compact system with twenty-seven unknowns, even before imposing the reality condition on the connection variable. Recently it has been proven [45] that such a system is symmetric hyperbolic if suitable combinations of the constraint equations are added to their evolution equations, thus effectively changing the flow outside the constraint sub-manifold of phase–space.

In Ashtekar's representation the basic variables are a densitized $SU(2)$ soldering form, $\tilde{\sigma}^a{}_A{}^B$ and a $SU(2)$ connection $A_{aA}{}^B$ which are tangent to a spacelike foliation of space time determined by given "lapse"–shift pair $\tilde{N} = N/det\sigma$, N^a.

The symmetric hyperbolic evolution equation system is:

$$\mathcal{L}_t\tilde{\sigma}^b = \frac{-i}{\sqrt{2}}\mathcal{D}_a(\tilde{N}[\tilde{\sigma}^a, \tilde{\sigma}^b]) + \frac{i}{\sqrt{2}}\tilde{N}[\tilde{C}, \tilde{\sigma}^b] + 2\mathcal{D}_a(N^{[a}\tilde{\sigma}^{b]}) + N^b\tilde{C} + [A_aN^a, \tilde{\sigma}^b]$$

$$\mathcal{L}_tA_b = \mathcal{D}_b(A_aN^a) + N^aF_{ab} + \frac{i}{\sqrt{2}}\tilde{N}[\tilde{\sigma}^a, F_{ba}] + \frac{i}{\sigma^2\sqrt{2}}\tilde{N}\tilde{\sigma}_bC + \frac{i}{\sigma^4}\tilde{N}\epsilon_b{}^{dc}\tilde{\sigma}_cC_d,$$

where $(\mathcal{D})$ is the $SU(2)$ derivative whose difference with respect to a flat connection is $A_{aA}{}^B$. C, C^a, and $\tilde{C}$ are the constraint equations,

$$
\begin{align}
C(\tilde{\sigma}, A) &:= tr(\tilde{\sigma}^a\tilde{\sigma}^bF_{ab}) \tag{5}\\
C_a(\tilde{\sigma}, A) &:= tr(\tilde{\sigma}^bF_{ab}) \tag{6}\\
\tilde{C}(\tilde{\sigma}, A) &:= \mathcal{D}_a\tilde{\sigma}^a, \tag{7}
\end{align}
$$

Note that here there is an extra vectorial constraint, a $SU(2)$ valued scalar, which corresponds to the fact that the system has extra degrees of freedom, the $SU(2)$ rotations. The extra constraint is just a strange way of asserting the symmetry of the second fundamental form, and is of the type of substitutions we made above to improve on the wave equation system. The constraints on themselves satisfy a symmetric hyperbolic system of equations.

Note that the principal part of the system is block diagonal and the eigenvectors–eigenvalues are very simple combinations of $\tilde{\sigma}$ with the elements of an orthogonal basis $\{\omega^a, m^a, \bar{m}^a\}$, where ω^a is the wave vector.

In this new system, the "lapse"-shift pair can be chosen arbitrarily. But in fact the "lapse" that appears here is a scalar density which has already incorporated the square root of h on it. So the freedom is actually the same as in the ADM representation. As in the frame representation, the lapse–shift pair appears only with derivatives up to first order. In this case it is relatively easy to see the freedom in making up evolution equations for the lapse–shift pair. As said above, the system is symmetric for Ashtekar's variables, since the lapse–shift pair enters as terms with up to first derivatives, one can take those terms from the non–principal part of the system and promote them into the principal part of a bigger system which incorporates the lapse–shift as extra variables. Thus, these terms constitute an off–diagonal block of the bigger principal part matrix. Imposing symmetry to the bigger matrix fixes the opposite off–diagonal part of the matrix. The only freedom left is on the lapse–shift block–diagonal part, which can be chosen to be any symmetric matrix we like. The non–principal part of the equation system on the lapse–shift sector can also be chosen arbitrarily. Of course, in contrast with the ADM representation results, one can also choose a gauge condition via elliptic equations on the lapse–shift pair. In this case, the elliptic system can be of first order in the lapse–shift or related variables. For instance, one could use Witten's equation to evolve them.

