Gauge and constraint degrees of freedom: from analytical to numerical approximations in General Relativity

C. Bona and Dana Alic
Departament de Fisica, Universitat de les Illes Balears.
Institute for Applied Computation with Community Code (IAC3).

Abstract. The harmonic formulation of Einstein’s field equations is considered, where the gauge conditions are introduced as dynamical constraints. The difference between the fully constrained approach (used in analytical approximations) and the free evolution one (used in most numerical approximations) is pointed out. As a generalization, quasi-stationary gauge conditions are also discussed, including numerical experiments with the gauge-waves testbed. The complementary 3+1 approach is also considered, where constraints are related instead with energy and momentum first integrals and the gauge must be provided separately. The relationship between the two formalisms is discussed in a more general framework (Z4 formalism). Different strategies in black hole simulations follow when introducing singularity avoidance as a requirement. More flexible quasi-stationary gauge conditions are proposed in this context, which can be seen as generalizations of the current 'freezing shift' prescriptions.

1. Harmonic formulation

Einstein’s field equations

\[ R_{ab} = 8 \pi (T_{ab} - \frac{1}{2} T g_{ab}), \]  

are usually expressed as a system of partial differential equations for the space-time metric \( g_{ab} \). Soon after Einstein’s 1915 paper, their mathematical structure was closely investigated, leading to a very convenient formulation (De Donder 1921, 1927), namely

\[ \frac{1}{2} g^{cd} \partial_{cd} g_{ab} + \partial_{(a} H_{b)} = \Gamma_{cab} H^c + 2 g^{cd} g^{ef} [\partial_e g_{ac} \partial_f g_{bd} - \Gamma_{ace} \Gamma_{bd} f] - 8 \pi (T_{ab} - \frac{T}{2} g_{ab}), \]  

where indices inside round brackets are symmetrized and we have noted

\[ H^a \equiv -g^{bc} \Gamma^a_{bc} = 1/\sqrt{g} \partial_{(a}(\sqrt{g} g^{ab}). \]  

One can now take advantage of the general covariance of the theory. Let us define spacetime coordinates by a set of four independent harmonic functions, namely

\[ \Box x^a = 1/\sqrt{g} \partial_{(a}(\sqrt{g} g^{ab}) = 0, \]  

where the box stands for the general-covariant wave operator acting on functions. In this harmonic coordinate system, the field equations get the simpler form

\[ \Box g_{ab} = \cdots - 16 \pi (T_{ab} - \frac{T}{2} g_{ab}), \]  

\[ (1) \]  

\[ (2) \]  

\[ (3) \]  

\[ (4) \]  

\[ (5) \]
where the dots stand for terms quadratic in the metric first derivatives. If we look at the principal part (the second derivatives terms), we see just a set of independent wave equations, one for every metric component. The coupling comes only through the right-hand-side terms.

System (5) is very convenient in analytical approximations. Let us assume for instance that the metric admits a development of the form
\[ g_{ab} = \eta_{ab} + h^{(1)}_{ab} + h^{(2)}_{ab} + \cdots, \]
where \( h^{(n)} \) is the \( n \)-th order perturbation. Then, one can express (5) in a recursive way:
\[ \eta^{cd} \partial^2_{cd} h^{(n+1)}_{ab} = F_{ab}(h^{(r)}, \partial h^{(s)}) \quad r, s \leq n, \]
which can by integrated just by inverting the standard (flat-space) wave operator.

System (5), however, is not equivalent to the original field equations (2). The coordinate conditions (4) can be interpreted as first-order constraints to be imposed on the solutions of the ‘relaxed’ system (5). The hard point in proving the well-posedness of the Cauchy problem for Einstein’s equations was precisely to prove that the harmonic constraints (4) where actually first integrals of the relaxed system (Choquet-Bruhat 1952). In the analytical perturbation framework, this translates into the fact that fulfilling the harmonic constraints at the \( n \)-th level does not imply the same thing at the next level. Obtaining a true solution of the Einstein equations implies adjusting the integration constants in such a way that
\[ \hat{H}^{(n+1)}_{a} \equiv 0, \]
and this must be done at every order in the perturbation development.

