Nonlinear impact of perturbation theory on numerical relativity

Edward Seidel

Max Planck Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Golm, Germany
and
Department of Physics and Astronomy, Louisiana State University, 202 Nicholson Hall, Baton Rouge, LA 70803-4001, USA

Received 23 October 2003
Published 13 January 2004
Online at stacks.iop.org/CQG/21/S339 (DOI: 10.1088/0264-9381/21/3/021)

Abstract
I discuss the impact of gauge-invariant perturbation theory, as developed originally by Vincent Moncrief, on numerical simulations of Einstein’s theory. Far from being replaced by numerical relativity, perturbative approaches remain essential for analysing, interpreting, extending and complementing fully nonlinear approaches. In the last decade, as computers have become ever more powerful tools for studying the full nonlinear equations, the power and application of perturbation theory has also grown. Its impact on numerical relativity is profound (decidedly nonlinear), and will surely continue to be for years to come.

PACS numbers: 04.25.Dm, 11.15.Bt

1. Introduction

Numerical relativity is a field that aims to solve Einstein’s equations through the use of computers, using numerical methods to solve the partial differential equations. These days, this generally involves developing large codes to run on monstrous computers, with hundreds or thousands of processors, and techniques involving cache utilization, message passing, parallel I/O, grid computing, etc. All of these are needed to obtain—in language that might more usually be associated with Moncrief—‘exact’ solutions to Einstein’s equations, or perhaps ‘arbitrarily accurate’ would be more appropriate. That is, any well-behaved, convergent code should produce solutions to Einstein’s equations with errors known to be proportional to $\Delta x^n$, where $\Delta x$ is the grid spacing and $n$ is known; most commonly $n = 2$. Given a set of computing resources, errors away from the true solution should be reduced by a factor of 4 if one doubles the resolution. Hence, with enough computer power, in an ideal world errors can be made arbitrarily small, and full, dynamical solutions to the nonlinear Einstein equations can be made arbitrarily close to the exact solution.
When I began as Moncrief’s student 20 years ago, the basic state of numerical relativity was the following: 2D calculations were advancing; Smarr had just made his pioneering attempt to study the simplest two black-hole problems in axisymmetry [55]. 3D calculations, although extremely limited, were starting; Nakamura and co-workers had begun a program aimed at tackling 3D systems [44]. The basic numerical method for evolution was explicit leapfrog [54] (on which, surprisingly, Moncrief wrote a paper in Phys. Rev. D [43]). Although fully numerical, nonlinear work was underway, perturbation theory was still one of the primary tools for studying systems of interest; Chandrasekhar and Detweiler, among others, were computing quasinormal modes of black holes [28] and, as I discuss in more detail below, Moncrief and colleagues had developed a program to compute gravitational radiation from core collapse via perturbative techniques. Although highly mathematically inclined, Moncrief had an unusual appreciation for and insight into numerical approaches, and guided most of his early students (among them Tom Moore, Roger Ove and myself) into numerical projects. In my case, this meant extending previously developed perturbative approaches in a more numerical setting.

Since that time, computer power has grown tremendously, by roughly a factor of $10^5$. (As Moncrief’s student I was lucky enough to be working on a 200MFlop machine, whereas my students today have access to $10^9$ TFlop machines!) Given the ‘brute force’ arguments above, one might expect a completely different world today, where numerical relativity, armed with vastly improved computers and numerical methods, could routinely tackle highly complex, fully 3D, highly nonlinear systems, such as coalescing binary black holes. In this fantasy world, basic numerical methods such as leapfrog would be replaced by some advanced, high-order, constraint preserving scheme, huge computers would produce ‘exact’ solutions to the fully nonlinear field equations, and perturbation theory would probably be relegated to classes on the history of numerical relativity.

I was among those who, although ‘brought up right’ with a firm grounding in mathematical perturbation theory, believed that fully nonlinear, numerical approaches would quickly overtake the very limited perturbative approaches. Although I had many plans to extend the perturbative approaches I had learned from Moncrief, and to apply them to an array of problems involving black holes, neutron stars, core collapse, and so on, like many others I quickly moved on to fully nonlinear numerical approaches. But I never left perturbation theory behind; like the death of Mark Twain, its (future) demise was highly exaggerated. To my repeated surprise, it kept coming back, in many different and unexpected ways, as a crucially important technique for analysing, interpreting, complementing and even in extending results in numerical relativity.

