Numerical Relativity @ Jena

Jacopo Tissino

2021-04-27

Introduction

Tuesday 2021-4-13, compiled 2021-04-27

The syllabus can be found here. There will also be exercise sessions; the exercises will be posted on the same webpage.

The hydrodynamic part of the program is more pertinent to the computational hydrodynamics course.

There will also be a final project. The most recommended books are those by Alcubierre and Baumgarte-Shapiro. A good standard GR reference is Wald. The notes by Ghourghoulon are good and complete. The numerical methods reference by Choptuik is quite good.

There are **four pillars**:

1. we need to **formulate GR as a set of PDEs** (a Cauchy problem, really), including relativistic hydrodynamics (also, at this point we should classify these equations — are they hyperbolic, elliptic?);
2. some issues we encounter are the problems with **coordinates and singularities**: after all, we are interested in black holes and other extreme objects, so we must be able to work around horizons — the gauge is arbitrary, but some choices are better than others, we can gauge away coordinate singularities, but we must also work around the physical singularities;
3. we should use **numerical methods for the solution of PDEs on adaptive grids**;
4. the calculations are quite expensive, so we need **high performance computing** and easily parallelizable code.

We will mostly discuss the first two pillars in this course.

A landmark paper is one by Pretorius in 2006; his code already had several very useful characteristics.

We need **excision** to ignore the interior of the BHs, but we need to track them since they move.

In the 1970s there were already precursors to NR, such as throwing test masses at black holes.

Do the horizons merge in a continuous fashion? Yes, but it's hard to write shock-free gauge equations. We shall explore this later.

The definition of an event horizon depends on global properties of the spacetime, so it cannot be done "at runtime": we must store the data for all the simulation and then do raytracing.

On the other hand, an apparent horizon is local and gauge dependent, and we can compute these at runtime. It is useful to find it since, if it exists, it is always inside or coincident with the event horizon.

The result of a BNS merger is not necessarily a BH immediately.

We also have gravitational collapse of a scalar field with a variable energy. Letting it evolve according to GR leads to BH formation. There is a phase transition, the parameter is p and we get $M_{BH} \sim \left|p - p_*\right|^{\beta}$.

Do we have self-similarity for all p or only for $p = p_*$?

NR can also be used to study stringy higher-dimensional BHs.

1 Setting up the geometry

The idea is to introduce a notion of time by foliating the manifold. We must restrict ourselves to globally hyperbolic spacetimes. There must be no ugly things like CTCs.

The lapse function controls the foliation: it defines the proper time of Eulerian observers.

How is the foliation Σ "bent" into the 4D manifold? How does $\hat{n}$ change as we transport it along Σ?

The "velocity" is defined by a curvature $K_{\mu\nu}$, which is defined as

$$K_{\mu\nu} = -\gamma^{\alpha}{}_{\mu} \nabla_{\alpha} n_{\nu} \,, \tag{1.1}$$

where γ is the metric restricted to the Σ_t foliation.

The fundamental variables are then: the three-metric γ_{ij}, the "velocity" K_{ij}, the lapse α and β^i, as well as j_i and S_{ij}.

The equations we will write for these by manipulating the Einstein equations can be decomposed into *spacetime dynamical equations*:

$$\left(\partial_t - \mathcal{L}_{\vec{\beta}}\right)\gamma_{ik} = -2\alpha K_{ik} \tag{1.2}$$

$$\left(\partial_t - \mathcal{L}_{\vec{\beta}}\right)K_{ik} = -D_i D_k \alpha + \alpha\left({}^{(3)}R_{ik} - 2K_{ik}K^{j}{}_{k} + K_{ik}\right) - 8\pi\alpha\left(S_{ik} - \frac{1}{2}\gamma_{ik}(S - E)\right), \tag{1.3}$$

as well as two *constraints*:

$${}^{(3)}R + K^2 - K_{ik}K^{ik} = 16\pi E \tag{1.4}$$

$$D_k(K\gamma^{k}{}_{i}) = 8\pi j_i \,, \tag{1.5}$$

and the latter do not involve any Lie derivatives: they are specific to a single Σ_t. Also, we will need matter dynamical equations: from the stress-energy tensor we define

$$S_{ik} = \gamma^{\mu}{}_{i}\gamma^{\nu}{}_{k}T_{\mu\nu} \tag{1.6}$$

$$S = S^i{}_i \tag{1.7}$$

$$j_i = -\gamma^\mu{}_i n^\nu T_{\mu\nu} \tag{1.8}$$

$$E = T^{\mu\nu} n_\mu n_\nu \,, \tag{1.9}$$

and from $\nabla_\mu T^{\mu\nu} = 0$ we can write

$$\partial_t q_\mu + \partial_i F^i_\mu(q) = s_\mu \,. \tag{1.10}$$

We will need to make sure that the problem is well-posed, so that solutions exist and depend continuously on the parameters.