1.1. Numerical Relativity applications

The relaxed system (5) is also very useful in numerical approximations. The usual practice is using explicit time-discretization algorithms. This means that the metric coefficients are computed at a given time slice, assuming that one knows their values at the previous ones. But again fulfilling the harmonic constraints (4) is not granted. Moreover, in numerical approximations one has no adjustable integration constants. This means that numerical errors make the contracted Christoffel symbols obtained from the relaxed system to depart from their assumed harmonic (zero) value:
\[ \langle \text{relaxed} \rangle \Gamma^{a}_{b} \neq 0. \]

In this ‘free evolution’ approach, one can use the non-zero values (9) to monitor the quality of the simulation. This can be done by introducing a ‘zero’ four-vector \( Z^{a} \) as the difference between the relaxed and the harmonic (zero) contracted Christoffel symbols, namely
\[ \langle \text{relaxed} \rangle \Gamma^{a}_{b} - \langle \text{harmonic} \rangle \Gamma^{a}_{b} = -2Z^{a}, \]
so that true Einstein’s solutions would correspond to \( Z^{a} = 0 \), which amounts to fulfilling the harmonic constraints. On the contrary, allowing for (10), solutions of the relaxed system would verify a generalized version of (2), in which (3) must be replaced by
\[ H^{a} \equiv -g^{bc} \Gamma^{a}_{bc} - 2Z^{a}. \]
Gauge and constraint degrees of freedom

The vector $Z^a$ provides a new degree of freedom which arises quite naturally in numerical simulations. A fully covariant description is provided by the 'Z4 system' (Bona et al. 2003)

$$R_{ab} + Z_{a;b} + Z_{b;a} = 8\pi (T_{ab} - T/2 g_{ab}) ,$$

(12)

where the semicolon stands for the covariant derivative. We can easily see that the true Einstein’s solutions are recovered when $Z^a = 0$. Moreover, allowing for the contracted Bianchi identities, the stress-energy tensor conservation implies

$$g^{bc} Z_{a;bc} + R_{ab} Z^b = 0 .$$

(13)

It is clear that the vanishing of $Z^a$, leading to true Einstein’s solutions, is a first integral of the 'subsidiary system' (13), which follows from the field equations just assuming the stress-energy tensor conservation.

1.2. Testing coordinate conditions

We have seen how harmonic coordinates can be preferred for the sake of simplifying the (approximate) integration of the field equations. But one would like to use instead the General Relativity coordinate freedom for simplifying the physical interpretation of the results. On the contrary, choosing a convoluted coordinate system can complicate even the simplest physical situations.

A good example of these unphysical gauge complications is the 'gauge waves' testbed (Alcubierre et al 2004). The Minkowsky (flat) metric can be written in some non-trivial harmonic coordinate system as:

$$ds^2 = F(x - t)(-dt^2 + dx^2) + dy^2 + dz^2 ,$$

(14)

where $F$ is an arbitrary function of its argument. One could naively interpret this as the propagation of an arbitrary wave profile with unit speed. But it is a pure gauge effect, because (14) is nothing but the Minkowsky metric in disguise. A more natural coordinate system should be adapted to the fact that flat spacetime is stationary, and this is not granted by the harmonic condition, as (14) shows dramatically.

The problem of finding 'quasi-stationary coordinates' (as stationary as possible) in a generic spacetime has been addressed recently (Bona et al 2005a). The idea is to find 'almost-Killing' vector fields $\xi^a$ by means of a standard variational principle

$$\delta S = 0 , \quad S = \int L \sqrt{g} d^4x ,$$

(15)

where the Lagrangian density $L$ is given by

$$L = \xi_{(a;b)}\xi^{(a;b)} - k\frac{2}{2} (\xi^c_c)^2$$

(16)

($k$ being an arbitrary parameter), and the variations of the vector field $\xi$ are considered in a fixed spacetime. The resulting Euler-Lagrange equations get the form

$$[ \xi^{a;b} + \xi^{b;a} - k\xi^c_c g^{ab} ; b ] = 0 .$$

(17)

('almost-Killing' equation), or the equivalent one

$$g^{bc}\xi_{a;bc} + R_{ab} \xi^b + (1 - k) \partial_a (\xi^c_c) = 0 .$$

(18)

