A convergent post-Newtonian approximation for the constraint equations in general relativity

Martin Lottermoser

(10th March 1991*)

Max-Planck-Institut für Astrophysik
Karl-Schwarzschild-Straße 1
W–8046 Garching bei München
Federal Republic of Germany

* This paper was later published in *Ann. Inst. Henri Poincaré*, 57(3), 279–317 (1992); the article is available through http://www.numdam.org/item?id=AIHPA_1992__57_3_279_0. The present file was created from version 1.4 of the TEX sources and differs from the original manuscript only in the fonts used, in the correction of seven errors in spelling or phrasing, a clarification in definition (4.9), and in the addition of a missing "−1" in the second equation after (4.9).

Abstract

Existing post-Newtonian approximation methods in general relativity lack a sound mathematical foundation. This paper takes the view of a space-like initial value problem and examines the constraint equations. It is proven that in suitable classes of function spaces and close to Newtonian values an iteration method can be devised which for every set of free data produces a sequence of approximations converging to a unique solution of the constraint equations. It is possible to obtain an upper limit on the difference between each step in the approximation and the unknown solution. In addition, one finds that for a matter tensor which is C^k in the metric and the matter variables the map from free initial data to solutions of the constraints is also C^k, where k is a positive integer, ∞, or ω. This provides a basis for approximation methods using Taylor expansions.

Résumé

Les méthodes, actuellement existantes, d'approximations post-newtoniennes à la relativité générale manquent d'une solide fondation mathématique. Le présent travail est une première étape dans cette direction. Il prend le point de vue du problème de Cauchy et examine les équations de contraintes. On prouve qu'il est possible de définir, dans des classes de fonctions appropriées, et au voisinage de données newtoniennes, une méthode d'itération qui associe à chaque ensemble de données libres une suite d'approximations convergeant vers une solution des contraintes. On peut aussi obtenir une borne sur l'écart entre la solution exacte et chaque étape de la suite d'approximations. De plus, si le tenseur d'énergie-impulsion est une fonctionelle de classe C^k de la métrique et des variables de matière, on trouve que l'application des données libres sur les solutions des contraintes est aussi de classe C^k, ou k est un entier naturel, ∞ ou ω. Ca produit une base pour des méthodes d'approximation utilisant des expansions de Taylor.

Table of contents

1. Introduction
2. Mathematical preliminaries
   2.1. Function spaces
   2.2. Functional analysis
3. Deriving the equations
   3.1. Variables
   3.2. Field equations
   3.3. The field equations at $\lambda = 0$
   3.4. The relativistic Cauchy problem
4. Properties of the parametrized constraint equations
   4.1. Properties of the metrics and of $E^{\alpha\beta}$
   4.2. An interlude on perfect fluids
   4.3. Solution of the constraints
   4.4. Application to approximations
5. Conclusions, speculations, and remaining problems

Acknowledgements
References

To Jürgen Ehlers and Bernd Schmidt

1. Introduction

As can be expected for every realistic theory, the theory of general relativity has the drawback that explicit solutions for physically interesting situations are almost impossible to find. Thus one is forced to employ approximation methods.

Generally speaking, an approximation scheme should consist of three parts:

(A1) A statement of the equation to be solved and a characterization of those solutions for which approximations are to be generated (*domain of definition*).

(A2) A rule by which one can generate the approximation to any of the solutions selected in (A1) (*algorithm*).

(A3) A statement describing in which sense the expression "approximation" is to be understood, i.e., what the relation is between the approximation and the solution (*accuracy limit*).

Note that the rule (A2) requires as input a specification of the solution one is looking for. As this "specification" will not usually take the form of an explicit statement of the solution, a prerequisite of an approximation scheme is the existence of a bijective map between a set of specifications ("data") and the set of solutions to be approximated.

Usually, one prefers schemes in which (A2) provides a method for generating an infinite sequence of approximations, and where an error limit in (A3) can be made arbitrarily small by choosing an approximation of sufficiently high "order" (index in the sequence). This means that the sequence of approximations converges to the solution (a rule for computing the error will usually define a topology). Of course it is not necessary for a method to produce a convergent sequence. In fact, ultimately divergent methods might give better results in the first few orders than convergent ones. Nevertheless, and although most of the ensuing remarks apply also to non-convergent methods, I shall for reasons of simplicity restrict myself to the convergent case. There, the order needed to attain a specified accuracy as a rule depends on how close the first approximation (0th order) is to the solution. One would therefore expect that for a good approximation method the set of 0th-order approximations has to be sufficiently dense in the set of functions where one is looking for solutions.

The observation of objects under the influence of gravity shows that for a large and important set of phenomena (such as the movement of planets in the solar system) the Newtonian theory of gravitation yields mathematical models which are in good or even excellent agreement with observations. If we assume that general relativity describes these phenomena correctly (or at least better than the Newtonian theory), we can deduce that for a certain kind of N-body system Newtonian solutions are very close to relativistic ones, where "close" is defined in terms of the experimental information accessible to us. This suggests that in these cases an approximation method starting with Newtonian solutions might exhibit a good rate of convergence. That is the motivation for the post-Newtonian approximation schemes in general relativity.

But what exactly *is* a post-Newtonian approximation? Apparently, the first procedure to be given this name was developed by S. Chandrasekhar [1] based on earlier work by others, in particular by Einstein, Infeld, and Hoffmann [2]. He started with a family of solutions of general relativity parametrized by $1/c$ (c being the velocity of light), assumed a particular form of the family of metrics in some coordinate system, and proceeded to do an expansion of everything in powers of $1/c$ around Minkowski space. The name "post-Newtonian" presumably arose because his first-order equations describe Newtonian solutions. I shall summarize the steps necessary for this procedure as follows:

(N1) Consider families of solutions of Einstein's equations, depending on a positive real parameter ε.

(N2) Select a set of variables describing a solution completely, and choose for each variable the leading power in ε. Now introduce new variables from which these powers are taken out.

(N3) Express the equations in terms of the new variables and rearrange them such that putting $\varepsilon = 0$ in the equations is possible and leads to Newtonian equations. For simplicity, I shall call the resulting set of equations the *ε-equations.* They have to be considered as functions of ε and the new variables.

(N4) Specify an approximation method for a solution of general relativity (situated at $\varepsilon = \varepsilon_1$, say) considered as a solution of the ε-equations. (Taking the rescaled variables to have the same meaning as the original ones we are led to $\varepsilon_1 = 1$.)

Several remarks are necessary here.

— The motivation behind step (N2) is that one expects Newtonian theory to be a good approximation to general relativity if certain expressions are "small", e.g., the gravitational potential, the matter velocities, and the ratio of temporal and spatial changes. One then selects the leading powers in ε such that they reflect the relative size of these expressions in those cases one is interested in. The mathematical requirement that the powers are chosen such that we end up with Newtonian equations for $\varepsilon = 0$ is not sufficient to fix these powers uniquely, even if we demand that at $\varepsilon = 0$ we shall obtain the full set of Newtonian equations with every Newtonian variable present. The reason is that general relativity allows its solutions more degrees of freedom than the Newtonian theory of gravitation (i.e., one has to specify more functions to identify a solution), and the only property required from the additional variables is that expressions describing the resulting new phenomena (gravitomagnetic effects and gravitational radiation) must have a positive power of ε in front.

As a rule one selects in this step an ε-dependent (flat) background metric and a vector field which is timelike and vorticity-free with respect to this background. They are used to distinguish between time and space coordinates, and hence it becomes possible to select different ε-powers for temporal and spatial components of tensors. Usually, one also adapts one's coordinates to this fictitious observer and metric. (Actually, people as a rule first choose their coordinates and then define the reference metric by specifying its components in that system in such a way that there is an orthogonal split between one timelike and three spacelike coordinates. The two methods are locally equivalent.)

— One important point in step (N3) is that transformations of the equations which are equivalent for $\varepsilon > 0$ might lead to different equations if one puts $\varepsilon = 0$. For example, consider a linear transformation for which the determinant vanishes at $\varepsilon = 0$ (take an equation and multiply it by ε). This doesn't change anything if ε is positive, but has rather drastic consequences for the limit. This is another point where different choices are possible, and therefore an explicit statement of the equations as functions of ε is essential.

The condition that $\varepsilon = 0$ should lead to Newtonian equations seems to exclude Chandrasekhar's method, because there Newtonian equations arise in the first-order perturbation around Minkowski space. But Minkowski space is a solution of general relativity, and therefore the $O(\varepsilon^0)$ equations are identically satisfied. Dividing by appropriate powers of ε—which is an equivalent transformation only for non-zero ε—the resulting ε-equations become Newtonian relations at the level ε^0.

It is important to realize that after step (N3) ε has changed its meaning: it is no longer parametrizing solutions of general relativity but has become part of the collection of objects specifying a solution of the ε-equations. A family of solutions will depend on a new parameter ε', and it will contain a function $\varepsilon(\varepsilon')$. Usually one considers only families with $\varepsilon = \varepsilon'$.

— The time-honoured method of approximation is to consider a family $S(\varepsilon)$ of solutions of the ε-equations, starting with a known Newtonian solution $S_0 = S(0)$ and ending with the desired solution $S_1 = S(\varepsilon_1)$ of general relativity. This family is assumed to depend at least differentiably on ε, in order to enable us to make a Taylor expansion with respect to ε around $\varepsilon = 0$. If the equations are also differentiable (considered as a function of ε and S), one can expand both at the same time and derive a sequence of equations to be solved successively for the expansion coefficients σ_j of $S(\varepsilon)$. The sequence $(T_i)_{i\in\mathbb{N}}$ of partial sums,

$$T_i := \sum_{j=0}^{i} \sigma_j \, \varepsilon_1^j,$$

then represents a sequence of approximations to S_1 with the difference being given by a suitable remainder term from Taylor's formula. If $S(\varepsilon)$ is analytic and ε_1 is sufficiently small, the approximation can be made arbitrarily good.

Of course, there are a number of points here which are unsatisfactory. For example, why should it be possible to go from a relativistic metric $g_{\mu\nu}$ to a Newtonian potential U, which is a totally different mathematical object? How much freedom is there to produce Newtonian equations, i.e., how many inequivalent post-Newtonian methods are there? And if $\varepsilon = 1/c$, what does it mean that c becomes infinite? To understand this and similar questions one needs a theory encompassing both, general relativity and Newtonian gravity, and using the same type of object to describe solutions in both theories. The axioms of such a "frame theory" have been stated by J. Ehlers [3,4] based on earlier work by a number of other authors. These axioms contain a real parameter λ which for positive values gives an upper bound $1/\sqrt{\lambda}$ on the relative velocity of time-like vectors, and hence is to be interpreted as a causality constant. Putting $\lambda = 0$ in the axioms produces a slightly generalized version of Newton's theory, whereas $\lambda = 1/c^2$ leads to the general theory of relativity. However, in this article I shall only use a restricted version of this theory which adheres to the scheme (N1–N3).

Taking (N1–N4) as the definition of the term "post-Newtonian approximation method", it is now obvious what an author of such a method has to do to satisfy the requirements (A1–A3) at the beginning.

(1) The ε-equations have to be determined, and sets of functions have to be selected for the solutions.

(2) For $\varepsilon = \varepsilon_1$ (when the ε-equations are Einstein's equations) one has to specify a set of "data" and prove that a 1-to-1 correspondence exists between these data and the solutions contained in the set chosen in step (1).

(3) Having selected an approximation method, one then has to prove that it is always applicable and really gives an approximation in some sense. If one chooses the method of expansion in ε, one has to prove that given the data for S_1 one can select a family $S(\varepsilon)$ of solutions of the ε-equations having a C^k-dependence on ε for a suitable k and satisfying $S(\varepsilon_1) = S_1$. Furthermore, one has to point out at least one way in which the sequence of equations one obtains can be solved for the coefficients σ_j.

I wish to emphasize that once one has made certain choices (which variables to take, where to look for solutions, what are the leading powers in ε, etc.) the remaining problem is purely mathematical and should be stated as such.

Of course, most full-blooded physicists would regard in particular point (3) above not as something to be proven, and would instead happily assume that such an expansion is always possible. This procedure has been very useful in physics, and people emphasizing the necessity of proofs are usually regarded as being hopelessly pedantic. It comes therefore as a pleasant surprise for the latter that in the business of post-Newtonian approximations this cavalier approach meets with disaster.

It seems to be a general experience with post-Newtonian methods that an expansion in powers of ε leads, for sufficiently high index j, to coefficients σ_j containing divergent integrals. In the case of the post-Newtonian method of Anderson and Decanio [5] this has for example been shown by Kerlick [6]. Why does this happen? By matching a post-Newtonian expansion in the near zone to a post-Minkowskian expansion in the far zone, Anderson *et al.* [7] and Blanchet and Damour [8,9] conclude that one has to expand with respect to a more general system of functions of ε, a system containing in particular $\ln \varepsilon$. However, if we look at post-Newtonian methods by themselves I do not believe that this is the correct answer.

Consider first a statement that a particular system of functions is *sufficient* for doing expansions. Such a general statement is almost certainly false, for every such system. The reason is that a solution of differential equations as a rule contains arbitrary constants, and the field equations alone can never fix the ε-dependence of these parameters. By choosing a function of ε which cannot be expanded in that system (and the "almost certainly" above refers to my belief that for every system such a function exists) one can construct counterexamples. The simplest way is to use a coordinate transformation involving such a function. This works even in the case of Minkowski space! Coordinate-independent counterexamples can be obtained by taking any known solution containing an arbitrary constant of integration with a coordinate-independent meaning, e.g., total mass.

These remarks show that the field equations alone are not a sufficient basis for the desired statement. To obtain a family $S(\varepsilon)$ of solutions with a particular dependence on ε we have to restrict in addition those parts of a solution not fixed by the equations. These parts are what I call "data". To prove that a particular system of functions of ε is sufficient for expansions one must first specify which ε-dependent families of data are permitted, second prove that for every ε the data determine a unique solution $S(\varepsilon)$, and finally show that the resulting family S

of solutions can be expanded with respect to the chosen system of functions of ε. Without restricting the families of data in this manner a proof is not possible.