As we have seen, the generalized harmonic time gauge seems to appear naturally in most attempts to get well posed evolution systems. Thus it seems to be really a key ingredient, perhaps with some physical content. One could argue in that direction from the circumstances in which it appears in [34], namely effectively improving the estimates of the behavior of solutions admitting a Newtonian limit that is in a way related to the longitudinal modes of the theory. This longitudinal modes are part of the evolution, although they are not expected to behave in a hyperbolic manner. See also the comments around equation (9) in [10].

5 Beyond the Prescribed Gauge

In this section we will look at several attempts to, once a hyperbolic system has been obtained for the evolution variables, extend it into a bigger well posed system where the lapse–shift pair, or more generally the gauge variables of a particular system, can be determined by some prescription which in general depends on the dynamical variables. The bigger system does not need to be hyperbolic, it only needs to be well posed, and so for instance it can be mixed elliptic–hyperbolic.

5.1 Trial and Error Method

One alternative would be to start with some arbitrary prescription for the gauge variables, evolve the solution for a while, stop, look for troublesome regions, and modify the gauge prescription there. One would do that only a finite number of times and choose smooth prescriptions and smooth transitions between them, so no problem of well posedness or numerical stability would be created by that procedure. With some experience, and luck, this procedure could work.

Perhaps the mayor drawback of this approach is the fact that the strict harmonic gauge, the most conspicuous choice, does not behave well under evolution. This has been known for a long time [14], and more recently new indications have been found in [5]. These findings are of course not valid for the generalized harmonic time gauge because one can trivially take any solution, draw a well behaved foliation on it, and identify the generalized gauge that *works* for that solution. Some attempts have been made in order to get loose from the generalized harmonic time condition, presumably with the intention of later imposing equations on the free variables, otherwise independent fields. In [3] a non-strictly hyperbolic system is found by taking yet another time derivative of the evolution equation for the momentum variable. In that way, they are able to prescribe in a completely free way the lapse–shift pair. In doing this, they obtain a non-strictly hyperbolic system, in the sense of Leray-Ohya, [55], which I presume in the language of first order systems means that it is a weakly hyperbolic, but with certain other properties which imply that the system is well posed in Gevrey classes of functions [9]. The resulting system, once brought to first order, has a rather big number of variables. It is not clear that one can stablish numerical stability and convergence for these types of systems, for at least in the continuum estimates an infinite number of derivatives are involved.

In [32], as said in the previous section, a system in the frame representation is made where the corresponding gauge variables can also be given arbitrarily. The importance of this freedom is that in this case one can prescribe directly the lapse. This is in contrast to the case of the generalized harmonic time condition, in which one prescribes the lapse up to the square root of the metric and so finds out what the lapse really was only after solving the problem.

---

[9] Basically one is able to bound norms on the solution by different norms on the initial data, with the last one involving more derivatives, but with smaller and smaller derivatives as the order of the derivatives in the norms increases.

5.2 Hyperbolic Extensions

One would also like to have recipes which could be used automatically during evolution, that is, algebraic or differential equations, which would not only fix uniquely the evolution of the gauge variables, but which would also result in a well posed evolutionary problem.

One approach has been taken in [11, 12], and [10] where the equation 4.1 has been modified to (in the notation of [19]):

$$\dot{N} = \frac{N^2}{2} f(N) P,$$

With this new evolution equation for the lapse function, they analyze the principal part of the equations and see that if $f > 0$ then it has imaginary eigenvalues and a complete set of eigenvectors. Thus, up to the smoothness requirement on the eigenvalues with respect to the wave vector, the system seems to be at least strongly hyperbolic and so well posed. This prescription enlarges the system a bit; as one also has to correctly include in it the corresponding equations for the first space derivatives of the lapse function, which become now a dynamical variable.

