2

THE 3+1 FORMALISM

2.1 Introduction

The Einstein field equations for the gravitational field described in the previous Chapter are written in a fully covariant way, where there is no clear distinction between space and time. This form of the equations is quite natural from the point of view of differential geometry, and has important implications in our understanding of the relationship between space and time. However, there are situations when we would like to recover a more intuitive picture where we can think of the dynamical evolution of the gravitational field in "time". For example, we could be interested in finding the future evolution of the gravitational field associated with an astrophysical system given some appropriate initial data. On the other hand, we might also be interested in studying gravity as a field theory similar to electrodynamics, and define a Hamiltonian formulation that can be used, for example, as the starting point for a study of the quantization of the gravitational field.

There exist several different approaches to the problem of separating the Einstein field equations in a way that allows us to give certain initial data, and from there obtain the subsequent evolution of the gravitational field. Specific formalisms differ in the way in which this separation is carried out. Here we will concentrate on the *3+1 formalism*, where we split spacetime into three-dimensional space on the one hand, and time on the other. The 3+1 formalism is the most commonly used in numerical relativity, but it is certainly not the only one. The two main alternatives to the 3+1 approach are known as the *characteristic formalism* where spacetime is separated into light-cones emanating from a central timelike world-tube, and the *conformal formalism* where we use hyperboloidal slices that are everywhere spacelike but intersect asymptotic null infinity, *plus* a conformal transformation that brings the boundary of spacetime to a finite distance in coordinate space. Both these alternatives will be discussed briefly in Section 2.9.

We should also mention yet another approach that is based on evolving the full four-dimensional spacetime metric directly by simply expanding out the Einstein equations is some adequate coordinate system. Indeed, this was the original approach taken by Hahn and Lindquist in their pioneering work on numerical relativity [158], and has also been recently used by Pretorius with considerable success in the context of the collision of two orbiting black holes [231]. The different formalisms have advantages and disadvantages depending on the specific physical system under consideration.

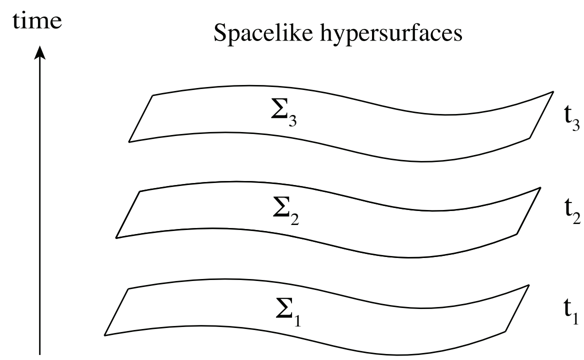

Fig. 2.1: Foliation of spacetime into three-dimensional spacelike hypersurfaces.

In the following sections I will introduce the 3+1 formalism of general relativity. The discussion found here can be seen in more detail in [206] and [305].

2.2 3+1 split of spacetime

In order to study the evolution in time of any physical system the first thing that needs to be done is to formulate such an evolution as an *initial value* or *Cauchy* problem: Given adequate initial (and boundary) conditions, the fundamental equations must predict the future (or past) evolution of the system.

When trying to write Einstein's equations as a Cauchy problem we immediately encounter a stumbling block: The field equations are written in such a way that space and time are treated on an equal footing. This covariance is very important (and quite elegant) from a theoretical point of view, but it does not allow us to think clearly about the evolution of the gravitational field in time. Therefore, the first thing we need to do in order to rewrite Einstein's equations as a Cauchy problem is to split the roles of space and time in a clear way. The formulation of general relativity that results from this splitting is known as the *3+1 formalism.*

Let us start by considering a spacetime with metric $g_{\alpha\beta}$. As already mentioned in Chapter 1, we will always assume that the spacetimes of interest are globally hyperbolic, that is, they have a Cauchy surface. Any globally hyperbolic spacetime can be completely foliated (*i.e.* sliced into three-dimensional cuts) in such a way that each three-dimensional slice is spacelike (see Figure 2.1). We can identify the foliation with the level sets of a parameter t which can then be considered a *universal time function* (but we should keep in mind that t will not necessarily coincide with the proper time of any particular observer). Because of this fact, such a foliation of spacetime into spatial hypersurfaces is often also called a *synchronization.*

Consider now a specific foliation, and take two adjacent hypersurfaces Σ_t and Σ_{t+dt}. The geometry of the region of spacetime contained between these

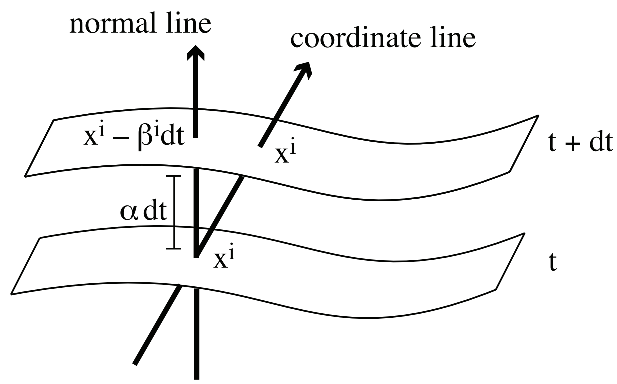

Fig. 2.2: Two adjacent spacelike hypersurfaces. The figure shows the definitions of the lapse function α and the shift vector β^i.

two hypersurfaces can be determined from the following three basic ingredients (see Figure 2.2):

- The three-dimensional metric γ_{ij} $(i, j = 1, 2, 3)$ that measures proper distances within the hypersurface itself:

$$dl^2 = \gamma_{ij}\, dx^i dx^j \, . \tag{2.2.1}$$

- The lapse of proper time $d\tau$ between both hypersurfaces measured by those observers moving along the direction normal to the hypersurfaces (the so-called *normal* or *Eulerian* observers):

$$d\tau = \alpha(t, x^i)\, dt \, . \tag{2.2.2}$$

  Here α is known as the *lapse function*.

- The relative velocity β^i between the Eulerian observers and the lines of constant spatial coordinates:

$$x^i_{t+dt} = x^i_t - \beta^i(t, x^j)\, dt \, , \qquad \text{(for Eulerian observers)} \tag{2.2.3}$$

  The 3-vector β^i is known as the *shift vector*.

Notice that both the way in which spacetime is foliated, and also the way in which the spatial coordinate system propagates from one hypersurface to the next, are not unique. The lapse function α and the shift vector β^i are therefore freely specifiable functions that carry information about our choice of coordinate system, and are known as the *gauge functions*.[23]

[23]The notation for lapse and shift used here is common, but certainly not universal. A frequently used alternative is to denote the lapse function by N, and the shift vector by N^i.

In terms of the functions $\{\alpha, \beta^i, \gamma_{ij}\}$, the metric of spacetime can be easily seen to take the following form:

$$ds^2 = \left(-\alpha^2 + \beta_i \beta^i\right) dt^2 + 2\beta_i \, dt dx^i + \gamma_{ij} \, dx^i dx^j \, , \tag{2.2.4}$$

where we have defined $\beta_i := \gamma_{ij} \, \beta^j$ (from here on we will assume that indices of purely spatial tensors are raised and lowered with the spatial metric γ_{ij}). The last equation is known as the 3+1 split of the metric.

More explicitly we have:

$$g_{\mu\nu} = \begin{pmatrix} -\alpha^2 + \beta_k \beta^k & \beta_i \\ \beta_j & \gamma_{ij} \end{pmatrix} , \tag{2.2.5}$$

$$g^{\mu\nu} = \begin{pmatrix} -1/\alpha^2 & \beta^i/\alpha^2 \\ \beta^j/\alpha^2 & \gamma^{ij} - \beta^i \beta^j/\alpha^2 \end{pmatrix} . \tag{2.2.6}$$

From the above expressions we can also show that the four-dimensional volume element in 3+1 language turns out to be given by

$$\sqrt{-g} = \alpha \sqrt{\gamma} \, , \tag{2.2.7}$$

with g and γ the determinants of $g_{\mu\nu}$ and γ_{ij} respectively.

Consider now the unit normal vector n^μ to the spatial hypersurfaces. It is not difficult to show that, in the coordinate system just introduced, this vector has components given by

$$n^\mu = \left(1/\alpha, -\beta^i/\alpha\right) \, , \qquad n_\mu = (-\alpha, 0) \, . \tag{2.2.8}$$

Note that this unit normal vector corresponds by definition to the 4-velocity of the Eulerian observers.

We can use the normal vector n^μ to introduce the 3+1 quantities in a more formal way that is not tied up with the choice of a coordinate system adapted to the foliation. The spatial metric γ_{ij} is simply defined as the metric induced on each hypersurface Σ by the full spacetime metric $g_{\mu\nu}$:

$$\gamma_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \, . \tag{2.2.9}$$

Notice that written in this way the spatial metric is a full four-dimensional tensor, but when written in the adapted coordinates its time components become trivial. Also, the last expression shows that the spatial metric is nothing more than the projection operator onto the spatial hypersurfaces.

Consider now our global time function t associated with the foliation. The lapse function is defined as

$$\alpha = (-\nabla t \cdot \nabla t)^{-1/2} \, . \tag{2.2.10}$$

(The vector ∇t is clearly timelike because the level sets of t are spacelike). The unit normal vector to the hypersurfaces can then be expressed in terms of α and t as

$$n^\mu = -\alpha \nabla^\mu t \, , \tag{2.2.11}$$

where the minus sign is there to guarantee that $\vec{n}$ is future pointing.