But for which system of functions of ε should one try to obtain a proof? Which system is *necessary?* This depends on the set of variables we choose to describe a solution, and on the usefulness of the set of solutions which can be reached by families which can be expanded in this system. For example, the ε-equations might permit families which do not depend on ε but which are solutions for all values of ε (the equations and variables to be used later are of this kind, because they permit the solution $\mathfrak{U}^{\alpha\beta} = 0$, $T^{\mu\nu} = 0$ independent of ε). In this case one might be tempted to say that the system $\{\varepsilon^0\}$ is the only necessary one. But the set of solutions of general relativity which can be reached in this manner (at $\varepsilon = \varepsilon_1$) will in general be insufficient to describe all situations one is interested in. The conclusions to be drawn from these observations are, first, that *every* system is permitted provided the resulting set of families of solutions is non-empty, and second, that the system chosen should be sufficiently flexible to enable us to bend families of solutions toward all solutions of general relativity we want to reach. Given any system, for example the set $\{ \varepsilon^n \mid n \in \mathbb{N} \}$ of non-negative integer powers of ε, the set of families of solutions which can be expanded in that system is determined, and the same is therefore true for the set of families of data generating these solutions. Having selected a basis for expansions in ε for the solutions, our first task is to characterize the resulting families of data. The list of properties derived will be complete if we can close the circle and prove that families of data satisfying these requirements lead to families of solutions which can be expanded in the system chosen at the outset.

I see only one possibility which could single out particular classes of systems. Depending on the kind of data chosen, the map from solutions to data will frequently preserve the property of having an expansion in that system (this is true for the Cauchy problem). This indicates the possibility that there exists a system of functions of ε which is "closed" under the ε-equations, in the sense that any family of data which can be expanded in that system leads to a family of solutions with the same property. Such a system would legitimately be called a "natural" one, and it would be of interest to find the smallest system of this kind containing the powers of ε.

Most existing post-Newtonian approximation methods ignore the necessity of explicitly fixing the solutions by specifying data, and I do not know of a single one which proves that a family of solutions exists and can be expanded, let alone shows what the relation is between approximation and solution. There is a paper by Futamase and Schutz [10] the title of which ("The post-Newtonian expansion is asymptotic to general relativity") and the abstract seem to indicate that they prove exactly this point (the expression "asymptotic" is used in the sense that there is a series with a remainder term of order ε^n). However, they state that they assume in particular (a) that the initial value problems they consider have solutions, and (b) that the ε-derivatives of their family $S(\varepsilon)$ for $\varepsilon \to 0$ "exist unless this assumption leads to contradictions". As this together with Taylor's theorem already completes the proof, their result does not contribute to a rigorous basis for post-Newtonian approximation methods. There is also a proof by Blanchet and Damour [8] that in the case of no incoming radiation the set $\{ \varepsilon^n \ln^m \varepsilon \mid n, m \in \mathbb{N} \}$ is sufficient for doing post-Newtonian expansions, but this is based on the assumptions (a) that the family $S(\varepsilon)$ exists, and (b) that it possesses a multipolar post-Minkowskian expansion.

Of course these remarks do not imply that previous versions of post-Newtonian approximations are useless. But they build on something which is still missing, and as long as we do not have a firm mathematical foundation for them they have to be treated as *unreliable,* because we do not know under which circumstances they give correct answers. That there is a limit to their validity can clearly be seen in the appearance of divergent integrals. Such a situation is only possible if the assumptions enabling us to derive these integrals contain a statement which for the set of solutions considered is false in general. As long as this statement is not identified, even the mathematical validity of the lower-order approximations is in doubt.

This doubt is however partially allayed by the fact that there is good agreement between the lower-order predictions and observations. This shows that although the mathematical derivations of these methods are questionable, the results are physically acceptable. But because we have not yet been able to establish a rigorous link between general relativity and these methods, we have to regard them as logically independent theories, and we do not know what the agreement between their predictions and the observations tells us about the theory of general relativity.

It is therefore necessary to establish a rigorous mathematical basis for post-Newtonian approximation methods. This paper is an attempt at a first step in this direction. It uses a Cauchy problem (space-like initial value problem) to fix the solution, and one result will be that under suitable circumstances the map from free initial data to solutions of the constraint equations is analytic. Therefore we know that by choosing a family of data which is analytic in ε we obtain a family of solutions of the constraints which can be expanded in a convergent power series in ε. In looking at the mathematical basis of the proof we will in addition realize that there is another method for generating a sequence of approximations, which is not based on an expansion of a family of solutions. This method will give a sequence converging to the solution of the constraints, and we shall see that it is in principle possible (although presumably tedious and possibly not very useful) to derive an explicit upper bound on the difference between approximations and the unknown solution.

2. Mathematical preliminaries

This section contains various general mathematical statements to be used in the remainder of the paper. They are given without proofs, as these can either be found in the literature or can be deduced by elementary methods. Physicists among the readers are warned that some familiarity with calculus on Banach spaces is assumed, but any introductory text should be sufficient for that [11,12].

2.1. Function spaces

It seems uneconomical to restrict oneself at the outset to a particular function space in which to search for solutions. Therefore I shall merely give a list of properties which will be used in the subsequent proofs, and leave it to the reader to check whether her or his favourite space fulfills these requirements. To ensure that the conditions are not contradictory, I shall however point out a space in which they are satisfied.

2.1.1. Abstract definition

Let Σ be some subset of $\mathbb{R}^3$ ($\mathbb{R}$ denotes the real numbers). I assume that there exist three real Banach spaces F_j, $j = 0$, 1, 2, of functions $f\colon \Sigma \to \mathbb{R}$ or of equivalence classes of such functions. They are assumed to have the following properties:

(2.1) *Addition and scalar multiplication are induced by pointwise operations.*

(2.2) *Multiplication, defined pointwise, results in the following bilinear bounded maps:*

$$\cdot\colon F_2 \times F_2 \to F_2, \quad \cdot\colon F_1 \times F_1 \to F_0, \quad \cdot\colon F_2 \times F_0 \to F_0.$$

(2.3) *There are three "differential operators"* $\partial/\partial x^i$,

$$\frac{\partial}{\partial x^i}\colon F_2 \to F_1, \qquad \frac{\partial}{\partial x^i}\colon F_1 \to F_0, \qquad i \in \{1,2,3\},$$

they are linear and bounded, and they commute on F_2.

(2.4) *The Laplacian*

$$\triangle\colon F_2 \to F_0, \qquad \triangle f := \sum_{i=1}^{3} \frac{\partial^2 f}{\partial\,(x^i)^2},$$

is bijective.

(2.5) F_2 *does not contain a constant function except* 0.

(2.6) F_2 *is a subset of the Banach algebra* $B := B(\Sigma,\mathbb{R})$ *of bounded functions on* Σ, *and the inclusion* $i\colon F_2 \to B$ *is continuous.*

(2.3) shows that the index "j" in "F_j" refers to a level of differentiability. Because of (2.3), (2.4), and the open mapping theorem (maps which are linear, continuous, and surjective are also open), the Laplacian is a homeomorphism. Condition (2.5) could also be derived from (2.1), (2.2), (2.4), and the following additional assumptions:

$$\cdot\colon F_2 \times F_1 \to F_1, \qquad \frac{\partial}{\partial x^i}(ab) = \frac{\partial a}{\partial x^i} b + a\frac{\partial b}{\partial x^i} \quad \text{for } a,\, b \in F_2. \tag{2.7}$$

Note as a curiosity that except in condition (2.7) nothing is assumed to justify the expression "differential operator" for $\partial/\partial x^i$.

One further condition remains to be stated:

(2.8) *There is a fourth Banach space* F_{-1}, *the operators* $\partial/\partial x^i$ *are also defined on* F_0 *with*

$$\frac{\partial}{\partial x^i}\colon F_0 \to F_{-1},$$

they commute on F_1, *and the Laplacian* $\triangle\colon F_1 \to F_{-1}$ *is injective.*

I have kept this requirement separate from the previous ones because it is not needed for developing the approximation method. It will merely be used to prove that solutions of the Newtonian theory in its usual form are also solutions of certain equations employed here.

From F_2 I construct an additional set:

$$K_2 := \{\, g\colon \Sigma \to \mathbb{R} \mid \exists a \in \mathbb{R},\ f \in F_2 :\ g = a + f \,\}. \tag{2.9}$$

Because of the properties (2.1) and (2.5), a and f are uniquely determined by g. Now define

$$\|g\|_{K_2} := |a| + m\|f\|_{F_2}, \tag{2.10}$$

where $m \in \mathbb{R}$ is a positive bound on the multiplication in F_2: $\|f_1 \cdot f_2\|_{F_2} \le m\|f_1\|_{F_2}\|f_2\|_{F_2}$. Thus K_2 is seen to be a Banach space, the direct sum $\mathbb{R} \oplus F_2$. Defining again multiplication by pointwise operations, we find by (2.1), (2.2), and the scalar multiplications in F_2 and F_0:

$$\cdot\colon K_2 \times K_2 \to K_2, \quad \cdot\colon K_2 \times F_0 \to F_0, \qquad \text{bilinear and continuous.} \tag{2.11}$$

The first property shows that K_2 is a Banach algebra (m has been included in the definition (2.10) in order to obtain $\|g \cdot f\|_{K_2} \le \|g\|_{K_2}\|f\|_{K_2}$). K_2 is commutative, and because the function equal to 1 everywhere belongs to it, it has a unit element. The operators $\partial/\partial x^i$ can be linearly and continuously extended to K_2 by defining

$$\frac{\partial}{\partial x^i}(a + f) := \frac{\partial}{\partial x^i} f, \tag{2.12}$$

as is obviously sensible. And finally, because of (2.6) the elements of K_2 are continuously embedded in the bounded functions. Because of (2.1) and (2.2), the resulting inclusion $i\colon K_2 \to B$ is a homomorphism for Banach algebras.

2.1.2. Example

For $\Sigma = \mathbb{R}^3$, a possible realization of the F_j is given by certain weighted Sobolev spaces [14]. They can be defined as follows [15]: given a non-negative integer s and a real number δ, consider the set of functions locally in L^2 with the property that all distributional derivatives up to the order s are again locally in L^2. The weighted Sobolev space $H_{s,\delta} := H_{s,\delta}(\mathbb{R}^3, \mathbb{R})$ is then the subset for which the following norm exists:

$$\|f\|_{s,\delta} := \sqrt{\sum_{\substack{\alpha,\\ |\alpha|\leq s}} \left(\left\|\sigma^{|\alpha|+\delta} D^\alpha f\right\|_{L^2}\right)^2}. \tag{2.13}$$

Here $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ is a multi-index, $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$, $\sigma := \sqrt{1 + |\vec{x}|^2}$ is a weight factor enforcing suitable fall-off at infinity,

$$D^\alpha f := \frac{\partial^{|\alpha|} f}{\partial^{\alpha_1} x\, \partial^{\alpha_2} y\, \partial^{\alpha_3} z}, \tag{2.14}$$

and

$$\|g\|_{L^2} := \sqrt{\textstyle\int |g|^2\, d\mu} \tag{2.15}$$

is the usual norm of $L^2(\mathbb{R}^3, \mathbb{R})$. These sets turn out to be real Banach spaces under pointwise addition and scalar multiplication. If one chooses

$$s \geq 2, \quad -\tfrac{3}{2} < \delta < -\tfrac{1}{2}, \qquad F_2 := H_{s,\delta}, \quad F_1 := H_{s-1,\delta+1}, \quad F_0 := H_{s-2,\delta+2}, \tag{2.16}$$

and $\partial/\partial x^i$ as the distributional derivative (or, equivalently, as the unique continuous extension of the usual differential operators on C^∞ functions), conditions (2.1) to (2.6) are satisfied. (2.1) is already valid in L^2, (2.2) can be found in [15], (2.3) is trivial, (2.4) is contained in [16,15], (2.5) holds because (2.7) is satisfied (the multiplication property is proven in [15]), and (2.6) is shown in [15]. In addition, the requirement (2.8) is also fulfilled, provided

$$s \geq 3, \qquad F_{-1} := H_{s-3,\delta+3}: \tag{2.17}$$

the injectivity of $\triangle$ on F_1 holds under these circumstances because for $-\frac{3}{2} < \delta < -\frac{1}{2}$ the Laplacian is injective on $H_{2,\delta+1}$ [16], and because for $s \geq 3$ the space $F_1 = H_{s-1,\delta+1}$ is a subset of $H_{2,\delta+1}$.

The degree s of differentiability of a function is usually not important for physical applications, but the behaviour at infinity is. It is therefore reassuring to find that for example the function

$$\left(1 + |\vec{x}|^2\right)^{-\alpha/2}, \qquad \alpha > \delta + \tfrac{3}{2},$$

which has a fall-off like $r^{-\alpha}$, is still in $H_{s,\delta}$. We shall see in a later section that F_0 is the space for components of the matter tensor, therefore one can by using $F_0 = H_{s-2,\delta+2}$ treat matter distributions with a fall-off like $r^{-\beta}$, $\beta > 2$, and that includes even cases with infinite total mass. One can therefore conclude that the spaces $H_{s,\delta}$ are sufficiently large to contain interesting solutions. On the other hand, knowing that a solution belongs to $H_{s,\delta}$ does not tell us much about it. If one needs solutions with more restricted properties (e.g., finite ADM mass) one can try to characterize these as subsets, or even look for other function spaces containing only elements satisfying the additional requirements. This is another reason for keeping proofs independent of a particular function space.

2.2. Functional analysis

2.2.1. Differentiability and analyticity

For any pair of Banach spaces E, F the sets of multilinear continuous maps from E to F are denoted by $\mathscr{L}_i(E,F)$, with i giving the number of arguments ($\mathscr{L}_0(E,F) := F$, $\mathscr{L}(E,F) := \mathscr{L}_1(E,F)$). Defining algebraic operations pointwise and taking the usual supremum norms these sets are Banach spaces.

For any open subset A of E the notation $C^k(A,F)$ refers to the set of functions $f\colon A \to F$ of differentiability class C^k. Here k is an element of $\overline{\mathbb{N}} := \mathbb{N}^* \cup \{\infty, \omega\}$, with $\mathbb{N}^*$ denoting the positive integers ($\mathbb{N}$ refers to the non-negative ones), C^∞ infinitely differentiable functions, and C^ω the analytic ones. I use the following definition of analyticity ([17], definition 15.1):

> **Definition.** *A function $f\colon A \to F$ is said to be of class $C^\omega(A,F)$, if for each $x_0 \in A$ there exists a positive $r \in \mathbb{R}$ and a sequence $(f_i)_{i\in\mathbb{N}}$ of continuous i-linear symmetric maps from E to F, such that*
>
> $$\sum_{i=0}^{\infty} \|f_i\|_{\mathscr{L}_i(E,F)}\, r^i \quad \textit{exists, and}$$
>
> $$f(x) = \sum_{i=0}^{\infty} f_i(\underbrace{x - x_0, \ldots, x - x_0}_{i \text{ times}}) \quad \textit{for all } x \textit{ such that } \|x - x_0\| < r.$$

If $E = \mathbb{R}$, we have $f_i(x-x_0,\ldots,x-x_0) = f_i(1,\ldots,1)(x-x_0)^i$, and the definition reduces to the statement that a function is called analytic if it can be represented by an absolutely convergent power series.