Although the system is well posed in the sense of the theory of partial differential equations, it has some instabilities from the point of view of the ordinary differential equations. A quick look at the toy model in [5] shows that if we take constant initial data for (in that paper's notation) $\alpha = \alpha_0$, $g = g_0$, and K_0, and null data for A, and D, then the resulting system is just a coupled set of ordinary equations. One can see that $g = g_0(K/K_0)^2$, and so

$$\frac{\dot{\alpha}}{K} = (\frac{\alpha}{K})^2 \frac{K_0^2}{g_0}(1 - f).$$

If $f = 1$, the harmonic time gauge in this notation, nothing happens at first sight. See nevertheless [5]. If $f > 1$ and $\frac{\alpha_0}{K_0} > 0$ then we have a singularity in a finite time. The same happens if $f < 1$ and initial data is taken so that $\frac{\alpha_0}{K_0}$ is negative. Thus we see that this gauge prescription can generate singularities which do not have much to do with the propagation modes, and so with the physics of the problem. In [5] and [6] numerical simulations have been carried out to study this problem. Needless to say, these instabilities would initially manifest themselves in numerical calculations via the forming of large gradients on the various fields coupled to the above fields, and the time at which they appear depends on the size of the trace of the momentum variable. In [6] a proposal to deal with this problem is made which consists of smoothing out the lapse via a parabolic term. In view of the fact that this problem already arises for constant data, it is doubtful that such a prescription can cure it.

Note that the above prescription for the evolution of the lapse for $f = f_0 > 0$ is identical to the one considered in [35], namely equation 8. It is most probably the case then that the same sort of instability would be present there, although the equations considered there are different, due to the inclusion

of terms proportional to the scalar constraint in order to render the system symmetric hyperbolic.

In [5] there is also a study of another type of singularity which is not ruled out with the choice of the harmonic gauge, $f = 1$. This singularity seems to be of a different nature, and is probably related to the instability of the harmonic gauge already mentioned. It clearly has to do with the non–linearities of the theory.

It should be mentioned that there are a wide variety of possibilities for making bigger hyperbolic systems out of those which are hyperbolic for a prescribed lapse-shift pair, or for the generalized harmonic gauge variant. In that respect, perhaps the systems which are more amenable to a methodological and direct study are the ones in the frame or in Ashtekar's representations, for there, as discussed in the previous section for the Ashtekar's representation systems, § 4.3, the possibilities to enlarge the system and keep it symmetric hyperbolic are quite clear and limited.

5.3 Elliptic Extensions

These are other types of approaches which take more into account the longitudinal modes of the theory, namely those which are related to the energy or matter content of the space time, and those which do not propagate as waves. The nature of these modes implies that these approaches seem to need a global knowledge of the solution, which in practice appears by imposing either elliptic or parabolic equations, the latter as a way to drive the solution close to satisfying an elliptic equation for larger times. Systems of that sort have already been used in applications: In [23], a hyperbolic system with lapse given by an elliptic equation is used in the proof of global existence of small data. The elliptic equation is used to impose the maximal slice condition during evolution, that is $P \equiv 0$. In that work, the *first order system* is for the electric and magnetic parts of the Weyl tensor, while the metric, connection, and extrinsic curvature tensor are obtained by *solving* elliptic equations on each slice. For their aims, obtaining a priori estimates, this suffices. For numerical simulations of evolution, it is better to solve, as much as possible, evolution equations, and not elliptic ones. Thus for this aim, equations –hopefully hyperbolic or at least parabolic– should be added to evolve the above mentioned (lower order in derivatives) variables. This has improved recently in [22], and [3] with a slight generalization to [23] in admitting arbitrarily prescribed P's. In particular, in [22] a complete proof of well posedness of mixed symmetric hyperbolic–elliptic systems is given. Such a proof must be implicit somewhere in [23], and a general argument has been given in [34]. Surprisingly, such a result, which has a rather simple argument based on the standard elliptic and hyperbolic estimates, has not before had the clean proof it deserves. This gauge has been used to show existence of near Newtonian solutions by [56].