For the definition of the shift vector we start by introducing three scalar functions β^i such that when we move from a given hypersurface to the next following the normal direction, the change in the spatial coordinates is given as before by

$$x^i_{t+dt} = x^i_t - \beta^i dt \ , \tag{2.2.12}$$

from which we can easily find

$$\beta^i = -\alpha \left(\vec{n} \cdot \nabla x^i \right) \ , \tag{2.2.13}$$

Thus defined, the β^i are scalars, but we can use them to define a 4-vector $\vec{\beta}$ by asking for its components in the adapted coordinate system to be given by $\beta^\mu = (0, \beta^i)$. The vector constructed in this way is clearly orthogonal to $\vec{n}$. We can then use the vectors $\vec{n}$ and $\vec{\beta}$ to construct a *time vector* $\vec{t}$ defined as

$$t^\mu := \alpha n^\mu + \beta^\mu \ . \tag{2.2.14}$$

The vector $\vec{t}$ is nothing more than the tangent vector to the *time lines*, *i.e.* the lines of constant spatial coordinates. Notice that, in general, we have $t^\mu \neq \nabla^\mu t$. From the above definition we find that $\vec{t}$ is such that $t^\mu n_\mu = -\alpha$, which implies

$$t^\mu \nabla_\mu t = 1 \ . \tag{2.2.15}$$

We then find that the shift is nothing more than the projection of $\vec{t}$ onto the spatial hypersurface

$$\beta_\mu := \gamma_{\mu\nu} t^\nu \ . \tag{2.2.16}$$

From this we see that we can introduce the shift vector in a completely coordinate-independent way by first choosing a vector field $\vec{t}$ satisfying (2.2.15), and then defining the shift through (2.2.16). It is important to stress the fact that the vector field $\vec{t}$ does not need to be timelike – it can easily be null or even spacelike (we will later see that this situation frequently arises in the case of black hole spacetimes). All we need to ask is that $\vec{t}$ is not tangent to the spatial hypersurfaces, and that it points to the future, which is precisely the content of equation (2.2.15). It might seem strange to allow $\vec{t}$ to be spacelike, since that would correspond to a faster than light motion of the coordinate lines, that is, a *superluminal* or *tachionic* shift. But we must remember that it is not a physical effect but only the coordinate lines that are moving "faster than light", and the coordinates can be chosen freely.

2.3 Extrinsic curvature

When considering the spatial hypersurfaces that constitute the foliation of spacetime, we need to distinguish between the intrinsic curvature of those hypersurfaces coming from their internal geometry, and the *extrinsic* curvature associated with the way in which those hypersurfaces are immersed in four-dimensional spacetime. The intrinsic curvature is given by the three-dimensional Riemann

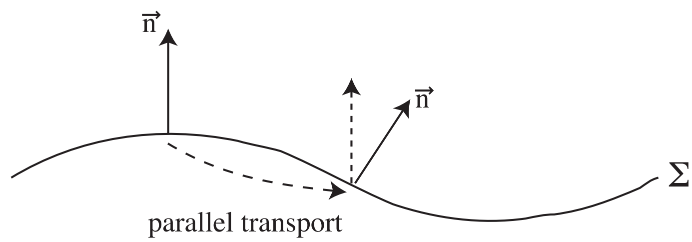

Fig. 2.3: The extrinsic curvature tensor is defined as a measure of the change of the normal vector under parallel transport.

tensor defined in terms of the 3-metric γ_{ij}. The extrinsic curvature, on the other hand, is defined in terms of what happens to the normal vector $\vec{n}$ as it is parallel-transported from one point in the hypersurface to another. In general, we will find that as we parallel transport this vector to a nearby point, the new vector will not be normal the hypersurface anymore. The *extrinsic curvature tensor* $K_{\alpha\beta}$ is a measure of the change of the normal vector under such parallel transport (see Figure 2.3).

In order to define the extrinsic curvature, we need to introduce the projection operator P^α_β onto the spatial hypersurfaces:

$$P^\alpha_\beta := \delta^\alpha_\beta + n^\alpha n_\beta \,, \tag{2.3.1}$$

which as we have seen is in fact nothing more than the induced spatial metric, $P_{\alpha\beta} = \gamma_{\alpha\beta}$. Using this projection operator, the extrinsic curvature tensor is defined as:

$$K_{\mu\nu} := -P^\alpha_\mu \, \nabla_\alpha n_\nu = -\left(\nabla_\mu n_\nu + n_\mu n^\alpha \nabla_\alpha n_\nu\right) \,, \tag{2.3.2}$$

As defined above, the tensor $K_{\mu\nu}$ is clearly a purely spatial tensor, that is, $n^\mu K_{\mu\nu} = n^\nu K_{\mu\nu} = 0$. This means, in particular, that in a coordinate system adapted to the foliation we will have $K^{00} = K^{0i} = 0$ (though in general we find that K_{00} and K_{0i} are not zero). Because of this, we will usually only consider the spatial components of K_{ij}. Moreover, the tensor $K_{\mu\nu}$ also turns out to be symmetric:

$$K_{\mu\nu} = K_{\nu\mu} \,. \tag{2.3.3}$$

A couple of remarks are important regarding the definition of $K_{\mu\nu}$. First, notice that the projection of $\nabla_\mu n_\nu$ is crucial in order to make $K_{\mu\nu}$ purely spatial. We could argue that because n^μ is unitary, its gradient is necessarily orthogonal to it. This is of course true in the sense that $n^\nu \nabla_\mu n_\nu = 0$, but $\nabla_\mu n_\nu$ is in general not symmetric and $n^\nu \nabla_\nu n_\mu \neq 0$ unless the normal lines are geodesic (which is not always the case). Let us now consider the symmetry of $K_{\mu\nu}$. As just mentioned $\nabla_\mu n_\nu$ is not in general symmetric even though n^μ is hypersurface orthogonal. The reason for this is that n^α is a unitary vector and thus in general is not equal to the gradient of the time function t, except when the lapse is unity. However, once we project onto the hypersurface, it turns out that $P_\mu^\alpha \nabla_\alpha n_\nu$ is indeed symmetric (the non-symmetry of $\nabla_\mu n_\nu$ has to do with the lapse, which is not intrinsic to the hypersurface). In order to see this, consider the congruence of timelike *geodesics* orthogonal to Σ, with unit tangent vector $\vec{\xi}$. In the neighborhood of Σ consider a new foliation of spacetime given by a time function $\tilde{t}$ such that $\xi_\mu = \nabla_\mu \tilde{t}$. Since $\vec{\xi}$ is the gradient of a scalar function we clearly will have $\nabla_\mu \xi_\nu = \nabla_\nu \xi_\mu$. Moreover, $\nabla_\mu \xi_\nu$ will be purely spatial without the need to project it since $\vec{\xi}$ is tangent to the timelike geodesics. Now, the vector field $\vec{n}$ will generally not coincide with $\vec{\xi}$ outside of Σ, but it will coincide within Σ, so its derivatives along directions tangential to Σ must be equal to those of $\vec{\xi}$, that is

$$P_\mu^\alpha \nabla_\alpha n_\nu = P_\mu^\alpha \nabla_\alpha \xi_\nu = \nabla_\mu \xi_\nu = \nabla_\nu \xi_\mu = P_\nu^\alpha \nabla_\alpha \xi_\mu = P_\nu^\alpha \nabla_\alpha n_\mu \ , \tag{2.3.4}$$

which proves the symmetry of $K_{\mu\nu}$.

Notice, in particular, that the symmetry of the projected part of $\nabla_\mu n_\nu$ also implies that when we contract it with a purely spatial antisymmetric tensor we will obtain zero, that is

$$T^{\mu\nu} \nabla_\mu n_\nu = 0 \ , \tag{2.3.5}$$

for all $T_{\mu\nu}$ such that $T_{\mu\nu} = -T_{\nu\mu}$, $n^\mu T_{\mu\nu} = 0$.

From its definition above, it is in fact not difficult to show that the extrinsic curvature $K_{\mu\nu}$ can be written in an entirely equivalent way as the Lie derivative of the spatial metric along the normal direction

$$K_{\mu\nu} = -\frac{1}{2} \, \pounds_{\vec{n}} \, \gamma_{\mu\nu} \ . \tag{2.3.6}$$

That this is so can be easily seen from the definition of the Lie derivative:

$$\begin{aligned}
\pounds_{\vec{n}} \, \gamma_{\mu\nu} &= n^\alpha \nabla_\alpha \gamma_{\mu\nu} + \gamma_{\mu\alpha} \nabla_\nu n^\alpha + \gamma_{\nu\alpha} \nabla_\mu n^\alpha \\
&= n^\alpha \nabla_\alpha \left(n_\mu n_\nu \right) + g_{\mu\alpha} \nabla_\nu n^\alpha + g_{\nu\alpha} \nabla_\mu n^\alpha \\
&= n^\alpha n_\mu \nabla_\alpha n_\nu + n^\alpha n_\nu \nabla_\alpha n_\mu + \nabla_\nu n_\mu + \nabla_\mu n_\nu \\
&= \left(\gamma_\mu^\alpha - g_\mu^\alpha \right) \nabla_\alpha n_\nu + \left(\gamma_\nu^\alpha - g_\nu^\alpha \right) \nabla_\alpha n_\mu + \nabla_\nu n_\mu + \nabla_\mu n_\nu \\
&= \gamma_\mu^\alpha \nabla_\alpha n_\nu + \gamma_\nu^\alpha \nabla_\alpha n_\mu = -2 K_{\mu\nu} \ ,
\end{aligned} \tag{2.3.7}$$

where we have used the fact that the covariant derivative of $g_{\mu\nu}$ is zero and also that $n^\alpha \nabla_\mu n_\alpha = 0$. We then find that the extrinsic curvature is essentially the "velocity" of the spatial metric as seen by the Eulerian observers. Notice that the extrinsic curvature depends only on the behavior of $\vec{n}$ within the slice Σ – it is therefore a geometric property of the slice itself.

Now, since $\vec{n}$ is normal to the hypersurface, it turns out that for any scalar function ϕ we have

$$\pounds_{\vec{n}} \, \gamma_{\mu\nu} = \frac{1}{\phi} \, \pounds_{\phi \vec{n}} \, \gamma_{\mu\nu} \ . \tag{2.3.8}$$

If, in particular, we take as our scalar function the lapse, we find that

$$K_{\mu\nu} = -\frac{1}{2\alpha} \pounds_{\alpha\vec{n}} \gamma_{\mu\nu} = -\frac{1}{2\alpha} \left(\pounds_{\vec{t}} - \pounds_{\vec{\beta}} \right) \gamma_{\mu\nu} \; , \tag{2.3.9}$$

which implies

$$\left(\pounds_{\vec{t}} - \pounds_{\vec{\beta}} \right) \gamma_{\mu\nu} = -2\alpha K_{\mu\nu} \; . \tag{2.3.10}$$

Concentrating now only on the spatial components, and remembering that in the adapted coordinate system we have $\pounds_{\vec{t}} = \partial_t$, we finally find

$$\partial_t \gamma_{ij} - \pounds_{\vec{\beta}} \gamma_{ij} = -2\alpha K_{ij} \; . \tag{2.3.11}$$

The last expression can also be rewritten as

$$\partial_t \gamma_{ij} = -2\alpha K_{ij} + D_i \beta_j + D_j \beta_i \; , \tag{2.3.12}$$

where here D_i represents the three-dimensional covariant derivative, that is, the one associated with the 3-metric γ_{ij}, which is in fact nothing more than the projection of the full four-dimensional covariant derivative: $D_\mu := P^\alpha_\mu \nabla_\alpha$.

