

Introducción a la relatividad numérica

Miguel Alcubierre
Instituto de Ciencias Nucleares
Universidad Nacional Autónoma de México
(malcubi@nucleares.unam.mx)

8 de noviembre de 2016

Índice

Resumen	3
1. La relatividad general	5
1.1. Introducción	5
1.2. Métrica y geodésicas	6
1.3. Curvatura	9
1.4. Base coordenada y derivadas covariantes	11
1.5. Las ecuaciones de campo de Einstein	14
1.6. Identidades de Bianchi y leyes de conservación	16
2. El formalismo de la relatividad numérica	18
2.1. Introducción	18
2.2. El formalismo 3+1	18
2.3. Curvatura extrínseca	21
2.4. Las ecuaciones de Einstein en el formalismo 3+1	24
2.5. El problema de los valores iniciales	26
2.6. Datos iniciales para agujeros negros	28
2.7. Condiciones de norma	29
2.7.1. Condiciones de foliación	30
2.7.2. Condiciones de corrimiento	33

3. Formulaciones alternativas e hiperbolicidad	36
3.1. Introducción	36
3.2. La formulación BSSN	37
3.3. El concepto de hiperbolicidad	41
3.4. Hiperbolicidad de las ecuaciones de evolución en 3+1	43
4. Aproximaciones en diferencias finitas	49
4.1. Introducción	49
4.2. Ideas fundamentales en las diferencias finitas	49
4.3. La ecuación de onda en una dimensión	51
4.4. Aproximaciones implícitas y moléculas computacionales	54
4.5. Consistencia, convergencia y estabilidad	56
4.6. Estabilidad de von Neumann	57
4.7. Ejemplos	61
5. Ejemplos de aplicaciones	64
5.1. Relatividad de juguete en 1+1 dimensiones	64
5.2. Relatividad en simetría esférica	68
Referencias	80

Resumen

La teoría de la relatividad general es una teoría altamente exitosa. No solo ha modificado de manera radical nuestra visión de la gravitación, el espacio y el tiempo, sino que también posee un enorme poder predictivo. A la fecha, ha pasado con extraordinaria precisión todas las pruebas experimentales y observacionales a que se le ha sometido. Entre sus logros mas importantes se encuentran la predicción de objetos exóticos como las estrellas de neutrones y los agujeros negros, y el modelo cosmológico de la Gran Explosión. Además, la relatividad general predice también la existencia de las ondas gravitacionales que tal vez sean observadas por primera vez en el transcurso de esta década.

La relatividad general, aunque conceptualmente sencilla y elegante, resulta ser en la práctica una teoría altamente compleja. Las ecuaciones de campo de Einstein son un sistema de diez ecuaciones diferenciales parciales en cuatro dimensiones, acopladas y no lineales. Escritas en su forma más general estas ecuaciones poseen miles de términos. Debido a esta complejidad, soluciones exactas de las ecuaciones de Einstein se conocen solo en situaciones con alto grado de simetría, ya sea en el espacio o en el tiempo: soluciones con simetría esférica o axial, soluciones estáticas o estacionarias, soluciones homogéneas e isotrópicas, etc. Si uno está interesado en estudiar situaciones con relevancia astrofísica, que involucren campos gravitacionales intensos, dinámicos y con poca o ninguna simetría, resulta imposible resolver las ecuaciones de Einstein de manera exacta. De la necesidad de estudiar este tipo de problemas ha surgido el área de la relatividad numérica, que intenta resolver las ecuaciones de Einstein utilizando aproximaciones numéricas y complejos códigos computacionales.

La relatividad numérica se desarrolló como un campo de investigación independiente a mediados de la década de los sesentas, comenzando con los esfuerzos pioneros de Hahn y Lindquist [32], pero no fue sino hasta finales de los setentas cuando las primeras simulaciones realmente exitosas fueron realizadas por Smarr y Eppley [54, 24, 25] en el contexto de la colisión de frente de dos agujeros negros. En esa época, sin embargo, el poder de las computadoras era aún muy modesto, y las simulaciones que podían realizarse se limitaban a problemas con simetría esférica, o a lo más simetría axial a muy baja resolución. Esta situación ha cambiado, durante las décadas de los ochenta y noventa una verdadera revolución tuvo lugar en el campo de la relatividad numérica. Problemas cada vez más complejos en muy diversos aspectos de la teoría de la gravitación se han estudiado, desde la simulación de estrellas en rotación, al estudio de defectos topológicos, el colapso gravitacional y las colisiones de objetos compactos. Quizá el resultado de mayor importancia que se ha obtenido a través de la relatividad numérica sea el descubrimiento por parte de Choptuik de los fenómenos críticos en el colapso gravitacional [20, 30]. Un resumen de la historia de la relatividad numérica y sus más recientes logros se pueden encontrar en [37].

La relatividad numérica esta alcanzando un estado de madurez. El desarrollo de poderosas super-computadoras, junto con el desarrollo de técnicas numéricas robustas han permitido que se realicen simulaciones de sistemas tridimensionales con campos gravitacionales intensos y dinámicos. Toda esta actividad esta ocurriendo en el momento justo, una nueva generación de detectores de ondas gravitacionales esta en estado avanzado de construcción, o incluso haciendo ya las primeras pruebas de calibración [39, 62, 29, 59]. Las señales esperadas, sin embargo, son tan débiles que incluso con la enorme sensibilidad de estos nuevos detectores resultará necesario extraerlas del ruido de fondo. Es mucho más sencillo extraer una señal enterrada en ruido si se sabe que buscar. Es aquí donde la relatividad numérica es requerida urgentemente, proporcionando a los observadores con predicciones precisas de los tipos de ondas generados por las fuentes astrofísicas más comunes. Vivimos en verdad un momento emocionante en el desarrollo de esta disciplina.

En estas notas se presenta una introducción a las principales ideas detrás de la relatividad numérica. Las notas están organizadas de la siguiente forma: En la sección 1 se presenta una breve introducción a la relatividad general. La sección 2 presenta el problema de Cauchy de la relatividad general e introduce el formalismo 3+1, que es el más comúnmente usado en la relatividad numérica (aunque no el único). En la sección 3 se presentan formulaciones alternativas de las ecuaciones de evolución en el formalismo 3+1 y se discute el concepto de hiperbolicidad. La sección 4 presenta una introducción a las aproximaciones en diferencias finitas.

Finalmente, en la sección 5 se presenta una serie de ejemplos de aplicaciones de las técnicas de la relatividad numérica al estudio de algunos sistemas sencillos. En particular, se estudia el caso de la relatividad en 1+1 dimensiones (1 dimensión espacial), y la relatividad en simetría esférica.

1. La relatividad general

1.1. Introducción

La teoría moderna de la gravitación es la “teoría de la relatividad general” postulada por Albert Einstein a fines de 1915 [22, 23]. De acuerdo a esta teoría, la gravitación no es una fuerza como se le consideraba en la física newtoniana, sino que es más bien una manifestación de la “curvatura” del espacio-tiempo. Un objeto masivo produce una distorsión en la geometría del espacio-tiempo, y a su vez esta distorsión controla o altera el movimiento de los objetos. Utilizando el lenguaje de John A. Wheeler, la materia le dice al espacio-tiempo como curvarse, y el espacio-tiempo le dice a la materia como moverse [64].

Ya al postular la teoría especial de la relatividad en 1905, Einstein sabía bien que la teoría de la gravitación de Newton debería ser modificada. La principal razón para esto era el que la teoría de Newton implicaba que la fuerza de gravedad se propagaba entre distintos objetos a velocidad infinita, lo que contradecía un principio fundamental de la relatividad especial: ninguna interacción física puede viajar más rápido que la velocidad de la luz. Es importante notar que al mismo Newton nunca le pareció convincente la existencia de esta “acción a distancia”, pero consideró que era una hipótesis necesaria hasta que se encontrara una mejor explicación de la naturaleza de la gravedad. En la década de 1905 a 1915, Einstein se dedicó a buscar esa explicación más adecuada.

Las ideas principales que guiaron a Einstein en su camino hacia la relatividad general fueron el “principio de covariancia general” que dice que las leyes de la física deben tomar la misma forma para todo observador, el “principio de equivalencia”, que dice que todos los objetos caen exactamente de la misma forma en un campo gravitacional, y el “principio de Mach”, llamado así en honor a Ernst Mach quién lo postuló a fines del siglo XIX, y que dice que la inercia local de un objeto debe ser producida por la distribución total de la materia en el Universo. El principio de covariancia general llevó a Einstein a pedir que las leyes de la física se escribieran en forma tensorial, el principio de equivalencia lo llevó a concluir que la manera natural de describir a la gravedad era identificándola con la geometría del espacio-tiempo, y el principio de Mach lo llevó a la idea de que dicha geometría debería ser alterada por la distribución de materia y energía.

En las siguientes secciones veremos como se expresan estas ideas en lenguaje matemático. Antes de seguir adelante, mencionare primero la convención de unidades que seguiré en estas notas. Utilizaré siempre las llamadas “unidades geométricas”, en las que la velocidad de la luz c y la constante de la gravitación universal de Newton G son ambas iguales a la unidad. Cuando se trabaja en este sistema de unidades, la distancia, el tiempo, la masa y la energía tienen siempre las mismas unidades (de distancia). Las unidades

convencionales (SI) del sistema internacional siempre pueden recuperarse añadiendo los factores de c y G que sean necesarios en cada caso para regresar a sus unidades correctas en el SI (por ejemplo, haciendo las substituciones $t \rightarrow ct$ y $M \rightarrow GM/c^2$).

Existen muchos libros introductorios a la relatividad general. La presentación que se da aquí está basada principalmente en las referencias [42, 50, 63].

1.2. Métrica y geodésicas

Como ya hemos mencionado, el principio de equivalencia llevó a Einstein a pensar que la gravitación podía identificarse con la curvatura del espacio-tiempo. Matemáticamente esto quiere decir que la teoría de la gravedad debería ser lo que se conoce como una “teoría métrica”, en la cual la gravedad se manifiesta única y exclusivamente a través de una distorsión en la geometría del espacio-tiempo.

Consideremos un espacio-tiempo continuo de cuatro dimensiones (tres dimensiones de espacio y una de tiempo), o en lenguaje matemático, una variedad diferenciable de cuatro dimensiones. Sean x^α las coordenadas de un evento en este espacio-tiempo, donde el índice α toma los valores $\{0, 1, 2, 3\}$: cuatro números que indican en que momento del tiempo, y en que lugar en el espacio ocurre ese evento (en estas notas siempre tomaremos la componente 0 como aquella que se refiere al tiempo, y las componentes $\{1, 2, 3\}$ como las que se refieren al espacio, también utilizaremos la convención de que índices griegos toman valores de 0 a 3, mientras que índices latinos toman valores de 1 a 3).

Entre dos eventos infinitesimalmente cercanos con coordenadas x^α y $x^\alpha + dx^\alpha$ es posible definir una “distancia invariante” ds^2 de la siguiente forma

$$ds^2 = \sum_{\alpha, \beta=1}^4 g_{\alpha\beta} dx^\alpha dx^\beta \equiv g_{\alpha\beta} dx^\alpha dx^\beta, \quad (1.1)$$

donde $g_{\alpha\beta}$ se conoce como el “tensor métrico”, o simplemente “la métrica”, y donde la última igualdad define la “convención de suma de Einstein”: índices que aparecen repetidos se suman. La distancia invariante es, como su nombre lo indica, una cantidad absoluta que no depende del sistema de coordenadas que se utilice para describir al espacio tiempo. En cada punto del espacio-tiempo, las componentes del tensor métrico forman una matriz simétrica no singular de 4×4 elementos, con eigenvalores cuyos signos son $(-, +, +, +)$, es decir, un eigenvalor negativo asociado al tiempo, y tres eigenvalores positivos asociados al espacio. En la relatividad especial, el tensor métrico se reduce a la llamada “métrica de Minkowski”

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \equiv \eta_{\alpha\beta} dx^\alpha dx^\beta, \quad (1.2)$$

que corresponde a un espacio-tiempo plano. En la relatividad general, por otro lado, el tensor métrico varía de un punto a otro.