In [34] a different elliptic condition is imposed in order to study near Newtonian solutions. An elliptic system is considered for both lapse and shift. It is similar, but not equal, to the above gauge, for in this work a much stronger

condition is required on the order of approach of relativistic solutions to the Newtonian limit. This implies globally controlling not only the lapse, but also the shift.

The last two works mentioned hint at some interplay between the problems of finding well behaving gauges for near Newtonian solutions and for long term evolution. The argument has been that, since in this gauge the principal part of the equations is well behaved near the –singular– Newtonian limit, and since the rest of the terms of the hyperbolic system go to zero on that limit, one expects for the time the solution exists to go to infinity as one approaches the Newtonian limit. Thus the gauge should be well behaved until then. I cite [43] for recent work on this and [44] for a well behaved system in asymptotically null slices amenable to study slow solutions near null infinity.

In the frame and in Ashtekar's representations one could even consider first order elliptic (spinorial) equations to fix the gauge variables. In the frame representation one can even fix gauge variables via an algebraic condition.

6   The Role of the Constraints

In this section, I analyze the role the constraint equations play in these new formulations. I will start by analyzing how the equations were handled in the original proof of well posedness. Then I discuss what has been said recently and what has to be done in the future, both for theoretical considerations and for numerical work. In particular, I discuss the role of the constraints in the initial–boundary value problem, a problem which has still to be fully solved, but which nevertheless is applied in most of the numerical simulations.

6.1   The Constraints in the Harmonic Gauge

In the harmonic gauge, the role of the constraint equations as such is hidden. This is because the constraint equations become evolution equations due to the gauge choice. Indeed, in the full harmonic gauge all components of the Einstein tensor become a big second order hyperbolic equation for the metric components. The need for the constraints to be satisfied at the initial surface enters the picture because it implies that the first time derivative of the gauge condition must vanish there in order to guarantee that the constraints vanish everywhere. Thus the question about the preservation of the constraint equations does not appear here. At most one can say it has been traded for the question of the gauge consistency.

6.2   The Constraints in the New Systems: Theoretical Considerations

In the new hyperbolic systems, where covariance is lost, one solves only for the evolution equations. Thus, the question of whether the constraint equations hold during evolution if they hold at the initial surface arises again. If the problem is about the evolution of the whole space–time, or about evolution on the domain of dependence of some space–like surface, then there is a good argument showing that the constraint equations would be satisfied as a consequence of the uniqueness of the system under consideration:

Assume initial data is given at some space–like hypersurface which satisfies the constraints there. We use the new evolution system and get a solution in the domain of dependence of the system, (which, if gauges propagate at speeds greater than light, might be smaller than the domain given by the metric). But using the harmonic gauge I know that there is a solution to the Einstein equation on a maximally extended domain of dependence. If one can diffeomorphically transform metric corresponding to that solution into one satisfying the gauge used for the evolution with the new system, then, since it satisfies all the equations, including the constraints, it follows that it will also satisfy the equations of the new system. Uniqueness of solutions of the new system implies it must be the one found initially and so it also satisfies the constraints. Thus we see that no particular consideration for the constraint equations is needed.

6.3 The Constraints in the New Systems: Numerical Considerations

For numerical simulations, the role of the constraint equations is delicate: Since there are always numerical errors, although the vector field defined by Einstein's equations in any of the above approaches is tangent to the constraint sub-manifold, we can only expect to be in the neighborhood of that sub-manifold. So, since one is effectively modifying the evolution equations outside the sub-manifold, that the vector field is tangent to it is not enough. For if that sub-manifold were unstable, it could very well be that a spurious numerical solutions could start growing during evolution and takes us completely away from it. This problem has been noticed by several people and has been considered in detail in [33]. There it is shown that, while in some evolution systems, the constraints themselves obey hyperbolic evolution equations, in others that is not the case so are presumably unstable.