We will consider here the particular 'harmonic' choice $k = 1$. Note that, in this case, the subsidiary system (13) is nothing but condition (15) for the $Z^a$ vector. We can then interpret that the combination $Z_{(a;b)}$ in the Z4 system (12) gets minimized, so
Gauge and constraint degrees of freedom

Figure 1. Gauge waves simulation with periodic boundary conditions and sinusoidal initial data for the metric component $g_{xx}$. The left panel corresponds to the harmonic coordinates case. After 100 round trips, the evolved profile (cross marks) nearly overlaps the initial one (continuous line), which corresponds also with the exact solution in the harmonic case. The right panel corresponds to the same simulation for the quasi-stationary coordinates defined by (19). We show the initial data and the evolved profiles after 1, 2, 5 and 100 round trips (continuous line). The solution clearly approaches the Minkowsky value ($g_{xx} = 1$), although a small residual profile remains.

that one gets as close as possible to the original Einstein system. The name 'harmonic' for the $k = 1$ choice can be justified if we take the integral curves of $\xi$ to be the time lines of our coordinate system. In this 'adapted coordinate system', condition (17) reads simply

$$g^{bc} \partial_t \Gamma^a_{bc} = 0,$$

which is a close generalization of the harmonic coordinates condition (see Bona et al 2005a for more details).

In order to test the behavior of these quasi-stationary conditions, we will consider the 'gauge waves' form (14) of the flat metric, with the following profile:

$$F = 1 - A \sin(2\pi(x - t)) \quad A = 0.1,$$

so that the resulting metric is periodic and we can identify for instance the points $-0.5$ and $0.5$ on the $x$ axis. This allows to set up periodic boundary conditions in numerical simulations, so that the initial profile keeps turning around along the $x$ direction. Note that the coordinate conditions (17) require the use of some damping term in order to ensure the stability of the solutions (see Bona et al 2005b for details).

The results of the numerical simulations are displayed in Fig. 1. The left panel shows the harmonic case, where the numerical results follow very closely the exact solution. Only a small amount of numerical dissipation is visible after 100 round trips (we are using here a third-order-accurate finite volume method in order to get rid of the dominant dispersion error). The right panel shows the behavior of the quasi-stationary condition (19) for the same initial data. We see that the amplitude is quickly decreasing, so that we get very close to the stationary (Minkowsky) value $g_{xx} = 1$ after only 5 round trips. Although a small residual profile remains, even after 100 round trips (continuous line), the initial amplitude has been reduced by an order of magnitude.
2. 3+1 formulation

About thirty years after the advent of the harmonic formalism, the 'evolution formalism' (Lichnerowicz 1944, Choquet 1956) provided a breakthrough in the understanding of Einstein’s equations structure. Four-dimensional spacetime was described as sliced by constant-time hypersurfaces. Space-time dynamics was then described as the time evolution of the three-dimensional geometry of these surfaces. This '3+1 formalism' was very convenient at that time for studying the initial value problem and it is widely used today in numerical applications.

The four-dimensional metric can be adapted to the 3+1 geometry in the following way:

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

where \( \gamma_{ij} \) is the three-dimensional metric on every slice. The 'lapse function' \( \alpha \) measures the proper-time versus coordinate-time ratio when moving along the lines normal to the slices (\( \beta^i = 0 \)). We will see how to take advantage of this degree of freedom in what follows. The 'shift' \( \beta^i \) measures in turn the deviation between these normal lines and the time lines. It does not affect the geometry of the slicing in any way. The extrinsic curvature of the slices is the proper time derivative of the space metric along the normal lines, namely:

\[
K_{ij} = -\frac{1}{2\alpha} \, (\partial_t - L_{\beta}) \, \gamma_{ij},
\]

where \( L \) stands for the Lie derivative.

In this framework, the set of ten Einstein’s field equations \( \Box \) can be decomposed into two different subsets (see Choquet 1967 for details). The space components provide six first-order evolution equations for \( K_{ij} \) (amounting to second-order evolution equations for \( \gamma_{ij} \)). The components

\[
n_b \, (G^{ab} = 8 \pi \, T^{ab})
\]

(where \( G_{ab} \equiv R_{ab} - \frac{1}{2} R g_{ab} \) is the Einstein tensor and \( n_a \) is the field of unit normals to the time slicing) provide instead four constraint equations for the pair \( \gamma_{ij}, K_{ij} \). These 'energy-momentum constraints' arise just from the field equations, independently of the coordinate gauge. Moreover, one does not get here any evolution equation for the lapse or the shift, which are just kinematical quantities. In order to determine them, one must specify four coordinate conditions, which must be added to the field equations in order to complete the evolution system. One can provide for instance four extra evolution equations

\[
(\partial_t - \beta^k \partial_k)\alpha = -\alpha^2 Q, \quad (\partial_t - \beta^k \partial_k)\beta^i = -\alpha Q^i,
\]

where \( Q \) and \( Q^i \) can be freely specified.