Las transformaciones de Lorentz garantizan que el intervalo ds^2 tiene el mismo valor para cualquier observador. Esto es consecuencia directa del postulado (o mejor dicho, del hecho empírico) de la invariancia de la velocidad de la luz. Nótese que debido a la presencia de un eigenvalor negativo, la distancia invariante no es positiva definida. De hecho, a partir de la métrica uno puede distinguir entre eventos relacionados entre sí de 3 formas distintas:

$$ds^2 > 0 \quad \text{intervalo espacialoide,} \quad (1.3)$$

$$ds^2 < 0 \quad \text{intervalo temporaloide,} \quad (1.4)$$

$$ds^2 = 0 \quad \text{intervalo nulo.} \quad (1.5)$$

Los intervalos espacialoides corresponden a eventos tales que un objeto tendría que moverse mas rápido que la luz para llegar de uno a otro (están separados principalmente en el “espacio”), Los temporaloides a eventos donde un objeto tiene que moverse más lento que la luz para llegar de uno a otro (están separados principalmente en el “tiempo”), y los nulos a eventos que pueden alcanzarse viajando precisamente a la velocidad de la luz (la frontera entre separación espacial y temporal). Todo objeto material se mueve siguiendo trayectorias de tipo temporaloide, y la luz se mueve siguiendo trayectorias nulas. Las trayectorias nulas definen lo que se conoce como el “cono de luz” (véase Figura 1). El cono de luz indica los eventos que pueden tener influencia física entre sí, y por lo tanto define la causalidad.

El tensor métrico también se puede utilizar para definir un producto escalar invariante entre dos vectores de cuatro dimensiones (“4-vectores”). Sean \vec{u} y \vec{v} dos 4-vectores con componentes u^α y v^α respectivamente, su producto escalar se define como:

$$\vec{u} \cdot \vec{v} := g_{\alpha\beta} u^\alpha v^\beta, \quad (1.6)$$

donde hemos utilizado de nuevo la convención de la suma. En particular, la magnitud cuadrada de un 4-vector se define como

$$u^2 := \vec{u} \cdot \vec{u} = g_{\alpha\beta} u^\alpha u^\beta. \quad (1.7)$$

Nótese que de nuevo, esta cantidad no es positiva definida, por lo que normalmente no sacamos su raíz cuadrada. Podemos entonces clasificar a los vectores de la misma manera que al intervalo en espacialoides, temporaloides y nulos. Es claro ahora que el intervalo no es otra cosa que la magnitud cuadrada del vector (infinitesimal) de desplazamiento.

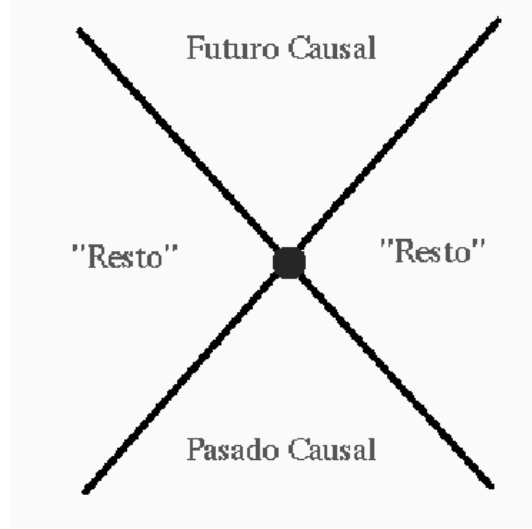


Figura 1: El cono de luz de un evento define las relaciones causales con otros eventos, y divide al espacio-tiempo en tres regiones: el pasado causal (aquellos eventos que pueden afectar al evento considerado), el futuro causal (aquellos eventos que pueden ser afectados por el evento considerado) y el “resto” (aquellos eventos con los que no hay contacto causal).

En relatividad especial, en ausencia de fuerzas externas los objetos se mueven en líneas rectas en el espacio-tiempo (correspondientes a movimiento rectilíneo-uniforme). La línea recta corresponde a una trayectoria de longitud extrema de acuerdo a la métrica de Minkowski (no necesariamente la más corta debido al signo negativo de uno de los eigenvalores). La idea de Einstein fue precisamente pensar que en la presencia de un campo gravitacional, los objetos aún se mueven siguiendo las trayectorias más rectas posibles, es decir trayectorias extremas, pero ahora en una métrica distinta. A esa trayectoria extrema se le llama “geodésica”. De esta forma, la gravedad no se ve como una fuerza externa, sino como una distorsión en la métrica. Dada esta distorsión, los objetos se mueven simplemente siguiendo una geodésica.

Se acostumbra parametrizar la trayectoria de un objeto utilizando el llamado “tiempo propio” $d\tau^2 := -ds^2$, que corresponde al tiempo medido por el objeto mismo (el tiempo que mide un reloj ideal atado al objeto). Nótese que este tiempo no tiene por qué ser igual a dt , pues t es solo una coordenada, es decir, una etiqueta arbitraria que asignamos a diferentes eventos. La ecuación para una geodésica está dada en general por

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} = 0, \quad (1.8)$$

donde las cantidades $\Gamma_{\beta\gamma}^\alpha$ se conocen como “símbolos de Christoffel” y están dados en términos de derivadas de la métrica como

$$\Gamma_{\beta\gamma}^\alpha := \frac{g^{\alpha\mu}}{2} \left[\frac{\partial g_{\beta\mu}}{\partial x^\gamma} + \frac{\partial g_{\gamma\mu}}{\partial x^\beta} - \frac{\partial g_{\beta\gamma}}{\partial x^\mu} \right]. \quad (1.9)$$

En la ecuación anterior introdujimos los coeficientes $g^{\alpha\beta}$ que se definen como las componentes de la matriz inversa a $g_{\alpha\beta}$: $g^{\alpha\mu}g_{\beta\mu} = \delta_\beta^\alpha$.

La ecuación geodésica puede escribirse también como:

$$\frac{du^\alpha}{d\tau} + \Gamma_{\beta\gamma}^\alpha u^\beta u^\gamma = 0, \quad (1.10)$$

donde hemos definido el 4-vector $u^\alpha := dx^\alpha/d\tau$. Este vector se conoce como la “4-velocidad”, y es la generalización relativista del concepto de velocidad ordinario. Es importante notar que:

$$u^2 = \vec{u} \cdot \vec{u} = g_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = \frac{ds^2}{d\tau^2} = -1, \quad (1.11)$$

es decir, la cuatro-velocidad es un vector unitario temporal.

Dado un campo gravitacional, es decir, dada la métrica del espacio-tiempo, la ecuación de las geodésicas (1.8) nos da la trayectoria de los objetos: el espacio-tiempo le dice a los objetos como moverse.

1.3. Curvatura

Como hemos visto, la métrica del espacio-tiempo nos permite obtener la trayectoria de los objetos en un campo gravitacional. Sin embargo, el tensor métrico no es la forma más conveniente de describir la presencia de un campo gravitacional. Para ver esto, basta con notar que incluso en un espacio plano, uno puede cambiar la forma del tensor métrico mediante una simple transformación de coordenadas. Por ejemplo, la métrica de un espacio plano de tres dimensiones en coordenadas cartesianas $\{x, y, z\}$ esta dada simplemente por la regla de Pitágoras:

$$dl^2 = dx^2 + dy^2 + dz^2, \quad (1.12)$$

mientras que la métrica del mismo espacio plano en coordenadas esféricas $\{r, \theta, \phi\}$ resulta ser

$$dl^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2, \quad (1.13)$$

lo puede verse fácilmente de la transformación de coordenadas

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta. \quad (1.14)$$

que implica

$$dx = dr \sin \theta \cos \phi - r \sin \theta \sin \phi d\phi + r \cos \theta \cos \phi d\theta, \quad (1.15)$$

$$dy = dr \sin \theta \sin \phi + r \sin \theta \cos \phi d\phi + r \cos \theta \sin \phi d\theta, \quad (1.16)$$

$$dz = dr \cos \theta - r \sin \theta d\theta. \quad (1.17)$$

A partir de la métrica (1.13) es posible calcular los símbolos de Christoffel para las coordenadas esféricas (no son todos iguales a cero!), y ver que la ecuación de una línea recta (geodésica) escrita en estas coordenadas ya no es trivial. Llevando a cabo transformaciones de coordenadas aún más elaboradas es posible terminar con una métrica mucho más compleja. Debemos encontrar entonces una forma de distinguir con certeza entre un espacio plano y uno que no lo es.

La manera de hacer esto es a través del llamado “tensor de curvatura de Riemann”. Este tensor mide el cambio de un vector al transportarlo alrededor de un circuito manteniéndolo siempre paralelo a sí mismo (“transporte paralelo”). En un espacio plano, el vector no cambia al hacer esto, en un espacio curvo, sí lo hace. Esto puede verse claramente si uno piensa en mover un vector en la superficie terrestre. Si comenzamos en un punto en el ecuador (punto A) con un vector apuntando al este, nos movemos hasta el polo norte (punto B) siguiendo un meridiano, bajamos siguiendo otro meridiano que haga un ángulo recto hacia el este con el primero hasta llegar de nuevo al ecuador (punto C), y luego seguimos el ecuador hasta volver al punto original, nos encontraremos con que el vector ahora apunta al sur (véase figura 2).

En estas notas no vamos a derivar al tensor de Riemann de primeros principios, nos limitaremos solo a escribirlo:

$$R^\sigma_{\mu\nu\rho} := \partial_\nu \Gamma^\sigma_{\mu\rho} - \partial_\mu \Gamma^\sigma_{\nu\rho} + (\Gamma^\alpha_{\mu\rho} \Gamma^\sigma_{\alpha\nu} - \Gamma^\alpha_{\nu\rho} \Gamma^\sigma_{\alpha\mu}), \quad (1.18)$$

donde ∂_μ es una abreviación de $\partial/\partial x^\mu$, y donde $\Gamma^\alpha_{\mu\nu}$ son los símbolos de Christoffel definidos anteriormente. Nótese que el tensor de Riemann tiene 4 índices, es decir $4 \times 4 \times 4 \times 4 = 256$ componentes. Sin embargo, tiene muchas simetrías, por lo que solamente tiene 20 componentes independientes (en 4 dimensiones). Es posible demostrar que el tensor de Riemann es igual a cero si y solo si el espacio es plano.

A partir del tensor de Riemann podemos definir el llamado “tensor de Ricci” sumando (contrayendo) dos de los índices libres:

$$R_{\mu\nu} := \sum_\lambda R^\lambda_{\mu\lambda\nu} = R^\lambda_{\mu\lambda\nu}. \quad (1.19)$$

El tensor de Ricci es simétrico en sus dos índices, por lo que tiene solamente 10 componentes independientes. Nótese que el hecho de que el tensor de Ricci sea cero *no* significa

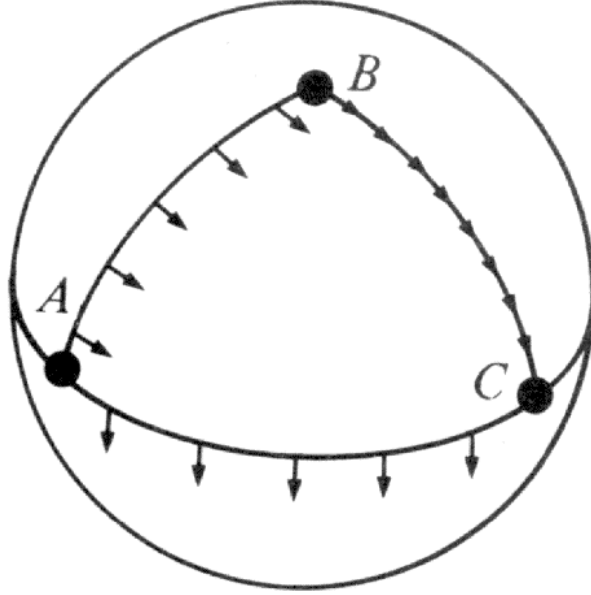


Figura 2: Transporte paralelo en la superficie terrestre.

que el espacio sea plano, ya que las 10 componentes restantes del Riemann pueden ser distintas de cero.