Slicing and coordinate choices

The gauge is defined by α and β^i: we can freely choose them. We want to have smoothness, to avoid singularities, to minimize grid distortion and for the problem to be well-posed.

The time is defined by the lapse: if we take $\alpha \equiv 1$ everywhere we have $\vec{a} = \nabla_n \vec{n} = D \log \alpha = 0$; this is called **geodesic slicing**, and we know that Eulerian observers follow geodesics. However, geodesics are "looking for" singularities, in that they easily fall inside them.

We could ask our hypersurfaces to bend as little as possible: we could minimize the trace of the extrinsic curvature, or even set $K = \nabla_a n^a = 0$. This allows our gauge not to create black holes. This translates to an elliptic equation for the lapse, to be solved together with the others.

This condition can be found to be equivalent to the maximization of the contained volume.

If we use geodesic slicing the simulation fails at time $\tau = \pi$ as the first gridpoint reaches the singularity. If we do excision by removing gridpoints as they fall in things are better but the simulation still fails.

If, instead, we impose $\partial_t \alpha = -\alpha K$, the foliation "freezes" as it passes the horizon, since the lapse function will prevent it. However, this means we will have large gradients in space as well as in time: as the grid points inside freeze the ones outside keep evolving. However, we can use a clever technique to minimize the distortion.

2 Introduction

What is the problem in NR? We have the Einstein equation $G_{ab} = 8\pi T_{ab}$, in geometric units $c = G = 1$, and we want to discuss its solutions which cannot be expressed analytically:

1. gravitational collapse;
2. BH or NS collisions;
3. dynamical stability of some stationary solutions — for example, Kerr black holes.

The commonalities among these are: strong gravity, the absence of symmetries (no Killing vectors), and the fact that these are dynamical.

We need to:

1. formulate the EFE as a PDE system;
2. check that it is well-posed;
3. simulate this using certain numerical algorithms;
4. extract information: we should use gauge-invariant quantities, such as gravitational waves or to-be-specified "energies".

Regarding the third point, we need to be able to solve nonlinear PDES: these can be elliptic or hyperbolic, deal with space-time decomposition, deal with singularities, and use high-performance computing effectively.

The PDE system

Let us start by treating the equations in vacuo: they reduce to $R_{\mu\nu} = 0$. The Ricci tensor can be expressed explicitly as

$$0 = R_{\mu\nu} = \underbrace{-\frac{1}{2} g^{\alpha\beta} \partial_\alpha \partial_\beta g_{\mu\nu} - g_{\alpha(\mu} \partial_{\nu)} H^\alpha}_{\text{principal part}} + Q_{\mu\nu}[g, \partial g] \tag{2.1}$$

$$H^\alpha = \partial_\mu g^{\alpha\mu} + \frac{1}{2} g^{\alpha\beta} g^{\rho\sigma} \partial_\beta g_{\rho\sigma}\,, \tag{2.2}$$

where the part denoted as the principal contains the highest derivatives of the metric, while the part denoted as Q is less important.

It seems like we have 10 equations for $g_{\mu\nu}$, so we are done! We actually do not, because of the Bianchi identities: $\nabla_a G^{ab} = 0$. These are four more equations, which we need to consider.

Our questions are:

1. what type of PDEs are these?
2. how do we formulate an initial/boundary value problem?
3. are these PDE problems well-posed?

Definition 2.1 (Well-posed PDE problem)**.** *It is a problem in which a unique solution exists, it is continuous, and it depends continuously on the boundary data.*

Maxwell equations in flat spacetime

Let us start with these as a reference. They are written in terms of the antisymmetric Maxwell-Faraday tensor $F_{\alpha\beta}$, and they read

$$0 = \partial^\alpha F_{\alpha\beta} = \partial^\alpha \left(\partial_\alpha A_\beta - \partial_\beta A_\alpha \right) , \tag{2.3}$$

where A^α is the vector potential. We could naively interpret these as four wave-like equations for the vector potential because of the $\Box A$ operator — however, this is not true. Let us show it by looking at the $\beta = 0$ equation:

$$\begin{aligned} 0 &= \partial^\alpha \partial_\alpha A_0 - \partial^\alpha \partial_0 A_\alpha = \Box A_0 - \partial_0 \partial^\alpha A_\alpha && (2.4) \\ &= -\partial_0^2 A_0 + \partial_i \partial^i A_0 + \partial_0^2 A_0 - \partial_0 \partial^i A_i && (2.5) \\ &= \partial^i (\partial_i A_0 - \partial_i A_0) = \partial^i F_{0i} = \partial^i E_i = C , && (2.6) \end{aligned}$$

where $E_\alpha = F_{\alpha 0} = F_{\alpha\beta} n^\beta$, where n^β is a unit four-vector in the time direction.