Moncrief’s influence is felt in many areas of numerical relativity. But in this paper I pay special tribute to his important contributions in perturbative methods and their impact on numerical relativity. This impact has been decidedly ‘nonlinear’ in at least two respects. First, perturbative approaches often make predictions that are at least qualitatively valid even in nonlinear regimes where one would not, a priori, expect them to work. Second, over time, they have become even more useful, and are more widely used, than ever before, even as fully nonlinear approaches are reaching maturity. In summary, perturbative methods, based on Moncrief’s early work, have formed the continuing foundation for much work in numerical relativity over the last several decades, and probably the next several as well!

---

1 For example, his original work with Ove [45]; extensive work with Berger; his recent work with Andersson [14]. The latter, through discussions with Moncrief at AEI, motivated and inspired my group to further develop the ‘BSSN’ system of equations, especially a class of very powerful gauge conditions [11]. These conditions have helped enable the EU numerical relativity community to make very important progress in fully 3D nonlinear evolutions [33].
2. Gauge–invariant perturbation theory

In this section I review Moncrief’s seminal approach to perturbation theory that has become the foundation for important tools in numerical relativity. The fundamental problem driving much of the work of several generations of relativists is the response of a Schwarzschild or Kerr black hole to its environment. As early as 1957, Regge and Wheeler [47] studied stability of the Schwarzschild singularity to small perturbations. The standard approach is to write the full metric as a sum of the static Schwarzschild metric $g_{\alpha\beta}^0$ and a perturbation $h_{\alpha\beta}$ (in Moncrief notation):

\[ g_{\alpha\beta} = g_{\alpha\beta}^0 + \epsilon h_{\alpha\beta}. \]  

One then plugs this expression into the vacuum Einstein equations and expands them up to first order in $\epsilon$, to get equations for the perturbation tensor $h_{\alpha\beta}$. The angular part of the solution can be separated off by expanding $h_{\alpha\beta}$ in tensor spherical harmonics. For each $\ell, m$ mode, one gets separate equations for the perturbed metric functions, which we can denote by $h^{(\ell m)}_{\alpha\beta}$.

As recognized by Regge and Wheeler, the perturbation equations break up into two linearly independent classes, even- and odd-parity. The odd-parity class is considerably simpler, having just three components $h^{(\text{odd}, \ell m)}_{\alpha\beta}$, while the even-parity class $h^{(\text{even}, \ell m)}_{\alpha\beta}$ has seven. However, the equations governing these functions are extremely complex. By using a clever choice of gauge, Regge and Wheeler were able to simplify the odd-parity system and derive the famous Regge–Wheeler equation

\[ \frac{\partial^2 \psi^{(\ell m)}}{\partial r^* \partial t} + V_{\text{RW}}^{(\ell)}(r) \psi^{(\ell m)} = 0, \]  

where $r^*$ is the so-called tortoise coordinate, given by $r^* = r + 2M \ln (r/2M - 1)$, and the Regge–Wheeler potential $V_{\text{RW}}^{(\ell)}(r)$ is a function of $r$ and $M$ and contains all the physics of the problem. This equation has been studied for the last half-century, and has revealed much of what is currently known about black holes in the Universe. In particular, studies of the equation revealed that black holes have quasi-normal modes, damped sinusoidal oscillations whose wavelength and damping time are determined by the mass of the black hole.

The even-parity case is so much complicated that it was not solved until 13 years later, when Zerilli was able to derive an amazingly similar wave equation, which differs only in the details of the potential $V_{\text{Zerilli}}^{(\ell)}(r)$. It was yet later shown by Chandrasekhar to have the same spectrum of quasi-normal modes as in the odd-parity case. The fact that such very different classes of perturbations would lead to such similar wave equations was a clue that there must be a fundamentally simpler, more transparent approach to the problem.