In order to see the relationship between the harmonic and the 3+1 formalisms, let us decompose the contracted Christoffel symbols \( \Gamma^a \equiv g^{bc} \Gamma^a_{bc}, \) namely:

\[
n_a \Gamma^a = \alpha \, \Gamma^0 = Q - t \, K, \quad \alpha \, \Gamma_i = Q_i - \partial_i \alpha + \alpha \, (3) \Gamma_i.
\]

It follows that fixing the value of \( \Gamma^0 \) amounts to provide the evolution equation for the lapse (which determines the time slicing), whereas fixing the value of \( \Gamma_i \) amounts to provide the evolution equation for the shift (which determines the time lines for a given slicing). The main difference is that the coordinate conditions \( \Box \) where introduced as constraints in the harmonic formalism, whereas the corresponding 3+1 conditions \( \Box \) are part of the evolution system. This will have important consequences in what follows.
2.1. Gravitational collapse scenarios

Gravitational collapse in General Relativity can lead to the arising of spacetime singularities, even from regular initial data (a massive star for instance). This is a serious issue in Numerical Relativity, as the computation can not proceed beyond singularity formation, even if the outside spacetime regions are regular. The lapse degree of freedom can be used in these cases to allow the computation to continue (in coordinate time) by locally diminishing the lapse (and then the proper time flow) in the collapsing regions, where the space volume element $\sqrt{\gamma}$ is diminishing. When done properly, the time slices do not reach the collapse singularity in a finite amount of coordinate time (singularity avoidance).

Let us illustrate this behavior by considering the harmonic slicing condition. Allowing for (25), this means to take $Q = \text{tr} K$. As far as singularity avoidance is a geometrical property of the slicing, the value of the shift is irrelevant, so we will use normal coordinates (zero shift) for simplicity. This condition can be easily integrated: allowing for the definitions (22, 24) we get

$$\partial_t \left( \sqrt{\gamma}/\alpha \right) = 0 \quad \Rightarrow \quad \alpha \sim \sqrt{\gamma}.$$  \hspace{1cm} (26)

It follows that the lapse tends to vanish locally (collapse of the lapse) where the space volume element goes to zero. Although one gets in this way arbitrarily close to the singularity, it can be shown that one does not actually reach it in a finite amount of coordinate time (Bona and Massó 1988).

The singularity avoidance of the harmonic slicing is really a limit case, vulnerable to numerical errors. This is a problem in gravitational collapse scenarios, specially in the harmonic formalism, where we must remember that the harmonic conditions where not part of the relaxed system: they were just dynamical constraints. This is why numerical codes based in the harmonic formalism currently excise the singularity-forming regions out of the computational domain. This amounts to set up an artificial internal boundary in the strong field region. Moreover, this must be a dynamical boundary, capable of changing shape or moving across the numerical grid. This is a difficult issue, but still feasible. This has been achieved for instance in the most recent Numerical Relativity breakthrough, where a binary black hole simulation lasted long enough for extracting a consistent gravitational wave signal for the first time (Pretorius 2005).