Es importante hacer notar que en algunas ocasiones hemos escrito tensores con los índices arriba y en otras con los índices abajo. Esto no es un error tipográfico, los tensores con índices arriba o abajo no son iguales, pero si están relacionados entre si. La regla es la siguiente: los índices de objetos geométricos se suben y bajan contrayendo con el tensor métrico $g_{\mu\nu}$ o su inversa $g^{\mu\nu}$. Por ejemplo:

$$\begin{aligned} v_\mu &= g_{\mu\nu} v^\nu, \\ v^\mu &= g^{\mu\nu} v_\nu, \\ T^{\mu\nu} &= g^{\mu\alpha} g^{\nu\beta} T_{\alpha\beta}, \\ R_{\sigma\mu\nu\rho} &= g_{\sigma\lambda} R^\lambda_{\mu\nu\rho} \end{aligned}$$

1.4. Base coordenada y derivadas covariantes

Cuando se considera el cambio de un campo vectorial (objetos con un índice) o tensorial (objetos con más de un índice) al moverse en el espacio-tiempo, debemos tomar en cuenta que al movernos de un punto a otro no solo las componentes de dichos objetos

pueden cambiar, sino que también la base en las que dichas componentes se miden cambia de un punto a otro. En efecto, cuando consideramos las componentes de un objeto geométrico (vector o tensor), dichas componentes siempre están dadas con respecto a una base específica.

En un espacio plano tridimensional en coordenadas cartesianas, la base es en general la llamada “base canónica”: $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$, $\vec{k} = (0, 0, 1)$. En un espacio curvo, o un espacio plano en coordenadas no triviales, es necesario elegir una base antes de poder hablar de las componentes de un objeto geométrico. Una base común (aunque desde luego no la única) es la llamada “base coordenada”, en la que se toman como elementos de la base a aquellos vectores que tienen una componente igual a 1 a lo largo de la dirección de una coordenada dada, y componentes iguales a 0 en las otras direcciones.

Por ejemplo, en un espacio plano en coordenadas esféricas $\{r, \theta, \phi\}$ la base coordenada es:

$$\vec{e}_r = (1, 0, 0), \quad \vec{e}_\theta = (0, 1, 0), \quad \vec{e}_\phi = (0, 0, 1). \quad (1.20)$$

Escritos en coordenadas cartesianas, estos vectores resultan ser:

$$\vec{e}_r = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (1.21)$$

$$\vec{e}_\theta = (r \cos \theta \cos \phi, r \cos \theta \sin \phi, -r \sin \phi), \quad (1.22)$$

$$\vec{e}_\phi = (-r \sin \theta \sin \phi, r \sin \theta \cos \phi, 0). \quad (1.23)$$

Nótese que estos vectores no son todos unitarios, sus magnitudes son:

$$|\vec{e}_r|^2 = (\sin \theta \cos \phi)^2 + (\sin \theta \sin \phi)^2 + (\cos \theta)^2 = 1 \quad (1.24)$$

$$|\vec{e}_\theta|^2 = (r \cos \theta \cos \phi)^2 + (r \cos \theta \sin \phi)^2 + (r \sin \phi)^2 = r^2, \quad (1.25)$$

$$|\vec{e}_\phi|^2 = (r \sin \theta \sin \phi)^2 + (r \sin \theta \cos \phi)^2 = r^2 \sin^2 \theta, \quad (1.26)$$

donde para calcular dichas magnitudes simplemente hemos sumado los cuadrados de las componentes cartesianas. Como hemos visto, la magnitud de un vector puede también calcularse directamente a partir de las componentes esféricas utilizando el tensor métrico:

$$|\vec{v}|^2 = g_{\alpha\beta} v^\alpha v^\beta. \quad (1.27)$$

No es difícil ver que si utilizamos la expresión anterior y la métrica del espacio plano en coordenadas esféricas (ecuación (1.13)) obtenemos las mismas magnitudes para los vectores de la base coordenada esférica que escribimos arriba.

Debido a que los vectores de la base coordenada tienen por definición componentes 1 para la coordenada correspondiente y 0 para las demás, uno encuentra para el producto escalar entre los vectores base:

$$\vec{e}_\alpha \cdot \vec{e}_\beta = g_{\alpha\beta}. \quad (1.28)$$

Esto muestra que las componentes del tensor métrico son precisamente el producto escalar de los vectores base.

Es importante notar que en las expresiones anteriores, v^α se refiere a las componentes del vector \vec{v} respecto a las coordenadas x^α , mientras que \vec{e}_α se refiere al vector de la base coordenada que apunta en la dirección x^α . Esto implica que en general:

$$\vec{v} = v^\alpha \vec{e}_\alpha, \quad (1.29)$$

ecuación que expresa al vector \vec{v} como una combinación lineal de los elementos de la base $\{\vec{e}_\alpha\}$ con coeficientes v^α .

La base coordenada, aunque no única, si es la más comúnmente utilizada, y las expresiones para el tensor de curvatura de Riemann que se presentaron en la sección anterior son precisamente las que se obtienen utilizando esta base.

Consideremos ahora el cambio de un vector a lo largo de una coordenada:

$$\frac{\partial \vec{v}}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha} (v^\beta \vec{e}_\beta) = \frac{\partial v^\beta}{\partial x^\alpha} \vec{e}_\beta + v^\beta \frac{\partial \vec{e}_\beta}{\partial x^\alpha}. \quad (1.30)$$

Esta ecuación muestra como la derivada de un vector es más que la simple derivada de sus componentes. Debemos también tomar en cuenta el cambio en los vectores de la base.

La derivada $\partial \vec{e}_\beta / \partial x^\alpha$ es a su vez un vector, por lo que puede expresarse como combinación lineal de los vectores de la base. Introducimos los símbolos $\Gamma_{\alpha\beta}^\mu$ para denotar los coeficientes de dicha combinación lineal:

$$\frac{\partial \vec{e}_\beta}{\partial x^\alpha} = \Gamma_{\alpha\beta}^\mu \vec{e}_\mu. \quad (1.31)$$

Los coeficientes $\Gamma_{\alpha\beta}^\mu$ son precisamente los símbolos de Christoffel que introdujimos anteriormente. Utilizando estos coeficientes tendremos:

$$\frac{\partial \vec{v}}{\partial x^\alpha} = \frac{\partial v^\beta}{\partial x^\alpha} \vec{e}_\beta + v^\beta \Gamma_{\alpha\beta}^\mu \vec{e}_\mu. \quad (1.32)$$

Cambiando el nombre de los índices sumados (lo que siempre puede hacerse), podemos reescribir esto como:

$$\frac{\partial \vec{v}}{\partial x^\alpha} = \left(\frac{\partial v^\beta}{\partial x^\alpha} + v^\mu \Gamma_{\alpha\mu}^\beta \right) \vec{e}_\beta. \quad (1.33)$$

Esta ecuación nos da directamente las componentes del vector $\partial \vec{v} / \partial x^\alpha$. Definimos ahora la “derivada covariante” del vector \vec{v} como:

$$v^\alpha_{;\beta} \equiv \nabla_\beta v^\alpha := \frac{\partial v^\alpha}{\partial x^\beta} + v^\mu \Gamma_{\beta\mu}^\alpha. \quad (1.34)$$

Nótese que hemos introducido dos notaciones distintas para la derivada covariante: En un caso con un punto y coma, y en otro con el símbolo de gradiente (nabla). Ambas notaciones son comunes y se utilizan indistintamente.

La derivada covariante nos dice como cambian las componentes de un vector al movernos en un espacio general, e incluye el cambio de los elementos de la base. Nótese que la derivada covariante se reduce a la derivada parcial cuando los símbolos de Christoffel son cero, lo que ocurre en un espacio plano en coordenadas cartesianas, pero no en coordenadas esféricas. Sin embargo, siempre es posible encontrar una transformación de coordenadas para la cual los símbolos de Christoffel son iguales a cero en un punto dado (pero no en otros puntos excepto si el espacio es plano). Esto se debe a que todo espacio curvo es “localmente plano”: en una región infinitesimalmente cercana a todo punto la geometría se aproxima a la plana (es por eso que los mapas de una ciudad en un plano son muy buenos, mientras que los mapas del mundo entero introducen distorsiones serias).

La derivada covariante de un vector con los índices abajo v_α resultan ser:

$$v_{\alpha;\beta} = \frac{\partial v_\alpha}{\partial x^\beta} - \Gamma_{\alpha\beta}^\mu v_\mu. \quad (1.35)$$

El mismo concepto de derivada covariante puede extenderse a tensores de muchos índices, la regla es añadir un término con símbolos de Christoffel por cada índice libre, con el signo adecuado dependiendo de si el índice esta arriba o abajo. Por ejemplo:

$$\begin{aligned} T^{\mu\nu}{}_{;\alpha} &= \partial_\alpha T^{\mu\nu} + \Gamma_{\alpha\beta}^\mu T^{\beta\nu} + \Gamma_{\alpha\beta}^\nu T^{\mu\beta}, \\ T_{\mu\nu}{}_{;\alpha} &= \partial_\alpha T_{\mu\nu} - \Gamma_{\alpha\mu}^\beta T_{\beta\nu} - \Gamma_{\alpha\nu}^\beta T_{\mu\beta}, \\ T^\mu{}_{\nu;\alpha} &= \partial_\alpha T^\mu{}_\nu + \Gamma_{\alpha\beta}^\mu T^\beta{}_\nu - \Gamma_{\alpha\nu}^\beta T^\mu{}_\beta. \end{aligned}$$

Utilizando esta regla es posible mostrar que la derivada covariante del tensor métrico es siempre cero:

$$g_{\mu\nu;\alpha} = 0, \quad g^{\mu\nu}{}_{;\alpha} = 0, \quad (1.36)$$

lo que implica que la operación de subir y bajar índices conmuta con las derivadas covariantes:

$$v^\mu{}_{;\alpha} = (g^{\mu\nu} v_\nu)_{;\alpha} = g^{\mu\nu} (v_{\nu;\alpha}). \quad (1.37)$$

1.5. Las ecuaciones de campo de Einstein

El elemento que aún nos falta en la teoría de la gravitación de Einstein es aquel que nos dice como se relaciona la geometría del espacio-tiempo con la distribución de materia y energía. Este elemento final esta contenido en las “ecuaciones de campo de Einstein”, o

simplemente las “ecuaciones de Einstein”. Estas ecuaciones pueden derivarse de diversas maneras, ya sea buscando una generalización relativista y consistente de la ley de la gravitación de Newton (el camino que siguió Einstein), o derivándolas de manera formal a partir de un principio variacional partiendo de un Lagrangiano adecuado (el camino que siguió Hilbert casi simultáneamente). Aquí nos limitaremos a escribirlas sin derivación. En su forma más compacta, las ecuaciones de Einstein tienen la forma

$$G_{\mu\nu} = 8\pi T_{\mu\nu}, \quad (1.38)$$

donde $G_{\mu\nu}$ es el llamado “tensor de Einstein” que está relacionado con el tensor de curvatura de Ricci (ver abajo), y $T_{\mu\nu}$ es el llamado “tensor de energía-momento” de la materia. Es decir, el lado izquierdo representa la geometría del espacio-tiempo, y el lado derecho la distribución de materia y energía. El factor de 8π es simplemente una normalización necesaria para obtener el límite newtoniano correcto. Nótese que hay 10 ecuaciones independientes en la expresión anterior, pues los índices μ y ν toman valores de 0 a 3, lo que da $4 \times 4 = 16$ ecuaciones, pero los tensores de Einstein y de energía-momento son simétricos en sus dos índices, lo que reduce a 10 el número de ecuaciones independientes.