When dealing with perturbations in relativity, one must be careful about interpreting the various metric components $h_{\alpha\beta}$ in terms of physics. Under a coordinate transformation of the form

\[ x^\mu \rightarrow x^\mu + \delta x^\mu, \]  

the metric coefficients will transform as well. Regge and Wheeler, and later Zerilli, were able to use this gauge freedom to eliminate certain metric functions to simplify the corresponding equations for the perturbations. However, considerable foresight is needed to pick the appropriate gauge for the equations to simplify, as exemplified by the fact that due to its sheer algebraic complexity the even-parity case was not solved for 13 years! And this was for the simplest, nontrivial spherical, static, vacuum spacetime, Schwarzschild. This overwhelming complexity would be far worse in the case of time-dependent systems, or systems with matter, such as collapsing stars.
Moncrief was the first to recognize and exploit the point that another approach might reveal the underlying physics of such systems much more transparently. The key idea is to consider linear combinations of the $h_{\alpha\beta}$ and their derivatives that are actually invariant under the transformation in equation (3), carried out only to first order in $\delta x$ (definition of gauge transformation). In his seminal 1974 paper [40], Moncrief developed this idea, assuming that if one constructs gauge-invariant quantities, which are clearly more closely connected to the underlying physics, then the analysis will proceed much more easily. This proved to be true, so that what had taken more than a decade to develop by the community, through ingenious choices of gauge and prodigious amounts of algebra, now could be derived very transparently in just a few pages!

The idea was very powerful. Far more than a trick to reduce a complex derivation down to size, it became a guiding principle for much work in rather different communities. The technique was now straightforward: first define a maximal set of gauge-invariant quantities on a given background, then compute their equations of motion, and entire theories of small perturbations on a given background could be (comparatively) easily derived. Although equations for gauge-dependent metric functions are highly tangled together, there is a considerable decoupling in the gauge-invariant approach. Physically relevant variables, which are naturally gauge invariant, are usually dominant in their own evolution equations. In an accompanying paper [41], Moncrief then applied this approach to the case of perturbations of a spherically symmetric, self-gravitating perfect fluid system, and again, the perturbation equations for the relevant physically important quantities are naturally developed using this approach.

This approach has been picked up and developed in various areas, including astrophysics and cosmology. But in the following sections I describe how it has impacted particularly the field of numerical relativity in the past, and continues to do so today.

3. Applications in numerical relativity

3.1. Perturbative time evolutions

The gauge-invariant approach devised by Moncrief could be applied not only to static backgrounds, but also to time-dependent ones (this would have been virtually impossible with the standard approach). This immediately gives rise to a clever form of numerical relativity, ‘numerical relativity light’. In this approach, the perturbation equations, transparent as they are with this gauge-invariant approach, can still be rather complex and need to be solved numerically. But because the different $(\ell, m)$ modes are decoupled, they can all be evolved independently as PDEs in one spatial dimension, reducing the 3D problem to 1D. Furthermore, the background, being time dependent, can be treated numerically as well, but it can be fully relativistic, nonlinear, strong field gravity. But again, it is reduced to a 1D (radial + time) PDE. Hence, this approach can be used in principle to treat virtually any fully relativistic, 3D problem, that is close to being spherical, as a set of 1 + 1 PDEs. This approach has proved very popular, and is still being used today, nearly 30 years after its development by Moncrief. I mention a few applications next.

Shortly after the publication of these first two papers on gauge-invariant perturbation theory, a series of papers studying the gravitational wave output from stellar core collapse was published by Cunningham, Price, and Moncrief (CPM) [30–32, 42]. Using the approach I just described, they took the background to be not a static star, but a dynamic, collapsing one. In this case, the stars were modelled by Oppenheimer–Snyder collapsing dust balls, and
the results showed for the first time that the quasinormal modes were strongly excited in the collapse to a black hole.

This work is not only still highly cited as evidence that collapse to a black hole produces quasinormal modes, the perturbative prediction has been shown to be remarkably robust. In virtually every system studied, where a black hole is formed, quasinormal ringing is observed, whether it is from nonlinear stellar collapse models [56, 52], black-hole collisions [15, 16, 10], or even the collapse of pure gravitational waves to form a black hole [9]. A key point to make is that predictions from perturbation theory turn out to be extremely robust, and often apply even in nonlinear regimes where they might not be expected to.