It is not clear to me that the condition that the constraint system be well posed is the one needed for considering a system free of this problem. First because well posedness as such is not enough to guarantee the possibility of a numerical scheme: The system could be well posed but still depart exponentially from the constraint sub-manifold, thus making impossible any reliable calculation. So the non–principal part of the system must also be considered, and probably suitably modified in the neighborhood of the constraint sub-manifold. Second, since one is never solving, or simulating, the constraint evolution equations, that is, they play no role in the scheme, why should one consider them at all? I think the emphasis should be put on guaranteeing a numerical scheme without spurious solutions; because, as argued above, uniqueness should imply that the constraints are satisfied. Thus, what seems to be needed is a connection between well posedness, or rather no exponential departure from the constraint sub-manifold, and lack of spurious solutions on the numerical schema.

6.4 The Constraints in the Initial–Boundary Value Problem

A completely different situation arises if one is considering an initial–boundary value problem for Einstein's equations. Although this problem has not been solved in its full generality for Einstein's equations, it is clear that in order for the constraints to be satisfied during evolution, some of the boundary values have to be chosen in a special way. It is here that the type of equations the constraints satisfy is most important. In particular, if they also form a hyperbolic system, then a study of its principal part at the boundary would tell which conditions are needed to guarantee uniqueness of the solutions, in particular the trivial solution, and so which are the boundary conditions we must force upon the evolution system for the dynamical fields. Since most numerical simulations are in fact initial–boundary value problems, the problem of well posedness and the problem of the propagation of the constraints are central.

References

[1] Abrahams, A., Anderson, A., Choquet-Bruhat, Y., and York Jr., J.W., "Hyperbolic formulation of general relativity", (March, 1997), [Online Los Alamos Archive Preprint]: cited on 18 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9703010`. 4.1

[2] Abrahams, A., Anderson, A., Choquet-Bruhat, Y., and York Jr., J.W., "Einstein and Yang-Mills theories in hyperbolic form without gauge fixing", *Phys. Rev. Lett.*, **75**, 3377–3381, (1995). For a related online version see: A. Abrahams, et al., "Einstein and Yang-Mills theories in hyperbolic form without gauge fixing", (June, 1995), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9506072`. 4.1

[3] Abrahams, A., Anderson, A., Choquet-Bruhat, Y., and York Jr., J.W., "A nonstrictly hyperbolic system for the Einstein equations with arbitrary lapse and shift", *C. R. Acad. Sci.*, **323**(serie IIb), 835–841, (1996). For a related online version see: A. Abrahams, et al., "A nonstrictly hyperbolic system for the Einstein equations with arbitrary lapse and shift", (July, 1996), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9607006`. 5.1, 5.3

[4] Abrahams, A., Anderson, A., Choquet-Bruhat, Y., and York Jr., J.W., "Geometrical hyperbolic systems for general relativity and gauge", *Class. Quantum Grav.*, **14**, A9–A22, (1997). For a related online version see: A. Abrahams, et al., "Geometrical hyperbolic systems for general relativity and gauge", (May, 1996), [Online Los Alamos Archive Preprint]: cited on 16 April 1997, `http://xxx.lanl.gov/abs/gr-qc/9605014`. 4.1

[5] Alcubierre, M., "Appearance of coordinate shocks in hyperbolic formalisms of general relativity", *Phys. Rev. D*, **55**, 5981–5991, (1997). 5, 7, 5.1, 5.2

[6] Alcubierre, M., and Masso, J., "Pathologies of hyperbolic gauges in general relativity and other field theories", (September, 1997), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9709024`. 5.2

[7] Ashtekar, A., "New Hamiltonian formulation of general relativity", *Phys. Rev. D*, **36**(6), 1587–1602, (1987). 4.3

[8] Ashtekar, A., *New Perspectives in Canonical Gravity*, (Bibliopolis, Napoli, 1988). 4.3

[9] Bona, C., and Massó, J., "Hyperbolic evolution system for numerical relativity", *Phys. Rev. Lett.*, **68**, 1097–1099, (1992). 4.1

[10] Bona, C., Massó, J., Seidel, E., and Stela, J., "First order hyperbolic formalism for Numerical Relativity", (September, 1997), [Online Los Alamos