Las ecuaciones de Einstein que acabamos de introducir no podrían parecer más simples. Esta simplicidad, sin embargo, es solo aparente pues cada término es en realidad una abreviación que representa objetos muy complejos. Escritas de su manera más extensa, en un sistema de coordenadas arbitrario, las ecuaciones de Einstein son 10 ecuaciones diferenciales parciales acopladas en 4 coordenadas, y tienen miles de términos.

Consideremos ahora cada término en las ecuaciones de Einstein por separado. El tensor de Einstein se define en términos del tensor de Ricci como

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \quad (1.39)$$

con $R := g^{\mu\nu} R_{\mu\nu}$ la traza del tensor de Ricci, también llamada el “escalar de curvatura”.

La segunda parte de las ecuaciones de Einstein, el tensor de energía-momento, describe la densidad de energía, la densidad de momento, y el flujo de momento de un campo de materia ($i, j = 1, 2, 3$):

$$T^{00} = \text{densidad de energía}, \quad (1.40)$$

$$T^{0i} = \text{densidad de momento}, \quad (1.41)$$

$$T^{ij} = \text{flujo de momento } i \text{ a través de la superficie } j. \quad (1.42)$$

Por ejemplo, para un fluido perfecto sin presión (llamado “polvo” en el lenguaje relativista) en un espacio plano tenemos:

$$T^{00} = \rho/(1 - v^2), \quad (1.43)$$

$$T^{0i} = \rho v^i / (1 - v^2), \quad (1.44)$$

$$T^{ij} = \rho v^i v^j / (1 - v^2), \quad (1.45)$$

donde ρ es la densidad de energía medida en el marco de referencia de un elemento de fluido, v^i es el campo de velocidad del fluido y $v^2 = v_x^2 + v_y^2 + v_z^2$.

Es importante recalcar que en relatividad todos los campos excepto el gravitacional se consideran como un tipo materia. Por ejemplo, un campo electromagnético tiene un tensor de energía-momento asociado que actúa como fuente de la gravedad, es decir, los campos eléctricos y magnéticos producen gravedad.

Antes de terminar con nuestra discusión de las ecuaciones de Einstein, hace falta un último comentario. En el caso del espacio vacío, el tensor de energía-momento es cero, y las ecuaciones de Einstein se reducen a

$$G_{\mu\nu} = 0, \quad (1.46)$$

o de forma equivalente

$$R_{\mu\nu} = 0. \quad (1.47)$$

Nótese que, como ya hemos mencionado, el hecho de que el tensor de Ricci sea cero no significa que el espacio sea plano. Esto es como debe ser, pues sabemos que el campo gravitacional de un objeto se extiende más allá del objeto mismo, por lo que la curvatura del espacio en una región del vacío cercana a un objeto masivo no puede ser cero. Las ecuaciones de Einstein en el vacío tienen otra aplicación importante, describen la forma en la que el campo gravitacional se propaga en el vacío y, de manera análoga a las ecuaciones de Maxwell, predicen la existencia de las ondas gravitacionales: perturbaciones del campo gravitacional que viajan a la velocidad de la luz. La predicción de la existencia de las ondas gravitacionales nos dice que en la teoría de Einstein las interacciones gravitacionales no se propagan a velocidad infinita, sino que lo hacen a la velocidad de la luz.

1.6. Identidades de Bianchi y leyes de conservación

A partir de la definición del tensor de curvatura de Riemann es posible demostrar que este tensor tiene la siguiente propiedad:

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0. \quad (1.48)$$

Las relaciones anteriores se conocen como las “identidades de Bianchi”. Estas identidades resultan muy importantes en la relatividad general. Una de sus consecuencias más

importantes es el hecho de que la divergencia covariante del tensor de Einstein es igual a cero:

$$G^{\mu\nu}{}_{;\nu} = 0. \quad (1.49)$$

El tensor de Einstein es la única combinación que puede obtenerse a partir del tensor de Ricci que tiene esta propiedad, y es precisamente por esto que las ecuaciones de Einstein involucran a este tensor y no al tensor de Ricci directamente. Si utilizamos ahora las ecuaciones de Einstein vemos que la propiedad anterior implica inmediatamente:

$$T^{\mu\nu}{}_{;\nu} = 0. \quad (1.50)$$

Esta última ecuación (4 ecuaciones en realidad) es de fundamental importancia, pues representa las leyes locales de conservación de la energía y el momento, y garantiza que la pérdida de energía y momento en una región esta compensada por el flujo de energía y momento fuera de dicha región. En el caso de un fluido, por ejemplo, la componente $\mu = 0$ de estas ecuaciones se transforma en la bien conocida “ecuación de continuidad”. Podemos ver, entonces, como las ecuaciones de Einstein implican automáticamente a las leyes de conservación de energía y momento.

2. El formalismo de la relatividad numérica

2.1. Introducción

Existen varios formalismos diferentes utilizados en la relatividad numérica. En cada caso, lo que debe hacerse es separar las ecuaciones de campo de Einstein de forma tal que podamos dar ciertas condiciones iniciales, y a partir de ellas obtener la evolución del campo gravitacional. Los diferentes formalismos difieren en la forma específica en que se hace esta separación. En estas notas nos limitaremos a estudiar el llamado “formalismo 3+1”, en donde se hace una separación entre las tres dimensiones del espacio, por un lado, y el tiempo, por otro. El formalismo 3+1 es el más comúnmente utilizado en la relatividad numérica, pero no es el único. Formalismos alternativos son el llamado “formalismo característico” [65], donde el espacio-tiempo se separa en conos de luz centrados en una cierta trayectoria temporaloide, y el “formalismo conforme” [27], donde se utiliza una transformación conforme que permite tener la frontera del espacio-tiempo, es decir el “infinito nulo asintótico”, a una distancia finita. Los diferentes formalismos tienen ventajas y desventajas dependiendo del tipo de sistema físico que se quiera estudiar.

Las siguientes secciones se introduce el formalismo 3+1 de la relatividad general. La discusión que aquí se encuentra puede verse con mayor detalle en las referencias [42, 67].

2.2. El formalismo 3+1

Para estudiar la evolución en el tiempo de cualquier sistema físico, lo primero que debe hacerse es formular dicha evolución como un “problema de valores iniciales” o “problema de Cauchy”: Dadas condiciones iniciales (y de frontera) adecuadas, las ecuaciones fundamentales deben poder predecir la evolución futura (o pasada) del sistema.

Al intentar escribir a las ecuaciones de Einstein como un problema de Cauchy nos enfrentamos a un problema: las ecuaciones de Einstein están escritas en forma tal que el espacio y el tiempo son simétricos y juegan papeles equivalentes. Esta “covariancia” de las ecuaciones es importante (y elegante) desde el punto de vista teórico, pero no permite pensar claramente en la evolución en el tiempo del campo gravitacional. Lo primero que debemos hacer para reescribir a las ecuaciones de Einstein como un problema de Cauchy es separar los papeles del espacio y el tiempo de forma clara. A la formulación de la relatividad general que resulta de esta separación se le conoce como el “formalismo 3+1”.

Consideremos un espacio-tiempo con una métrica $g_{\alpha\beta}$. Si dicho espacio-tiempo puede ser totalmente “foliado” (es decir, separado en cortes tridimensionales) de tal forma que cada hoja tridimensional sea de tipo espacialoide, entonces se dice que dicho espacio-tiempo es “globalmente hiperbólico”. Un espacio-tiempo globalmente hiperbólico no posee

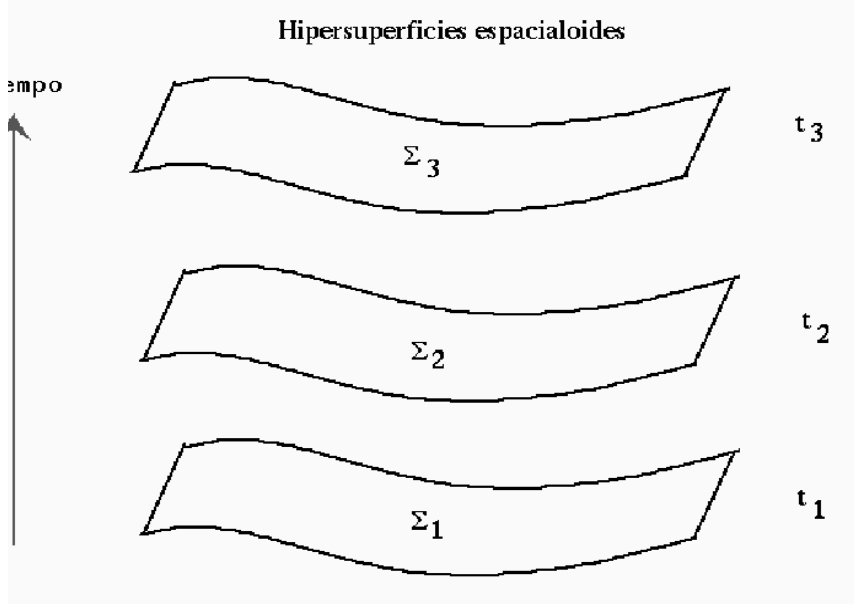


Figura 3: Foliación del espacio-tiempo en hiper-superficies tridimensionales de tipo espacialoide.

curvas temporaloides cerradas, lo que significa que no permite viajes hacia atrás en el tiempo. No todos los espacio-tiempos posibles tienen esta propiedad, pero en el formalismo 3+1 se asume que los espacio-tiempos físicamente razonables son de este tipo.

Una vez que se tiene un espacio-tiempo globalmente hiperbólico, por definición podemos foliarlo en una serie de hiper-superficies de tipo espacialoide (ver figura 3). Esta foliación en general no es única. Definimos el parámetro t como aquel que identifica a las distintas hojas de la foliación; t puede considerarse entonces como un “tiempo universal” (pero cuidado, t no tiene por qué coincidir con el tiempo propio de nadie). Consideremos ahora una cierta foliación, y tomemos dos hiper-superficies infinitesimalmente cercanas Σ_t y Σ_{t+dt} . La geometría de la región del espacio-tiempo contenida entre ambas hiper-superficies puede determinarse a partir de los siguientes 3 ingredientes (véase figura 4):

- La métrica tridimensional γ_{ij} ($i, j = 1, 2, 3$) que mide las distancias dentro de la hiper-superficie misma:

$$dl^2 = \gamma_{ij} dx^i dx^j. \quad (2.1)$$

- El lapso de tiempo propio entre ambas hiper-superficies que mide un observador que se mueve en la dirección normal a ellas (observador de Euler):

$$d\tau = \alpha(t, x^i) dt. \quad (2.2)$$

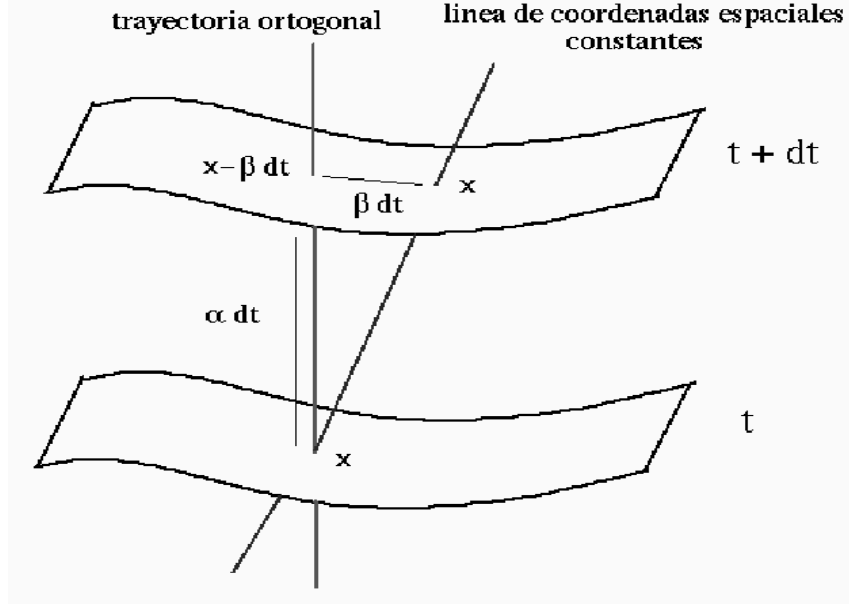


Figura 4: Dos hiper-superficies espacialoides infinitesimalmente cercanas. La figura muestra las definiciones de la función de lapso α y el vector de corrimiento β^i .