The approach of CPM was generalized in my own PhD thesis to the case of more general, time-dependent spherical backgrounds, with perfect fluid interiors matched to vacuum exteriors [51], and used with sophisticated (at the time!) core collapse/bounce codes to study gravitational wave emission from supernovae [49, 50]. Not recognizing how important this approach would continue to be, I later moved on to other work, but in the meantime the same approach was further (and better) developed by others, especially in recent years in [35, 36]. It continues to be applied to a series of problems, most recently by Shibata and colleagues [37]. Similar approaches are also underway at present in a large collaborative EU project [33].

3.2. Extraction

Throughout the 1970s and 1980s, numerical relativists struggled with the problem of extracting meaningful gravitational wave signals out of their numerical results. Even assuming the codes were faithfully solving Einstein’s equations, they were all written in a particular gauge that made identification of gravitational waves difficult. Many different measures of gravitational wave indicators were proposed, but all suffered from theoretical ambiguities, gauge dependences and other problems.

In the late 1980s, Abrahams realized that Moncrief’s gauge-invariant approach to perturbation theory could be applied directly to fully nonlinear numerical simulations [5–7]. The original idea of Abrahams was based on a flat space background, which was generalized later to include a Schwarzschild background [1]. But the basic idea is the same: one considers the numerically generated metric $g_{\alpha\beta, \text{num}}$ to be the sum of a background part and a perturbation:

$$g_{\alpha\beta, \text{num}} = g_{\alpha\beta} + h_{\alpha\beta}. \quad (4)$$

In the original treatment, the background was considered flat space and in the revised treatment [1] it was taken to be a time-dependent, generalized spherical background (e.g., not necessarily in Schwarzschild coordinates). The perturbation $h_{\alpha\beta}$ is expanded in tensor spherical harmonics as always.

To compute the elements of both $g_{\alpha\beta}$ and $h_{\alpha\beta}$, we simply integrate the numerically evolved metric components $g_{\alpha\beta, \text{num}}$, taken directly from the output of the numerical code, against appropriate spherical harmonics over a carefully chosen coordinate 2—sphere surrounding the source of the waves. The orthogonality of the $Y_{\ell m}$ allows one to ‘project’ the contributions of the general wave signal into individual modes. The background parts $g_{\alpha\beta}$ are computed by integrating $g_{\alpha\beta, \text{num}}$ against $Y_{00}$, while the perturbative components $h_{\alpha\beta}$ are computed by integrating $g_{\alpha\beta, \text{num}}$ against $Y_{\ell m}$ (technical complications arising from the tensor spherical harmonics are glossed over here).

The resulting functions can then be combined in a gauge-invariant way, following the prescription given by Moncrief. The beauty of this approach is that it is a simple prescription, that allows one to pick out various modes of waves, in different codes, written in different
gauges. It is incredibly effective, and has been shown to be able to accurately extract very small waves buried in the metric functions that are computed in a numerical simulation. In the next section I show an example.

A few further remarks are in order about this technique. First, it is not truly a perturbative approach in a systematic sense. It simply provides a procedure for removing the spherical background from a fully nonlinear numerical evolution, and then for decomposing what remains (‘everything nonspherical’) in $\ell, m$ modes, treating it as if it were a perturbation. In reality, what remains may be rather large in practice, and for example, it may not obey perturbative evolution equations. Secondly, the procedure assumes a spherical background, whereas in many systems of interest the background would be better treated as an axisymmetric, rotating system. Nonetheless, the extraction procedure has been used on what are effectively Kerr backgrounds with reasonable results [26, 10]. Alternate approaches involving the Weyl scalars are also sometimes used for rotating systems (see, e.g., [55, 16, 12]) and extensions of the gauge-invariant procedure for this case are under development by some groups.

This basic method has by now become the standard wave extraction technique in numerical relativity, having been used in dozens of works by groups in the US, Europe and Japan (see, e.g., a recent article by Shibata and Uryu [53]).

3.3. Combined extraction and evolution

The first two applications above show how the gauge-invariant perturbation theory is used in numerical relativity (a) to treat perturbative stellar systems by evolving relativistic perturbation equations on nonlinear, time-dependent backgrounds, and (b) to decompose a fully nonlinear numerical evolution into a spherical and nonspherical part, applying the gauge-invariant construction to the nonspherical pieces to extract a waveform. In this section, I show how groups have combined these two approaches to take nonlinear solutions to the Einstein equations (generated numerically or analytically), extract gauge-invariant quantities and use their evolution equations to advance the solution in time.