Archive Preprint]: cited on 19 January 1998, http://xxx.lanl.gov/abs/gr-qc/9709016. 4.1, 4.3, 5.2

[11] Bona, C., Massó, J., Seidel, E., and Stela, J., "New Formalism for Numerical Relativity", *Phys. Rev. Lett.*, **75**, 600–603, (1995). For a related online version see: C. Bona, et al., "New formalism for numerical relativity", (December, 1994), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, http://xxx.lanl.gov/abs/gr-qc/9412071. 4.1, 5.2

[12] Bona, C., Massó, J., and Stela, J., "Numerical black holes: a moving grid approach", *Phys. Rev. D*, **51**, 1639–1645, (1995). For a related online version see: C. Bona, et al., "Numerical black holes: a moving grid approach", (December, 1994), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, http://xxx.lanl.gov/abs/gr-qc/9412070. 5.2

[13] Bruhat, Y., "Theoreme d' existence pour certain systemes d'equations aux deriveés partielles nonlinaires", *Acta Math.*, **88**, 141–225, (1952). 1.1

[14] Bruhat, Y., "Un theoreme d'inestabilite pour certain equations hyperboliques nonlinaires", *C. R. Acad. Sci.*, **276A**, 281, (1973). 7, 5.1

[15] Choquet-Bruhat, Y., "The Cauchy Problem", in Witten, L., ed., *Gravitation: an introduction to current research*, 68–103, (Wiley, New York, USA, 1962).

[16] Choquet-Bruhat, Y., "Espaces temps eineiniens généraux, chocs gravitationnels", *Ann. Inst. Henri Poincare*, **8**, 327–338, (1968). 1.1

[17] Choquet-Bruhat, Y., and Christodoulou, D., "Elliptic systems in $H_{s,\delta}$ spaces on manifolds which are Euclidean at infinity", *Acta Math.*, **145**, 129–150, (1981). 1.1

[18] Choquet-Bruhat, Y., Christodoulou, D., and Francaviglia, M., "Cauchy data on a manifold", *Ann. Inst. Henri Poincare, A*, **29**, 241–255, (1978). 1.1, 3

[19] Choquet-Bruhat, Y., and Ruggeri, T., "Hyperbolicity of the 3+1 System of Einstein Equations", *Commun. Math. Phys.*, **89**, 269–275, (1983). 4.1, 8, 5.2

[20] Choquet-Bruhat, Y., and York Jr., J.W., "The Cauchy Problem", in Held, A., ed., *General Relativity and Gravitation 1*, (Plenum, New York, USA, 1980).

[21] Choquet-Bruhat, Y., and York Jr., J.W., "Well posed reduced systems for the Einstein equations", *Comptes Ren. Acad. Paris*, **231**(I), 1089–1095, (1995). For a related online version see: Y. Choquet-Bruhat, et al., "Well posed reduced systems for the Einstein equations", (June, 1996), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, http://xxx.lanl.gov/abs/gr-qc/9606001.

[22] Choquet-Bruhat, Y., and York Jr., J.W., "Mixed elliptic and hyperbolic system for the Einstein equations". Editions Pitagora, (1996). For a related online version see: Y. Choquet-Bruhat, et al., "Mixed elliptic and hyperbolic system for the Einstein equations", (January, 1996), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, http://xxx.lanl.gov/abs/gr-qc/9601030. 5.3

[23] Christodoulou, D., and Klainerman, S., *The Global Nonlinear Stability of the Minkowski Space*, (Princeton University Press, Princeton, USA, 1993). 1, 3, 5.3

[24] Christodoulou, D., and O'Murchadha, N., "The Boost Problem in General Relativity", *Commun. Math. Phys.*, **80**, 271–300, (1981). 1.1

[25] Cutler, C., and Wald, R., "Existence of radiating Eintein-Maxwell solutions which are C^∞ on all of $\mathcal{J}^+$ and $\mathcal{J}^-$", *Class. Quantum Grav.*, **6**, 453–466, (1989). 1

[26] DeTurk, D., "The Cauchy problem for Lorentz metrics with prescribed Ricci curvature", *Comp. Math.*, **48**, 327–349, (1983). 3.1