This brings us half-way to our goal of writing Einstein's equations as a Cauchy problem: We already have an evolution equation for the spatial metric γ_{ij}. In order to close the system we still need an evolution equation for K_{ij}. It is important to notice that until now we have only worked with purely geometric concepts, and we have not used the Einstein field equations at all. It is precisely from the field equations that we will obtain the evolution equations for K_{ij}. In other words, the evolution equation (2.3.11) for the 3-metric is purely kinematic, while the dynamics of the system will be contained in the evolution equations for K_{ij}.

2.4 The Einstein constraints

The true dynamics of the gravitational field are contained in the Einstein field equations. In order to proceed further, we need to rewrite these equations in 3+1 language. This can be done by considering contractions of these equations with the normal vector $\vec{n}$ and with the projector operator onto the hypersurface P^μ_ν. We will proceed with this in two stages. In this Section we will consider those contractions involving the normal vector, and will leave the rest of the equations until the next Section.

The starting point is to express the four-dimensional Riemann curvature tensor $R^\alpha_{\beta\mu\nu}$ in terms of the intrinsic three-dimensional Riemann tensor of the hypersurface itself ${}^{(3)}R^\alpha_{\beta\mu\nu}$, and the extrinsic curvature tensor $K_{\mu\nu}$. The derivation of these relations is straightforward but rather long, so we will just state the final results here (the interested reader can see *e.g.* [295]). The full projection of the Riemann tensor onto the spatial hypersurfaces turns out to be given by the so-called *Gauss–Codazzi* equations

$$P^\delta_\alpha P^\kappa_\beta P^\lambda_\mu P^\sigma_\nu \, R_{\delta\kappa\lambda\sigma} = {}^{(3)}R_{\alpha\beta\mu\nu} + K_{\alpha\mu} K_{\beta\nu} - K_{\alpha\nu} K_{\beta\mu} \; , \tag{2.4.1}$$

Similarly, the projection onto the hypersurfaces of the Riemann tensor contracted once with the normal vector results in the *Codazzi–Mainardi* equations

$$P^\delta_\alpha P^\kappa_\beta P^\lambda_\mu n^\nu \, R_{\delta\kappa\lambda\nu} = D_\beta K_{\alpha\mu} - D_\alpha K_{\beta\mu} \; , \tag{2.4.2}$$

where as before $D_\mu = P^\alpha_\mu \nabla_\alpha$.

In order to start rewriting the Einstein field equations in 3+1 language, notice first that

$$\begin{aligned} P^{\alpha\mu} P^{\beta\nu} R_{\alpha\beta\mu\nu} &= \left(g^{\alpha\mu} + n^\alpha n^\mu\right)\left(g^{\beta\nu} + n^\beta n^\nu\right) R_{\alpha\beta\mu\nu} \\ &= R + 2n^\mu n^\nu R_{\mu\nu} \\ &= 2n^\mu n^\nu G_{\mu\nu} \; , \end{aligned} \tag{2.4.3}$$

with $G_{\mu\nu}$ the Einstein tensor. On the other hand, the Gauss–Codazzi relations imply that

$$P^{\alpha\mu} P^{\beta\nu} R_{\alpha\beta\mu\nu} = {}^{(3)}R + K^2 - K_{\mu\nu} K^{\mu\nu} \tag{2.4.4}$$

where $K := K^\mu_\mu$ is the trace of the extrinsic curvature. We then find

$$2 \, n^\mu n^\nu G_{\mu\nu} = {}^{(3)}R + K^2 - K_{\mu\nu} K^{\mu\nu} \; , \tag{2.4.5}$$

which through the Einstein equations becomes

$${}^{(3)}R + K^2 - K_{\mu\nu} K^{\mu\nu} = 16\pi\rho \; , \tag{2.4.6}$$

where we have defined the quantity $\rho := n^\mu n^\nu T_{\mu\nu}$ that corresponds to the local energy density as measured by the Eulerian observers. Notice that this equation involves no explicit time derivatives, so it is not an evolution equation but rather a constraint that must be satisfied at all times. This equation is known as the *Hamiltonian* or *energy* constraint.

Consider now the mixed contraction of the Einstein tensor. We find,

$$P^{\alpha\mu} n^\nu G_{\mu\nu} = P^{\alpha\mu} n^\nu R_{\mu\nu} \; . \tag{2.4.7}$$

The Codazzi–Mainardi equations then imply

$$\gamma^{\alpha\mu} n^\nu G_{\mu\nu} = D^\alpha K - D_\mu K^{\alpha\mu} \; , \tag{2.4.8}$$

which again through the field equations becomes

$$D_\mu \left(K^{\alpha\mu} - \gamma^{\alpha\mu} K\right) = 8\pi j^\alpha \; , \tag{2.4.9}$$

where now $j^\alpha := -P^{\alpha\mu} n^\nu T_{\mu\nu}$ corresponds to the momentum density as measured by the Eulerian observers. Notice that there are in fact three equations in the last expression, as the index α is free, but the case $\alpha = 0$ is trivial. As before, there are no time derivatives in these equations so they are also constraints. They are known as the *momentum* constraints.

In a coordinate system adapted to the foliation, the Hamiltonian and momentum constraints take the final form

$$^{(3)}R + K^2 - K_{ij}K^{ij} = 16\pi\rho \; , \qquad (2.4.10)$$

$$D_j \left(K^{ij} - \gamma^{ij} K \right) = 8\pi j^i \; , \qquad (2.4.11)$$

with

$$\rho := n^\mu n^\nu T_{\mu\nu} \; , \qquad j^i := -P^{i\mu} n^\nu T_{\mu\nu} \; . \qquad (2.4.12)$$

It is important to notice that the constraints not only do not involve time derivatives, but they are also completely independent of the gauge functions α and β^i. This indicates that the constraints are relations that refer purely to a given hypersurface.

Notice that having a set of constraint equations is not a feature of general relativity alone. In electrodynamics we have the Maxwell equations which in three-dimensional vector calculus notation, and in Gaussian units, take the form

$$\nabla \cdot \mathbf{E} = 4\pi\rho \; , \qquad \nabla \cdot \mathbf{B} = 0 \; , \qquad (2.4.13)$$

$$\partial_t \mathbf{E} = \nabla \times \mathbf{B} - 4\pi \mathbf{j} \; , \qquad \partial_t \mathbf{B} = -\nabla \times \mathbf{E} \; , \qquad (2.4.14)$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields respectively, ρ is the charge density and $\mathbf{j}$ the current density (here ∇ stands for the ordinary flat space gradient operator and should not be confused with a four-dimensional covariant derivative). The first two equations involving the divergence of the electric and magnetic fields do not involve time derivatives, so they are in fact constraints, just as in general relativity. The main difference is that Maxwell's theory has only two constraint equations, while general relativity has four (one Hamiltonian constraint and three momentum constraints). The remaining two Maxwell equations (or rather six since they are vector-valued equations) are the true evolution equations for electrodynamics. The corresponding equations for gravity will be derived in the next Section.

The existence of the constraints implies, in particular, that in the 3+1 formulation it is not possible to specify arbitrarily all 12 dynamical quantities $\{\gamma_{ij}, K_{ij}\}$ as initial conditions. The initial data must already satisfy the constraints, otherwise we will not be solving Einstein's equations. We will come back to this issue in Chapter 3 when we discuss how to find initial data. The constraints also play other important roles. For example, they are crucial in the Hamiltonian formulation of general relativity (see Section 2.7). They are also very important in the study of the well-posedness of the system of evolution equations, something we will have a chance to discuss briefly in Section 2.8 and also in Chapter 5.

2.5 The ADM evolution equations

The Hamiltonian and momentum constraints give us four of the ten independent Einstein field equations, and as we have seen they do not correspond to the evolution equations of the gravitational field, but rather to relations between the dynamical variables that must be satisfied at all times. The evolution of the gravitational field is contained in the remaining six field equations.

In order to find these equations we still need the projection onto the hypersurfaces of the Riemann tensor contracted twice with the normal vector. This will give us the last six independent components of Riemann (the Gauss–Codazzi and Gauss–Mainardi equations give us 14 components of Riemann). These projections turn out to be given by

$$P^\delta_\mu P^\kappa_\nu n^\lambda n^\sigma R_{\delta\lambda\kappa\sigma} = £_{\vec{n}} K_{\mu\nu} + K_{\mu\lambda} K^\lambda_\nu + \frac{1}{\alpha} D_\mu D_\nu \alpha \; . \tag{2.5.1}$$

The first thing to notice is that these relations do involve the lapse function α. Also, they make reference to the Lie derivative of the extrinsic curvature along the normal direction, which clearly corresponds to evolution in time.

Now, from the Gauss–Codazzi equations (2.4.1) we also find

$$P^\delta_\mu P^\kappa_\nu \left(n^\lambda n^\sigma R_{\delta\lambda\kappa\sigma} + R_{\delta\kappa} \right) = {}^{(3)}R_{\mu\nu} + K K_{\mu\nu} - K_{\mu\lambda} K^\lambda_\nu \; , \tag{2.5.2}$$

which, together with (2.5.1), implies

$$\begin{aligned} £_{\vec{t}} K_{\mu\nu} - £_{\vec{\beta}} K_{\mu\nu} &= - D_\mu D_\nu \alpha \\ &\quad + \alpha \left(- P^\delta_\mu P^\kappa_\nu R_{\delta\kappa} + {}^{(3)}R_{\mu\nu} + K K_{\mu\nu} - 2 K_{\mu\lambda} K^\lambda_\nu \right) \; , \end{aligned} \tag{2.5.3}$$

Using now the Einstein equations written in terms of $R_{\mu\nu}$, equations (1.13.4), we find

$$\begin{aligned} £_{\vec{t}} K_{\mu\nu} - £_{\vec{\beta}} K_{\mu\nu} &= - D_\mu D_\nu \alpha + \alpha \left[{}^{(3)}R_{\mu\nu} + K K_{\mu\nu} - 2 K_{\mu\lambda} K^\lambda_\nu \right] \\ &\quad + 4\pi\alpha \left[\gamma_{\mu\nu} \left(S - \rho \right) - 2 S_{\mu\nu} \right] \; , \end{aligned} \tag{2.5.4}$$

where ρ is the same as before, and where we have defined $S_{\mu\nu} := P^\alpha_\mu P^\beta_\nu T_{\alpha\beta}$ as the spatial stress tensor measured by the Eulerian observers (with $S := S^\mu_\mu$). Concentrating on the spatial components and again writing $£_{\vec{t}} = \partial_t$, the last expression becomes

$$\begin{aligned} \partial_t K_{ij} - £_{\vec{\beta}} K_{ij} &= - D_i D_j \alpha + \alpha \left[{}^{(3)}R_{ij} + K K_{ij} - 2 K_{ik} K^k_j \right] \\ &\quad + 4\pi\alpha \left[\gamma_{ij} \left(S - \rho \right) - 2 S_{ij} \right] \; , \end{aligned} \tag{2.5.5}$$

or expanding the Lie derivative along the shift vector

$$\begin{aligned} \partial_t K_{ij} &= \beta^k \partial_k K_{ij} + K_{ki} \partial_j \beta^k + K_{kj} \partial_i \beta^k - D_i D_j \alpha \\ &\quad + \alpha \left[{}^{(3)}R_{ij} + K K_{ij} - 2 K_{ik} K^k_j \right] + 4\pi\alpha \left[\gamma_{ij} \left(S - \rho \right) - 2 S_{ij} \right] \; . \end{aligned} \tag{2.5.6}$$

These equations give us the dynamical evolution of the six independent components of the extrinsic curvature K_{ij}. Together with equations (2.3.11) for the evolution of the spatial metric they finally allow us to write down the field equations for general relativity as a Cauchy problem. It is important to notice that we do not have evolution equations for the gauge quantities α and β^i. As we have mentioned before, these quantities represent our coordinate freedom and can therefore be chosen freely.