La función α que aparece en esta expresión se conoce como la “función de lapso”.

- La velocidad relativa β^i entre los observadores de Euler y las líneas con coordenadas espaciales constantes:

$$x_{t+dt}^i = x_t^i - \beta^i(t, x^j) dt, \quad \text{para observadores de Euler} \quad (2.3)$$

Al 3-vector β^i se le llama el “vector de corrimiento” (nótese el signo menos en la definición anterior).

Nótese que la manera en la que se hace la foliación no es única, y de la misma forma, la manera en la que se propaga el sistema de coordenadas espacial de una superficie a otra tampoco lo es. Esto significa que tanto la función de lapso α como el vector de corrimiento β^i son funciones que pueden especificarse libremente. Estas funciones determinan nuestro sistema de coordenadas y se conocen como “funciones de norma”.

En términos de las funciones $\{\alpha, \beta^i, \gamma_{ij}\}$, la métrica del espacio-tiempo toma la siguiente forma:

$$ds^2 = (-\alpha^2 + \beta_i \beta^i) dt^2 + 2 \beta_i dt dx^i + \gamma_{ij} dx^i dx^j, \quad (2.4)$$

donde hemos definido $\beta_i := \gamma_{ij} \beta^j$ (de aquí en adelante asumimos que los índices de tensores espaciales se suben y bajan con la 3- métrica γ_{ij}).

En forma explícita tenemos:

$$g_{\mu\nu} = \begin{pmatrix} -\alpha^2 + \beta_k \beta^k & \beta_i \\ \beta_j & \gamma_{ij} \end{pmatrix}, \quad (2.5)$$

$$g^{\mu\nu} = \begin{pmatrix} -1/\alpha^2 & \beta^i/\alpha^2 \\ \beta^j/\alpha^2 & \gamma^{ij} - \beta^i \beta^j/\alpha^2 \end{pmatrix}. \quad (2.6)$$

De la misma forma, no es difícil mostrar que las componentes del vector normal unitario n^μ a las hiper-superficies son:

$$n^\mu = (1/\alpha, -\beta^i/\alpha), \quad n_\mu = (-\alpha, 0), \quad n^\mu n_\mu = -1. \quad (2.7)$$

Nótese que este vector normal unitario corresponde por definición a la cuatro-velocidad de los observadores de Euler.

2.3. Curvatura extrínseca

Al hablar de las hiper-superficies espaciales que forman la foliación del espacio-tiempo, debemos distinguir entre la curvatura “intrínseca” de dichas hiper-superficies proveniente de su geometría interna, y su curvatura “extrínseca” relacionada con la forma en que éstas se encuentran inmersas en el espacio-tiempo de 4 dimensiones.

La curvatura intrínseca estará dada por el tensor de Riemann tridimensional que se define en términos de la métrica espacial γ_{ij} . La curvatura extrínseca, por otro lado, se define en términos de lo que le ocurre al vector normal n^α al transportarlo paralelamente de un sitio a otro de la superficie. En general encontraremos que al transportar paralelamente este vector a un punto cercano, el nuevo vector ya no será normal a la superficie. El “tensor de curvatura extrínseca” $K_{\alpha\beta}$ es una medida del cambio en el vector normal bajo transporte paralelo.

Antes de definir la curvatura extrínseca, resulta necesario introducir el operador de “proyección” P_β^α a las hipersuperficies espaciales:

$$P_\beta^\alpha := \delta_\beta^\alpha + n^\alpha n_\beta, \quad (2.8)$$

donde n^α es el vector unitario normal a las hiper-superficies. Es fácil mostrar que para cualquier vector v^α tenemos:

$$(P_\beta^\alpha v^\beta) n_\alpha = 0, \quad (2.9)$$

(basta recordar que $n^\alpha n_\alpha = -1$). Esto significa que todo vector proyectado a la hipersuperficie es ortogonal a n^α . Es posible también proyectar tensores con varios índices, basta contraer todos los índices libres con el operador de proyección:

$$P T_{\alpha\beta} \equiv P_\alpha^\mu P_\beta^\nu T_{\mu\nu} . \quad (2.10)$$

Usando el operador de proyección, el tensor de curvatura extrínseca se define como:

$$K_{\alpha\beta} := -P \nabla_\alpha n_\beta = -(\nabla_\alpha n_\beta + n_\alpha n^\mu \nabla_\mu n_\beta) , \quad (2.11)$$

donde ∇_α es la derivada covariante que como hemos visto esta dada en términos de los símbolos de Christoffel como

$$\nabla_\alpha n_\beta = \partial_\alpha n_\beta - \Gamma_{\alpha\beta}^\lambda n_\lambda . \quad (2.12)$$

Dada la definición anterior, es posible demostrar que $K_{\alpha\beta}$ es un tensor simétrico: $K_{\alpha\beta} = K_{\beta\alpha}$. También es claramente un tensor puramente espacial, es decir, $n^\alpha K_{\alpha\beta} = n^\alpha K_{\beta\alpha} = 0$ (nótese que esto significa que $K^{00} = K^{0i} = 0$, pero K_{00} y K_{0i} en general *no* son cero). Debido a esto de ahora en adelante utilizaremos solamente sus componentes espaciales K_{ij} .¹

Substituyendo la forma explícita del vector normal (2.7) en la definición de la curvatura extrínseca es posible demostrar que K_{ij} está dado en términos de la métrica espacial como:

$$K_{ij} = \frac{1}{2\alpha} [-\partial_t \gamma_{ij} + D_i \beta_j + D_j \beta_i] , \quad (2.13)$$

donde D_i representa la derivada covariante tridimensional, es decir, asociada a la 3-métrica γ_{ij} . La ecuación anterior puede reescribirse como:

$$\partial_t \gamma_{ij} = -2\alpha K_{ij} + D_i \beta_j + D_j \beta_i . \quad (2.14)$$

¹Es necesario hacer dos comentarios acerca de la definición de $K_{\alpha\beta}$. En primer lugar, la proyección de $\nabla_\alpha n_\beta$ es crucial si se desea que $K_{\alpha\beta}$ sea puramente espacial. Podría argumentarse que debido a que n^α es unitario, su gradiente es siempre ortogonal a si mismo. Esto es correcto en el sentido de que $n^\alpha \nabla_\beta n_\alpha = 0$, pero ∇_α no es en general simétrico y $n^\alpha \nabla_\alpha n_\beta \neq 0$ a menos que las líneas normales sean geodésicas (lo que no siempre ocurre). Consideremos ahora la simetría de $K_{\alpha\beta}$, como hemos mencionado $\nabla_\alpha n_\beta$ en general no es simétrico pese a que el vector n^α es ortogonal a una hipersuperficie. La razón por la que esto ocurre es que n^α es unitario, por lo que en general no es igual al gradiente de la función de tiempo t , excepto cuando el lapso es igual a 1. Sin embargo, una vez habiendo proyectado hacia la hipersuperficie, resulta que $P \nabla_\alpha n_\beta$ sí es simétrico (la asimetría de $\nabla_\alpha n_\beta$ tiene que ver con el lapso, que no es intrínseco a la hipersuperficie).

Podemos ver entonces como la curvatura extrínseca K_{ij} está relacionada con el cambio en el tiempo de la métrica espacial. Esta asociación se puede hacer más precisa utilizando el concepto de “derivada de Lie”, que para escalares, vectores y co-vectores se define como:

$$\mathcal{L}_{\vec{u}} \phi = u^\mu \partial_\mu \phi , \quad (2.15)$$

$$\mathcal{L}_{\vec{u}} v^\alpha = u^\mu \partial_\mu v^\alpha - v^\mu \partial_\mu u^\alpha , \quad (2.16)$$

$$\mathcal{L}_{\vec{u}} \omega_\alpha = u^\mu \partial_\mu \omega_\alpha + \omega_\mu \partial_\alpha u^\mu . \quad (2.17)$$

La derivada de Lie es un concepto de derivada diferente a la diferenciación covariante, independiente de la métrica y los símbolos de Christoffel, y que se define en términos de una congruencia de líneas con vector tangente \vec{u} . La congruencia define un mapeo de los vectores base de la siguiente forma: 1) en un punto dado, considérense las líneas que tienen como tangente a los vectores base, 2) ahora arrástrense esas líneas a lo largo de la congruencia moviéndolas la misma “distancia”, medida como el cambio en el valor del parámetro que define la congruencia, 3) los nuevos vectores base son los vectores tangentes a las líneas arrastradas. Generalmente, los vectores base arrastrados no coinciden con los vectores base en el nuevo punto. La derivada de Lie de un tensor se define como la diferencia entre el valor del tensor aplicado a los vectores base arrastrados, y el valor del tensor aplicado a los vectores base originales en ese punto.

La derivada de Lie también puede entenderse de otra manera. Si construimos un sistema de coordenadas tal que las líneas coordenadas de una coordenada particular coincidan con nuestra congruencia, entonces la derivada de Lie es simplemente la derivada parcial respecto a esa coordenada. La derivada de Lie es entonces una forma de definir las derivadas parciales independientemente de las coordenadas.

Para un tensor con dos índices abajo, la derivada de Lie resulta ser:

$$\mathcal{L}_{\vec{u}} T_{\alpha\beta} = u^\mu \partial T_{\alpha\beta} + T_{\alpha\mu} \partial_\beta u^\mu + T_{\mu\beta} \partial_\alpha u^\mu . \quad (2.18)$$

En particular, para el tensor métrico se tiene:

$$\mathcal{L}_{\vec{u}} g_{\alpha\beta} = u^\mu \partial g_{\alpha\beta} + g_{\alpha\mu} \partial_\beta u^\mu + g_{\mu\beta} \partial_\alpha u^\mu . \quad (2.19)$$

Por otro lado:

$$\begin{aligned} \nabla_\alpha u_\beta + \nabla_\beta u_\alpha &= \partial_\alpha u_\beta + \partial_\beta u_\alpha - 2\Gamma_{\alpha\beta}^\mu u_\mu \\ &= \partial_\alpha u_\beta + \partial_\beta u_\alpha + u^\mu \partial_\mu g_{\alpha\beta} - u^\mu \partial_\alpha g_{\beta\mu} - u^\mu \partial_\beta g_{\alpha\mu} \\ &= u^\mu \partial_\mu g_{\alpha\beta} + g_{\beta\mu} \partial_\alpha u^\mu + g_{\alpha\mu} \partial_\beta u^\mu , \end{aligned} \quad (2.20)$$

de donde obtenemos ²:

$$\mathcal{L}_{\vec{u}} g_{\alpha\beta} = \nabla_{\alpha} u_{\beta} + \nabla_{\beta} u_{\alpha} . \quad (2.21)$$

Comparando este último resultado con (2.14) vemos que esta ecuación se puede reescribir como:

$$\left(\partial_t - \mathcal{L}_{\vec{\beta}} \right) \gamma_{ij} = -2\alpha K_{ij} . \quad (2.22)$$

De nuestra discusión la derivada de Lie sabemos que $\partial_t = \mathcal{L}_{\vec{t}}$, con \vec{t} el vector tangente a las líneas coordinadas de t . De la figura 4 vemos que este vector está dado por $\vec{t} = \alpha\vec{n} + \vec{\beta}$. El resultado final es entonces:

$$\mathcal{L}_{\alpha\vec{n}} \gamma_{ij} = -2\alpha K_{ij} . \quad (2.23)$$

Esta ecuación muestra que la curvatura extrínseca es precisamente la derivada temporal de la métrica espacial vista por los observadores de Euler.