The following statements for C^k-functions will be used frequently:

— Concatenation of two C^k-functions results in a C^k-function.

— A linear or bilinear continuous function is C^ω.

— $C^\omega(A,F) \subset C^\infty(A,F)$.

If f is a differentiable function, $Df(x_0)$ is used to denote its derivative at x_0. One should never forget that derivatives at a point are maps: an expression like $Df(x_0)^{-1}$ refers to the inverse map, not to something one might write as $1/Df(x_0)$. If f has several arguments, $D_j f(x_1,\ldots,x_n)$ denotes the partial derivative with respect to the jth argument.

2.2.2. The implicit function theorem

An important tool will be the following theorem [12,17]:

> **Implicit Function Theorem.** *Let E, F, G be Banach spaces, A an open subset of $E \times F$, $(x_0,y_0) \in A$, $f \in C^k(A,G)$ for some $k \in \overline{\mathbb{N}}$, $f(x_0,y_0) = 0$, and assume that $D_2 f(x_0,y_0)$ is bijective.*
>
> *Then there exist open neighbourhoods V of (x_0,y_0) in A and L of x_0 in E, and a function $l \in C^k(L,F)$, such that for all $(x,y) \in E \times F$ the following is true:*
>
> $$x \in L, \quad y = l(x) \quad \Longleftrightarrow \quad (x,y) \in V, \quad f(x,y) = 0. \tag{2.18}$$

The theorem tells us that on L there is a solution l of $f(x,l(x))=0$, and that on V there are no other solutions (x,y) of $f(x,y)=0$.

The proof of this theorem involves the application of Banach's fixed point theorem, i.e., one constructs certain sequences converging to values of the solution. This is of course an approximation method, and closer inspection shows that one can even get an upper bound on the difference between approximation and solution. One possible summary of this result is the following

Prescription: *Consider the function*

$$g \in C^k(A,F), \qquad g(x,y) := y - D_2 f(x_0,y_0)^{-1}\big[f(x,y) - D_1 f(x_0,y_0)(x)\big]. \tag{2.19}$$

Choose a positive real number $a<1$ and find a positive $r \in \mathbb{R}$ such that the closed ball with radius r around (x_0,y_0) is contained in

$$V(a) := \{\,(x,y) \in A \mid \|Dg(x,y)\|_{\mathscr{L}(E\times F,F)} \le a\,\}. \tag{2.20}$$

Such an r exists, because the derivative Dg is continuous and we have $Dg(x_0,y_0)=0$. Now let

$$s := \frac{1-a}{\|k\|_{\mathscr{L}(E,E\times F)}} r, \qquad k(x) := \big(x,\, -[D_2 f(x_0,y_0)^{-1} \circ D_1 f(x_0,y_0)](x)\big)\,. \tag{2.21}$$

Then every $x \in E$ satisfying $\|x-x_0\|_E \le s$ belongs to L, and for each such x the sequence $(l_j(x))_{j\in\mathbb{N}}$,

$$\begin{aligned} l_0(x) &:= y_0 - \big[D_2 f(x_0,y_0)^{-1} \circ D_1 f(x_0,y_0)\big]\,(x-x_0), \\ l_{j+1}(x) &:= l_j(x) - \big[D_2 f(x_0,y_0)^{-1} \circ f\big]\,(x, l_j(x)), \qquad j \in \mathbb{N}, \end{aligned} \tag{2.22}$$

is well-defined and converges to the value of the solution at x:

$$\|l_j(x) - l(x)\|_F \le \frac{a^j}{1-a} Q(x), \qquad j \in \mathbb{N}. \tag{2.23}$$

Here $Q(x)$ describes the quality of $l_0(x)$ as an approximation:

$$Q(x) := \|D_2 f(x_0,y_0)^{-1}\|_{\mathscr{L}(G,F)}\,\|f(x,l_0(x))\|_G. \tag{2.24}$$

Other versions are possible; this particular one is merely intended to give an indication of the kind of statement we can obtain. It can be derived by applying an appropriate version of the inverse function theorem to $\tilde{f}(x,y) := (x, f(x,y))$.

Sometimes it is possible to extend the set L in the implicit function theorem on which the solution is defined by matching various local solutions. The following lemma gives conditions under which two local solutions agree on a common domain of definition.

Lemma 2.1. *E, F, G are Banach spaces; $V \subset E \times F$ is open; $f\colon V \to G$. f has the property that for every (x_0,y_0) in V there is a neighbourhood V_0 in V such that for all (x,y_1), (x,y_2) in V_0 we have:*

$$f(x,y_1) = 0 = f(x,y_2) \quad \Longrightarrow \quad y_1 = y_2. \tag{2.25}$$

Let W be a connected subset of E and let $l_i \in C^0(W, F)$, $i = 1, 2$, be solutions of:

$$(x, l(x)) \in V, \quad f(x, l(x)) = 0 \qquad \textit{for all } x \in W. \tag{2.26}$$

Finally, assume that there is an $\hat{x} \in W$ such that $l_1(\hat{x}) = l_2(\hat{x})$.

It then follows that $l_1 = l_2$.

The proof consists in showing that the subset Q of W on which both solutions agree is open and closed. Because W is connected, Q must then either be empty or identical with W, and the first possibility is excluded by the existence of $\hat{x}$.

The condition imposed on f can be paraphrased by saying that f is "locally injective for solutions". Note that this is a consequence of the property

$$(x, y) \in V, \quad f(x, y) = 0 \quad \Longrightarrow \quad x \in L, \quad y = l(x) \tag{2.27}$$

ensured on the set V appearing in the implicit function theorem.

2.2.3. Banach algebras

Finally, I shall need a few facts about a Banach algebra K with unit element e.

— The set of invertible elements

$$\mathrm{Inv}(K) := \{\, a \in K \mid \exists b \in K\colon a \cdot b = e = b \cdot a \,\} \tag{2.28}$$

is open in K, and contains the open ball $W := \{\, a \in K \mid \|a - e\| < 1 \,\}$. For an $a \in \mathrm{Inv}(K)$, the inverse b is uniquely determined and also an element of $\mathrm{Inv}(K)$, and the map $a \mapsto b$ is in $C^\omega(\mathrm{Inv}(K), K)$. (Start from [11], section 4.8. This is basically due to the convergence of the geometric series.)

— If K is commutative, then for every $a \in W$ there is a unique element $\sqrt{a}$ of W satisfying $(\sqrt{a})^2 = a$. The map $a \mapsto \sqrt{a}$ is in $C^\omega(W, K)$. (The binomial series of index $\frac{1}{2}$ is responsible for existence and analyticity.)

3. Deriving the equations

3.1. Variables

Given a metric tensor $g_{\mu\nu}$ (Greek indices take values in $\{0, 1, 2, 3\}$, Latin indices in $\{1, 2, 3\}$) one can define the tensor density

$$\bar{\mathfrak{g}}^{\alpha\beta} := \sqrt{g}\, g^{\alpha\beta}, \qquad g := |\det(g_{\mu\nu})| \,. \tag{3.1}$$

In the case of Minkowski space and signature $+2$ (or $-{+}{+}{+}$) there exist coordinates x^μ (I also write t for x^0) such that $(g_{\mu\nu}) = (\underset{0}{g}_{\mu\nu}) := \mathrm{diag}(-c^2, 1, 1, 1)$ and hence

$$(\bar{\mathfrak{g}}^{\alpha\beta}) = (\underset{0}{\bar{\mathfrak{g}}}^{\alpha\beta}) := \begin{pmatrix} -\varepsilon & 0 & 0 & 0 \\ 0 & 1/\varepsilon & 0 & 0 \\ 0 & 0 & 1/\varepsilon & 0 \\ 0 & 0 & 0 & 1/\varepsilon \end{pmatrix}, \qquad \varepsilon := \frac{1}{c}, \tag{3.2}$$

where c is the velocity of light. Using a particular coordinate system to define such a reference metric, we can form the tensor density [18]

$$\mathfrak{U}^{\alpha\beta} := \frac{1}{4\varepsilon^3}\left(\bar{\mathfrak{g}}^{\alpha\beta} - \underset{0}{\bar{\mathfrak{g}}}^{\alpha\beta}\right). \tag{3.3}$$

It can be used as a variable instead of $g_{\mu\nu}$, as the latter can be recovered from it:

$$\bar{\mathfrak{g}}^{\sigma\tau} = \underset{0}{\bar{\mathfrak{g}}}^{\sigma\tau} + 4\varepsilon^3\mathfrak{U}^{\sigma\tau}, \qquad g = |\det(\bar{\mathfrak{g}}^{\mu\nu})|, \qquad g^{\alpha\beta} = \frac{1}{\sqrt{g}}\bar{\mathfrak{g}}^{\alpha\beta}. \tag{3.4}$$

I shall frequently use separate names for particular components:

$$U := \mathfrak{U}^{00}, \qquad W^a := \mathfrak{U}^{0a}, \qquad Z^{ab} := \mathfrak{U}^{ab}. \tag{3.5}$$

The relativistic space-times $(M, g_{\mu\nu})$ considered in this paper can now be characterized by the following properties:

The manifold M is a subset of $\mathbb{R}^4$, and there exist on M global coordinates (t, x^a) satisfying:

(3.6) $\Box\, x^\mu = 0$ *or equivalently* $\bar{\mathfrak{g}}^{\mu\nu}{}_{,\nu} = 0$ *(harmonic coordinates).*

(3.7) *The hypersurfaces Σ_t of constant t are space-like and diffeomorphic to Σ.*

(3.8) $\mathfrak{U}^{\alpha\beta}$ *restricted to Σ_t belongs to $S_4(F_2)$, where $S_n(F)$ denotes the $n \times n$ symmetric matrices with components in the Banach space F.*

(3.9) *The time derivative $\partial/\partial t$ (or $\dot{}$) satisfies on Σ_t:*

(a) $\dot{\mathfrak{U}}^{\alpha\beta}$ *belongs to $S_4(F_1)$,* $\ddot{\mathfrak{U}}^{\alpha\beta}$ *to $S_4(F_0)$.*

(b) $\partial/\partial t$ commutes with the $\partial/\partial x^a$.

The immediate introduction of the x^μ simplifies the necessary statements but obscures the underlying geometric structures. I have introduced here on an orientable manifold a background Lorentz metric $\underset{0}{g}_{\mu\nu}$ and a vector field $B = \partial/\partial t$ which is time-like and vorticity-free with respect to this background. (dt can be derived from $\partial/\partial t$. The converse is not true in Newtonian theory, hence $\partial/\partial t$ is the fundamental object.) Together these also define a Euclidean metric

$$e_{\mu\nu} := \underset{0}{g}_{\mu\nu} + \varepsilon^2(1+\varepsilon^2)\underset{0}{g}_{\mu\sigma}\underset{0}{g}_{\nu\tau}B^\sigma B^\tau = \delta_{\mu\nu} \tag{3.10}$$

which will play a rôle later on. (Strictly speaking, "$1+\varepsilon^2$" should be "$1+v^2\varepsilon^2$" with a positive constant v having the dimension of a velocity. As different v result in equivalent metrics, however, we are free to choose $v = 1$ in whichever system of units we use.) The function spaces F_j are based on these background objects. Given the metric $g_{\mu\nu}$, the background has to be chosen such that the resulting $\mathfrak{U}^{\alpha\beta}$ has the properties desired.

From now on, all fields on space-time are considered as maps from t into function spaces on Σ, and the field equations are equations for these maps. In particular, the constraints are equations for values at $t = 0$. By "hiding" thus the argument (x^a) which varies over Σ, we are able to obtain statements which are valid on all of Σ.

3.2. Field equations

The equations to be satisfied by the unknown relativistic solution S_1 are of course Einstein's equations

$$G^{\alpha\beta} = \frac{8\pi G}{c^4}\,T^{\alpha\beta}, \tag{3.11}$$

with $G^{\alpha\beta} = R^{\alpha\beta} - \frac{1}{2}R\,g^{\alpha\beta}$ being the Einstein tensor, $R_{\mu\nu} = R^{\sigma}{}_{\mu\sigma\nu}$ the Ricci tensor, $R = g^{\sigma\tau}R_{\sigma\tau}$ the Ricci scalar, $R^{\alpha}{}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\delta\beta,\gamma} - \Gamma^{\alpha}_{\gamma\beta,\delta} + \Gamma^{\alpha}_{\gamma\sigma}\Gamma^{\sigma}_{\delta\beta} - \Gamma^{\alpha}_{\delta\sigma}\Gamma^{\sigma}_{\gamma\beta}$ the Riemann tensor, and G the gravitational constant. The matter tensor $T^{\alpha\beta}$ should be thought of as a function ("matter model") of the metric and of certain unspecified matter variables. It does not explicitly depend on x^{μ}.