The evolution equations (2.5.5) are known in the numerical relativity community as the Arnowitt–Deser–Misner (ADM) equations. However, as written above, these equations are in fact not in the form originally derived by ADM [31], but they are instead a non-trivial rewriting due to York [305]. It is important to mention exactly what the difference is between the original ADM equations and the ADM equations *à la* York, which we will call from now on *standard* ADM. The two groups of equations differ in two main aspects. In the first place, the original ADM variables are the spatial metric γ_{ij} and its canonical conjugate momentum π_{ij} coming from the Hamiltonian formulation of general relativity (see the following Section), and which is related to the extrinsic curvature as [24]

$$K_{ij} = -\frac{1}{\sqrt{\gamma}} \left(\pi_{ij} - \frac{1}{2}\, \gamma_{ij}\, \pi \right) \;, \tag{2.5.7}$$

with $\pi = \pi^i_i$ and γ the determinant of γ_{ij}. This change of variables is, of course, a rather minor detail. However, even if we rewrite the original ADM evolution equations in terms of K_{ij}, they still differ from (2.5.5) and have the form

$$\begin{aligned} \partial_t K_{ij} - \pounds_{\vec{\beta}} K_{ij} = {} & -D_i D_j \alpha + \alpha \left[{}^{(3)}R_{ij} + K K_{ij} - 2 K_{ik} K^k_j \right] \\ & + 4\pi\alpha \left[\gamma_{ij} \left(S - \rho \right) - 2 S_{ij} \right] - \frac{\alpha \gamma_{ij}}{2}\, \mathcal{H} \,, \end{aligned} \tag{2.5.8}$$

with $\mathcal{H}$ the Hamiltonian constraint (2.4.10) written as:

$$\mathcal{H} := \frac{1}{2} \left({}^{(3)}R + K^2 - K_{ij} K^{ij} \right) - 8\pi\rho = 0 \;, \tag{2.5.9}$$

and where the factor $1/2$ in the definition of $\mathcal{H}$ is there for later convenience.

The difference between the ADM and York evolution equations can be traced back to the fact that the version of ADM comes from the field equations written in terms of the Einstein tensor $G_{\mu\nu}$, whereas the version of York was derived instead from the field equations written in terms of the Ricci tensor $R_{\mu\nu}$. It is clear that both sets of evolution equations for K_{ij} are physically equivalent since they only differ by the addition of a term proportional to the Hamiltonian constraint, which must vanish for any physical solution. However, the different evolution equations for K_{ij} are not *mathematically* equivalent. There are basically two reasons why this is so:

[24] One must remember that the original goal of ADM was to write a Hamiltonian formulation for general relativity that could be used as a basis for quantum gravity, and not a system of evolution equations for dynamical simulations.

1. In the first place, the space of solutions to the evolution equations is different in both cases, and only coincides for physical solutions, that is, those that satisfy the constraints. In other words, both systems are only equivalent in a subset of the full space of solutions. This subset is called the *constraint hypersurface* (but notice that this is not a hypersurface in spacetime, but instead a hypersurface in the space of solutions to the evolution equations). Of course, we could always argue that since in the end we are only interested in physical solutions, this distinction is irrelevant. This is strictly true only if we can solve the equations exactly. But in the case of numerical solutions there will always be some error that will take us out of the constraint hypersurface, and the issue then becomes not only relevant but crucial: If we move slightly off the constraint hypersurface, does the subsequent evolution remain close to it, or does it diverge rapidly away from it?
2. The second reason why both systems of evolution equations differ mathematically is related to the last point and is of greater importance. Since the Hamiltonian constraint has second derivatives of the spatial metric (hidden inside the Ricci scalar), then by adding a multiple of it to the evolution equations we are in fact altering the very structure of the differential equations.

These types of considerations take us to a fundamental observation that has today become one of the most active areas of research associated with numerical relativity: The 3+1 evolution equations are highly non-unique since we can always add to them arbitrary multiples of the constraints. The different systems of evolution equations will still coincide in the physical solutions, but might differ dramatically in their mathematical properties, and particularly in the way in which they react to small violations of the constraints (inevitable numerically). This observation is crucial, and we will come back to it both in Section 2.8, and in Chapter 5.

A final consideration about the 3+1 evolution equations has to do with the propagation of the constraints: If the constraints are satisfied initially, do they remain satisfied during the evolution? The answer to this question is, not surprisingly, yes, but it is interesting to see how this comes about. In fact it is through the Bianchi identities that the propagation of the constraints during the evolution is guaranteed. To see this, we will follow an analysis due to Frittelli [136], and define the following projections of the Einstein field equations

$$\mathcal{H} := n^\alpha n^\beta G_{\alpha\beta} - 8\pi\rho \,, \tag{2.5.10}$$

$$\mathcal{M}_\mu := -n^\alpha P^\beta_\mu G_{\alpha\beta} - 8\pi j_\mu \,, \tag{2.5.11}$$

$$\mathcal{E}_{\mu\nu} := P^\alpha_\mu P^\beta_\nu G_{\alpha\beta} - 8\pi S_{\mu\nu} \,, \tag{2.5.12}$$

with ρ, j^μ and $S_{\mu\nu}$ defined as before in terms of the stress-energy tensor $T_{\mu\nu}$. Notice that here $\mathcal{H} = 0$ corresponds precisely to the Hamiltonian constraint (2.5.9) (with the correct 1/2 factor), while $\mathcal{M}^\mu = 0$ corresponds to the momentum constraints. On the other hand, $\mathcal{E}_{\mu\nu} = 0$ reduces to the original ADM evolution equations (multiplied by 2). An important observation is that in this context York's version of the evolution equations in fact corresponds to $\mathcal{E}_{\mu\nu} - \gamma_{\mu\nu}\mathcal{H} = 0$. The Einstein field equations in terms of these quantities can then be written as

$$G_{\mu\nu} - 8\pi T_{\mu\nu} = \mathcal{E}_{\mu\nu} + n_\mu \mathcal{M}_\nu + n_\nu \mathcal{M}_\mu + n_\mu n_\nu \mathcal{H} = 0 \, , \tag{2.5.13}$$

and the (twice contracted) Bianchi identities imply

$$\nabla^\mu \left(\mathcal{E}_{\mu\nu} + n_\mu \mathcal{M}_\nu + n_\nu \mathcal{M}_\mu + n_\mu n_\nu \mathcal{H} \right) = 0 \, . \tag{2.5.14}$$

By taking the normal projection and the projection onto the hypersurface of these identities and rearranging terms, we obtain the following system of evolution equations for the constraints

$$n^\nu \nabla_\nu \mathcal{H} = -D^\nu \mathcal{M}_\nu - \mathcal{E}_{\mu\nu} D^\mu n^\nu + \mathcal{L}_{\mathcal{H}}(\mathcal{H}, \mathcal{M}_\sigma) \, , \tag{2.5.15}$$

$$n^\nu \nabla_\nu \mathcal{M}_\mu = -D^\nu \mathcal{E}_{\mu\nu} - \mathcal{E}_{\mu\nu} n^\lambda \nabla_\lambda n^\nu + \mathcal{L}_{\mathcal{M}_\mu}(\mathcal{H}, \mathcal{M}_\sigma) \, , \tag{2.5.16}$$

where the $\mathcal{L}$'s are shorthand for terms proportional to $\mathcal{H}$ and $\mathcal{M}_\mu$ that have no derivatives of these quantities. That these are evolution equations for the constraints can be seen from the fact that the terms on the left hand side are derivatives along the normal direction, *i.e.* out of the hypersurface. Notice that, if we have initial data such that $\mathcal{H} = \mathcal{M}_\mu = 0$, and we use the ADM evolution equations $\mathcal{E}_{\mu\nu} = 0$, the above equations guarantee that on the next hypersurface the constraints will still vanish. This proves that the constraints will remain satisfied. If, on the other hand, we use York's evolution equations $\mathcal{E}_{\mu\nu} = \gamma_{\mu\nu}\mathcal{H}$, the same result clearly follows. In fact, it is clear that we could take $\mathcal{E}_{\mu\nu}$ to be equal to any combination of constraints.

There is, however, an important difference in the structure of the constraint evolution equations when taking either the ADM or York's version of the 3+1 evolution equations, and we will come back to it later, in Chapter 5. For the moment, it is sufficient to mention that the constraint evolution equations are mathematically well-posed for York's system, but they are not well posed for the original ADM system, so York's system should be preferred.

2.6 Free versus constrained evolution

In the previous Sections we have separated the Einstein field equations into evolution equations and constraints equations. As already mentioned, the Bianchi identities guarantee that if the constraints are initially satisfied they will remain satisfied during the subsequent evolution. This is, however, an exact statement, and since here we are ultimately interested in numerical simulations it is important to consider the problem of how to evolve the geometric quantities γ_{ij} and K_{ij} while keeping the constraints satisfied. Ideally, we would like to have some discretized version of the 3+1 evolution equations and constraints that would guarantee that the discrete constraints remain satisfied during evolution. Unfortunately, such a discretized form of the 3+1 equations is not known to exist at this time. We must then live with the fact that not all 10 Einstein equations will remain satisfied at the discrete level during a numerical simulation. Of course, we expect that a good numerical implementation would be such that as we approach the continuum limit we would recover a solution of the full set of equations.