Importantly, this equation ($C = 0$) does not contain second time derivatives (they cancelled out). Can we get a wave equation from $\partial^0 C = 0$, maybe? No: the computation goes like

$$\begin{aligned} \partial^0 C &= \partial^0 (\partial^\alpha \partial_\alpha A_0 - \partial_0 A_\alpha) && (2.7) \\ &= \partial^i \left[\partial^\alpha (\partial_\alpha A_i - \partial_i A_\alpha) \right] , && (2.8) \end{aligned}$$

where we used the fact that $\partial^\alpha \partial^\beta F_{\alpha\beta} = 0$. The thing we are taking a derivative of is the left-hand side of the Maxwell equations for $\beta = i$! If A is a solution, the equation $\partial^0 C = 0$ is an *identity*, it does not constrain the solution in any way.

In summary, the Maxwell equations for A^α are 3 evolution equations (which contain second time derivatives) and the one $C = 0$ equation, which is a *constraint.*

What this means is that the Maxwell equations are **undetermined**! Three evolution equations, four unknowns. They are not well posed according to our mathematical definition.

However, from a physical point of view this is not a problem: we know that there is gauge freedom in electromagnetism, so the couple E_α, B_α are calculated from A_α up to a gauge transformation. A_α and $A_\alpha + \partial_\alpha \phi$ for any scalar field ϕ represent the same electric and magnetic fields.

What we can then do is to fix the gauge: we can choose, for example the Lorentz gauge $\partial_\alpha A^\alpha = 0$.

After doing so, the Maxwell equations reduce to $0 = \partial_\alpha \partial^\alpha A_\beta = 0$. These are 4 dynamical equations (containing $\partial_0 \partial_0$) for 4 unknowns A_α. This, then is a well-posed Cauchy problem.

Does the Lorentz Gauge hold for all time? Yes, because

$$0 = \partial^\beta \left(\Box A_\beta \right) = \Box (\partial^\beta A_\beta) , \tag{2.9}$$

so the quantity in parentheses is zero initially and its derivative is always zero: then, it remains zero.

What happened to the $C = 0$ constraint? If $C = 0$ initially then the condition $C = 0$ remains:

$$C = \Box A_\alpha - \partial^\alpha \partial_0 A_\alpha = \Box A_\alpha - \partial_0 \partial^\alpha A_\alpha = 0\,, \tag{2.10}$$

which is usually stated as: "**the constraints are transported along the dynamics**".

Finally then we can say that the Maxwell equations, in an appropriate gauge, are well-posed.

The EFE

In the EFE case, the Bianchi identities have the same role as our $C = 0$ constraint.

If n^b is a timelike vector field, then the projection of the Einstein equations

$$0 = G_{ab} n^a = C_b[\partial_i^2 g, \partial g, g]\,, \tag{2.11}$$

are *constraint* equations. On the other hand, the Bianchi identities guarantee that these are transported along the dynamics.

If we pick a gauge such that $0 = C^\mu = G^{0\mu}$, then the Bianchi identities read

$$\nabla_\alpha G^{\alpha\mu} = 0 \tag{2.12}$$

$$\partial_0 G^{0\mu} = -\partial_k G^{k\mu} - \Gamma^\mu_{\alpha\beta} G^{\alpha\beta} - \Gamma^\alpha_{\alpha\rho} G^{\mu\rho}\,, \tag{2.13}$$

and they contain at most $\partial_0^2 g$.

Can we write the Einstein equations in a way that makes them explicitly well-posed? Let us ignore the term $Q_{\mu\nu}$ — the important part is the principal one. The first term is already $g^{\alpha\beta}\partial_\alpha\partial_\beta$, the Dalambertian. The part we do not like is the derivative of H^α.

Can we say that $H^\alpha = 0$? Yes! This is the Hilbert / Lorentz / Harmonic gauge. Then the EFE read

$$\Box_g g_{\mu\nu} \simeq 0\,, \tag{2.14}$$

where the $\simeq$ sign denotes the fact that this is only considering the principal part.

Causality and globally hyperbolic spacetime

What we have seen so far is still not enough to show that the PDE system is, as a whole, hyperbolic. For that, we need to consider the eingenvalues of the problem.