This has several very important uses:

- **Testbeds.** If the system is truly in a perturbative regime at the point where the gauge-invariant quantities are extracted, the evolution equations obtained by Moncrief’s procedure can be used to evolve the extracted quantities. These results can be compared with the results obtained from fully nonlinear evolutions of the system to test the accuracy of both the numerical codes and the extraction procedure itself.
- **Analysis.** Evolutions of the gauge-invariant equations can be used in comparison to the fully nonlinear, numerical evolutions to give insight into the physics obtained, which pure numerical evolution alone cannot provide.
- **Extension.** Building on these ideas, an interesting hybrid approach has been developed, whereby fully nonlinear evolutions are carried out to the point where the system reaches a perturbative regime. Numerical difficulties prevent the nonlinear evolution codes from continuing forever. But at some point, meaningful perturbation quantities can still be extracted and evolved according to their evolution equations, extending the solution forward in time.

I give examples of each of these approaches in turn.

In the first paper where the gauge-invariant extraction procedure was applied to black-hole spacetimes [1], it was shown that waveforms could be extracted from fully nonlinear distorted black-hole evolutions, and that quasinormal ringing modes were generated even in very strong distorted black-hole systems. But the idea of using the gauge-invariant evolution
Nonlinear impact of perturbation theory on numerical relativity

The technique was to extract gauge-invariant data from the numerical simulation not only to examine waveforms on a given timelike line, but also along a spacelike or null surface. This could then be used as initial data for the perturbation equations, which could be evolved independently of the nonlinear code. The perturbatively evolved waves could be compared with waves extracted at later times from the full nonlinear code, or used to extend the evolution to times beyond the point at which the nonlinear evolution was stopped. The results were in spectacular agreement, showing the validity and accuracy of both the nonlinear code and the wave extraction process.

This development paved the way for a host of other application of the basic hybrid perturbative/nonlinear evolution technique. After a number of developments in vacuum spacetimes (discussed further below), early examples of this technique applied to evolutions of collapsing matter can be found in [8], where a failing fully nonlinear numerical simulation was extended further in time by extracting gauge-invariant waves as Cauchy data for a perturbative evolution, so that waveforms could be computed at times much later than the nonlinear numerical code was able to run.

In another example, a similar technique was applied to full 3D distorted black-hole simulations [13]. Distorted black-hole initial data were evolved in a 3D numerical code written in Cartesian coordinates, and waves were computed using the gauge-invariant waveform extraction procedure, showing the development of quasinormal ringing at late times. But are the results correct? As a test of this, wave content was extracted from the numerically generated initial data, and evolved independently using a Zerilli equation code, and waveforms were compared. The agreement was spectacular, even for modes that carry as little as $10^{-7} M$ in energy! Results are shown in figure 1. Not only do such studies provide important testbeds for numerical relativity codes, but they also show the extent to which waveform extraction can be done in 3D simulations. In spite of the fact that very small waves are propagating through a ‘gauge-infested’ background, and in spite of numerical errors, grid stretching, and other difficulties, it turns out that numerical codes can very accurately evolve
the spacetimes, and that the procedure is capable of faithfully extracting even tiny waves in the system.

I now turn to a spectacular example of how a hybrid numerical-perturbative approach not only provided an unexpectedly strong validity check on numerical results, but also provided unusual insight into the physics of black-hole collisions. Smarr studied Misner data for two black holes in the 1970s, but the first time long-term evolutions were successfully carried out accurately enough that detailed waveforms could also be studied was 20 years later [15, 16]. Shortly afterwards, Price and Pullin [46] applied these gauge-invariant extraction and evolution ideas to the same initial data. It might seem to be impossible to treat colliding black holes perturbatively, but if the two holes are so close together initially that they have actually already merged into one, they might be considered as a single perturbed Schwarzschild hole (the so-called close limit). And maybe the technique could work if they were not already merged initially, but merely nearby.