In practice, we can take two different approaches to the problem of choosing which set of equations to solve numerically. The first approach is known as *free evolution*, and corresponds to the case when we start with a solution of the constraint equations as initial data (see Chapter 3), and later advances in time by solving all 12 evolution equations for γ_{ij} and K_{ij}. The constraints are then only monitored to see how much they are violated during evolution, which gives a rough idea of the accuracy of the simulation. Alternatively, we can choose to solve some or all of the constraint equations at each time step for a specific subset of the components of the metric and extrinsic curvature, and evolve the remaining components using the evolution equations. This second approach is known as *constrained evolution*.

Constrained evolution is in fact ideal in situations with high degree of symmetry, like spherical symmetry for example, but is much harder to use in the case of fully three-dimensional systems. Also, the mathematical properties of a constrained scheme are much more difficult to analyze in the sense of studying the well-posedness of the system of equations (see Chapter 5). And finally, since the constraints involve elliptic equations, a constrained scheme is also far slower to solve numerically in three dimensions than a free evolution scheme. Because of all these reasons, in the remainder of the book we will always assume that we are using a free evolution scheme.

For completeness, we should mention a third alternative known as *constrained metric evolution*. In this approach, we choose some extra condition on the metric tensor, such as for example conformal flatness, and impose this condition during the whole evolution, thus simplifying considerably the equations to be solved. Such an approach has been used with some success by Wilson and Mathews in hydrodynamical simulations [300]. However, imposing extra conditions on the metric is not in general compatible with the Einstein field equations, and the results from such simulations should therefore be regarded just as approximations even in the continuum limit. For example, the condition of conformal flatness essentially eliminates the gravitational wave degrees of freedom. Whether such an approximation is good or not will depend on the specific physical system under study, and the physical information we wish to extract from the simulation.

2.7 Hamiltonian formulation

A field theoretical formulation of general relativity starts from the Hilbert Lagrangian which was already introduced in the previous Chapter:

$$L = R \,, \tag{2.7.1}$$

with R the Ricci scalar of the spacetime. As already mentioned, from a variational principle we can obtain the field equations taking this Lagrangian as the starting point.

The Lagrangian formulation of a field theory takes a covariant approach. In the first place, the Lagrangian itself must be a scalar function, and also the field equations derived from the variational principle come out in fully covariant form. A different approach is to take instead a Hamiltonian formulation of the theory. This approach has important advantages, and in particular is the starting point of quantum field theory. However, a Hamiltonian formulation requires a clear distinction to be made between space and time, so it is therefore not covariant. In field theories other than general relativity, and particularly when working on a flat spacetime background, there is already a natural way in which space and time can be split. In general relativity, on the other hand, no such natural splitting exists. However, we can take the 3+1 perspective and use this splitting as a basis to construct a Hamiltonian formulation of the theory. Of course, we can not interpret the time function t directly as a measure of the proper time of any given observer, since the spacetime metric needed in order to do this is the unknown dynamical variable under study.

The first step in a Hamiltonian formulation is to identify the configuration variables that describe the state of the field at any given time. For this purpose we will choose the spatial metric variables γ_{ij}, together with the lapse α and the co-variant shift vector β_i. We now need to rewrite the Hilbert Lagrangian in terms of these quantities and their derivatives. Notice that, from the definition of the Einstein tensor we have

$$n^\mu n^\nu G_{\mu\nu} = n^\mu n^\nu R_{\mu\nu} + \frac{1}{2}\, R \quad \Rightarrow \quad R = 2 \left(n^\mu n^\nu G_{\mu\nu} - n^\mu n^\nu R_{\mu\nu} \right) \, . \tag{2.7.2}$$

The first term on the right hand side of the last equation was already obtained from the Gauss–Codazzi relations and is given by equation (2.4.5). For the second term we use the Ricci identity that relates the commutator of covariant derivatives to the Riemann tensor (equation (1.9.3)):

$$\begin{aligned} n^\mu n^\nu R_{\mu\nu} &= n^\mu n^\nu R^\lambda{}_{\mu\lambda\nu} \\ &= n^\nu \left(\nabla_\lambda \nabla_\nu n^\lambda - \nabla_\nu \nabla_\lambda n^\lambda \right) \\ &= \nabla_\lambda \left(n^\nu \nabla_\nu n^\lambda \right) - \nabla_\nu \left(n^\nu \nabla_\lambda n^\lambda \right) - \nabla_\lambda n^\nu \nabla_\nu n^\lambda + \nabla_\nu n^\nu \nabla_\lambda n^\lambda \\ &= \nabla_\lambda \left(n^\nu \nabla_\nu n^\lambda \right) - \nabla_\nu \left(n^\nu \nabla_\lambda n^\lambda \right) - K_{\lambda\nu} K^{\lambda\nu} + K^2 \, . \end{aligned} \tag{2.7.3}$$

In the previous expression, we have directly identified some terms with the extrinsic curvature even though no projection operator is present. We can readily verify that the contractions in those expressions guarantee that the result follows. The Ricci scalar then takes the form

$$R = {}^{(3)}R + K_{\mu\nu} K^{\mu\nu} - K^2 - 2\, \nabla_\lambda \left(n^\nu \nabla_\nu n^\lambda - n^\lambda \nabla_\nu n^\nu \right) \, . \tag{2.7.4}$$

The last term in this equation is a total divergence, and since in the end we are only interested in the action S which is an integral of the Lagrangian over a given volume Ω, this term can be transformed into an integral over the boundary of Ω and can therefore be ignored. The Lagrangian of general relativity in 3+1 language can then be written as

$$L = {}^{(3)}R + K_{ij}K^{ij} - K^2 \; . \tag{2.7.5}$$

Notice that L has a similar structure to that of the Hamiltonian constraint, but with the sign of the quadratic terms in the extrinsic curvature reversed.

To obtain the Lagrangian density $\mathcal{L}$ we first need to remember that the four-dimensional volume element in 3+1 adapted coordinates is given by $\sqrt{-g} = \alpha\sqrt{\gamma}$, so the Lagrangian density takes the form

$$\mathcal{L} = \alpha\sqrt{\gamma}\left({}^{(3)}R + K_{ij}K^{ij} - K^2\right) \; . \tag{2.7.6}$$

The canonical momenta conjugate to the dynamical field variables are defined now as derivatives of the Lagrangian density with respect to the velocities of those fields. For the spatial metric we have the following conjugate momenta

$$\pi^{ij} := \frac{\partial \mathcal{L}}{\partial \dot{\gamma}_{ij}} \; , \tag{2.7.7}$$

which, using the fact that $\dot{\gamma}_{ij} = -2\alpha K_{ij} + \pounds_{\vec{\beta}}\gamma_{ij}$, can be reduced to

$$\pi^{ij} = -\sqrt{\gamma}\left(K^{ij} - \gamma^{ij}K\right) \; . \tag{2.7.8}$$

By taking now the trace, we can invert this relation between π_{ij} and K_{ij} to recover equation (2.5.7):

$$K_{ij} = -\frac{1}{\sqrt{\gamma}}\left(\pi_{ij} - \frac{1}{2}\,\gamma_{ij}\,\pi\right) \; , \tag{2.7.9}$$

Since the Lagrangian density is independent of any derivatives of the lapse and shift, it is clear that the momenta conjugate to these variables are zero.

The Hamiltonian density is now the defined as

$$\mathscr{H} = \pi^{ij}\dot{\gamma}_{ij} - \mathcal{L} \; , \tag{2.7.10}$$

which in this case implies

$$\mathscr{H} = \left[-\alpha\sqrt{\gamma}\left({}^{(3)}R + K^2 - K_{ij}K^{ij}\right) + 2\pi^{ij}D_i\beta_j\right] \; . \tag{2.7.11}$$

After ignoring a total divergence, this can be rewritten as

$$\begin{aligned} \mathscr{H} &= -\sqrt{\gamma}\left[\alpha\left({}^{(3)}R + K^2 - K_{ij}K^{ij}\right) + 2\beta_i D_j\left(K^{ij} - \gamma^{ij}K\right)\right] \\ &= -2\sqrt{\gamma}\left[\alpha\mathcal{H} + \beta_i\mathcal{M}^i\right] \; , \end{aligned} \tag{2.7.12}$$

with $\mathcal{H}$ and $\mathcal{M}^i$ defined as before, but without the matter contributions (which would arise when we add the Hamiltonian density for the matter). The total Hamiltonian is now defined as

$$\mathbf{H} := \int \mathscr{H} d^3x \; . \tag{2.7.13}$$

Variation of this Hamiltonian with respect to α and β^i immediately yields the Hamiltonian and momentum constraints for vacuum, $\mathcal{H} = 0$ and $\mathcal{M}^i = 0$. In other words, the lapse and shift behave as Lagrange multipliers for a constrained system.

The evolution equations for the gravitational field now simply follow from Hamilton's equations:

$$\dot{\gamma}_{ij} = \frac{\delta \mathbf{H}}{\delta \pi^{ij}} \; , \qquad \dot{\pi}^{ij} = -\frac{\delta \mathbf{H}}{\delta \gamma_{ij}} \; . \tag{2.7.14}$$

From the first of these equations we find

$$\dot{\gamma}_{ij} = \frac{2\alpha}{\sqrt{\gamma}} \left(\pi_{ij} - \frac{1}{2} \gamma_{ij} \pi \right) + \pounds_{\vec{\beta}} \gamma_{ij} \; , \tag{2.7.15}$$

which is nothing more than the standard evolution equation for γ_{ij} written in terms of π_{ij} instead of K_{ij}, and from the second equation we recover the ADM evolution equations for π_{ij}, which are equivalent to (2.5.8).

There is an interesting observation we can make at this point, due to Anderson and York [22]. When using the Hamilton equations to derive the 3+1 evolution equations, we usually take the lapse function α as an independent quantity to be kept constant during the variation with respect to γ_{ij}. However, we might take a different point of view and assume that the independent gauge function is not the lapse as such, but rather the *densitized lapse* defined as

$$\tilde{\alpha} := \alpha / \sqrt{\gamma} \; . \tag{2.7.16}$$

This change is far from trivial, as keeping $\tilde{\alpha}$ constant alters the dependency of the Hamiltonian density on the metric during the variation (we have effectively multiplied it by a factor of $\sqrt{\gamma}$). The resulting evolution equations for π_{ij} are now different, and correspond precisely to those of York, equations (2.5.5). This observation points to the fact that perhaps the densitized lapse $\tilde{\alpha}$ is a more fundamental free gauge function than the lapse α itself. We will encounter the densitized lapse several times throughout the text.