Rewriting the equations in terms of the $\mathfrak{U}^{\alpha\beta}$ we obtain

$$E^{\alpha\beta} = 4\pi G\,|d|\,T^{\alpha\beta}, \tag{3.12}$$

where (compare with [13] §20.3):

$$\begin{aligned}
E^{\alpha\beta} &:= H^{\alpha\beta} + A^{\alpha\beta} + \lambda B^{\alpha\beta} + \lambda^2 C^{\alpha\beta},\\
H^{\alpha\beta} &:= \mathfrak{g}^{\mu\nu}\mathfrak{U}^{\alpha\beta}{}_{,\mu\nu} - 2\mathfrak{g}^{\mu(\alpha}\mathfrak{U}^{\beta)\nu}{}_{,\mu\nu} + \mathfrak{g}^{\alpha\beta}\mathfrak{U}^{\mu\nu}{}_{,\mu\nu},\\
A^{\alpha\beta} &:= 2\left(\tfrac{1}{2}\mathfrak{g}_{\mu\nu}\mathfrak{g}_{\sigma\tau} - \mathfrak{g}_{\tau\mu}\mathfrak{g}_{\sigma\nu}\right)\left(\mathfrak{g}^{\alpha\eta}\mathfrak{g}^{\beta\kappa} - \tfrac{1}{2}\mathfrak{g}^{\alpha\beta}\mathfrak{g}^{\eta\kappa}\right)\mathfrak{U}^{\mu\nu}{}_{,\eta}\mathfrak{U}^{\sigma\tau}{}_{,\kappa},\\
B^{\alpha\beta} &:= 4\mathfrak{g}_{\mu\nu}\left(2\mathfrak{g}^{\eta(\alpha}\mathfrak{U}^{\beta)\nu}{}_{,\kappa}\mathfrak{U}^{\mu\kappa}{}_{,\eta} - \tfrac{1}{2}\mathfrak{g}^{\alpha\beta}\mathfrak{U}^{\mu\sigma}{}_{,\tau}\mathfrak{U}^{\nu\tau}{}_{,\sigma} - \mathfrak{g}^{\sigma\tau}\mathfrak{U}^{\alpha\mu}{}_{,\sigma}\mathfrak{U}^{\beta\nu}{}_{,\tau}\right),\\
C^{\alpha\beta} &:= 4\left(\mathfrak{U}^{\alpha\beta}{}_{,\mu}\mathfrak{U}^{\mu\nu}{}_{,\nu} - \mathfrak{U}^{\alpha\mu}{}_{,\nu}\mathfrak{U}^{\beta\nu}{}_{,\mu}\right),
\end{aligned} \tag{3.13}$$

and I have defined:

$$\lambda := \varepsilon^2, \qquad \mathfrak{g}^{\alpha\beta} := \varepsilon\bar{\mathfrak{g}}^{\alpha\beta}, \qquad (\mathfrak{g}_{\mu\nu}) := \varepsilon(\bar{\mathfrak{g}}^{\alpha\beta})^{-1}, \qquad d := \lambda\det(\bar{\mathfrak{g}}^{\alpha\beta}). \tag{3.14}$$

By direct computation we find:

$$\begin{aligned}
\mathfrak{g}^{\alpha\beta} &= \underset{0}{\mathfrak{g}}^{\alpha\beta} + 4\lambda^2\mathfrak{U}^{\alpha\beta}, \qquad (\underset{0}{\mathfrak{g}}^{\alpha\beta}) = \operatorname{diag}(-\lambda,1,1,1), \qquad (\mathfrak{g}_{\mu\nu}) = \frac{\operatorname{adj}(\mathfrak{g}^{\alpha\beta})}{d},\\
d &= -1 + \lambda\,4U - \lambda^2\,4\operatorname{tr}Z + \lambda^3\,16(U\operatorname{tr}Z - W^2) - \lambda^4\,8(\operatorname{tr}^2 Z - \operatorname{tr}Z^2)\\
&\quad + \lambda^5\,64\big(U\tfrac{1}{2}(\operatorname{tr}^2 Z - \operatorname{tr}Z^2) - W^2\operatorname{tr}Z + (\vec{W}, Z\vec{W})\big)\\
&\quad - \lambda^6\,64\det Z + \lambda^7\,256\big(U\det Z - (\vec{W}, \operatorname{adj}Z\,\vec{W})\big).
\end{aligned} \tag{3.15}$$

Here I have employed the usual notation for vectors and matrices on $\mathbb{R}^n$. In particular, tr is the trace operator, $(\ldots,\ldots)$ the scalar product, and adj denotes the adjoint matrix (the transpose of the matrix of cofactors).

Now assume that ε is allowed to vary from its previously fixed value of $\varepsilon = \varepsilon_1 := 1/c$. Then (3.13) defines $E^{\alpha\beta}$ in terms of λ, $\mathfrak{U}^{\alpha\beta}$, $\mathfrak{g}_{\sigma\tau}$, and $\mathfrak{g}^{\mu\nu}$. The last two can be expressed in terms of $\mathfrak{U}^{\alpha\beta}$ by means of (3.4) and (3.14), but that relation is not defined for $\varepsilon = 0$. In contrast, (3.15) characterizes $\mathfrak{g}^{\alpha\beta}$ and $\mathfrak{g}_{\mu\nu}$ for arbitrary values of ε, provided the additional variable d is non-zero. Therefore, (3.15) will now be considered as the definition of $\mathfrak{g}^{\alpha\beta}$, $\mathfrak{g}_{\mu\nu}$, and d. $E^{\alpha\beta}$ is then a function of λ and $\mathfrak{U}^{\mu\nu}$, defined for every argument resulting in a non-vanishing d.

The rôle of d can be understood by going to the frame theory. Although there is no 4-dimensional non-degenerate metric at $\lambda = 0$, it is still possible to define for every value of λ a 4-dimensional invariant volume form, unique up to orientation. Apart from the restrictions imposed by anti-symmetry it has only one non-zero component, and that is described by d.

As can now be seen immediately, $E^{\alpha\beta}$ contains ε only in the form of non-negative integer powers of λ. Hence it seems safe to take λ instead of ε as our basic variable. It is also not difficult to check that a solution of (3.12) for any positive λ is a solution of Ehlers' frame theory with λ being the causality constant. If $T^{\alpha\beta}$ is defined for $\lambda = 0$, the first fact shows that it is possible to put $\lambda = 0$ in (3.12), and the second suggests that this might result in Newtonian equations.

The frame theory shows that the equations of motion ${T^{\alpha\beta}}_{;\beta} = 0$, which for $\lambda \neq 0$ follow from the field equations, have to be imposed as an additional axiom if $\lambda = 0$ (the equation of continuity and Euler's equation do not follow from the Poisson equation for the Newtonian potential). One should therefore instead of (3.12) and the conditions (3.6) for harmonic coordinates consider the following set of equations [19]:

$$E^{\alpha\beta} = 4\pi G\,|d|\,T^{\alpha\beta}, \qquad {T^{\alpha\beta}}_{;\beta} = 0, \qquad {\mathfrak{U}^{\mu\nu}}_{,\nu} = 0. \tag{3.16}$$

These together with (3.13), (3.15), and the relations

$$\Gamma^{\alpha}_{\beta\gamma} = \mathfrak{g}^{\alpha\mu}(2\mathfrak{g}_{\beta\sigma}\mathfrak{g}_{\gamma\tau} - \mathfrak{g}_{\beta\gamma}\mathfrak{g}_{\sigma\tau}){\mathfrak{U}^{\sigma\tau}}_{,\mu} + 2\lambda(\mathfrak{g}_{\sigma\tau}\delta^{\alpha}_{(\beta}{\mathfrak{U}^{\sigma\tau}}_{,\gamma)} - 2\mathfrak{g}_{\sigma(\beta}{\mathfrak{U}^{\alpha\sigma}}_{,\gamma)}) \tag{3.17}$$

are the "ε-equations" for the method proposed here (although the parameter is called λ and not ε). The relativistic solution S_1 to be approximated is obviously a solution of (3.16) for $\lambda = \lambda_1 := \varepsilon_1^2$. But is a solution of (3.16) for $\lambda = 0$ a Newtonian solution?

3.3. The field equations at $\lambda = 0$

Putting $\lambda = 0$ in (3.16) we obtain the following equations:

$$\begin{aligned} &\triangle U = 4\pi G\varrho, \qquad \triangle W^a = 4\pi G j^a, \qquad \triangle Z^{ab} = 4\pi G S^{ab} + U_{,a}U_{,b} - \tfrac{1}{2}\,|\nabla U|^2\,\delta^{ab}, \\ &\varrho_{,0} + {j^b}_{,b} = 0, \qquad {j^a}_{,0} + {S^{ab}}_{,b} = -\varrho U_{,a}, \qquad {\mathfrak{U}^{\mu\nu}}_{,\nu} = 0. \end{aligned} \tag{3.18}$$

Here I have introduced density, current, and stress tensor:

$$\varrho := T^{00}, \qquad j^a := T^{0a}, \qquad S^{ab} := T^{ab}. \tag{3.19}$$

(For a perfect fluid with velocity field v^a and pressure p, the last two are $j^a = \varrho v^a$ and $S^{ab} = \varrho v^a v^b + p\delta^{ab}$.) Note that U is opposite in sign to the potential used in some post-Newtonian methods. The choice made here agrees with the convention in classical mechanics that the force is equal to minus the gradient of the potential. This introduces a minus sign in front of the Poisson integral, and methods making heavy use of the latter therefore prefer to shift the minus sign to Poisson's equation.

An important point in (3.18) is that all the equations for the gravitational field are elliptic and not hyperbolic, in contrast to the situation in general relativity. This has consequences for the nature of observable effects, but also for the treatment of the Cauchy problem. In the last section, I shall try to connect this fact with the breakdown of post-Newtonian methods.

By dropping the harmonicity condition and the equations for W^a and Z^{bc} we see that any solution $(\mathfrak{U}^{\alpha\beta}, T^{\mu\nu})$ of (3.18) gives also a solution $(U, \varrho, j^a, S^{bc})$ of the Newtonian theory of gravitation. Now suppose that a Newtonian solution is given. Because the Laplacian $\triangle\colon F_2 \to F_0$ is bijective, the Poisson equations for W^a and Z^{bc} can be solved uniquely. Hence, a Newtonian solution determines a pair $(\mathfrak{U}^{\alpha\beta}, T^{\mu\nu})$ of which we know that all the equations above are satisfied, possibly except ${\mathfrak{U}^{\mu\nu}}_{,\nu} = 0$. We now assume in addition (2.8) and obtain by insertion of the $\mathfrak{U}^{\alpha\beta}$ in the equation of continuity and Euler's equation:

$$0 = \varrho_{,0} + {j^b}_{,b} = \frac{\triangle {\mathfrak{U}^{0\beta}}_{,\beta}}{4\pi G}, \qquad 0 = {j^a}_{,0} + {S^{ab}}_{,b} + \varrho U_{,a} = \frac{\triangle {\mathfrak{U}^{a\beta}}_{,\beta}}{4\pi G}. \tag{3.20}$$

Because the Laplacian is injective on F_1 we can conclude that $\mathfrak{U}^{\alpha\beta}{}_{,\beta} = 0$ and hence that there is a one-to-one correspondence between Newtonian solutions and solutions of (3.18).

Note that W^a and Z^{bc} are physically irrelevant if $\lambda = 0$ (they disappear from the metric and the connection). Nevertheless, (3.18) determines them uniquely. This is not a deficiency but instead an indication of the suitability of (3.16)! No system of equations for the $\mathfrak{U}^{\alpha\beta}$ could be well-behaved at $\lambda = 0$ if some of the variables were not contained in it any more at that point.

3.4. The relativistic Cauchy problem

It is well known (see most textbooks on general relativity) that the solution of the Cauchy problem for Einstein's equations has to be divided into two steps. In the first, one prescribes *free data* on an initial hypersurface Σ_0 and solves *constraint equations* for some additional functions on this surface. Together with the free data, these form the input for step two in which one solves *evolution equations* for a solution in the neighbourhood of the initial surface. Because of the tensorial nature of the equations and the absence of tensorial theorems about their solutions it is also necessary to add a coordinate condition. Experience shows that harmonic coordinates are a useful choice, and this is enforced here by the observation that for a suitable choice of the spaces F_j Newtonian solutions are automatically harmonic if expressed in the usual coordinates of classical mechanics. (There are actually two notions of harmonicity here: $\Box\, x^\mu = 0$ and $\mathfrak{U}^{\alpha\beta}{}_{,\beta} = 0$. For $\lambda \neq 0$ they are equivalent, but in the case $\lambda = 0$ the first is much weaker than the second. The latter is the one used here.)

The constraint equations are geometrical in origin and therefore must be present in (3.16). It only remains to identify them in this formulation by looking for four relations not containing second-order time derivatives. First note that the harmonicity condition can be written

$$\dot{U} = -W^b{}_{,b}, \qquad \dot{W}^a = -Z^{ab}{}_{,b}, \tag{3.21}$$

and can therefore be used to eliminate all t-derivatives of U and W^a from the equations. This identifies $\dot{U}$ and $\dot{W}^a$ as functions which cannot be part of the set of free data. Furthermore, inspection of the second-order derivatives in $H^{0\alpha}$ reveals that after substitution of (3.21) these four expressions depend only on λ, U, W^a, Z^{bc}, $\dot{Z}^{de}$, and their spatial derivatives, but not on $\ddot{Z}^{ab}$:

$$\begin{aligned} H^{00} &= \triangle\, U - \lambda Z^{cd}{}_{,cd} + 4\lambda^2 \left(Z^{cd} U_{,cd} - 2W^c{}_{,cd} W^d + U Z^{cd}{}_{,cd}\right), \\ H^{0a} &= \triangle\, W^a + \lambda \dot{Z}^{ac}{}_{,c} + 4\lambda^2 \left(Z^{cd} W^a{}_{,cd} - 2Z^{ac}{}_{,cd} W^d - U \dot{Z}^{ac}{}_{,c}\right). \end{aligned} \tag{3.22}$$

Therefore, if one can choose a set of matter data such that $T^{0\alpha}$ is also independent of $\ddot{Z}^{ab}$, the constraint equations should be:

$$E^{0\alpha} = 4\pi G\,|d|\,T^{0\alpha} \qquad \text{after substitution of (3.21) and evaluated on } \Sigma_0. \tag{3.23}$$

For this reason I shall adopt the following prescription [20] for the first step in the solution of the Cauchy problem at $\lambda = \lambda_1$:

— Choose on the initial hypersurface as free data Z^{ab}, $\dot{Z}^{cd}$, and a suitable set m of matter variables.

— Solve the equations (3.23) for U and W^a.

— Determine $\dot{U}$ and $\dot{W}^a$ from (3.21).

The description of the evolution equations will be omitted here. As I have no statement about the existence of solutions to these equations in the function spaces considered, I can't be sure of having a sensible form for them anyway.

Note that I have been concerned here with the relativistic case, i.e., positive λ. This is not the correct way to approach the Cauchy problem for $\lambda = 0$, as can be seen from the equation for Z^{ab} in (3.18). I shall come back to this point in the last section. Until then, the reader should not be unduly bothered by this problem, because our goal is after all to obtain the relativistic and not a Newtonian solution.

4. Properties of the parametrized constraint equations

I use the expression "parametrized constraints" as an abbreviation for "those equations parametrized by λ which for $\lambda = \lambda_1$ are the relativistic constraints", i.e., (3.23). The aim of this section is to prove that there exists a unique and well-behaved map from λ and the relativistic free data $(Z^{ab}, \dot{Z}^{cd}, m)$ to solutions of (3.23) in a neighbourhood of $\lambda = 0$, and to show how approximations for solutions can be generated.

4.1. Properties of the metrics and of $E^{\alpha\beta}$

The purpose of the following lemma is to identify a set A_1 on which the metric functions $\mathfrak{g}^{\alpha\beta}$, d, and $\mathfrak{g}_{\mu\nu}$ behave as desired, when considered as functions of λ and $\mathfrak{U}^{\alpha\beta}$.