We have initially assumed a "global motion of time".

Note that we do not mean a specific global time: in SR we had such a notion, because of the light-cone structure. An equation like $\Box_f \phi = 0$ is always "clear", its character is always determined by the lightcone structure — the condition at a point is affected by the past light cone, and it affects the future lightcone.

In GR this is not the case in general! An example are closed timelike curves.

In order to not have this problem, we need to restrict ourselves to a smaller class of spacetimes.

Definition 2.2. *An achronal set $S \subset M$ if made of events which are not connected by timelike curves.*

Definition 2.3. *The future domain of dependence $D_+(S)$ is the set of events such that every causal curve starting from a point in $D_+(S)$ intersects S in the past.*

Definition 2.4. *The future Cauchy horizon $H_+(S)$ is the boundary of $D_+(S)$.*

We can make an analogous definition by substituting the past for the future, and the plus for a minus.

Definition 2.5. *The domain of dependence of S is $D(S) = D_+(S) \cup D_-(S)$.*

Definition 2.6. *A **Cauchy surface** is a hypersurface $\Sigma \subset M$ of spatial character such that $D(\Sigma) = M$, or $H(\Sigma) = 0$.*

A property of Cauchy surface is that every causal curve intersects Σ exactly once.
Then we can say that

Definition 2.7. *A manifold (M, g) is **globally hyperbolic** iff there exists a Cauchy surface $\Sigma \subset M$.*

This is a key hypothesis in numerical relativity. We restrict our class of solutions to globally hyperbolic spacetimes.

There is a theorem ensuring that the initial value problem $\Box_G \phi = 0$ is well-posed in globally hyperbolic spacetimes.

3 3 + 1 geometry

We take our 4D spacetime (M, g, ∇) and equip it with a scalar field $t \colon M \to \mathbb{R}$ such that the vector Tuesday 2021-4-27, compiled 2021-04-27

$$(\mathrm{d}t)^a = g^{ab} (\mathrm{d}t)_b \tag{3.1}$$

is timelike.

This defines 3D spatial hypersurfaces, whose normal vector is $n^a = -\alpha(\mathrm{d}t^a)$.

These surfaces are denoted as Σ_t, and their tangent vectors $v^a \in T_P(\Sigma_t)$ are all spacelike.

Embedding is a bijective map ϕ from an $n-1$ dimensional manifold $\hat{\Sigma}$ to a subset Σ of an n dimensional manifold M.

We can move tensor fields from $\hat{\Sigma}$ to M and back: these are known as the pullback and pushforward operations. Not all fields can be moved in this way.

The idea is to identify Σ with the manifold $\hat{\Sigma}$ — however, they are distinct conceptually.

A metric γ on Σ is given by the pullback of g on M: this is the **induced** metric $\gamma = \phi^* g$. The components of this metric are calculated as

$$\gamma_{ij} = \frac{\partial x^\alpha}{\partial x^i} \frac{\partial x^\beta}{\partial x^j} g_{\alpha\beta} \,. \tag{3.2}$$

Also, we can define a connection D and a Riemann tensor $\mathcal{R}$ on the submanifold; its components are $\mathcal{R}_{ijkl}$. This is the "internal", or "intrinsic curvature" of the submanifold.

There is another curvature we could define: how does Σ deform into M? This is the **extrinsic curvature**.

For example, a 1D curve has zero intrinsic curvature, but it can bend. This is determined by the map

$$K\colon T_P(\Sigma) \times T_P(\Sigma) \to \mathbb{R}\,, \tag{3.3}$$

which acts on two vectors $u, v \in T_P(\Sigma)$ as $K(u, v) = K_{ab} u^a v^b = -u_a v^b \nabla_b n^a$.

As an example, take $M = \mathbb{R}^3$ and $\Sigma = \mathbb{R}^2$; the metric on M is the identity and the Riemann tensor vanishes, so we get $\mathcal{R} = K = 0$.

If, instead, we take $\Sigma = C^2$, the surface of a cylinder, we will have $\mathcal{R}_{ijkl} = 0$ since we can deform it into a plane, while

Claim 3.1. *the extrinsic curvature of this objects has a nonvanishing component $k_{\varphi\varphi} = -a$, where a is the radius, while the trace reads $k = k^i_i = -1/a$.*

On the other hand, if we take a sphere we have both $\mathcal{R}_{ijkl} \neq 0$ and $k_{ij} = 0$; also, the trace reads $k = -2/a$.