Price and Pullin extracted the gauge-invariant wave content in the initial data, which in this case is given analytically, by expanding it around a Schwarzschild background. They were thus able to formally generate an analytic, one-parameter family of data, with the parameter being the distance between the holes. Clearly, if the holes were taken to be very far apart, the approximation breaks down, but formally the expression still gives initial data which can then be evolved with the Zerilli equation.

The results were spectacular! For cases where the holes were already surrounded by a common horizon, the Price–Pullin waveform results agreed beautifully with the nonlinearly computed numerical results, confirming the validity of both approaches. But incredibly, the agreement was excellent even for a range of cases where the holes did not have common apparent or event horizons! Naively one would not expect this! (For holes that are too widely separated, the results started to diverge, indicating the expected failure of the perturbation expansion.) This very unexpected result had an explanation, which was basically that even for holes that are separated by some distance, if the potentials of the Zerilli equation are still overlapping, they look from the perturbative point of view to be a single hole, at least outside the potential barriers.

These results were so startling, and so impressive, that an industry developed around these ideas, and many applications followed (see, e.g., [2, 4, 3, 18, 34, 17, 39, 27, 38, 19]).

The success of these techniques suggests, among other things, that these are very powerful methods that can be used hand-in-hand with fully nonlinear numerical evolutions, and can be applied in a variety of black-hole spacetimes where one might naively think they would not work. For these reasons, many researchers are continuing to apply these techniques to more and more complicated black-hole spacetimes.

Lastly, I mention an important extension of these ideas that has been developed in the last few years, starting in my group at AEI. This is the aptly named ‘Lazarus’ approach, which generalizes and combines approaches of [1, 8] on extending evolutions, and the industry of perturbative approaches to black-hole collisions. The idea is similar, but the approach is much more sophisticated. First one evolves orbital binary black-hole data, with a fully nonlinear numerical code. If one is able to evolve it close to the point of merger, in principle, one should be able to define a background, extract wave perturbations and evolve them with the relevant perturbation theory to extract waveforms. In this case, the Zerilli approach is inadequate, and a Teukolsky approach, governing perturbations about a Kerr black hole, must be used. The approach has produced the first waveforms from a series of orbital black-hole initial data sets, based on work of Cook [29] and Baumgarte [25], and is an important step forward in the quest of numerical relativists to determine waveforms from coalescing black holes [23, 24, 20, 21, 22].
4. Summary

I have tried to present a partial survey of a rather narrow segment of numerical relativity, to show the profound and unexpected impact that Moncrief’s work has had\(^2\). It may seem especially surprising that a highly mathematical approximation technique would become such a force, with so many independent threads, in a field that aims to find ‘exact’ solutions to Einstein’s equations in the absence of symmetries. It is tribute to the depth of Moncrief’s work, and the insight it can provide, that his work has found so many important and lasting uses in a field peripheral to his primary lines of research.

It is clear that the influence of perturbative methods will continue for years, even as fully nonlinear numerical relativity—and the computers it depends on—become still more powerful. After all, the goal is to divine physics from the numbers so produced, and to provide intuition about the underlying processes that govern the systems numerical relativity tries to study. In these ways, Moncrief has been a magician, and numerical relativists are trying to borrow a bit of his magic for themselves.

Acknowledgments

It is a pleasure, most importantly, to thank my thesis advisor, Vincent Moncrief, for the privilege of being his student, and for giving me a wonderful foundation to build my career upon. I also wish to thank Gabrielle Allen, Marco Bruni and Frank Herrmann for helpful suggestions in preparing this manuscript.

References


\(^2\) For space reasons, I have had to be very selective; not only have I omitted many references that have clear signs of Moncrief’s influence all over them, I have left out entire developments, e.g., boundary treatments based on gauge-invariant approaches [48].


[27] Campanelli M, Krivan W and Lousto C O 1998 The imposition of Cauchy data to the Teukolsky equation: II. numerical comparison with the Zerilli–Moncrief approach to black hole perturbations \textit{Phys. Rev. D} \textbf{58} 024016


[33] Decoherence factor \textit{Preprint gr-qc/0305058}


[38] Carlos Lousto O 2001 Perturbative evolution of nonlinear initial data for binary black holes Zerilli vs. Teukolsky equation \textit{Phys. Rev. D} \textbf{63} 047504


Nonlinear impact of perturbation theory on numerical relativity