Lemma 4.1. *Let $\mathfrak{U}^{\mu\nu}$ be in $S_4(F_2)$, and define by (3.15) $\mathfrak{g}^{\alpha\beta}$ and d for every $(\lambda, \mathfrak{U}^{\alpha\beta})$, and $\mathfrak{g}_{\mu\nu}$ for every $(\lambda, \mathfrak{U}^{\alpha\beta})$ for which $d(\lambda, \mathfrak{U}^{\alpha\beta})$ belongs to* $\mathrm{Inv}(K_2)$.

Then we have:

(a) The map $(\lambda, \mathfrak{U}^{\mu\nu}) \mapsto \mathfrak{g}^{\alpha\beta}$ is in $C^\omega(\mathbb{R} \times S_4(F_2), S_4(K_2))$, and $(\lambda, \mathfrak{U}^{\mu\nu}) \mapsto d$ belongs to $C^\omega(\mathbb{R} \times S_4(F_2), K_2)$.

(b) The set

$$A_1 := \{ (\lambda, \mathfrak{U}^{\alpha\beta}) \in \mathbb{R} \times S_4(F_2) \mid d(\lambda, \mathfrak{U}^{\alpha\beta}) \in \mathrm{Inv}(K_2) \text{ and in each point of } \Sigma \\ \mathfrak{g}^{\alpha\beta}(\lambda, \mathfrak{U}^{\mu\nu}) \text{ can be diagonalized to } \mathrm{diag}(-\lambda, 1, 1, 1) \} \tag{4.1}$$

is open and contains all pairs with $\lambda = 0$. (It also contains all points with $\mathfrak{U}^{\alpha\beta} = 0$, but that is only of interest for post-Minkowskian methods.) On A_1 we have:

$$d(\lambda, \mathfrak{U}^{\alpha\beta}) < 0. \tag{4.2}$$

(c) The map $(\lambda, \mathfrak{U}^{\mu\nu}) \mapsto \mathfrak{g}_{\alpha\beta}$ belongs to $C^\omega(A_1, S_4(K_2))$.

This definition of A_1 ensures that in the case of positive λ the metric has Lorentz signature.

Proof: (a) Consider as an example the component $\mathfrak{g}^{00}(\lambda, \mathfrak{U}^{\alpha\beta}) = -\lambda + 4\lambda^2 U$. The projections $(\lambda, \mathfrak{U}^{\alpha\beta}) \mapsto \lambda$ and $(\lambda, \mathfrak{U}^{\alpha\beta}) \mapsto U$ are linear and continuous, therefore C^ω. $\lambda \mapsto \lambda^2$ is C^ω. $(\lambda^2, U) \mapsto 4\lambda^2 U$ is bilinear and continuous (scalar multiplication in F_2), therefore C^ω. The canonical inclusions $\mathbb{R} \to K_2$ and $F_2 \to K_2$ are linear and continuous, hence C^ω. Finally, addition in K_2 is bilinear and continuous, and therefore $(\lambda, 4\lambda^2 U) \mapsto \mathfrak{g}^{00}$ is C^ω. The chain rule then gives the result that $(\lambda, \mathfrak{U}^{\alpha\beta}) \mapsto \mathfrak{g}^{00}$ is an element of $C^\omega(\mathbb{R} \times S_4(F_2), K_2)$. The other components can be treated similarly.

For d, one also has to exploit the fact that multiplication in F_2 is bilinear and continuous with values in F_2.

(b) Consider the set

$$C := \{\, (\lambda, \mathfrak{U}^{\alpha\beta}) \in \mathbb{R} \times S_4(B) \mid d(\lambda, \mathfrak{U}^{\alpha\beta}) \in \mathrm{Inv}(B) \text{ and in each point of } \Sigma \\ \mathfrak{g}^{\alpha\beta}(\lambda, \mathfrak{U}^{\mu\nu}) \text{ can be diagonalized to } \mathrm{diag}(-\lambda, 1, 1, 1) \,\}, \tag{4.3}$$

where B again refers to the Banach algebra of bounded functions on Σ. Suppose we had already shown that this set is open with respect to the norm induced from $\mathbb{R}$ and B and that it contains all pairs $(0, \mathfrak{U}^{\alpha\beta})$. Then let $i \colon \mathbb{R} \times S_4(F_2) \to \mathbb{R} \times S_4(B)$ be the continuous embedding existing by virtue of (2.6). Because $\mathrm{Inv}(K_2)$ is a subset of $\mathrm{Inv}(B)$ (remember that the embedding of K_2 in B is also a homomorphism for the multiplication), we can rewrite the definition of A_1 as:

$$A_1 = d^{-1}(\mathrm{Inv}(K_2)) \cap i^{-1}(C).$$

But this is an intersection of two open sets (pre-images of open sets under continuous maps) and therefore also open. In addition, all pairs $(0, \mathfrak{U}^{\alpha\beta})$ belong to $d^{-1}(\mathrm{Inv}(K_2))$, because $d(0, \mathfrak{U}^{\alpha\beta}) = -1 \in \mathrm{Inv}(K_2)$. Thus it is sufficient to restrict our attention to C, and to prove that it is open in $\mathbb{R} \times S_4(B)$ and contains all pairs $(0, \mathfrak{U}^{\alpha\beta})$.

Because $\mathfrak{g}^{\alpha\beta}$ is a tensor density of weight $+1$, it is not obvious what I mean by "diagonalization". I define it to mean that in each point of Σ there is a collection of 4 tensor densities $E^{(\alpha)}{}_\beta$ of rank 1 and weight $-\frac{1}{2}$ such that:

$$\det(E^{(\alpha)}{}_\beta) > 0, \qquad \big(\mathfrak{g}^{\sigma\tau}(\lambda, \mathfrak{U})\, E^{(\alpha)}{}_\sigma\, E^{(\beta)}{}_\tau\big) = \mathrm{diag}(-\lambda, 1, 1, 1). \tag{4.4}$$

Hence the right-hand side denotes a collection of scalars, not a rank-2 tensor or a tensor density.

Putting $\lambda = 0$ and $E^{(\alpha)}{}_\beta = \delta^\alpha_\beta$ we see that (4.4) is satisfied. Together with $d(0, \mathfrak{U}) = -1 \in \mathrm{Inv}(B)$ this means that all $(0, \mathfrak{U})$ belong to C. Taking the determinant of (4.4) and using $\det(\mathfrak{g}^{\alpha\beta}) = \lambda d$ we obtain (4.2) for non-zero λ, $d(0, \mathfrak{U}) = -1$ completing the statement for all of $C \supset A_1$. It remains therefore to show that C is open.

Most people will believe or even know that in the set of symmetric matrices the non-degenerate matrices of a particular signature form an open subset. The statement to be proven here goes beyond this in two respects: first we need that C is open in the norm of bounded functions, i.e., uniformly over Σ. The second point is due to our desire to obtain an open set containing all points with $\lambda = 0$ and leading to a Lorentz metric for $\lambda > 0$. Hence it is unavoidable to require a λ-dependent rank and signature of $\mathfrak{g}^{\alpha\beta}$, and therefore the fact that C is open is a property of this particular map $(\lambda, \mathfrak{U}) \mapsto \mathfrak{g}^{\alpha\beta}$, and cannot be reduced to a property of the set of all $\mathfrak{g}^{\alpha\beta}$. This unfortunately entails a certain amount of uninteresting detail.

For the ensuing estimates I must select one of the norms inducing the usual topology on $S_4(B)$. I choose the "operator norm"

$$\|\mathfrak{U}\|_{B^{4\times 4}} := \sup\Big\{ \frac{\|\mathfrak{U}^{\alpha\beta} V_\alpha W_\beta\|_B}{\|V\|_{B^4} \|W\|_{B^4}} \;\Big|\; V,\, W \in B^4 \text{ and non-zero} \Big\} \tag{4.5}$$

induced from $B^{4\times 4}$, because it leads directly to the estimate

$$\|\mathfrak{U}^{\alpha\beta} V_\alpha W_\beta\|_B \le \|\mathfrak{U}\|_{B^{4\times 4}} \|V\|_{B^4} \|W\|_{B^4},$$

and that will be used frequently. The norm $\|V\|_{B^4}$ on bounded tensor densities V_α of weight $-\frac{1}{2}$ is defined by means of the tensor density of weight $+1$ associated with the Euclidean metric $e_{\mu\nu}$

defined in (3.10). In the coordinates chosen, however, tensor and tensor density both have the components $\delta^{\mu\nu}$, so this is a somewhat pedantic distinction:

$$\|V\|_{B^4} = \left(\left\|\sum_{\mu=0}^{3}(V_\mu)^2\right\|_B\right)^{1/2}.$$

Let there now be given an element $(\lambda_0, \mathfrak{U}_0)$ of C. My aim is to prove existence of positive numbers δ_1, $\delta_2 \in \mathbb{R}$ such that for all $(\lambda, \mathfrak{U})$ in $\mathbb{R} \times S_4(B)$ we have:

$$|\lambda - \lambda_0| < \delta_1, \quad \|\mathfrak{U} - \mathfrak{U}_0\|_{B^{4\times 4}} < \delta_2 \quad \Longrightarrow \quad (\lambda, \mathfrak{U}) \in C.$$

I shall first show that there is a set of *bounded* $E^{(\alpha)}$ diagonalizing $\mathfrak{g}^{\sigma\tau}(\lambda_0, \mathfrak{U}_0)$, and then that for sufficiently small δ_1, δ_2 these $E^{(\alpha)}$ can be modified to give a set of $F^{(\alpha)}$ diagonalizing $\mathfrak{g}^{\sigma\tau}(\lambda, \mathfrak{U})$.

The first step is trivial for $\lambda_0 = 0$, because the ${E^{(\alpha)}}_\beta = \delta^\alpha_\beta$ satisfying (4.4) in this case are obviously bounded. Let therefore λ_0 be non-zero. There is a theorem of algebra which tells us that because $\mathfrak{g}^{\alpha\beta}(\lambda_0, \mathfrak{U}_0)$ is symmetric there exists (at each point of Σ) a basis $\{H^{(\alpha)}\}$ which is at the same time orthogonal with respect to $\mathfrak{g}^{\alpha\beta}(\lambda_0, \mathfrak{U}_0)$ and orthonormal with respect to $\delta^{\mu\nu}$. Because of the second property, these $H^{(\alpha)}$ are bounded ($\|H\| = 1$). As the signature is basis-independent it can be arranged that the three diagonal elements of $\mathfrak{g}^{\sigma\tau}(\lambda_0, \mathfrak{U}_0)$ belonging to the $H^{(a)}$ are positive, and that the value belonging to $H^{(0)}$ has the sign opposite to that of λ_0. Putting

$$h^{(0)} := \frac{1}{\lambda_0}\, \mathfrak{g}^{\sigma\tau}(\lambda_0, \mathfrak{U}_0) {H^{(0)}}_\sigma {H^{(0)}}_\tau, \qquad h^{(a)} := \mathfrak{g}^{\sigma\tau}(\lambda_0, \mathfrak{U}_0) {H^{(a)}}_\sigma {H^{(a)}}_\tau$$

and

$$E^{(\alpha)} := H^{(\alpha)} / \sqrt{|h^{(\alpha)}|}$$

we obtain a basis $\{E^{(\alpha)}\}$ satisfying (4.4), but we do not yet know whether it is bounded. However, because the $H^{(\alpha)}$ are orthonormal with respect to $\delta^{\mu\nu}$ the coefficient matrix ${H^{(\alpha)}}_\beta$ is orthogonal. Taking the determinant of

$$\left(\mathfrak{g}^{\sigma\tau}(\lambda_0, \mathfrak{U}_0) {H^{(\alpha)}}_\sigma {H^{(\beta)}}_\tau\right) = \operatorname{diag}\left(\lambda_0 h^{(0)}, h^{(1)}, h^{(2)}, h^{(3)}\right)$$

and using $\det(\mathfrak{g}^{\mu\nu}) = \lambda d$ we obtain:

$$\prod_{\alpha=0}^{3} h^{(\alpha)} = d(\lambda_0, \mathfrak{U}_0).$$

As $d(\lambda_0, \mathfrak{U}_0)$ belongs to $\operatorname{Inv}(B)$ its inverse is defined and bounded. Hence there exists a positive number a such that $\prod_{\alpha=0}^{3} \left|h^{(\alpha)}\right| = |d(\lambda_0, \mathfrak{U}_0)| \geq a$. On the other hand there is an upper bound $b \geq \left|h^{(\alpha)}\right|$ because $\mathfrak{g}^{\alpha\beta}$ and the $H^{(\alpha)}$ are bounded. We obtain:

$$\left|h^{(\alpha)}\right| = \frac{\prod_{\mu=0}^{3} \left|h^{(\mu)}\right|}{\prod_{\mu=0, \mu\neq\alpha}^{3} \left|h^{(\mu)}\right|} \geq \frac{a}{b^3}.$$

Therefore each $1/\left|h^{(\alpha)}\right|$ is bounded, and hence the $E^{(\alpha)}$ are bounded as desired.

The remainder of this proof is concerned with showing that there exist positive real numbers δ_1, δ_2, ζ, and η such that $\zeta < 1$, $\eta < 1$, and for all $(\lambda, \mathfrak{U})$ satisfying $|\lambda - \lambda_0| < \delta_1$ and $\|\mathfrak{U} - \mathfrak{U}_0\| < \delta_2$ there exists a collection of $F^{(\alpha)}$ such that:

$$\begin{aligned} &\det({F^{(\alpha)}}_\beta) = \det({E^{(\alpha)}}_\beta), \qquad \mathfrak{g}^{\sigma\tau}(\lambda, \mathfrak{U})\, {F^{(\alpha)}}_\sigma\, {F^{(\beta)}}_\tau = 0 \quad \text{if } \alpha \neq \beta, \\ &\left|f^{(a)} - 1\right| \leq \zeta, \qquad \left|f^{(0)} + \lambda\right| \leq \eta\, |\lambda|. \end{aligned} \tag{4.6}$$

Here I have defined:

$$f^{(\alpha)} := \mathfrak{g}^{\sigma\tau}(\lambda,\mathfrak{U})F^{(\alpha)}{}_\sigma F^{(\alpha)}{}_\tau. \tag{4.7}$$

From (4.6) we can conclude that the $F^{(\alpha)}$ form an orthogonal basis with respect to $\mathfrak{g}^{\alpha\beta}(\lambda,\mathfrak{U})$, and that the diagonal elements $f^{(\alpha)}$ have the proper signs. Hence $\mathfrak{g}^{\alpha\beta}$ has the correct signature. For non-zero λ ($\lambda = 0$ is trivial) we can also conclude that $d(\lambda,\mathfrak{U})$ belongs to $\mathrm{Inv}(B)$, because the estimates ensure that the $f^{(\alpha)}$ belong to $\mathrm{Inv}(B)$ and because we have:

$$d(\lambda,\mathfrak{U}) = \frac{\prod f^{(\alpha)}}{\lambda\det^2(F^{(\mu)}{}_\nu)} = \frac{\prod f^{(\alpha)}}{\lambda\det^2(E^{(\mu)}{}_\nu)} = -\frac{\prod f^{(\alpha)}}{\lambda} d(\lambda_0,\mathfrak{U}_0).$$

Here I have once more used $\det(\mathfrak{g}^{\alpha\beta}) = \lambda d$. Therefore, (4.6) ensures that all these $(\lambda,\mathfrak{U})$ belong to C. It remains to prove that statement.