2.8 The BSSNOK formulation

We have already mentioned that the 3+1 evolution equations are in fact non-unique since we can always add to them arbitrary multiples of the constraints to obtain new evolution equations that are just as valid. In fact, a large number of alternative formulations to ADM have been proposed in the literature. In Chapter 5, when we discuss the concept of well-posedness of a Cauchy problem and the hyperbolicity of a system of partial differential equations, we will have the opportunity to discuss some of the different reformulations and why some should be expected to be better than others in practical applications.

Here, however, we will introduce a specific reformulation that has proven to be particularly robust in the numerical evolution of a large variety of spacetimes, both with and without matter. Over the past few decades, and particularly since researchers started to work with full three-dimensional evolution codes for numerical relativity in the early 1990s, it was realized that the ADM equations lacked the necessary stability properties for long-term numerical simulations, something which is now known to be related to the fact that these equations are only weakly hyperbolic (see Chapter 5). In 1987, Nakamura, Oohara and Kojima presented a reformulation of the ADM evolution equations based on a conformal transformation that showed improved stability properties when compared to ADM [215]. This formulation evolved over the following years, though it remained largely unnoticed by the majority of researchers in numerical relativity until in 1998 Baumgarte and Shapiro systematically compared it to ADM in a series of spacetimes showing that the new formulation had far superior stability properties in all cases considered [50]. It was at this point that this reformulation became more widely noticed, and today it is used in one form or another by most large three-dimensional codes in numerical relativity. The more common version of this formulation is based on the work of Shibata and Nakamura [268], and Baumgarte and Shapiro [50], and is commonly known as the BSSN (Baumgarte, Shapiro, Shibata and Nakamura) formulation. This formulation has also been called "conformal ADM", though this name is not particularly good as it fails to make reference to the most important difference between the new formulation and ADM. A better name would probably be "conformal Γ formulation", since as we will see below the crucial element is the introduction of the auxiliary variables Γ^i. However, here we will use the name "BSSNOK formulation" (BS + Shibata, Nakamura, Oohara and Kojima) in order to make reference to the more commonly used name, and at the same time give due credit to the original authors.

In order to introduce the BSSNOK formulation, consider first a conformal rescaling of the spatial metric of the form

$$\tilde{\gamma}_{ij} := \psi^{-4} \gamma_{ij} \ . \tag{2.8.1}$$

Here ψ is a conformal factor that can in principle be chosen in a number of different ways. For example, when evolving black hole spacetimes with conformally flat initial data, we can simply take ψ to be the initial singular conformal factor and then ask for this conformal factor to remain fixed in time, something that allows us to evolve only the non-singular part of the metric and is known as the *puncture* method for evolutions (see Chapter 6). Alternatively, we can take the conformal factor to be initially given by some arbitrarily chosen scalar function and then propose some convenient evolution equation for this scalar function.

In the BSSNOK formulation, we choose the conformal factor in such a way that the conformal metric $\tilde{\gamma}_{ij}$ has unit determinant, that is

$$\psi^4 = \gamma^{1/3} \qquad \Rightarrow \qquad \psi = \gamma^{1/12} \, , \tag{2.8.2}$$

with γ the determinant of γ_{ij}. Furthermore, we ask for this relation to remain satisfied during the evolution. Now, from (2.3.11) we find that the evolution equation for the determinant of the metric is

$$\partial_t \gamma = \gamma \left(-2\alpha K + 2 D_i \beta^i \right) = -2\gamma \left(\alpha K - \partial_i \beta^i \right) + \beta^i \partial_i \gamma \, , \tag{2.8.3}$$

which implies

$$\partial_t \psi = -\frac{1}{6} \, \psi \, \left(\alpha K - \partial_i \beta^i \right) + \beta^i \partial_i \psi \, . \tag{2.8.4}$$

In practice we usually works with $\phi = \ln \psi = \frac{1}{12} \ln \gamma$, so that $\tilde{\gamma}_{ij} = e^{-4\phi} \gamma_{ij}$ and

$$\partial_t \phi = -\frac{1}{6} \, \left(\alpha K - \partial_i \beta^i \right) + \beta^i \partial_i \phi \, . \tag{2.8.5}$$

Recently, however, it has been suggested by Campanelli *et al.* [93] that evolving instead $\chi = 1/\psi^4 = \exp(-4\phi)$ is a better alternative when considering black hole spacetimes for which ψ typically has a $1/r$ singularity (so that ϕ has a logarithmic singularity), while χ is a C^4 function at $r = 0$. For regular spacetimes, of course, it should make no difference if we evolve ϕ, ψ or χ.

The BSSNOK formulation also separates the extrinsic curvature into its trace K and its tracefree part

$$A_{ij} = K_{ij} - \frac{1}{3} \gamma_{ij} K \, . \tag{2.8.6}$$

We further make a conformal rescaling of the traceless extrinsic curvature of the form[25]

$$\tilde{A}_{ij} = \psi^{-4} A_{ij} = e^{-4\phi} A_{ij} \, . \tag{2.8.7}$$

A crucial point is that BSSNOK also introduces three auxiliary variables known as the *conformal connection functions* and defined as

[25] As we will see in Chapter 3 when we discuss initial data, the "natural" conformal rescaling of the traceless extrinsic curvature is in fact $\tilde{A}^{ij} = \psi^{10} A^{ij}$, which implies $\tilde{A}_{ij} = \psi^2 A_{ij}$ (assuming we raise and lower indices of conformal quantities with the conformal metric). Since I wish to present the standard form of the BSSNOK equations, here I will continue to use the rescaling $\tilde{A}_{ij} = \psi^{-4} A_{ij}$. However, in order to avoid possible confusion later, the reader is advised to keep in mind that this rescaling is different from the one we will use in the next Chapter. It is also important to mention that if we choose to use $\tilde{A}_{ij} = \psi^2 A_{ij}$ instead, some of the equations in the BSSNOK formulation in fact simplify (most notably the momentum constraints and the evolution equations for $\tilde{\Gamma}^i$ and $\tilde{A}_{ij}$ itself), and it also becomes clear that the densitized lapse $\tilde{\alpha} = \alpha \gamma^{-1/2} = \alpha \psi^{-6}$ plays an important role (the BSSNOK equations with the natural conformal rescaling can be found in Appendix C).

$$\tilde{\Gamma}^i := \tilde{\gamma}^{jk}\tilde{\Gamma}^i_{jk} = -\partial_j \tilde{\gamma}^{ij} \; , \tag{2.8.8}$$

where $\tilde{\Gamma}^i{}_{jk}$ are the Christoffel symbols of the conformal metric, and where the second equality comes from the definition of the Christoffel symbols in the case when the determinant $\tilde{\gamma}$ is equal to 1 (which must be true by construction). So, instead of the 12 ADM variables γ_{ij} and K_{ij}, BSSNOK uses the 17 variables ϕ, K, $\tilde{\gamma}_{ij}$, $\tilde{A}_{ij}$ and $\tilde{\Gamma}^i$.[26] We can take the point of view that there are only 15 dynamical variables since $\tilde{A}_{ij}$ is traceless and $\tilde{\gamma}_{ij}$ has unit determinant, but here we will take the point of view that we are freely evolving all components of $\tilde{A}_{ij}$ and $\tilde{\gamma}_{ij}$ (however, enforcing the constraint $\tilde{A} = 0$ during a numerical calculation does seem to improve the stability of the simulations significantly, so it has become standard practice in most numerical codes).

Up to this point all we have done is redefine variables and introduce three additional auxiliary variables. The evolution equation for ϕ was already found above, while those for $\tilde{\gamma}_{ij}$, K and $\tilde{A}_{ij}$ can be obtained directly from the standard ADM equations. The system of evolution equations then takes the form[27]

$$\frac{d}{dt}\tilde{\gamma}_{ij} = -2\alpha \tilde{A}_{ij} \; , \tag{2.8.9}$$

$$\frac{d}{dt}\phi = -\frac{1}{6}\alpha K \; , \tag{2.8.10}$$

$$\begin{aligned} \frac{d}{dt}\tilde{A}_{ij} &= e^{-4\phi}\left\{-D_i D_j \alpha + \alpha R_{ij} + 4\pi\alpha\left[\gamma_{ij}\left(S-\rho\right) - 2S_{ij}\right]\right\}^{\mathrm{TF}} \\ &\quad + \alpha\left(K\tilde{A}_{ij} - 2\tilde{A}_{ik}\tilde{A}^k{}_j\right) \; , \end{aligned} \tag{2.8.11}$$

$$\frac{d}{dt}K = -D_i D^i \alpha + \alpha\left(\tilde{A}_{ij}\tilde{A}^{ij} + \frac{1}{3}K^2\right) + 4\pi\alpha\left(\rho + S\right) \; , \tag{2.8.12}$$

with $d/dt := \partial_t - £_{\vec{\beta}}$, and where TF denotes the tracefree part of the expression inside the brackets. In the previous expressions we have adopted the convention that indices of conformal quantities are raised and lowered with the conformal metric so that, for example, $\tilde{A}^{ij} = e^{4\phi}A^{ij}$. It is also important to notice that, in the evolution equation for K, the Hamiltonian constraint has been used in order to eliminate the Ricci scalar:

$$R = K_{ij}K^{ij} - K^2 + 16\pi\rho = \tilde{A}_{ij}\tilde{A}^{ij} - \frac{2}{3}\,K^2 + 16\pi\rho \; . \tag{2.8.13}$$

We then see how we have already started to add multiples of constraints to evolution equations.

Notice that in the evolution equations for $\tilde{A}_{ij}$ and K there appear covariant derivatives of the lapse function with respect to the physical metric γ_{ij}. These

[26] It should be noted that the formulation of [268] uses instead of the $\tilde{\Gamma}^i$ the auxiliary variables $F_i := -\sum_j \partial_j \tilde{\gamma}_{ij}$.