[27] Fischer, A., and Marsden, J., "The Einstein Evolution Equations as a First-Order Symmetric Hyperbolic Quasilinear System", *Commun. Math. Phys.*, **28**, 1–38, (1972). 1.1

[28] Fischer, A. Marsden, J., "General relativity, partial differential equations, and dynamical systems", *AMS Proc. Symp. Pure Math.*, **23**, 309–327, (1973). 1.1

[29] Friedrich, H., "The asymptotic characteristic initial value problem for Einstein's vacuum field equations as an initial value problem for a first-order quasilinear symmetric hyperbolic system", *Proc. R. Soc. London, Ser. A*, **378**, 401–421, (1981). 1.1, 4.2

[30] Friedrich, H., "On the regular and the asymptotic characteristic initial value problem for Einstein's vacuum field equations", *Proc. R. Soc. London, Ser. A*, **375**, 169–184, (1981). 1.1, 4.2

[31] Friedrich, H., "On the hyperbolicity of Einstein's and other gauge field equations", *Commun. Math. Phys.*, **100**, 525–543, (1985). 7, 4.2

[32] Friedrich, H., "Hyperbolic reductions for Einstein's equations", *Class. Quantum Grav.*, **13**, 1451–1469, (1996). 1.2, 3, 7, 8, 4.2, 9

[33] Frittelli, S., "Note on the propagation of the constraints in standard 3+1 general relativity", *Phys. Rev. D*, **55**, 5992–5996, (1997). 3.2, 6.3

[34] Frittelli, S., and Reula, O., "On the newtonian limit of general relativity", *Commun. Math. Phys.*, **166**, 221–235, (1994). For a related online version see: S. Frittelli, et al., "On the Newtoninan limit of general relativity",

(June, 1995), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9506077`. 4.1, 4.3, 5.3

[35] Frittelli, S., and Reula, O., "First-order symmetric-hyperbolic Einstein equations with arbitrary fixed gauge", *Phys. Rev. Lett.*, **76**, 4667–4670, (1996). For a related online version see: S. Frittelli, et al., "First-order symmetric-hyperbolic Einstein equations with arbitrary fixed gauge", (May, 1996), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9605005`. 8, 5.2

[36] Geroch, R., "Partial Differential Equations of Physics", (February, 1996), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9602055`. Scottish Summer School in Theoretical Physics. 1.2, 6, 3

[37] Geroch, R., "The Local Nonsingularity Theorem", *J. Math. Phys.*, **24**(7), 1851–1858, (1983). 1.1

[38] Geroch, R., and Xanthopolous, B.C., "Asymptotic Simplicity Is Stable", *J. Math. Phys.*, **19**, 714–719, (1978). 1

[39] Gustaffson, B., Kreiss, H.-O., and Oliger, J., *Time-Dependent Problems and Difference Methods*, (Wiley, New York, USA, 1995). 1.2, 2, 2.4, 6

[40] Hadamard, J., *Lectures on Cauchy's problem in linear partial differential equations*, (Yale University Press, New Haven, Connecticut, 1921). 2.1

[41] Hawking, S.W., and Ellis, G.F.R., *The Large Scale Structure of Space-Time*, (Cambridge University Press, Cambridge, United Kingdom, 1973). 2, 3.1, 7

[42] Hughes, T., Kato, T., and Marsden, J., "Well-posed quasi-linear second order hyperbolic systems with applications to nonlinear elastodynamics and general relativity", *Arch. Rat. Mech. Anal.*, **63**, 273–294, (1976). 1.1, 3

[43] Iriondo, M., Leguizamón, E., and Reula, O., "Fast and Slow Solutions in General Relativity: The Initialization Procedure", (September, 1997), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, `http://xxx.lanl.gov/abs/gr-qc/9709078`. 5.3

[44] Iriondo, M., Leguizamón, E., and Reula, O., "The Newtonian Limit on Asymptotically Null Foliations", (September, 1997), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, `http://xxx.lanl.gov/gr-qc/970907`. 5.3