Esto nos permite llegar a la mitad del camino necesario para reescribir la relatividad general como un problema de Cauchy: ya tenemos una ecuación de evolución para γ_{ij} . Pero para cerrar el sistema aún nos falta una ecuación de evolución para K_{ij} . Es importante notar que hasta ahora hemos trabajado solo con conceptos geométricos, y no hemos utilizado las ecuaciones de campo de Einstein. Es precisamente a partir de las ecuaciones de Einstein de donde obtendremos la ecuación de evolución para K_{ij} . Dicho de otra forma, la ecuación de evolución (2.14) para γ_{ij} es meramente cinemática, mientras que la ecuación de evolución para K_{ij} contendrá la información dinámica.

2.4. Las ecuaciones de Einstein en el formalismo 3+1

Para escribir las ecuaciones de Einstein en lenguaje 3+1, vamos a utilizar al operador de proyección, junto con el vector normal, para separar las ecuaciones de Einstein en 3 grupos:

- Proyección normal (1 ecuación):

$$n^{\alpha} n^{\beta} (G_{\alpha\beta} - 8\pi T_{\alpha\beta}) = 0 . \quad (2.24)$$

- Proyección a la hiper-superficie (6 ecuaciones):

$$P (G_{\alpha\beta} - 8\pi T_{\alpha\beta}) = 0 . \quad (2.25)$$

²Este resultado se puede derivar mucho más rápido si notamos que para el cálculo de derivadas de Lie, las derivadas parciales se pueden substituir por derivadas covariantes pues los símbolos de Christoffel siempre se cancelan. Recordando que la derivada covariante de la métrica es cero encontramos el resultado deseado

- Proyección mixta (3 ecuaciones):

$$P [n^\alpha (G_{\alpha\beta} - 8\pi T_{\alpha\beta})] = 0. \quad (2.26)$$

Para expresar estos conjuntos de ecuaciones en el lenguaje 3+1, es necesaria un álgebra larga que no haremos en estas notas. Aquí nos limitaremos a señalar los resultados finales.

De la proyección normal obtenemos la siguiente ecuación:

$$R^{(3)} + (\text{tr } K)^2 - K_{ij}K^{ij} = 16\pi\rho. \quad (2.27)$$

donde $R^{(3)}$ es el escalar de curvatura de la métrica espacial, $\text{tr } K \equiv \gamma^{ij}K_{ij}$ es la traza del tensor de curvatura extrínseca, y ρ es la densidad de energía de la materia medida por los observadores de Euler:

$$\rho := n_\alpha n_\beta T^{\alpha\beta}. \quad (2.28)$$

La ecuación (2.27) no tiene derivadas temporales (aunque sí tiene derivadas espaciales de γ_{ij} dentro del escalar de Ricci). Debido a esto no es una ecuación de evolución sino una “constricción” o “ligadura” (a veces también llamada “vínculo”) del sistema. Como está relacionada con la densidad de energía ρ , a esta constricción se le conoce como la “constricción de energía”, o la “constricción hamiltoniana”.

De la proyección mixta de las ecuaciones de Einstein obtenemos:

$$D_j [K^{ij} - \gamma^{ij}\text{tr } K] = 8\pi j^i, \quad (2.29)$$

donde j^i es el “flujo de momento” medido por los observadores de Euler:

$$j^i := P_\beta^i (n_\alpha T^{\alpha\beta}). \quad (2.30)$$

La ecuación (2.29) tampoco tiene derivadas temporales, por lo que es otra constricción (de hecho 3 constricciones). Estas ecuaciones se conocen como “constricciones de momento”.

La existencia de las constricciones implica que en la relatividad general no es posible especificar de manera arbitraria las 12 cantidades dinámicas $\{\gamma_{ij}, K_{ij}\}$ como condiciones iniciales. Las constricciones deben satisfacerse ya desde el inicio, o no estaremos resolviendo las ecuaciones de Einstein.

Las últimas 6 ecuaciones se obtienen de la proyección a la hiper-superficie y contienen la verdadera dinámica del sistema. Estas ecuaciones son:

$$\begin{aligned} \partial_t K_{ij} &= \beta^a D_a K_{ij} + K_{ia} D_j \beta^a + K_{ja} D_i \beta^a \\ &\quad - D_i D_j \alpha + \alpha [R_{ij}^{(3)} - 2K_{ia} K_j^a + K_{ij} \text{tr } K] \\ &\quad + 4\pi\alpha [\gamma_{ij} (\text{tr } S - \rho) - 2S_{ij}], \end{aligned} \quad (2.31)$$

donde S_{ij} es el “tensor de esfuerzos” de la materia definido como:

$$S_{ij} := P T_{ij} . \quad (2.32)$$

Nótese que, de la misma manera que ocurrió con las ecuaciones de evolución para la 3-métrica, las ecuaciones anteriores pueden reescribirse como:

$$\begin{aligned} \left(\partial_t - \mathcal{L}_{\vec{\beta}} \right) K_{ij} = & -D_i D_j \alpha + \alpha \left[R_{ij}^{(3)} - 2K_{ia} K_j^a + K_{ij} \operatorname{tr} K \right] \\ & + 4\pi\alpha \left[\gamma_{ij} (\operatorname{tr} S - \rho) - 2S_{ij} \right] , \end{aligned} \quad (2.33)$$

Las ecuaciones (2.14) y (2.31) forman un sistema cerrado de ecuaciones de evolución. A estas ecuaciones se les conoce como las ecuaciones de Arnowitt-Deser-Misner, o simplemente las ecuaciones ADM [13, 67]. Estas ecuaciones finalmente nos permiten escribir las ecuaciones de campo de la relatividad general como un problema de Cauchy.

Es importante notar que no tenemos ecuaciones de evolución para las variables de norma $\{\alpha, \beta^i\}$. Como hemos dicho antes, las variables de norma se pueden elegir libremente.

Es posible demostrar también, usando las identidades de Bianchi, que las ecuaciones de evolución garantizan que si las constricciones se satisfacen al tiempo inicial, entonces se satisfecerán a todo tiempo posterior: Las ecuaciones de evolución propagan las constricciones.

2.5. El problema de los valores iniciales

La existencia de las constricciones en la relatividad general implica que no es posible elegir de manera arbitraria las 12 cantidades dinámicas $\{\gamma_{ij}, K_{ij}\}$ como condiciones iniciales. Los datos iniciales deben elegirse de tal modo que las constricciones se satisfagan desde el principio. Esto significa que antes de comenzar cualquier evolución, es necesario primero resolver el problema de los valores iniciales para obtener valores apropiados de $\{\gamma_{ij}, K_{ij}\}$ que representen la situación física de interés.

Las constricciones forman un conjunto de 4 ecuaciones diferenciales de tipo elíptico, y en general son difíciles de resolver. Existen, sin embargo, varios procedimientos conocidos para resolver estas ecuaciones en situaciones específicas. Hasta hace poco, el procedimiento más común era la llamada “descomposición conforme” de York-Lichnerowicz [38, 67]. Más recientemente, el llamado “formalismo del sándwich delgado” [66] también se ha vuelto muy popular para construir datos iniciales. En estas notas nos concentraremos en el primero de estos procedimientos debido a que es el mejor conocido y por lo mismo mejor entendido. Una reseña reciente sobre los diferentes procedimientos puede encontrarse en [21].

La descomposición conforme de York-Lichnerowicz parte de considerar una transformación de la 3-métrica del tipo:

$$\gamma_{ij} = \phi^4 \tilde{\gamma}_{ij} , \quad (2.34)$$

donde la métrica $\tilde{\gamma}_{ij}$ se considera como dada. Una transformación de este tipo se conoce como “ transformación conforme ” debido a que preserva los ángulos. A la métrica $\tilde{\gamma}_{ij}$ se le llama la “ métrica conforme ”.

Por otro lado, la curvatura extrínseca se separa primero en su traza:

$$\text{tr}K = \gamma^{ij} K_{ij} , \quad (2.35)$$

y su parte de traza cero A_{ij} :

$$A_{ij} := K_{ij} - \frac{1}{3} \gamma_{ij} \text{tr}K . \quad (2.36)$$

También se lleva a cabo una transformación conforme de A_{ij} de la forma:

$$A_{ij} = \phi^{-10} \tilde{A}_{ij} , \quad (2.37)$$

donde la décima potencia se elige por conveniencia, ya que simplifica las ecuaciones que se obtienen después. Ahora podemos utilizar el hecho de que cualquier tensor simétrico con traza cero puede separarse en dos partes de la siguiente manera:

$$\tilde{A}_{ij} = \tilde{A}_{ij}^* + (\hat{L}W)_{ij} , \quad (2.38)$$

donde \tilde{A}_{ij}^* es un tensor de divergencia cero $\tilde{D}^i \tilde{A}_{ij}^* = 0$ (y traza cero también), W_i es un vector, y \hat{L} es un operador diferencial definido como:

$$(\hat{L}W)_{ij} = \tilde{D}_i W_j + \tilde{D}_j W_i - \frac{2}{3} \tilde{\gamma}_{ij} \tilde{D}_k W^k , \quad (2.39)$$

con \tilde{D}_i la derivada covariante asociada con la métrica conforme γ_{ij} .

Para encontrar nuestros datos iniciales, asumimos ahora que las cantidades $\tilde{\gamma}_{ij}$, $\text{tr}K$ y \tilde{A}_{ij}^* están dadas, y utilizamos las cuatro ecuaciones de constricción para determinar las cuatro cantidades $\{\phi, W_i\}$.

No es difícil mostrar que la constricción hamiltoniana toma la forma:

$$8\tilde{D}^2\phi - \tilde{R}\phi + \phi^{-7} (\tilde{A}_{ij}\tilde{A}^{ij}) - \frac{2}{3} \phi^5 (\text{tr}K)^2 + 16\pi\tilde{\phi}^{-3}\rho = 0 , \quad (2.40)$$

lo que nos da una ecuación elíptica para ϕ (esencialmente la ecuación de Poisson). Es importante recordar que el operador laplaciano \tilde{D}^2 que aparece en esta ecuación debe ser el correspondiente a la métrica conforme.

Por otro lado, las constricciones de momento se convierten en

$$\tilde{D}^2 W^i - \frac{2}{3}\phi^6 \tilde{D}^i \text{tr} K - 8\pi \tilde{j}^i = 0, \quad (2.41)$$

que son tres ecuaciones elípticas acopladas para W^i .

Las ecuaciones anteriores son la forma más general de escribir las constricciones en la descomposición de York-Lichnerowicz. Nótese que las cuatro ecuaciones están acopladas entre si.

Una manera de simplificar notablemente el problema y desacoplar las ecuaciones es simplemente elegir $\text{tr} K = 0$. Si, además, escogemos la métrica conforme como la correspondiente a un espacio plano $\tilde{\gamma}_{ij} = \delta_{ij}$, las constricciones se reducen (en el vacío) a:

$$8 D_{\text{plano}}^2 \phi + \phi^{-7} (\tilde{A}_{ij} \tilde{A}^{ij}) = 0, \quad (2.42)$$

$$D_{\text{plano}}^2 W^i = 0, \quad (2.43)$$

donde D_{plano}^2 es simplemente el operador laplaciano estándar. La segunda ecuación es lineal y puede resolverse de forma analítica en muchos casos. Una vez teniendo la solución para W^i , se reconstruye \tilde{A}_{ij} y se resuelve la ecuación de Poisson para ϕ .