First consider the case $\lambda_0 = 0$. I can choose $E^{(\alpha)}{}_\beta = \delta^\alpha_\beta$ and:

$$\begin{aligned} &\eta \in \mathbb{R}, \qquad 0 < \eta < 1, \qquad 0 < \delta_2 \in \mathbb{R}, \\ &\delta_1 := \frac{1}{6\eta}\left(\sqrt{1 + \frac{3\eta^2}{\|\mathfrak{U}_0\| + \delta_2}} - 1\right), \qquad \gamma := 4\delta_1(\|\mathfrak{U}_0\| + \delta_2), \qquad \zeta := 3\delta_1\gamma. \end{aligned} \tag{4.8}$$

This implies $\zeta = 1 - \gamma/\eta < 1$. Now assume $|\lambda| < \delta_1$, $\|\mathfrak{U} - \mathfrak{U}_0\| < \delta_2$. As a first step I prove by induction over a:

For each $a \in \mathbb{N}^*$, $a \leq 3$, there exists a collection $\{\, F^{(j)} \mid 1 \leq j \leq a \,\}$ such that:

(i) $F^{(j)} - E^{(j)}$ is a linear combination of the $E^{(k)}$ with $k < j$, and:

$$\|F^{(j)} - E^{(j)}\| \leq \frac{\gamma_j}{1-\gamma_j},$$

where:

$$\gamma_l := |\lambda|\,(l-1)\gamma < \zeta < 1, \qquad 1 \leq l \leq 4. \tag{4.9}$$

(ii) If $j \neq k$, $F^{(j)}$ is orthogonal to $F^{(k)}$ with respect to $\mathfrak{g}^{\alpha\beta}(\lambda,\mathfrak{U})$.

(iii)

$$\left|f^{(j)} - 1\right| \leq \frac{|\lambda|\,\gamma}{1-\gamma_j}.$$

For $a = 1$, (i) leaves no choice but to take $F^{(1)} = E^{(1)}$, and (ii) is always true.

$$\left|f^{(1)} - 1\right| = \left|\mathfrak{g}^{\sigma\tau}E^{(1)}{}_\sigma E^{(1)}{}_\tau - 1\right| = 4\lambda^2\left|Z^{11}\right| \leq 4\lambda^2\|\mathfrak{U}\| \leq |\lambda|\,\gamma,$$

and that gives (iii). Now assume that the conditions (i)–(iii) are true for some a. Because of $|\lambda|\,\gamma < \delta_1\gamma = \zeta/3$ we have:

$$\left|f^{(j)} - 1\right| < \frac{\delta_1\gamma}{1 - 2\delta_1\gamma} < \zeta < 1$$

and hence:

$$\left|f^{(j)}\right| \geq 1 - \left|f^{(j)} - 1\right| \geq \frac{1-\gamma_{j+1}}{1-\gamma_j} > 0.$$

Therefore I can use the Gram-Schmidt method of orthogonalization:

$$F^{(a+1)} := E^{(a+1)} - \sum_{j=1}^{a} \frac{q_j^{(a+1)}}{f^{(j)}} F^{(j)}, \qquad a < 3,$$

where:
$$q_k^{(\alpha)} := \mathfrak{g}^{\sigma\tau}(\lambda,\mathfrak{U})\, E^{(\alpha)}{}_\sigma\, F^{(k)}{}_\tau, \qquad \alpha > k \text{ or } \alpha = 0. \tag{4.10}$$
(We shall need $\alpha = 0$ later.) This definition for $F^{(a+1)}$ ensures (ii) and also the first part of (i). For the second part of (i) we need a few estimates:

$$\|F^{(k)}\| \leq \|E^{(k)}\| + \|F^{(k)} - E^{(k)}\| \leq \frac{1}{1-\gamma_k},$$
$$\left|q_k^{(\alpha)}\right| = 4\lambda^2 \left|\mathfrak{U}^{\alpha b} F^{(k)}{}_b\right| \leq 4\lambda^2 \|\mathfrak{U}\| \|F^{(k)}\| \leq \frac{|\lambda|\,\gamma}{1-\gamma_k}.$$

Hence:

$$\|F^{(a+1)} - E^{(a+1)}\| \leq \sum_{j=1}^{3} \frac{\left|q_j^{(a+1)}\right|}{f^{(j)}} \|F^{(j)}\| \leq |\lambda|\,\gamma \sum_{j=1}^{a} \frac{1}{(1-\gamma_j)(1-\gamma_{j+1})} = \frac{\gamma_{a+1}}{1-\gamma_{a+1}}.$$

This proves (i), leaving (iii):

$$\begin{aligned}
\left|f^{(a+1)} - 1\right| &\leq \left|\mathfrak{g}^{\sigma\tau} E^{(a+1)}{}_\sigma E^{(a+1)}{}_\tau - 1\right| + \sum_{j=1}^{a} \frac{\left|q_j^{(a+1)}\right|^2}{f^{(j)}} \\
&\leq 4\lambda^2 \|\mathfrak{U}\| + \sum_{j=1}^{a} \frac{\lambda^2\gamma^2}{(1-\gamma_j)(1-\gamma_{j+1})} \\
&\leq \frac{|\lambda|\,\gamma}{1-\gamma_{a+1}}.
\end{aligned}$$

This completes the induction, and (i)–(iii) are thus shown to be valid for $a = 3$. The next step is to choose
$$F^{(0)} := E^{(0)} - \sum_{k=1}^{3} \frac{q_k^{(0)}}{f^{(k)}} F^{(k)},$$

in order to obtain the final vector for an orthonormal basis. Checking (4.6) we find that only one statement remains to be proven, the estimate for $f^{(0)}$:

$$\begin{aligned}
\left|f^{(0)} + \lambda\right| = \left|\mathfrak{g}^{\sigma\tau} E^{(0)}{}_\sigma F^{(0)}{}_\tau + \lambda\right| &\leq \left|\mathfrak{g}^{\sigma\tau} E^{(0)}{}_\sigma E^{(0)}{}_\tau + \lambda\right| + \sum_{j=1}^{3} \frac{\left|q_j^{(0)}\right|^2}{f^{(j)}} \leq \\
\leq 4\lambda^2 \|\mathfrak{U}\| + \sum_{j=1}^{3} \frac{\lambda^2\gamma^2}{(1-\gamma_j)(1-\gamma_{j+1})} &\leq \frac{|\lambda|\,\gamma}{1-\zeta} = |\lambda|\,\eta,
\end{aligned}$$

and thus the proof for (4.6) is complete in the case $\lambda_0 = 0$.

I am not going to prove the case $\lambda_0 \neq 0$. The proof proceeds on exactly the same lines as for $\lambda_0 = 0$, the main difference being that $|\lambda|\,\gamma$ can be replaced by a suitable ψ satisfying $\|\mathfrak{g}(\lambda,\mathfrak{U}) - \mathfrak{g}(\lambda_0,\mathfrak{U}_0)\| \leq \psi$. The parameters δ_1, δ_2, ζ, and η will of course be different, also the norms $\|E^{(\alpha)}\|$ will now enter explicitly. For example, all the factors 3 in (4.8) are really $\sum_{j=1}^{3}\|E^{(j)}\|^2$. If a reader does not believe the statement for this case, she or he should replace "A_1" by "open kernel of A_1" in the remainder of the paper. The proof as far as it is given here shows that this contains all points with $\lambda = 0$, and that is all which will be used.

(c) The components of the adjoint matrix adj $\left(\mathfrak{g}^{\alpha\beta}\right)$ are sub-determinants of $\left(\mathfrak{g}^{\alpha\beta}\right)$, i.e., they are built by multiplication and addition in K_2. These are bilinear and continuous operations and therefore C^ω. d maps A_1 into $\mathrm{Inv}(K_2)$, so $1/d$ is defined and belongs also to K_2. Moreover, the map $d \mapsto 1/d$ is C^ω, and multiplying $1/d$ with $\mathrm{adj}\,(\mathfrak{g}^{\alpha\beta})_{\mu\nu}$ in K_2 we find by the chain rule that $\mathfrak{g}_{\mu\nu}$ is a C^ω function of $(\lambda, \mathfrak{U}^{\alpha\beta})$ on A_1. ■

Having ensured that the functions defined in (3.15) are well-behaved, we can now approach (3.13).

Lemma 4.2. *Let*

$$A_2 := A_1 \times S_4(F_1) \times S_4(F_0) \subset \mathbb{R} \times S_4(F_2) \times S_4(F_1) \times S_4(F_0). \tag{4.11}$$

Then A_2 is open, contains all points with $\lambda = 0$, and the maps $(\lambda, \mathfrak{U}^{\alpha\beta}, \dot{\mathfrak{U}}^{\gamma\delta}, \ddot{\mathfrak{U}}^{\varepsilon\zeta}) \mapsto E^{\alpha\beta}$ belong to $C^\omega(A_2, F_0)$.

Proof: A_2 is open and contains all points with $\lambda = 0$ because this is true of A_1. It therefore remains only to show analyticity. Looking at the form of $E^{\alpha\beta}$ in (3.13) and keeping in mind the lessons of the previous lemma we find that we merely have to check that a number of operations are (bi-)linear and continuous with results in the correct function spaces.

Consider first the derivatives. The $\mathfrak{U}^{\alpha\beta}{}_{,\mu}$ are either $\dot{\mathfrak{U}}^{\alpha\beta}$ ($\mu = 0$) or $\mathfrak{U}^{\alpha\beta}{}_{,i}$. In the first case they are among the independent variables, in the second they are obtained by a linear continuous map from $\mathfrak{U}^{\alpha\beta}$. In both cases they belong to F_1. Second order derivatives are in the same way seen to be C^ω functions with values in F_0.

Having established this, $H^{\alpha\beta}$ is seen to be C^ω because of the properties of the multiplication $\cdot : K_2 \times F_0 \to F_0$ and of the addition in F_0. $A^{\alpha\beta}$, $B^{\alpha\beta}$, and $C^{\alpha\beta}$ follow suit with the multiplications $\cdot : K_2 \times K_2 \to K_2$, $\cdot : F_1 \times F_1 \to F_0$, and $\cdot : K_2 \times F_0 \to F_0$. ■

4.2. An interlude on perfect fluids

The main theorem later in this paper will require certain properties of the matter tensor. The following lemma is intended as a side remark to show that these properties can be satisfied in the case of an isentropic perfect fluid with a prescribed equation of state.

Lemma 4.3. *The matter tensor for a perfect fluid in the frame theory is*

$$T^{\alpha\beta} := (\varrho + \lambda p) U^\alpha U^\beta + p g^{\alpha\beta} \tag{4.12}$$

with density ϱ, pressure p, and 4-velocity U^α. Choose $m := (\varrho, v^a) \in M := F_0 \times K_2^3$ as matter variables,

$$U^a = U^0 v^a, \qquad \sqrt{|d|}\, \mathfrak{g}_{\mu\nu} U^\mu U^\nu = -1, \tag{4.13}$$

and assume an equation of state

$$p = z(\lambda, \varrho), \qquad z \in C^k(A_3, F_0), \tag{4.14}$$

for some $k \in \overline{\mathbb{N}}$ and some open $A_3 \subset \mathbb{R} \times F_0$. Finally, let

$$A_4 := \{\, (\lambda, \mathfrak{U}^{\alpha\beta}, m) \in A_1 \times M \mid |d| \in W,\ \mathfrak{g}_{\mu\nu} v^\mu v^\nu \in \mathrm{Inv}^-(K_2),\ (\lambda, \varrho) \in A_3 \,\}, \tag{4.15}$$

with $v^0 := 1$, $W := \{\, a \in K_2 \mid \|a - 1\| < 1 \,\}$, and $\mathrm{Inv}^-(K_2)$ denoting the everywhere negative elements in $\mathrm{Inv}(K_2)$.

Then A_4 is open and contains all $(0, \mathfrak{U}^{\alpha\beta}, m)$ for which $(0, \varrho)$ belongs to A_3. The maps $(\lambda, \mathfrak{U}^{\alpha\beta}, m) \mapsto T^{\mu\nu}$ belong to $C^k(A_4, F_0)$, and we have:

$$T^{00} = \varrho, \qquad T^{0a} = \varrho v^a \qquad \textit{for } \lambda = 0. \tag{4.16}$$

Before I embark on the proof, a word of explanation seems necessary regarding the normalization condition for U^α. It is equivalent to $\hat{g}_{\mu\nu} U^\mu U^\nu = 1$ where $\hat{g}_{\mu\nu} := -\lambda g_{\mu\nu} = -\sqrt{|d|}\, \mathfrak{g}_{\mu\nu}$ is the "time metric" of Ehlers' frame theory, used to measure time intervals, whereas proper lengths are determined with the help of the "space metric" $g^{\alpha\beta} = \mathfrak{g}^{\alpha\beta}/\sqrt{|d|}$. In the case $\lambda = 0$, these two separate and it is no longer possible to determine one from the other. But the $\mathfrak{U}^{\alpha\beta}$ determine values for both.

Proof: The set $\mathrm{Inv}^-(K_2)$ is open because it can be written as $\mathrm{Inv}(K_2) \cap i^{-1}(\mathrm{Inv}^-(B))$, where the inclusion $i\colon K_2 \to B$ is continuous, and because $\mathrm{Inv}^-(B)$ is open:

$$f \in \mathrm{Inv}^-(B), \quad g \in B, \quad \|g - f\|_B < \frac{1}{\|1/f\|_B} \quad \Longrightarrow \quad g \in \mathrm{Inv}^-(B). \tag{4.17}$$

Because of (4.2), we have $|d| = -d$ on $A_1 \times M$. A_4 is thus the intersection of four open sets: the first is $A_1 \times M$, and the other three are pre-images of open sets under continuous maps and therefore also open. Thus A_4 is open, and because the first three sets contain all values with $\lambda = 0$, it contains all $(\lambda = 0)$-points for which $(0, \varrho)$ belongs to A_3.