[27] From now on, and where there is no possibility of confusion, we will simply drop the index (3) from the three-dimensional Ricci tensor.

can be easily calculated by using the fact that the Christoffel symbols are related through:

$$\begin{aligned}\tilde{\Gamma}^k_{ij} &= \Gamma^k_{ij} - \frac{1}{3}\,\left(\delta^k_i \Gamma^m_{jm} + \delta^k_j \Gamma^m_{im} - \gamma_{ij}\gamma^{kl}\Gamma^m_{lm}\right) \\ &= \Gamma^k_{ij} - 2\left(\delta^k_i \partial_j \phi + \delta^k_j \partial_i \phi - \gamma_{ij}\gamma^{kl}\partial_l \phi\right)\ ,\end{aligned} \tag{2.8.14}$$

where $\tilde{\Gamma}^k_{ij}$ are the Christoffel symbols of the conformal metric, and where we have used the fact that $\partial_i \phi = \frac{1}{12}\,\partial_i \ln \gamma = \frac{1}{6}\,\Gamma^m_{im}$. This implies, in particular, that

$$\tilde{\Gamma}^i = e^{4\phi}\Gamma^i + 2\tilde{\gamma}^{ij}\partial_j \phi\ . \tag{2.8.15}$$

In the evolution equation for $\tilde{A}_{ij}$ we also need to calculate the Ricci tensor associated with the physical metric, which can be separated into two contributions in the following way:

$$R_{ij} = \tilde{R}_{ij} + R^{\phi}_{ij}\ , \tag{2.8.16}$$

where $\tilde{R}_{ij}$ is the Ricci tensor associated with the conformal metric $\tilde{\gamma}_{ij}$:

$$\begin{aligned}\tilde{R}_{ij} &= -\frac{1}{2}\tilde{\gamma}^{lm}\partial_l \partial_m \tilde{\gamma}_{ij} + \tilde{\gamma}_{k(i}\partial_{j)}\tilde{\Gamma}^k + \tilde{\Gamma}^k \tilde{\Gamma}_{(ij)k} \\ &+ \tilde{\gamma}^{lm}\left(2\tilde{\Gamma}^k_{l(i}\tilde{\Gamma}_{j)km} + \tilde{\Gamma}^k_{im}\tilde{\Gamma}_{klj}\right)\ .\end{aligned} \tag{2.8.17}$$

and where R^{ϕ}_{ij} denotes additional terms that depend on ϕ:

$$R^{\phi}_{ij} = -2\tilde{D}_i \tilde{D}_j \phi - 2\tilde{\gamma}_{ij}\tilde{D}^k \tilde{D}_k \phi + 4\tilde{D}_i \phi\, \tilde{D}_j \phi - 4\tilde{\gamma}_{ij}\tilde{D}^k \phi\, \tilde{D}_k \phi\ , \tag{2.8.18}$$

with $\tilde{D}_i$ the covariant derivative associated with the conformal metric.

We must also be careful with the fact that in the evolution equations above we are computing Lie derivatives with respect to $\vec{\beta}$ of *tensor densities*, that is tensors multiplied by powers of the determinant of the metric γ. If a given object is a tensor times $\gamma^{w/2}$, then we say that it is a tensor density of weight w. The Lie derivative of a tensor density of weight w is simply given by

$$\pounds_{\vec{\beta}} T = \left[\pounds_{\vec{\beta}} T\right]_{w=0} + w\, T\, \partial_i \beta^i\ , \tag{2.8.19}$$

where the first term denotes the Lie derivative assuming $w = 0$, and the second is the additional contribution due to the density factor. The density weight of $\psi = e^{\phi} = \gamma^{1/12}$ is clearly $1/6$, so the weight of $\tilde{\gamma}_{ij}$ and $\tilde{A}_{ij}$ is $-2/3$, and the weight of $\tilde{\gamma}^{ij}$ is $2/3$. In particular we have

$$\pounds_{\vec{\beta}} \phi = \beta^k \partial_k \phi + \frac{1}{6}\,\partial_k \beta^k\ , \tag{2.8.20}$$

$$\pounds_{\vec{\beta}} \tilde{\gamma}_{ij} = \beta^k \partial_k \tilde{\gamma}_{ij} + \tilde{\gamma}_{ik}\partial_j \beta^k + \tilde{\gamma}_{jk}\partial_i \beta^k - \frac{2}{3}\tilde{\gamma}_{ij}\partial_k \beta^k\ . \tag{2.8.21}$$

There are several motivations for the change of variables introduced above. First, the conformal transformation and the separating out of the trace of the extrinsic curvature are done in order to have better control over the slicing conditions that, as we will see in Chapter 4, are generally related with the trace of K_{ij}. On the other hand, the introduction of the conformal connection variables $\tilde{\Gamma}^i$ has the important consequence that when these functions are considered as independent variables, then the second derivatives of the conformal metric that appear on the right hand side of equation (2.8.11) (contained in the Ricci tensor (2.8.17)) reduce to the simple scalar Laplace operator $\tilde{\gamma}^{lm}\partial_l\partial_m\tilde{\gamma}_{ij}$. All other terms with second derivatives of $\tilde{\gamma}_{ij}$ have been rewritten in terms of first derivatives of the $\tilde{\Gamma}^i$.

If the $\tilde{\Gamma}^i$ are to be considered as independent variables, we are of course still missing an evolution equation for them. This equation can be obtained directly from (2.8.8) and (2.3.11):

$$\partial_t\tilde{\Gamma}^i = -\partial_j\left(\pounds_{\vec{\beta}}\tilde{\gamma}^{ij}\right) - 2\left(\alpha\partial_j\tilde{A}^{ij} + \tilde{A}^{ij}\partial_j\alpha\right) \,, \tag{2.8.22}$$

which after expanding out the Lie derivative term becomes

$$\begin{aligned}\partial_t\tilde{\Gamma}^i = {} & \tilde{\gamma}^{jk}\partial_j\partial_k\beta^i + \frac{1}{3}\,\tilde{\gamma}^{ij}\partial_j\partial_k\beta^k + \beta^j\partial_j\tilde{\Gamma}^i - \tilde{\Gamma}^j\partial_j\beta^i + \frac{2}{3}\,\tilde{\Gamma}^i\partial_j\beta^j \\ & - 2\left(\alpha\partial_j\tilde{A}^{ij} + \tilde{A}^{ij}\partial_j\alpha\right) \,.\end{aligned}$$

The last three terms of the first line clearly form the Lie derivative for a vector density of weight 2/3, while the extra terms involving second derivatives of the shift arise from the fact that the $\tilde{\Gamma}^i$ are not really components of a vector density, but are rather contracted Christoffel symbols. Bearing this in mind we can rewrite the last equation in more compact form as

$$\frac{d}{dt}\,\tilde{\Gamma}^i = \tilde{\gamma}^{jk}\partial_j\partial_k\beta^i + \frac{1}{3}\,\tilde{\gamma}^{ij}\partial_j\partial_k\beta^k - 2\left(\alpha\partial_j\tilde{A}^{ij} + \tilde{A}^{ij}\partial_j\alpha\right) \,. \tag{2.8.23}$$

We are still missing one key element of the BSSNOK formulation. In practice it turns out to be that, in spite of the motivations mentioned above, if we use equations (2.8.9), (2.8.10), (2.8.11), (2.8.12), and (2.8.23) in a numerical simulation the system turns out be violently unstable. In order to fix this problem we need to consider the momentum constraints, which in terms of the new variables take the form

$$\partial_j\tilde{A}^{ij} = -\tilde{\Gamma}^i_{jk}\tilde{A}^{jk} - 6\tilde{A}^{ij}\partial_j\phi + \frac{2}{3}\tilde{\gamma}^{ij}\partial_j K + 8\pi\tilde{j}^i \,, \tag{2.8.24}$$

with $\tilde{j}^i := e^{4\phi}j^i$. We can now use this equation to substitute the divergence of $\tilde{A}^{ij}$ that appears in the evolution equation for the $\tilde{\Gamma}^i$. We find:

$$\begin{aligned}\frac{d}{dt}\tilde{\Gamma}^i = {} & \tilde{\gamma}^{jk}\partial_j\partial_k\beta^i + \frac{1}{3}\,\tilde{\gamma}^{ij}\partial_j\partial_k\beta^k - 2\tilde{A}^{ij}\partial_j\alpha \\ & + 2\alpha\left(\tilde{\Gamma}^i_{jk}\tilde{A}^{jk} + 6\tilde{A}^{ij}\partial_j\phi - \frac{2}{3}\tilde{\gamma}^{ij}\partial_j K - 8\pi\tilde{j}^i\right) \,.\end{aligned} \tag{2.8.25}$$

The final system of evolution equations is then (2.8.9), (2.8.10), (2.8.11), (2.8.12), and (2.8.25). This new system not only does not present the violent instability mentioned before, but it also turns out to be far more stable than ADM in all cases studied until now. That this is so was first shown empirically (that is, through direct comparison of numerical simulations) by Baumgarte and Shapiro [50], and was later put on somewhat firmer ground by Alcubierre *et al.* [6] by considering linear perturbations of flat space. However, a full understanding of why BSSNOK is much more robust that ADM will have to wait until Chapter 5 when we discuss the concept of hyperbolicity.

The use of the momentum constraints to modify the evolution equation for the $\tilde{\Gamma}^i$ is the key ingredient of the BSSNOK formulation, and it is used in one way or another in all current implementations of this system of equations. There are some additional tricks that have been developed by different groups to make the numerical simulations even better behaved, like actively forcing the trace of the conformal-traceless extrinsic curvature $\tilde{A}_{ij}$ to remain zero during the evolution and using the independently evolved $\tilde{\Gamma}^i$ only in terms where derivatives of these functions appear, but we will not go into such details here.

2.9 Alternative formalisms

In this Chapter we have discussed the 3+1, or Cauchy, formalism for studying the evolution of spacetime. However, it is also important to mention other approaches to the problem of constructing a solution of Einstein's equations given adequate initial data. I will consider here two alternative approaches that have important strengths, namely the characteristic formulation and the conformal approach. The discussion here will be very brief, focusing mainly on the basic ideas behind each approach.

2.9.1 *The characteristic approach*

The characteristic formulation originates from the idea of using a foliation of spacetime based not on spacelike but rather on null hypersurfaces. Here we will discuss the main ideas behind this approach only briefly; a more detailed discussion can be found in the review paper by Winicour [301].

The idea of using null hypersurfaces is very attractive if we are interested in extracting gravitational wave information from an astrophysical system. When working with spatial hypersurfaces, this extraction typically requires the use of perturbative expansions around some Schwarzschild background (see Chapter 8) that in theory only work well at infinity. On the other hand, there is a completely unambiguous and rigorous description of gravitational waves on null hypersurfaces even in the non-linear context. This is the main motivation for the use of the characteristic approach, but it has yet another important advantage. In practice, in numerical simulations we can only evolve a finite region of spacetime, so when using spatial hypersurfaces we are forced to introduce some artificial boundary condition at a finite distance. Rigorous *outgoing wave* boundary conditions can

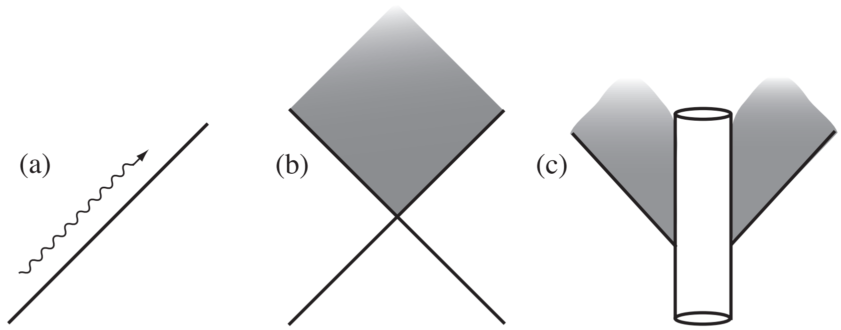

Fig. 2.4: Domain of dependence for the characteristic formulation. (a) A single null hypersurface has an empty domain of dependence as there are light rays coming from infinity that pass arbitrarily close to it but never intersect it. (b) The double null approach uses two null hypersurfaces. (c) A different approach is to use a central timelike world-tube (or world-line) to have a non-trivial domain of dependence.

in fact only be imposed at asymptotic null infinity, so the use of hypersurfaces that are null far away is ideal for this purpose.