2.6. Datos iniciales para agujeros negros

Como primer ejemplo de la solución del problema de valores iniciales, consideremos el siguiente caso: una métrica conformemente plana $\tilde{\gamma}_{ij} = \delta_{ij}$ y un momento de simetría temporal, es decir $K_{ij} = 0$. En ese caso, las constricciones de momento se satisfacen trivialmente, y la condición hamiltoniana toma la forma sencilla:

$$D^2 \phi = 0, \quad (2.44)$$

que no es otra cosa que la ecuación de Laplace. Las condiciones de frontera corresponden a un espacio asintóticamente plano (muy lejos el campo gravitacional tiende a cero), por lo que en infinito debemos tener $\phi = 1$. La solución mas sencilla de esta ecuación que satisface la condición de frontera es entonces:

$$\phi = 1, \quad (2.45)$$

lo que implica que la métrica espacial es simplemente:

$$dl^2 = dx^2 + dy^2 + dz^2. \quad (2.46)$$

Es decir, hemos recuperado los datos iniciales para el espacio de Minkowski (aunque por una vía bastante elaborada).

La siguiente solución de interés es:

$$\phi = 1 + k/r, \quad (2.47)$$

con k una constante arbitraria, por lo que la métrica espacial es ahora (en coordenadas esféricas)

$$dl^2 = (1 + k/r)^4 [dr^2 + r^2 d\Omega^2]. \quad (2.48)$$

No es difícil mostrar que esto no es otra cosa que la métrica de un agujero negro de Schwarzschild en las llamadas “coordenadas isotrópicas”, donde la masa del agujero negro esta dada por $M = 2k$. Hemos encontrado la solución para el problema de valores iniciales correspondiente a un agujero negro de Schwarzschild.

Como la ecuación de Laplace es lineal, podemos superponer soluciones para obtener nuevas soluciones. Por ejemplo, la solución:

$$\phi = 1 + \sum_{i=1}^N \frac{M_i}{2|r - r_i|}, \quad (2.49)$$

representa a N agujeros negros en los puntos r_i inicialmente en reposo, y se conoce como los datos iniciales de Brill-Lindquist. Es posible generalizar dicha solución al caso en que los agujeros negros no estén en reposo, construyendo así, por ejemplo, datos iniciales para dos agujeros negros en órbita. Sin embargo en ese caso la solución de la constricción hamiltoniana ya no es tan sencilla y debe obtenerse por métodos numéricos.

Es importante notar que este tipo de datos iniciales para agujeros negros son modelos de tipo topológico: los agujeros negros están representados por agujeros de gusano (túneles de Einstein-Rosen) que unen nuestro Universo con otros. Esto ocurre en la solución de Schwarzschild y corresponde a un agujero negro “eterno” (no formado por colapso). En el mundo real, uno espera que los agujeros negros se formen por un colapso y el túnel de Einstein-Rosen no estará ahí.

2.7. Condiciones de norma

Como hemos mencionado, en el formalismo 3+1 la elección del sistema de coordenadas está dada en términos de las “variables de norma”: la función de lapso α y el vector de corrimiento β^i . Estas funciones aparecen en las ecuaciones de evolución (2.14) y (2.31) para la métrica y la curvatura extrínseca. Sin embargo, las ecuaciones de Einstein no nos dicen que valores deben tomar las variables de norma. Esto es lo que deberíamos esperar, pues es claro que las coordenadas deben poder elegirse libremente.

La libertad en la elección de las variables de norma es una bendición a medias. Por un lado, nos permite elegir las cosas de manera que las ecuaciones se simplifiquen, o de manera que la solución sea mejor comportada. Por otro lado, surge inmediatamente la pregunta: ¿Cuál es una buena elección de las funciones α y β^i ? Recuérdese que esta decisión debe tomarse incluso antes de comenzar la evolución. Existen, desde luego, infinitas elecciones posibles. En estas notas nos limitaremos a mencionar algunas de las más conocidas.

2.7.1. Condiciones de foliación

Consideremos en primer lugar la elección de la función de lapso α . La elección de esta función determina como avanza el tiempo propio entre distintas hiper-superficies espaciales, es decir, determina como esta rebanado el espacio-tiempo en superficies de coordenada t constante. Debido a esto, a la elección del lapso se le llama comúnmente una “condición de foliación”.

Antes de considerar diferentes condiciones de foliación, debemos mencionar un par de resultados que son muy importantes al hablar sobre este tema. En primer lugar, consideremos el movimiento de los observadores de Euler, es decir, aquellos que se mueven siguiendo las líneas normales a las hiper-superficies que forman la foliación. Nótese que no hay razón para suponer que estos observadores están en caída libre, por lo que en general pueden tener lo que se conoce como una “aceleración propia” que mide la fuerza requerida para mantenerlos en una trayectoria diferente a la caída libre. Esta aceleración está dada por la derivada direccional de la cuatro-velocidad de los observadores de Euler (el vector normal n^μ) a lo largo de si misma:

$$a^\mu = n^\nu \nabla_\nu n^\mu . \quad (2.50)$$

Nótese que esto involucra a la derivada covariante en 4 dimensiones ∇_μ . No es difícil mostrar que esta ecuación implica que la aceleración está dada en términos de la función de lapso por

$$a_i = \partial_i \ln \alpha . \quad (2.51)$$

Otra relación importante está relacionada con la evolución de los elementos de volumen asociados a los observadores de Euler. El cambio de estos elementos de volumen en el tiempo esta dado por la divergencia de las velocidades de estos observadores $\nabla_\mu n^\mu$. Usando la definición de la curvatura extrínseca encontramos ahora

$$\nabla_\mu n^\mu = -\text{tr} K , \quad (2.52)$$

es decir, el cambio de los elementos de volumen en el tiempo es la traza de la curvatura extrínseca.

Regresemos ahora al problema de como elegir la función de lapso. La manera más obvia de elegir esta función es simplemente decidir que queremos que el tiempo coordinado t coincida en todo punto con el tiempo propio de los observadores de Euler. Después de todo, ¿qué podría ser más natural? Esta elección corresponde a tomar $\alpha = 1$ en todo el espacio, y de la ecuación (2.51) vemos que en este caso la aceleración de los observadores de Euler es cero, es decir, están en caída libre y se mueven a lo largo de geodésicas. Debido a esto, a esta elección se le conoce como “foliación geodésica”.

La foliación geodésica, pese a parecer muy natural, es en la práctica una muy mala elección. La razón de esto es fácil de ver si pensamos que les ocurre observadores en caída libre en un campo gravitacional no uniforme. Al ser el campo no uniforme, diferentes observadores caen en distintas direcciones y nada impide que choquen entre ellos. Cuando esto sucede, nuestro sistema de coordenadas que esta atado a estos observadores deja de ser uno a uno: un mismo punto tiene ahora más de un valor de las coordenadas asociado a él. Esto significa que el sistema de coordenadas se ha vuelto singular. Debido a esto, la foliación geodésica no se utiliza.

Para buscar una mejor condición de lapso, debemos pensar cuál fue el problema con la foliación geodésica. El problema básico es que observadores en caída libre son “enfocados” por el campo gravitacional. Esto significa que se acercan unos a otros, por lo que los elementos de volumen disminuyen hasta hacerse cero. Podemos entonces intentar construir una foliación donde se exija que los elementos de volumen se mantengan constantes. De la ecuación (2.52) vemos que esto es equivalente a pedir:

$$\text{tr}K = \partial_t \text{tr}K = 0 . \quad (2.53)$$

Como se ve en esta expresión, es claro que debemos exigir no solo que $\text{tr}K$ sea cero inicialmente, sino que se mantenga cero durante la evolución. Por otro lado, a partir de las ecuaciones de evolución (2.14) y (2.31) encontramos

$$\partial_t \text{tr}K = \beta^i \partial_i \text{tr}K - D^2 \alpha + \alpha \left[K_{ij} K^{ij} + \frac{1}{2} (\rho + \text{tr}S) \right] , \quad (2.54)$$

donde se utilizó la la constricción hamiltoniana para eliminar el escalar de curvatura. Imponiendo ahora la condición (2.53) encontramos que la función de lapso debe satisfacer la siguiente ecuación elíptica:

$$D^2 \alpha = \alpha \left[K_{ij} K^{ij} + \frac{1}{2} (\rho + \text{tr}S) \right] . \quad (2.55)$$

Esta condición se conoce como “foliación maximal” [56]. Este nombre se debe a que puede demostrarse que esta es la condición que garantiza que el volumen de una región cualquiera de la hiper-superficie espacial se maximiza ante pequeñas variaciones.

La condición de foliación maximal se ha utilizado en muchas simulaciones numéricas de diferentes sistemas, incluyendo agujeros negros, y sigue siendo popular a la fecha. Sus ventajas están en que es una ecuación sencilla, cuya solución es suave (por ser producto de una ecuación elíptica), y que garantiza que los observadores de Euler no pueden enfocarse. En particular, la foliación maximal tiene la propiedad de evadir singularidades, tanto singularidades de coordenadas debidas al enfoque de observadores, como singularidades físicas como aquellas que se encuentran en el interior de un agujero negro. En el caso de un agujero negro, la solución de la condición maximal hace que la función de lapso se haga exponencialmente cero en la región interior al agujero negro, fenómeno que se conoce como el “colapso del lapso” (ver Fig. 5). El colapso del lapso resulta muy útil ya que de otra forma las hiper-superficies marcharían directo a la singularidad física en un tiempo muy corto, donde se encontrarían valores infinitos del campo gravitacional, lo que causaría que el código computacional se detenga.

Una de las desventajas de utilizar una foliación maximal es el hecho de que en tres dimensiones la solución de la ecuación elíptica (2.55) resulta muy costosa en tiempo de máquina. Debido a esto se han buscado condiciones de foliación alternativas que tengan propiedades similares a la foliación maximal pero que a la vez sean más fáciles de resolver. Una familia de condiciones de lapso que tiene estas propiedades es la llamada condición de Bona-Masso [16]. Esta familia especifica una ecuación de evolución para la función de lapso se la forma:

$$\partial_t \alpha - \beta^i \partial_i \alpha = -\alpha^2 f(\alpha) K , \quad (2.56)$$

donde $f(\alpha)$ es una función positiva pero por lo demás arbitraria de α (la razón por la que debe ser positiva tiene que ver con el concepto de hiperbolicidad que se discute en la Sección 3). El término con el vector de corrimiento β^i se incluye para formar la “derivada total” respecto al tiempo, es decir, la derivada a lo largo de la dirección normal.

La condición de Bona-Masso incluye a sub-familias muy conocidas. Por ejemplo, si se toma $f = 1$ se obtiene la llamada “foliación armónica”, para la cual la coordenada t resulta obedecer la ecuación de onda $g^{\mu\nu} \nabla_\mu \nabla_\nu t = 0$. En este caso la ecuación de evolución del lapso se puede integrar para obtener:

$$\alpha = h(x^i) \gamma^{1/2} . \quad (2.57)$$

donde $h(x^i)$ es una función independiente del tiempo, y γ es el determinante de la métrica espacial, es decir, $\gamma^{1/2}$ es el elemento de volumen.