Now consider $T^{\alpha\beta}$. The main point to be shown is that $U^\alpha U^\beta$ is a C^ω function on A_4. This is due to the following facts: (a) $-d$ is an analytic map from A_4 to W, (b) the square root is C^ω on W with values in W, (c) W is a subset of $\mathrm{Inv}(K_2)$, (d) $\mathfrak{g}_{\mu\nu} v^\mu v^\nu$ belongs to $\mathrm{Inv}(K_2)$, (e) multiplication in $\mathrm{Inv}(K_2)$ is C^ω with values in $\mathrm{Inv}(K_2)$, and (f) division on $\mathrm{Inv}(K_2)$ is C^ω. All this tells us that

$$(U^0)^2 = -\frac{1}{\sqrt{|d|}\, \mathfrak{g}_{\mu\nu} v^\mu v^\nu} \tag{4.18}$$

is well-defined, positive, and in $C^\omega(A_4, K_2)$, and the rest is simple. ■

4.3. Solution of the constraints

The independent variable (representing λ and the relativistic free data) in the parametrized constraint equations is

$$x := (\lambda, Z^{ab}, \dot{Z}^{cd}, m) \in E := \mathbb{R} \times S_3(F_2) \times S_3(F_1) \times M, \tag{4.19}$$

where M is some Banach space for the matter variables m. The solution to be determined is

$$y := (U, W^a) \in F := F_2^4, \tag{4.20}$$

and we shall also need the space

$$G := F_0^4. \tag{4.21}$$

Eliminating $\dot{y}$ and $\ddot{y}$ by means of the harmonicity condition (3.21) we obtain the parametrized constraint equations (3.23) as a function of x and y,

$$f^\alpha(x, y) := E^{0\alpha} - 4\pi G\, |d|\, T^{0\alpha} \qquad \text{on } \Sigma_0, \tag{4.22}$$

provided $T^{0\alpha}$ does not contain $\ddot{Z}^{ab}$.

Lemma 4.4. *Let the maps $(x, y) \mapsto T^{0\alpha}$ be in $C^k(A_5, F_0)$, $k \in \overline{\mathbb{N}}$, for some open $A_5 \subset E\times F$ with the property:*

$$(x, y) \in A_5 \text{ and } \lambda(x) = 0 \quad \Longrightarrow \quad \forall y' \in F : \ (x, y') \in A_5. \tag{4.23}$$

On $N \times F$,

$$N := \{\, x \in E \mid \lambda(x) = 0 \text{ and } (x, y) \in A_5 \text{ for some } y \in F \,\}, \tag{4.24}$$

$T^{0\alpha}$ is assumed not to depend on y. (This means: the matter variables must be chosen such that this is the case.) Further, let:

$$A := \{\, (x, y) \in A_5 \mid (\lambda, \mathfrak{U}^{\mu\nu}) \in A_1 \,\}. \tag{4.25}$$

Then the following statements hold:

(a) A is open and contains $N \times F$. The constraint function f is defined on A and belongs to $C^k(A, G)$.

(b) For every $x_0 \in N$ there is exactly one $y_0 \in F$ with $f(x_0, y_0) = 0$.

(c) For every $(x_0, y_0) \in N \times F$ with $f(x_0, y_0) = 0$ there exist open sets $L \subset E$, $V \subset A$ with $\{x_0\} \times F \subset V$, and a function $l \in C^k(L, F)$ such that for all $(x, y) \in E \times F$ one has:

$$x \in L, \quad y = l(x) \quad \Longleftrightarrow \quad (x, y) \in V, \quad f(x, y) = 0. \tag{4.26}$$

Proof: (a) A is the intersection of A_5 with the pre-image of A_1 under a continuous map and therefore open. By assumption, A_5 contains $N \times F$, and A_1 contains all points with $\lambda = 0$ (see lemma 4.1). Hence $N \times F$ is a subset of A.

By lemma 4.2 we know that the $E^{\alpha\beta}$ belong to $C^\omega(A_2, F_0)$. The elimination of the time derivatives of the $\mathfrak{U}^{0\alpha}$ is described by the function:

$$\begin{aligned} &s\colon S_4(F_2) \times S_3(F_1) \to S_4(F_2) \times S_4(F_1) \times F_0^4, \\ &(\mathfrak{U}^{\mu\nu}; \dot{\mathfrak{U}}^{0\alpha}, \dot{Z}^{ab}; \ddot{U}, \ddot{W}^a) = s(\mathfrak{U}^{\sigma\tau}, \dot{Z}^{ab}) := (\mathfrak{U}^{\mu\nu}; -\mathfrak{U}^{\alpha c}{}_{,c}, \dot{Z}^{ab}; Z^{cd}{}_{,cd}, -\dot{Z}^{ac}{}_{,c}). \end{aligned} \tag{4.27}$$

This map is C^ω, because its components are built from continuous linear or bilinear operations. Hence the $E^{0\alpha}$ as appearing in f,

$$E^{0\alpha}\big(\lambda, s(\mathfrak{U}^{\mu\nu}, \dot{Z}^{ab})\big),$$

are seen to be in $C^\omega(A, F_0)$. The remaining term in $f(x, y)$, $4\pi G\,|d|\,T^{0\alpha}$, is first converted to $-4\pi G\,d\,T^{0\alpha}$ by taking into account (4.2), and the rest is left as an exercise for the reader.

(b) Putting $\lambda = 0$ and using $A^{0\alpha}\big|_{\lambda=0} = 0$, $H^{0\alpha}\big|_{\lambda=0} = \triangle\,\mathfrak{U}^{0\alpha}$ (from (3.22)), and $d|_{\lambda=0} = -1$ we see that for any $x_0 \in N$, $y \in F$ we have:

$$f^\alpha(x_0, y) = \triangle\,\mathfrak{U}^{0\alpha} - 4\pi G\,T^{0\alpha}(x_0). \tag{4.28}$$

The resulting four Poisson equations can be solved uniquely for the $\mathfrak{U}^{0\alpha}$, i.e., for y.

(c) For this point, we use the implicit function theorem. If we check its assumptions we find that it remains only to verify that $D_2 f(x_0, y_0)$ is bijective. But from (4.28) it follows immediately that:

$$D_2 f^\alpha(x_0, y_0)(\eta) = \triangle\,\eta^\alpha, \qquad \eta \in F. \tag{4.29}$$

Hence $D_2 f(x_0, y_0): F_2^4 \to F_0^4$ consists of 4 independent Laplacians $\triangle: F_2 \to F_0$ and is therefore bijective. Copying the statement of the implicit function theorem we obtain (c) with a set V' satisfying only $(x_0, y_0) \in V'$ instead of $\{x_0\} \times F \subset V'$. But because $G - \{0\}$ is open and f is continuous, the set $f^{-1}(G - \{0\})$ of all pairs (x, y) not satisfying the constraints is an open subset of A. Because of (b), it contains all (x_0, y), $y \in F$, with the exception of (x_0, y_0). Putting

$$V := V' \cup f^{-1}(G - \{0\})$$

we obtain an open subset of A, containing V' and $\{x_0\} \times F$, and having the property that any $(x, y) \in V$ with $f(x, y) = 0$ must belong to V'. This completes the proof. ■

We now come to the final result:

Theorem (Solution of the Parametrized Constraint Equations)
Let M be a Banach space for the matter variables m,

$$\begin{aligned} x = (\lambda, Z^{ab}, \dot{Z}^{ab}, m) &\in E = \mathbb{R} \times S_3(F_2) \times S_3(F_1) \times M, \\ y = (U, W^a) &\in F = F_2^4, \\ G &= F_0^4, \end{aligned}$$

and assume for the matter tensor:

$$\begin{aligned} &(x, y) \mapsto T^{0\alpha} \in C^k(A_5, F_0), \qquad k \in \overline{\mathbb{N}}, \qquad A_5 \subset E \times F \text{ open}, \\ &(x, y) \in A_5 \text{ and } \lambda(x) = 0 \quad \Longrightarrow \quad (x, y') \in A_5 \text{ for all } y' \in F, \\ &T^{0\alpha}(x, y) = T^{0\alpha}(x) \quad \text{if } \lambda(x) = 0. \end{aligned}$$

Then the parametrized constraints $f(x, y) = 0$,

$$f: A \to G, \qquad A = \{\, (x, y) \in A_5 \mid (\lambda, \mathfrak{U}^{\mu\nu}) \in A_1 \,\}, \qquad f^\alpha(x, y) = E^{0\alpha} - 4\pi G\, |d|\, T^{0\alpha},$$

with A_1 from (4.1), are well-defined and satisfy:

There exist open neighbourhoods $L \subset E$ of

$$N = \{\, x \in E \mid \lambda(x) = 0 \text{ and } (x, y) \in A_5 \text{ for some } y \in F \,\}$$

and $V \subset A$ of $N \times F$ and a map $l \in C^k(L, F)$ such that for all (x, y) in $E \times F$ we have:

$$x \in L, \quad y = l(x) \qquad \Longleftrightarrow \qquad (x, y) \in V, \quad f(x, y) = 0.$$

This theorem tells us that (a) in a neighbourhood L of N there exists a C^k map from data x to solutions y of the parametrized constraints, and (b) that this map picks up all solutions (x, y) in a neighbourhood V of the solutions belonging to data in N.

Proof: Take any $x_0 \in N$. By lemma 4.4(b) there is a unique $y_0 \in F$ such that $f(x_0, y_0) = 0$, and lemma 4.4(c) gives existence of $L(x_0)$, $V(x_0)$, and l_{x_0} with the properties:

$$\begin{aligned} &L(x_0) \subset E \text{ open}, \qquad V(x_0) \subset A \text{ open}, \\ &l_{x_0} \in C^k(L(x_0), F), \qquad \{x_0\} \times F \subset V(x_0), \\ &\forall (x, y) \in E \times F: \quad x \in L(x_0), \quad y = l_{x_0}(x) \qquad \Longleftrightarrow \qquad (x, y) \in V(x_0), \quad f(x, y) = 0. \end{aligned} \tag{4.30}$$

To prove the statement of the theorem one has to show that on a suitable set L the various l_{x_0} are simply restrictions of a map l defined on L. This, of course, is what lemma 2.1 is for. To use it we have to make sure that for any pair $L(x_0)$, $L(x_0')$ with non-empty intersection both contain a point where l_{x_0} and $l_{x_0'}$ agree. We know that they have to agree on N by lemma 4.4(b), therefore one possibility is to restrict these neighbourhoods such that if they have a non-empty intersection, it must contain a point of N.

First, select for each $L(x_0)$ an open convex subset $L'(x_0)$ which is still a neighbourhood of x_0. The projection onto the Newtonian plane,

$$\pi: E \to E, \qquad \pi(\lambda, \mathfrak{U}, \ldots) := (0, \mathfrak{U}, \ldots), \tag{4.31}$$

is linear and continuous, therefore $K(x_0) := \pi^{-1}(L'(x_0))$ is open and convex (it is helpful to draw a picture of E at this point). The set $L''(x_0) := L'(x_0) \cap K(x_0)$ then has the following properties:

— it is open and convex,

— it contains x_0,

— it is a subset of $L(x_0)$, and therefore l_{x_0} is in $C^k(L''(x_0), F)$,

— it satisfies $\pi(L''(x_0)) \subset L''(x_0) \cap N$.

This leads to:

$$L := \bigcup_{x_0 \in N} L''(x_0), \qquad V := \Big(\bigcup_{x_0 \in N} V(x_0)\Big) \cap (L \times F). \tag{4.32}$$

These sets are open, L contains N, and V contains $N \times F$.

f is “locally injective for solutions” on V as required by lemma 2.1: take any pair $(x, y) \in V$; by construction, there is an open neighbourhood $V(x_0) \cap (L \times F) \subset V$ of (x, y) for some $x_0 \in N$. On $V(x_0)$ we have by (4.30):

$$(x', y') \in V(x_0), \quad f(x', y') = 0 \quad \Longrightarrow \quad x' \in L(x_0), \quad y' = l_{x_0}(x'),$$

and hence:

$$(x', y_1'),\ (x', y_2') \in V(x_0), \quad f(x', y_1') = 0 = f(x', y_2') \quad \Longrightarrow \quad y_1' = y_2'. \tag{4.33}$$

Now take any $x \in L$. I claim that the definition

$$l: L \to F, \qquad l(x) := l_{x_0}(x) \quad \text{for some } x_0 \in N \text{ with } x \in L''(x_0), \tag{4.34}$$

is free of contradictions. To prove this I have to show that any two possible l_{x_0}, $l_{x_0'}$ agree in x. Let $\widehat{L} := L''(x_0) \cap L''(x_0')$. $\widehat{L}$ is open and connected, and it contains x and $\hat{x} := \pi(x) \in N$, with $l_{x_0}(\hat{x}) = l_{x_0'}(\hat{x})$ by lemma 4.4(b). For every $x' \in \widehat{L}$ we have:

$$x' \in \widehat{L} \subset L''(x_0) \subset L(x_0) \quad \Longrightarrow \quad (x', l_{x_0}(x')) \in V(x_0),$$

and the same for x_0'. Hence $(x', l_{x_0}(x')) \in V$ and $(x', l_{x_0'}(x')) \in V$. We conclude from lemma 2.1 $l_{x_0} = l_{x_0'}$ on $\widehat{L}$ and therefore that $l_{x_0}(x) = l_{x_0'}(x)$. Besides that we have also obtained $(x, l(x)) \in V$ and $f(x, l(x)) = 0$.

This leaves only uniqueness still unproven. Let $(x, y) \in V$, $f(x, y) = 0$. By definition of V there is an x_0 such that $(x, y) \in V(x_0)$, and (4.30) gives $x \in L(x_0)$, $y = l_{x_0}(x)$. But $(x, y) \in V$ also implies (again by definition of V) $x \in L$, and on $L \cap L(x_0)$ we have $l_{x_0} = l$, hence $y = l(x)$. ■

This theorem is proven here because of the statement it contains about the dependence of solutions of the constraints on λ. Regarding existence and uniqueness of relativistic solutions ($\lambda = \lambda_1$), it presents nothing fundamentally new [20,21,22,23], although usually other variables are employed and the equations are inverted in a neighbourhood of Minkowski space. It seems, however, to have gone unnoticed so far that the proofs can be freed from their dependence on particular function spaces and on the topological properties of the initial hypersurface. The theorem given here covers local as well as global results, provided we can find suitable function spaces for a given Σ. Strictly speaking this has been shown only for $\Sigma \subset \mathbb{R}^3$, because this paper is concerned with post-Newtonian methods, and in the Newtonian case Σ must carry a Euclidean metric [3]. But the extension to other manifolds in general relativity should be straightforward. Apart from continuity for multiplications and derivatives, the key requirement seems to be that a certain elliptic operator for a Riemannian background metric is bijective (for Minkowski space the flat Laplacian), if necessary paired with restriction to a boundary. It would be desirable to approach the evolution equations in a similar spirit. There, a hyperbolic operator for a background Lorentz metric should play a corresponding rôle (the wave operator for Minkowski space). This might give a well-structured and comparatively simple proof for the existence of solutions of Einstein's equations in a neighbourhood of a known solution.