There are, however, two important points to consider when using null hypersurfaces. The first is that even relatively small perturbations can cause the light rays generating such a hypersurface to cross, leading to the formation of caustics which are very difficult to handle numerically. The other point to consider is that the domain of dependence of a single regular null hypersurface is always empty. This is easy to understand as there can always be light rays coming from infinity that will come arbitrarily close to any point in the hypersurface without ever touching it, no matter how far the hypersurface extends. It turns out then that we must always add a second boundary surface in order to have a non-trivial future domain of dependence. Typical choices are to use either a second null hypersurface (the *double null* approach), or a central timelike world-tube that in some cases is collapsed to a single world-line (see Figure 2.4).

In the characteristic approach, we consider a foliation of null hypersurfaces corresponding to the level sets of a coordinate function u. On each hypersurface, different generating light rays are identified with angular coordinates x^A $(A = 2, 3)$, and an extra radial coordinate λ is used to parameterize these rays. In terms of these coordinates, the Einstein field equations take the following schematic form:

$$\partial_\lambda F = H_F\left(F, G\right) \; , \tag{2.9.1}$$

$$\partial_u \partial_\lambda G = H_G\left(F, G, \partial_u G\right) \; , \tag{2.9.2}$$

where F is a set of geometric variables related to the null hypersurface and G a set of evolution variables. The functions H_F and H_G are non-linear operators acting on the values of the variables on the hypersurface. The specific form of the equations depends on the choice of the form of the metric and other details.

A very common choice is to use the *Bondi–Sachs* null coordinate system, which in the general three-dimensional case corresponds to a spacetime metric of the form [71, 247]

$$\begin{aligned} ds^2 = & -\left(e^{2\beta}\frac{V}{r} - r^2 h_{AB} U^A U^B\right) du^2 \\ & - 2e^{2\beta} du dr - 2r^2 h_{AB} U^B du dx^A + r^2 h_{AB} dx^A dx^B \ , \end{aligned} \tag{2.9.3}$$

where here the radial coordinate r is used instead of λ. We then find hypersurface equations (*i.e.* involving only derivatives inside the hypersurface) for the metric functions $\{\beta, V, U^A\}$, and evolution equations (involving derivatives with respect to the null coordinate u) for the metric functions h_{AB}.

One important advantage of this approach is the fact that there are no elliptic constraints on the data, so the initial data is free. Additionally, there are no second derivatives in time (*i.e.* along the direction u), so there are fewer variables than in a 3+1 approach. Plus, null infinity can be compactified and brought to a finite distance in coordinate space, so that no artificial boundary conditions are required.[28]

A lot of work has been devoted to developing characteristic codes in spherical and axial symmetry, and today there are also well-developed three-dimensional codes that have been used to study, for example, scattering of waves by a black hole and even simulations of stars orbiting a black hole. A crucial development was the evolution of a black hole spacetime in a stable way for an essentially unlimited time by turning the problem around and considering a foliation of ingoing null hypersurfaces interior to an *outer* timelike world-tube [147].

The characteristic formalism has a series of advantages over traditional 3+1 approaches, but is has one serious drawback. As already mentioned, caustics can easily develop in null hypersurfaces, particularly in regions with strong gravitational fields. In those regions, a 3+1 approach should be much better behaved. This has led to an idea known as *Cauchy-characteristic matching* (see *e.g.* [57]), which uses a standard 3+1 approach based on a timelike hypersurface in the interior strong field region, matched to a null hypersurface in the exterior to carry the gravitational radiation to infinity (see Figure 2.5). Cauchy-characteristic matching has been shown to work well in simple test cases, but the technique has not yet been fully developed for the three-dimensional case for reasons related mainly to finding a stable and consistent way of injecting boundary data coming from the null exterior to the 3+1 interior.

[28] Compactifying spatial infinity is generally not a good idea since it implies a reduction in resolution at large distances: Wave packets get compressed as seen in coordinate space as they move outward. This gradual reduction in resolution acts on numerical schemes essentially as a change in the refraction index and causes waves to be "back-scattered" by the numerical grid. Compactifying null infinity, however, does not have this problem. Nevertheless, some numerical implementations do compactify spatial infinity but require the use of strong artificial damping of the waves as they travel outward to avoid this numerical back-scattering (see *e.g.* [231]).

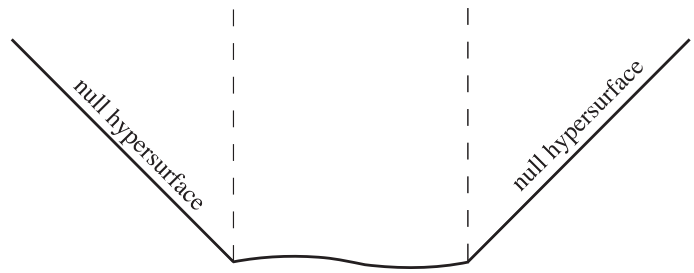

Fig. 2.5: Cauchy-characteristic matching. An interior region uses the standard 3+1 decomposition, while the exterior region uses a characteristic approach.

2.9.2 *The conformal approach*

As mentioned in the previous section, the idea of Cauchy-characteristic matching would seem to be a very promising way of having a well-behaved 3+1 interior, and at the same time a null exterior that allows us to compactify null infinity in order to both read off gravitational radiation directly and also to avoid the need to impose unphysical boundary conditions. There is, however, another approach based on a similar idea that uses a smooth hypersurface that is spacelike everywhere but nevertheless intersects null infinity. This approach evolves a conformal metric $\tilde{\gamma}_{\mu\nu}$ related to the physical metric of spacetime $g_{\mu\nu}$ through $\tilde{\gamma}_{\mu\nu} = \Omega^{-2} g_{\mu\nu}$. The conformal transformation is chosen is such a way that, in the conformal space, infinity corresponds to the boundary of a finite region where the conformal factor vanishes $\Omega = 0$.

In this approach, spacetime is foliated into spacelike hypersurfaces that reach null infinity. These hypersurfaces are called *hyperboloidal*, a name that can be understood if we consider for a moment the geometry of the hypersurface given by the hyperbola $t^2 - x^2 = a^2$ in Minkowski spacetime (we consider only the upper branch):

$$ds^2 = -dt^2 + dx^2 = \left(-\frac{x^2}{x^2 + a^2} + 1 \right) dx^2 = \frac{a^2}{x^2 + a^2} \, dx^2 \; . \tag{2.9.4}$$

The metric of this hypersurface is clearly positive definite for all x, *i.e.* it is spatial everywhere, but it can be shown to reach null infinity. Figure 2.6 shows how such hypersurfaces look in a conformal diagram.

If we attempt to write down the Einstein field equations for the conformal metric in a straightforward way, it turns out that they are singular at places where $\Omega = 0$, that is, at null infinity. However, a regular form of the conformal field equations has been derived by Friedrich (see *e.g.* [132, 165] and references therein). The variables involved are the connection coefficients, the trace and tracefree parts of the Riemann curvature tensor (the so-called Weyl tensor, see Chapter 8), plus the conformal factor Ω and its derivatives. Friedrich's system of

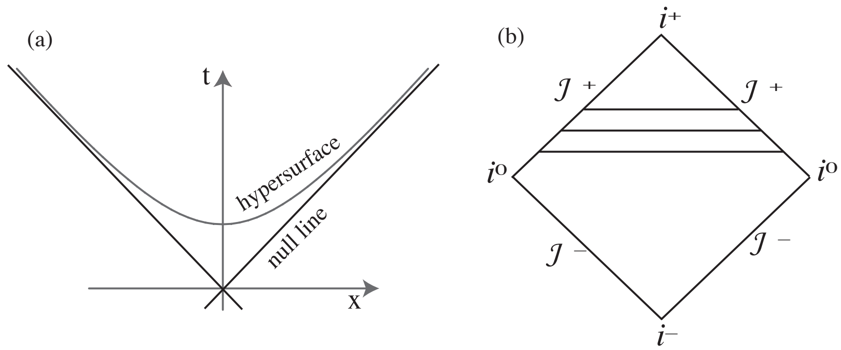

Fig. 2.6: Hyperboloidal hypersurfaces. (a) A hyperboloidal hypersurface in Minkowski spacetime. (b) A foliation of such hypersurfaces as seen in a conformal diagram.

equations can also be shown to be *symmetric hyperbolic* (see Chapter 5), which guarantees that it is mathematically well posed.

Since in the conformal formulation we are evolving the conformal factor Ω as an independent function, the position of the boundary of spacetime at null infinity $\mathscr{J}$ is not known *a priori* (except at $t = 0$). We then need to extend the physical initial data in some suitably smooth way and evolve the dynamical variables "beyond infinity". This has one important advantage, namely that it is possible to put an arbitrary (but well behaved) boundary condition at the outer boundary of the computational region without affecting the physical spacetime, as anything beyond $\mathscr{J}$ is causally disconnected from the interior.

The conformal formulation would seem to be an ideal solution to the weaknesses of both the standard 3+1 approach and the characteristic formulation. Being based on spatial hypersurfaces, it does not have to deal with the problem of caustics associated with the characteristic formulation. At the same time, by reaching null infinity, it allows clean extraction of gravitational radiation and other physical quantities such as total mass and momentum. The main problem faced by the conformal formulation today is related to the problem of constructing hyperboloidal initial data. Also, being based on spatial hypersurfaces, it will have to solve many of the same problems that standard 3+1 formulations are currently faced with, namely the choice of a good gauge and the stability of the evolutions against constraint violation. Though important progress has been made in recent years and numerical simulations of weak data in the full three-dimensional case have been carried out successfully, the conformal formulation is still considerably less developed than the standard 3+1 formulation. However, its conceptual elegance and fundamental strengths mean that this approach represents a very important promise for the future development of numerical relativity.