An obvious alternative to dynamical excision is to consider time slicing conditions with stronger singularity-avoidance properties. This can be done by generalizing the harmonic slicing condition in the following way (Bona et al 1995):

$$Q = f \text{ tr} K,$$  \hspace{1cm} (27)

where $f$ is an arbitrary function of the lapse, so that the original harmonic slicing is recovered for $f = 1$. Most of the current binary-black-hole simulations in the BSSN formalism actually use the slicing condition (27) with the choice $f = 2/\alpha$. This corresponds to the ‘1+log’ slicing condition (Bernstein 1993), the name coming from the resulting form of the lapse in normal coordinates (zero shift):

$$\alpha = \alpha_0 + \ln\left(\gamma/\gamma_0\right).$$  \hspace{1cm} (28)

It follows from (28) that the coordinate time evolution stops before even getting close to the collapse singularity. It is easy to see in this case that the time evolution has a fixed point where the lapse vanishes (a limit surface). In normal coordinates, this happens when

$$\gamma = \gamma_0 \exp\left(-\alpha_0\right),$$  \hspace{1cm} (29)
that is well before the vanishing of the space volume element (the initial lapse value being usually close to one). The appearance of a limit surface provides a safety margin for black hole simulations, as far as numerical errors have no chance to result into hitting the singularity.

Unfortunately, singularity-avoidance does not come for free. Time slices get distorted in the process, mainly by a length increase along the radial direction (slice stretching, see Reimann and Brügmann 2004), which is nevertheless compatible with the overall collapse. This causes a progressive loss of resolution that can produce high-frequency noise in numerical simulations where the grid gets too coarse to resolve the profiles of the dynamical fields. This is analogous to the well known Gibbs phenomenon, and usually leads to code crashing as far as the noise starts growing quickly.

Slice stretching is inherent to singularity avoidant slicing conditions, but the appearance of high-frequency noise can be delayed, even avoided, in many ways. One can just increase the grid resolution or use instead more specific tools: artificial dissipation terms (Gustavson et al 1995) or advanced finite volume methods, like the ones currently used in Computational Fluid Dynamics (for a recent implementation, see Alic et al 2007). Just after Pretorius paper, many groups published improved binary-black-hole simulations, obtained with 3+1 codes using the ‘1+log’ slicing condition (Campanelli et al 2006, Baker et al 2006, Diener et al 2006).

2.2. Quasi-stationary gauge conditions and singularity avoidance

At this point, it is interesting to ask whether the quasi-stationary coordinate conditions derived from the variational principle (15) can be made compatible with the singularity avoidance requirement. Note that, when adapting our spacetime coordinates to a solution \( \xi \) of the almost-Killing equation (17), we are demanding two different things: the time lines are chosen to be the integral curves of \( \xi \) and the time coordinate is chosen to be the preferred affine parameter associated to these lines. Although the first requirement fits perfectly into the idea of getting a quasi-stationary gauge condition, the second one lacks of a clear physical motivation. It seems that singularity avoidance is not enforced in this way.

A better strategy is to choose a priori the time coordinate, that is a spacetime slicing given by

\[
\phi(x^a) = \text{constant},
\]

with a view to enforce singularity avoidance. Then, we can use this time coordinate as a parameter along the integral lines of the almost-Killing vector \( \xi \), by requiring

\[
\xi^a \partial_a \phi = 1.
\]

This amounts to constraint the vector \( \xi \) to fulfill (31) in the minimization process. In other words, we introduce a Lagrange multiplier and minimize

\[
L' \equiv L + \lambda (\xi^a \partial_a \phi - 1)
\]

(for any \( \phi \) given a priori) instead of the original Lagrangian (18).

The resulting Euler-Lagrange equations include now the constraint (31), plus the system

\[
\left[ \xi^{a;b} + \xi^{b;a} - k \xi^{c}_{\ c} g^{a b} \right];_b = \lambda \partial^a \phi.
\]
which generalizes the almost-Killing equation (17). Using adapted coordinates,
\[ \phi = t , \quad \xi = \partial_t , \]
it can be written as
\[ g^{bc} \partial_t \Gamma^a_{bc} + (1 - k) g^{ab} \partial_t \Gamma^c_{bc} = \lambda g^{at} . \] (35)

Let us now split condition (35) into its 3+1 components. The time component can be ignored, because it just provides the value of the Lagrange multiplier itself. Remember that the time slicing was chosen a priori, so that we do not need any further condition for the lapse. The (downstairs) space component,
\[ g_{ia} g^{bc} \partial_t \Gamma^a_{bc} + (1 - k) \partial_t \Gamma^c_{ic} = 0 , \] (36)
provides a (second order) evolution equation for the shift. In this way, we have completely splitted the gauge conditions: the shift equation (36) should work with any time slicing, that can be freely chosen a priori.