4.4. Application to approximations

Suppose one wishes to obtain a post-Newtonian approximation for a relativistic solution S_1, characterized by means of its free initial data $(Z_1^{ab}, \dot{Z}_1^{cd}, m_1)$. The theorem in section 4.3 then tells us that if (a) the matter variables have been chosen such that the matter tensor in the Newtonian case does not depend on the gravitational potentials U and W^a, (b) the matter tensor is at least C^1 in its arguments, and (c) $x_1 = (\lambda_1, Z_1^{ab}, \dot{Z}_1^{cd}, m_1)$ belongs to a certain set L, there exists a solution $y_1 = l(x_1)$ of the constraint equations, unique in a set V. The theorem does not tell us how large L or V are. However, a lower bound s on the radius of L around any $x_0 \in N$ can be obtained from the prescription given in section 2, so provided we can find a suitable collection (x_0, a, r) (y_0 is determined by x_0) with $\|x_1 - x_0\|_E \leq s$, we know that y_1 exists. For this step it is in particular necessary to know a bound on $\|\triangle^{-1}\|_{\mathscr{L}(F_0,F_2)}$ for the function spaces one has chosen. Having thus established the existence of the solution we can try to approximate it. I consider two possibilities: expansion in powers and iteration.

To use expansion in powers of a real parameter ε as a method of approximation we have first to choose a family $x = A(\varepsilon)$ of initial data. It should have the properties:

— $A(0) = x_0$, $A(\varepsilon_1) = x_1$, $A(\varepsilon) \in L$ for all $0 \leq \varepsilon \leq \varepsilon_1$.

— A is $C^{k'}$ in ε for some $k' \in \overline{\mathbb{N}}$. This includes the requirement that $\lambda(\varepsilon)$ must be $C^{k'}$, admitting in particular $\lambda = \varepsilon$ and $\lambda = \varepsilon^2$.

Let m be the minimum of k and k', where k is the degree of differentiability of the matter tensor. By the theorem in section 4.3 we know that A generates a family $B(\varepsilon) := (l \circ A)(\varepsilon)$ of solutions of the parametrized constraints with l being C^k, and by the chain rule we conclude that B is a C^m function of ε. It can therefore be expanded in a Taylor series, and for $m = \omega$ this series will even converge for sufficiently small ε. We can insert A and B into the equations and differentiate up to m times with respect to ε, because the equations are C^k by lemma 4.4(a). Putting $\varepsilon = 0$ we obtain a sequence of equations to be solved successively for the Taylor coefficients of B. The quality of this approximation will depend on the choice of A, and in fact A can easily be chosen such as to lead to no useful results at all (for example one can include a sine function of sufficiently high frequency in order to force the radius of convergence to drop below ε_1). This method of expansion should therefore be supplemented by rules for choosing A, preferably by error estimates depending on properties of A.

In contrast, the method of iteration proposed here does not require the introduction of a family of solutions. All that is needed is a suitable Newtonian starting point x_0. The prescription in section 2 then tells us that the sequence

$$z_j := l_j(x_1), \qquad l_j \text{ from (2.22)}, \tag{4.35}$$

is guaranteed to satisfy

$$\|z_j - y_1\|_F \le \frac{a^j}{1-a} Q(x_1), \qquad Q \text{ from (2.24)}, \tag{4.36}$$

and will therefore converge to y_1. Because of $D_2 f(x_0, y_0) = \triangle$ on F_2^4, one merely has to solve successive Poisson equations to generate this sequence.

Presumably, other methods of iteration could also be used. If for example $D_2 f(x_1, y)$ can be easily inverted on V, Newton's method is an obvious choice, and will probably give a better rate of convergence.

5. Conclusions, speculations, and remaining problems

The purpose of this paper was to provide a sound mathematical basis for the investigation and justification of post-Newtonian approximation methods in general relativity.

Starting with general requirements for an approximation method and giving a definition of "post-Newtonian approximation" we were led to a set of requirements to be satisfied by such a method. This included in particular the necessity of proving that unknown solutions can be characterized by specifying data, the justification of the method thereby being reduced to identifying those data for which the method works. I have then tried to point out in which respect I consider previous approaches unsatisfactory.

Probably most people will agree that the requirements I have listed are desirable properties of an approximation method, the only question being whether these conditions can be satisfied. The main body of the paper was therefore concerned with the proof that in the context of a space-like initial value problem it is possible to define a post-Newtonian approximation method for solutions of the constraints which has the desired properties.

In the course of that proof we also obtained the statement that families of data which are C^k in a parameter ε lead to families of solutions of the constraints having the same property, provided the matter tensor is well-behaved. This shows in particular that whatever problems might be responsible for the breakdown of the usual sort of approximation method, they must have their source in the evolution equations. This is easily seen in the present framework: writing out the H^{ab} term in (3.16) after substituting (3.21) we find:

$$H^{ab} = (-\lambda + 4\lambda^2 U)\ddot{Z}^{ab} + 8\lambda^2 \dot{Z}^{ab}{}_{,c} W^c + \triangle Z^{ab} + 4\lambda^2 Z^{cd} Z^{ab}{}_{,cd}. \tag{5.1}$$

The highest t-derivative has a factor λ in front and will disappear at $\lambda = 0$ (compare (3.18)), leading to an additional constraint equation. If we approach the time evolution as a sort of ordinary differential equation for initial data on hypersurfaces of constant t, we have to divide by λ to solve for these t-derivatives: the equations are thus seen to be singular at $\lambda = 0$. In such a situation only some of all possible ε-dependent families of initial data lead to solutions which are regular at $\lambda = 0$, because at least $Z^{ab}(0)$ can not be specified arbitrarily.

Let us assume that we want to obtain a post-Newtonian approximation based on expansion in powers of ε. As described in the introduction, the first step is to look at the families of data belonging to such C^ω families of solutions. By adding a few technical assumptions (the Laplacian is bijective, derivatives commute, and the equations of motion can be solved for the time derivatives of the matter variables), it can be shown that $Z^{ab}(\varepsilon)$ and $\dot{Z}^{cd}(\varepsilon)$ on the initial hypersurface are already completely determined by the matter variables. I conjecture that this property is the only requirement to be fulfilled by the data:

> **Conjecture:** *There is a special kind of initial value problem with additional constraints for Z^{ab} and $\dot{Z}^{cd}$, in which the only free data are the matter variables, and in which a C^ω family (in ε) of initial data will generate a C^ω family of solutions.*

This initial value problem would obviously be an extension of the corresponding Newtonian problem, in which Z^{ab} and $\dot{Z}^{cd}$ do not appear. One might speculate that its solutions would turn out to have in some sense no degrees of freedom for gravitational radiation. If this conjecture should turn out to be true, the breakdown of post-Newtonian methods could from the point of view of the Cauchy problem be entirely attributed to the simultaneous and incompatible assumptions of analyticity in ε and freedom to choose Z^{ab} and $\dot{Z}^{cd}$.

However, as we have seen it is not really necessary to consider families of solutions, and therefore the question of their ε-dependence seems not very fruitful. The approximation method given by the prescription in section 2 requires as input merely the initial data set x_1 of the relativistic solution to be approximated and a suitable Newtonian starting point x_0. This of course does not imply that a similar approach to the evolution equations would circumvent the singularity problem, because the implicit function theorem requires that the *equations* have to be at least C^1 in all the variables, including λ. But this problem rests entirely on the $(\lambda = 0)$-plane in the space of solutions. It might therefore happen that one merely has to find a suitable way to choose x_0 and to move from there into the $(\lambda > 0)$-region and into a sufficiently small neighbourhood of the solution in order to obtain existence, uniqueness, and a sequence converging to the solution. From the point of view of ε-dependent families of solutions this seems to be identical with choosing the first few Taylor coefficients of $Z^{ab}(\varepsilon)$ and $\dot{Z}^{cd}(\varepsilon)$ such that the equations are well-behaved if restricted to this subset. Presumably one can by properly choosing a sufficient number of Taylor coefficients obtain every level of differentiability for the resulting family of solutions. The conjecture above is a statement about what happens if we choose all coefficients suitably. There are indications [24] that a proof along these lines is possible for the equations described here.

But even if we can solve these problems satisfactorily, we cannot expect that post-Newtonian approximations are possible for all relativistic space-times. The range of applicability will be locally limited by the strength of relativistic phenomena, but we also have to expect that it will in general always be bounded in space and in time. The statement derived here for the constraints suggests that the function spaces can be chosen such that for asymptotically flat space-times there is no spatial limitation, and perhaps a corresponding result can be obtained for the evolution equations. However, it seems likely that the resulting space-times will not be as general as we could wish, particularly in the point of gravitational radiation. Of course, Newtonian theory is mainly good at describing the motion of gravitationally interacting bodies, and it is unfair to demand of it a basis for approximating an effect it does not know. Nevertheless, it would be nice to have an approximation method yielding good results for the motion of stars as well as for the gravitational radiation generated by them. This is the motivation for matching a post-Newtonian approximation in the near zone to a post-Minkowskian approximation in the far zone. But there is also another approach which perhaps is even better suited to this problem: the post-Newtonian method based on the characteristic initial value problem and proposed by Winicour [25]. It would seem profitable to attack this problem in a similar spirit as done here for the Cauchy problem.

Acknowledgements

This article is based on a Ph. D. thesis prepared under the supervision of Jürgen Ehlers. It was considerably revised after criticism expressed by Thibault Damour on a previous version (MPA preprint 515, March 1990).

References

[1] S. Chandrasekhar
"The post-Newtonian equations of hydrodynamics in general relativity"
Ap. J. 142, 1488-1512 (1965)

[2] A. Einstein, L. Infeld, B. Hoffmann
"The gravitational equations and the problem of motion"
Ann. Math. 39, 65–100 (1938)

[3] Jürgen Ehlers
"Über den Newtonschen Grenzwert der Einsteinschen Gravitationstheorie"
In: J. Nitsch, J. Pfarr, E.-W. Stachow (eds.): *Grundlagenprobleme der modernen Physik,* Mannheim, Wien, Zürich: Bibliographisches Institut, 1981. Pp. 65–84.

[4] Jürgen Ehlers
"On Limit Relations Between, And Approximative Explanations Of, Physical Theories"
In: R. Barcan Marcus, G.J.W. Dorn, P. Weingartner (eds.): *Logic, Methodology and Philosophy of Science VII*, Amsterdam, etc.: North Holland, 1986. (Studies in Logic and the Foundations of Mathematics; 114). Pp. 387–403.

[5] James L. Anderson, Thomas C. Decanio
"Equations of hydrodynamics in general relativity in the slow motion approximation"
Gen. Rel. Grav. 6, 197–237 (1975)

[6] G. David Kerlick
"Finite Reduced Hydrodynamic Equations in the Slow-Motion Approximation to General Relativity. Part II. Radiation Reaction and Higher-Order Divergent Terms"
Gen. Rel. Grav. 12, 521–543 (1980)

[7] James L. Anderson, Ronald E. Kates, Lawrence S. Kegeles, Richard G. Madonna
"Divergent integrals of post-Newtonian gravity: Nonanalytic terms in the near-zone expansion of a gravitationally radiating system found by matching"
Phys. Rev. D 25, 2038–2048 (1982)

[8] L. Blanchet, T. Damour
"Radiative Gravitational Fields in General Relativity. I. General Structure of the Field Outside the Source"
Phil. Trans. R. Soc. Lond. A 320, 379–430 (1986)

[9] Luc Blanchet, Thibault Damour
"Tail-transported temporal correlations in the dynamics of a gravitating system"
Phys. Rev. D 37, 1410–1435 (1988)

[10] T. Futamase, Bernard F. Schutz
"The Newtonian and post-Newtonian approximations are asymptotic to general relativity"
Phys. Rev. D 28, 2363-2372 (1983)

[11] Lynn H. Loomis, Shlomo Sternberg
Advanced Calculus
Reading/Mass., etc.: Addison-Wesley, 1968

[12] Henri Cartan
Differential Calculus
Paris: Hermann/London: Kershaw, 1971

[13] C.W. Misner, K.S. Thorne, J.A. Wheeler
Gravitation
San Francisco: W.H. Freeman, 1973

[14] M. Cantor
"Spaces of Functions with Asymptotic Conditions on $\mathbb{R}^n$"
Indiana Univ. Math. J. 24(9), 897–902 (1975)

[15] Y. Choquet-Bruhat, D. Christodoulou
"Elliptic systems in $H_{s,\delta}$ spaces on manifolds which are Euclidean at infinity"
Acta Math. 146(1–2), 129–150 (1981)

[16] Robert C. McOwen
"The Behavior of the Laplacian on Weighted Sobolev Spaces"
Comm. Pure Appl. Math. 32, 783–795 (1979)

[17] Klaus Deimling
Nonlinear Functional Analysis
Berlin, etc.: Springer, 1985

[18] V.A. Fock
Theory of Space, Time and Gravitation
London, New York: Pergamon Press, 1960

[19] Jürgen Ehlers
Private communication
1985

[20] Yvonne Bruhat
"The Cauchy Problem"
In: Louis Witten (ed.): *Gravitation: an introduction to current research*, New York, London: John Wiley & Sons, 1962. Pp. 130–168.

[21] Yvonne Choquet-Bruhat, James W. York, Jr.
"The Cauchy Problem"
In: A. Held (ed.): *General Relativity and Gravitation*, Volume 1. New York, London: Plenum Press, 1980. Pp. 99–172.

[22] A. Chaljub-Simon, Y. Choquet-Bruhat
"Global solutions of the Lichnerowicz equation in general relativity on an asymptotically Euclidean complete manifold"
Gen. Rel. Grav. 12(2), 175–185 (1980)

[23] Y. Choquet-Bruhat
"Recent Results on the Cauchy Problem in General Relativity"
In: *Aspetti Matematici Della Teoria Della Relatività*, Roma: Accademia Nazionale Dei Lincei, 1983. (Atti Dei Convegni Lincei; 57). Pp. 17–25.

[24] Oscar Reula
Private communication
June 1990

[25] J. Winicour
"Newtonian gravity on the null cone"
J. Math. Phys. 24, 1193–1198 (1983)