[45] Iriondo, M., Leguizamón, E., and Reula, O., "Einstein's equations in Ashtekar variables constitute a symmetric hyperbolic system", *Phys. Rev. Lett.*, **79**, 4732–4735, (1997). For a related online version see: M. Iriondo, et

al., "Einstein's equations in Ashtekar variables constitute a symmetric hyperbolic system", (October, 1997), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, http://xxx.lanl.gov/abs/gr-qc/9710004. 4.3

[46] John, F., *Formation of Singularities in Elastic Waves. Lecture Notes in Phys.*, (Springer-Verlag, 1984). 2, 4, 4.2

[47] Klainerman, S., "Uniform decay estimates and the Lorentz invariance of the classical wave equation", *Commun. Pure Appl. Math.*, **38**, 321–332, (1985). 1

[48] Klainerman, S., "The null condition and global existence to nonlinear wave equations", *Lect. Appl. Math.*, **23**, 293–326, (1986). 1

[49] Klainerman, S., "Remarks on the global Sobolev inequalities in Minkowski Space", *Commun. Pure Appl. Math.*, **40**, 111–117, (1987). 1

[50] Kreiss, H., and Lorentz, J., *Initial-Boundary Value Problems and the Navier-Stokes Equations*, (Acad. Press, San Diego, USA, 1989). 1.2, 2, 6

[51] Kreiss, H. O., "Über sachgemässe Cauchyprobleme", *Math. Scand.*, **7**, 71–80, (1959). 2

[52] Kreiss, H.-O., Nagy, G., Ortiz, O., and Reula, O., "Global Existence and Exponential Decay for Hyperbolic Dissipative Relativistic Fluid Theories", *J. Math. Phys.*, **38**, 5272–5279, (1997). For a related online version see: H.-O. Kreiss, et al., "Global Existence and Exponential Decay for Hyperbolic Dissipative Relativistic Fluid Theories", (February, 1997), [Online Los Alamos Archive Preprint]: cited on 19 January 1998, http://xxx.lanl.gov/abs/gr-qc/9702008. 6

[53] Lanczos, C., "Ein vereinfachendes Koordinatensystem für die Einsteinschen Gravitationsgleichungen", *Phys. Z.*, **23**, 537–539, (1922).

[54] Leray, J., *Hyperbolic Differential Equations*, (Institute for Advanced Studies, Princeton, New Jersey, 1953). 2.1, 2.2

[55] Leray, J., and Ohya, Y., "Equations et Systemes Non-Lineaires hyperboliques non-stricts", *Math. Ann.*, **170**, 167, (1967). 2.1, 2.2, 5.1

[56] Rendall, A., "The Newtonian limit for asymptotically flat solutions of the Vlasov-Einstein system", *Commun. Math. Phys.*, **163**, 89–112, (1994). For a related online version see: A. Rendall, "The Newtonian limit for asymptotically flat solutions of the Vlasov-Einstein system", (March, 1993), [Online Los Alamos Archive Preprint]: cited on 7 January 1998, http://xxx.lanl.gov/abs/gr-qc/9303027. 5.3

[57] Sideris, T., "Formation of Singularities in 3-d Compressible Fluids", *Commun. Math. Phys.*, **101**, 475–485, (1985).   2, 4

[58] Taylor, M., *Pseudodifferential operators and nonlinear PDE - Progress in Mathematics 100*, (Birkhäuser, Boston, USA, 1991).   3, 3, 6

[59] van Putten, M.H.P.M., and Eardley, D.M., "Nonlinear wave equations for relativity", *Phys. Rev. D*, **53**, 3056–3063, (1996). For a related online version see: M.H.P.M. van Putten, et al., "Nonlinear wave equations for relativity", (May, 1995), [Online Los Alamos Archive Preprint]: cited on 17 January 1998, http://xxx.lanl.gov/abs/gr-qc/9505023.   4.2

[60] Wald, R., *General Relativity*, (The University of Chicago Press, Chicago, USA, 1984).   2