The quasi-stationary shift condition (36) is completely independent of the value of the Lagrange multiplier. This means that we would get the same condition from the original unconstrained Lagrangian. We can conclude that the slicing constraint (31) does not affect the minimization process in the shift sector.

Let us test these conditions in a spherically symmetric black-hole simulation. We will write the Schwarzschild line element in the ‘wormhole’ form:
\[ ds^2 = - (\tanh \eta)^2 dt^2 + 4m^2 (\cosh \eta/2)^4 (d\eta^2 + d\Omega^2) , \] (37)
which can be obtained from the isotropic form by the following coordinate transformation
\[ r = m/2 \exp(\eta) . \] (38)

We will combine the ‘1+log’ slicing prescription with the quasi-stationary shift condition (36) with \( k = 1/2 \). We run the simulation up to 1000\( m \), as in the zero
shift case (normal coordinates). We see in Fig. 2 that the lapse collapses as usual, avoiding the singularity in both cases as expected. The effect of the shift is adding some outgoing speed to the grid nodes, so that the advance of the collapse front across the grid is delayed. We have added to the shift condition a standard damping term (Bona et al. 2005b) in order to avoid the shift values to grow without control. Note that the slopes at the collapse front are smoothed out, so they can be better resolved. The logarithmic character of the grid makes these effects less apparent for larger values of the \( \eta \) coordinate.

Our results prove that the quasi-stationary shift conditions are actually compatible with standard singularity-avoidant slicing conditions. We have found, however, that the results can depend crucially of particular value of the parameter \( k \). It is suggestive that, at least in the particular case shown here, the best results are obtained with the \( k = 1/2 \) choice. This is a very special value, because the minimum principle leads in this case to a minimisation of the conformal-Killing equation: a quasi-conformal shift condition. This opens an interesting perspective for future work.

**Acknowledgements**

This work has been supported by the Spanish Ministry of Science and Education through the research project number FPA2004-03666 and by the Balearic Conselleria d’Economia Hissenda i Innovació through the project PRDIB-2005GC2-06.

**References**

Alcubierre, M. et al. 2004, *Class. Quantum Grav.* 21(2), 589-613.
Alic, D. Bona, C. Bona-Casas, C. and Massó, J. 2007, *Phys. Rev.* D (in press), arXiv:0706.1189
Baker, J.G. et al. 2006, *Phys. Rev. Lett.* 96, 111102.
Bernstein, D. 1993, *Ph. D. Thesis* (Dept. of Physics, Univ. of Illinois at Urbana-Champaign).
Bona, C. and Massó, J. 1988, *Phys. Rev.* D38, 2419.
Bona, C. Massó, J. Seidel, E. and Stela, J. 1995, *Phys. Rev. Lett.* 75, 600.
Bona, C. Ledvinka, T. Palenzuela, C. and Žáček, M. 2003, *Physical Review* D67, 104005.
Bona, C. Carot, J. and Palenzuela-Luque, C. 2005, *Phys. Rev.* D72, 124010.
Bona, C. Lehner, L. and Palenzuela-Luque, C. 2005, *Phys. Rev.* D72, 104009.
Campanelli, M. Lousto, C.O. Marronetti, P. and Zlochower, Y. 2006, *Phys. Rev. Lett.* 96, 111101.
Fournès-Bruhat, Y. 1952, *Acta Math.* 88, 141.
Choquet-Bruhat, Y. 1956, *J. Rat. Mec. Analysis* 5, 951.
Choquet-Bruhat, Y. 1967, in *Gravitation: An Introduction to Current Research* (L. Witten, editor). John Wiley, New York.
De Donder, T. 1921, *La Gravifique Einsteinienne*, Gauthier-Villars, Paris.
De Donder, T. 1927, *The Mathematical Theory of Relativity*, Massachusetts Institute of Technology, Cambridge, MA.
Diener, P. et al. 2006, *Phys. Rev. Lett.* 96, 121101.
Gustafsson, B. Kreiss, H.O. and Oliger, J. 1995, *Time dependent problems and difference methods* (New York: Wiley).
Lichnerowicz, A. 1944, *J. Math. Pures Appl.* 23, 37.
Pretorius, F. 2005, *Phys. Rev. Lett.* 95, 121101.
Reimann, B. and Brügmann, B. 2004, *Phys. Rev.* D69 044006.