We can use **projectors** to decompose tensors on M into tensors on Σ and "pieces" along n. This is because we can decompose the tangent space of the ambient manifold into

$$T_P(M) = V_P(n) \oplus T_P(\Sigma)\,, \tag{3.4}$$

so that a vector v^α is written as $v_\perp n^a + v^a_\parallel$, where $v_\parallel \in T_P(\Sigma)$.

We can define a projector P that maps a vector v^a to $v^a_\parallel$. Explicitly, this will look like

$$P^a{}_b = \delta^a{}_b + n^a n_b\,. \tag{3.5}$$

The induced metric can alternatively be written as the projection of the four-metric into Σ:

$$\gamma_{ab} = P^c_a P^d_b g_{cd} = g_{ab} + n_a n_b\,. \tag{3.6}$$

This is coordinate independent! Also, we can write the projectors as

$$P^a_b = \gamma^a_b = g^{ac}\gamma_{cb}\,, \tag{3.7}$$

therefore we will not write P anymore — we can just use γ.

We also need the covariant derivative: schematically, it is written as

$$DT = P \dots P \nabla T\,. \tag{3.8}$$

Also, the extrinsic curvature reads

$$K_{ab} = -\gamma^c_a \gamma^d_b \nabla_{(c} n_{d)}\,. \tag{3.9}$$

Using P on every tensor is what "putting into 3+1 form" means.

Eulerian observers

These are observers whose worldlines are defined by n: Σ_t is the set of all events which are simultaneous to the Eulerian observers.

Definition 3.1. *The acceleration of E. observers is defined as*

$$a_a = n^b \nabla_b n_a \,, \tag{3.10}$$

therefore $a_a n^a = 0$.

Definition 3.2. *The normal evolution vector $m^a = \alpha n^a$ is defined so that*

$$\nabla_m t = m^a (\mathrm{d}t)_a = +1 \,. \tag{3.11}$$

Claim 3.2. *3+1 geometry defines the kinematic of 3+1 GR. Specifically, the claims we make are:*

1. *The vector m carries points from Σ_t to $\Sigma_{t+\delta t}$.*
2. *The lapse function α relates t to the proper time of Eulerian observers.*
3. *The Lie derivative along m, $\mathscr{L}_m$, transports tensors from Σ_t to $\Sigma_{t+\delta t}$.*
4. *$\mathscr{L}_m \gamma = -2\alpha K$.*

Let us give some hints: if P is on Σ_t and P' is on $\Sigma_{t+\delta t}$, then

$$t(P') = t(P + \delta t \cdot m) = t(p) + \delta t \underbrace{m^a (\mathrm{d}t)_a}_{=1} = t(P) + \delta t \,, \tag{3.12}$$

while for the second point, the change in proper time reads

$$\mathrm{d}\tau^2 = -g(\delta t m, \delta t m) = -m^a m_a \,\mathrm{d}t^2 = \alpha^2 \,\mathrm{d}t^2 \,, \tag{3.13}$$

which is the reason for the term "lapse function".

The third point follows from the first and the definition of the Lie derivative.

As for the fourth, we have

$$\begin{aligned} \mathscr{L}_n \gamma_{ab} &= \mathscr{L}_n \left(g_{ab} + n_a n_b \right) && (3.14) \\ &= 2\nabla_{(a} n_{b)} + n_a \mathscr{L}_n n_b + n_b \mathscr{L}_n n_a && (3.15) \\ &= 2\Big[\nabla_{(a} n_{b)} + n_{(a} a_{b)} \Big] = -2K_{ab} \,. && (3.16) \end{aligned}$$

The last passage is left as an exercise: we start by carrying out the projections, then simplify terms in the form $nn\nabla n$ by substituting the acceleration.

The Lie derivative along n can always be written as

$$\mathscr{L}_n \gamma_{ab} = \phi^{-1} \mathscr{L}_{\phi n} \gamma_{ab} \tag{3.17}$$

for any scalar field.

Now we have three expressions for K_{ab} — each of them can be taken to be the definition, and the others can be derived from it.

What is the physical interpretation of these expressions? They tell us several things:

1. Σ_t and $\Sigma_{t+\delta t}$ are identified by the diffeomorphism generated by m;
2. the spacetime (M, g) is the "time" development of (Σ, γ), where the "time evolution" is governed by $\mathscr{L}_m$;
3. we can identify γ as the "main variable" of (3+1) GR, and also we identify K as the "velocity" of γ: the equation $\mathscr{L}_m \gamma = -2\alpha K$ is in the form "time derivative of variable = velocity", a kinematic equation.

This all hinges on the possibility to define these non-intersecting hypersurfaces.