Otra familia interesante se obtiene tomando $f(\alpha) = N/\alpha$ con N una constante positiva. Esto corresponde a la familia “1+log” [1, 12], cuyo nombre proviene de que en este caso la función del lapso resulta tener la forma:

$$\alpha = h(x^i) + \ln \left(\gamma^{N/2} \right) . \quad (2.58)$$

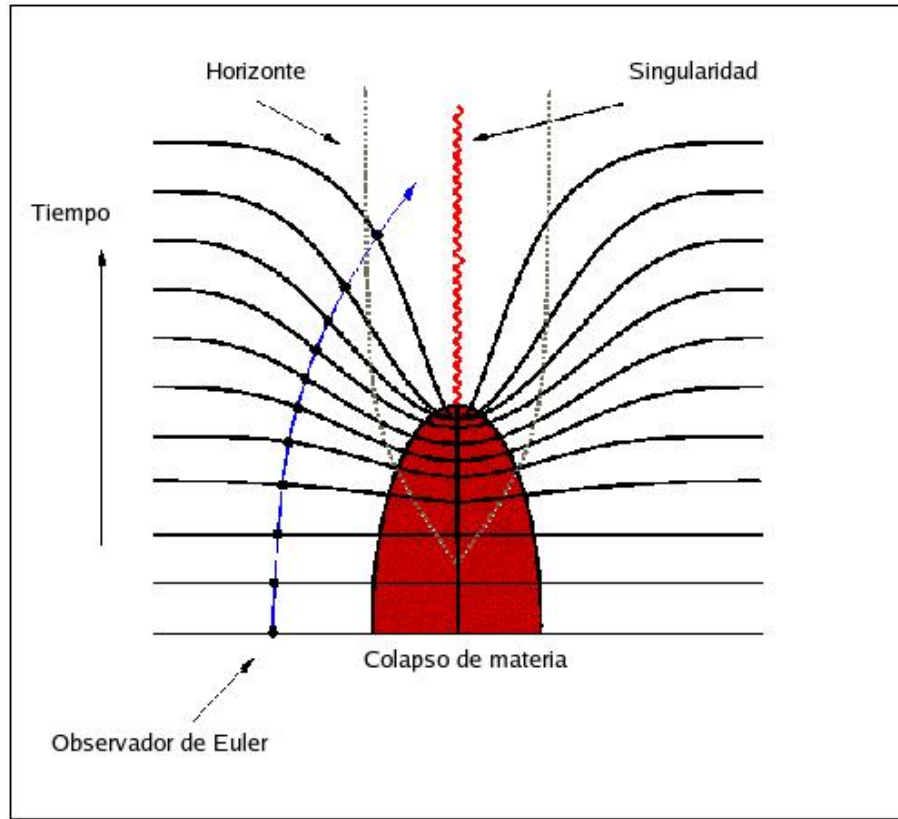


Figura 5: Diagrama esquemático del colapso del lapso.

La familia $1+\log$ conduce a foliaciones muy similares a las que se obtienen con el lapso maximal, y también tiene la propiedad de evadir singularidades. En la práctica se ha encontrado que el caso $N = 2$ es particularmente bien comportado [8].

2.7.2. Condiciones de corrimiento

Es menos lo que se conoce sobre elecciones del vector de corrimiento que sobre condiciones de foliación. La razón es que tomar el vector de corrimiento igual a cero funciona bien en muchos casos. Sin embargo, si existen propuestas de como elegir el vector de corrimiento en diversas situaciones, y además se sabe que en ciertas circunstancias tomar el vector de corrimiento igual a cero es una muy mala elección.

La propuesta mejor conocida para elegir el vector de corrimiento se conoce como “distorsión mínima” [55]. La idea básica es utilizar al vector de corrimiento para eliminar lo más posible la distorsión de los elementos de volumen durante la evolución. Para ello se define el llamado “tensor de distorsión” Σ_{ij} como:

$$\Sigma_{ij} = \frac{1}{2} \gamma^{1/2} \partial_t \tilde{\gamma}_{ij} , \quad (2.59)$$

donde γ es el determinante de la métrica espacial y $\tilde{\gamma}_{ij} = \gamma^{-1/3} \gamma_{ij}$. Es decir, $\tilde{\gamma}_{ij}$ es una transformación conforme de la métrica en la que se ha factorizado el elemento de volumen ($\tilde{\gamma} = 1$). El tensor de distorsión puede entenderse entonces como el cambio en la *forma* de los elementos de volumen durante la evolución (y no en su volumen interno).

A partir de la ecuación de evolución para la métrica (2.14), puede demostrarse que:

$$\Sigma_{ij} = -\alpha \left(K_{ij} - \frac{1}{3} \gamma_{ij} \text{tr} K \right) + \frac{1}{2} \left(\hat{L} \beta \right)_{ij} , \quad (2.60)$$

donde \hat{L} es el mismo operador que definimos cuando discutimos como encontrar datos iniciales (ecuación (2.39)).

La condición de distorsión mínima se obtiene de pedir que la integral sobre todo el espacio del cuadrado no-negativo del tensor de distorsión, $\Sigma_{ij} \Sigma^{ij}$, sea lo más pequeña posible. En este sentido, esta condición minimiza la distorsión de manera global. La condición que acabamos de describir se expresa matemáticamente como un principio variacional, que tiene como solución:

$$D^j \Sigma_{ij} = 0 , \quad (2.61)$$

o de manera equivalente:

$$D^j \left(\hat{L} \beta \right)_{ij} = 2 D^j \left[\alpha \left(K_{ij} - \frac{1}{3} \gamma_{ij} \text{tr} K \right) \right] . \quad (2.62)$$

Esta es la forma final de la condición de distorsión mínima.

Hay varias cosas que es importante notar de esta condición. En primer lugar, se trata de tres ecuaciones diferenciales (el índice i esta libre), que en principio nos permiten calcular las tres componentes del vector β^i . En segundo lugar, es un sistema de ecuaciones de tipo elíptico (generalizaciones de la ecuación de Poisson), acopladas entre sí. Debido a esto, la condición de distorsión mínima es muy difícil de resolver en el caso tri-dimensional (aunque no imposible), lo que ha hecho que sea poco utilizada en la práctica.

Condiciones de corrimiento relacionadas con la distorsión mínima pero mucho más fáciles de resolver se han propuesto en los últimos años. En particular, propuestas del tipo:

$$\partial_t^2 \beta^i \propto D^j \Sigma_j^i , \quad (2.63)$$

que transforman el sistema de ecuaciones elípticas en un sistema hiperbólico se han propuesto en [9]. Estas condiciones resultan ser fáciles de aplicar y muy robustas, pero en estas notas no hablaremos sobre ellas.

Para terminar esta sección es importante comentar que, si bien es cierto que en muchos casos tomar el vector de corrimiento igual a cero funciona muy bien, hay casos en los que es claro que esto no puede ser así. En particular, sistemas con momento angular (agujeros negros o estrellas de neutrones en rotación) producen un efecto físico llamado “arrastre de sistemas inerciales”, que en pocas palabras significa que observadores cercanos al objeto que rota se ven arrastrados en la rotación de éste. En este tipo de situaciones, la única forma de impedir que el sistema de coordenadas sea arrastrado, y termine completamente enredado alrededor del objeto, es utilizando un vector de corrimiento que se oponga al arrastre. Estas situaciones también se dan en el caso de la órbita de objetos compactos, por lo que el estudio de condiciones adecuadas de corrimiento se ha vuelto un tema de gran actualidad.

3. Formulaciones alternativas e hiperbolicidad

3.1. Introducción

En la sección anterior se introdujeron las ecuaciones de evolución de ADM, ecuaciones (2.14) y (2.31). De hecho, estas ecuaciones no están escritas en la forma original de Arnowitt, Deser y Misner [13], sino que son una reescritura no trivial por parte de York [67]. Aún así, en la comunidad numérica las ecuaciones (2.14) y (2.31) se conocen como ecuaciones ADM.

Es importante señalar precisamente en que difieren las ecuaciones de ADM originales de las ecuaciones ADM *a la* York (que llamaremos de ahora en adelante “ADM estándar”). Ambos grupos de ecuaciones difieren en dos aspectos distintos. En primer lugar, las variables ADM originales son la métrica espacial γ_{ij} y su momento canónico conjugado π_{ij} que se obtiene a partir de la formulación lagrangiana de la relatividad general, y que está relacionado con la curvatura extrínseca como ³:

$$K_{ij} = -\gamma^{-1/2} \left(\pi_{ij} - \frac{1}{2} \gamma_{ij} \pi \right), \quad (3.1)$$

con $\pi = \text{tr } \pi_{ij}$.

Este cambio de variables es un detalle más bien menor. Sin embargo, incluso si se reescriben en términos de la curvatura extrínseca, las ecuaciones de evolución que ADM obtienen para K_{ij} difieren de (2.31) y tienen la forma:

$$\begin{aligned} \partial_t K_{ij} &= \beta^a D_a K_{ij} + K_{ia} D_j \beta^a + K_{ja} D_i \beta^a \\ &\quad - D_i D_j \alpha + \alpha \left[R_{ij}^{(3)} - 2K_{ia} K_j^a + K_{ij} \text{tr } K \right] \\ &\quad + 4\pi\alpha \left[\gamma_{ij} (\text{tr } S - \rho) - 2S_{ij} \right] - \frac{\alpha}{4} \gamma_{ij} H, \end{aligned} \quad (3.2)$$

donde H es la constricción hamiltoniana (2.27):

$$H := R^{(3)} + (\text{tr } K)^2 - K_{ij} K^{ij} - 16\pi\rho = 0. \quad (3.3)$$

Es claro que ambos grupos de ecuaciones de evolución para K_{ij} son físicamente equivalentes ya que difieren solo en la adición de un término proporcional a la constricción hamiltoniana que debe ser cero para cualquier solución física. Sin embargo, las diferentes ecuaciones de evolución para K_{ij} no son *matemáticamente* equivalentes. Hay básicamente dos razones por las que esto ocurre:

³Es importante recordar que la motivación original de ADM era el buscar una formulación lagrangiana que pudiera usarse como base para la gravedad cuántica, y no un sistema de ecuaciones de evolución para simulaciones dinámicas.

1. En primer lugar, el espacio de soluciones de las ecuaciones de evolución es distinto, y solo coincide para soluciones físicas, es decir, aquellas que satisfacen las constricciones. Dicho de otro modo, ambos sistemas solo son equivalentes en un subconjunto del espacio de soluciones, conocido como la “hiper-superficie de constricción”.

Desde luego, uno podría argumentar que dado que solo estamos interesados en soluciones físicas, la distinción que hemos hecho es irrelevante. Esto es estrictamente cierto si uno puede resolver las ecuaciones de forma exacta. Pero en el caso de soluciones numéricas siempre habrá un elemento de error que nos llevará fuera de la hiper-superficie de constricción, y la pregunta de que ocurre en ese caso se vuelve no solo relevante sino de fundamental importancia: Si nos salimos ligeramente de la hiper-superficie de constricción, ¿Nos mantenemos a una distancia pequeña, o nos alejamos de ella cada vez más rápido?

2. La segunda razón por la que ambos sistemas difieren matemáticamente esta relacionada con la anterior y es de mayor importancia. Debido a que la constricción hamiltoniana tiene derivadas de la métrica (dentro del escalar de curvatura R), al añadir un múltiplo de esta a las ecuaciones de evolución estamos cambiando la estructura de las ecuaciones, es decir, podemos pasar de un sistema hiperbólico a uno elíptico, o incluso de un sistema bien comportado a uno que no lo es.

Estas consideraciones nos llevan a una observación fundamental que es hoy en día una de las áreas de investigación más importantes en la teoría asociada a la relatividad numérica: las ecuaciones de evolución son altamente no únicas pues siempre es posible añadirles múltiplos arbitrarios de las constricciones. Diferentes ecuaciones coinciden en las soluciones físicas, pero pueden diferir de manera dramática en sus propiedades matemáticas y en particular en la manera en la que responden a pequeñas violaciones de las constricciones (inevitables numéricamente). Cual es la mejor manera de escribir las ecuaciones de evolución es un problema abierto de gran actualidad.

3.2. La formulación BSSN

Aún cuando las ecuaciones de ADM son el punto de partida de la relatividad numérica (en el caso del formalismo 3+1), en la práctica no resultan ser muy bien comportadas frente a pequeñas violaciones de las constricciones. En la siguiente sección discutiremos una posible explicación de este hecho, pero de momento lo tomaremos como una observación empírica: las ecuaciones ADM no son muy estables ante violaciones de las constricciones.

Desde principios de la década de los 90's, se han propuesto una gran cantidad de formulaciones alternativas de las ecuaciones de evolución con diversas motivaciones. En

estas notas no podemos verlas todas, ni siquiera una porción representativa, así que nos limitaremos a discutir una formulación particular que se ha vuelto muy popular en los últimos años debido a que en la práctica ha resultado ser muy bien comportada. Esta es la formulación conocida como BSSN (Baumgarte y Shapiro [14], Shibata y Nakamura [52]).

La formulación BSSN difiere de la formulación ADM en varios aspectos. En primer lugar, BSSN introduce nuevas variables similares a las utilizadas en el problema de valores iniciales que discutimos en la Sección 2.5. La métrica espacial γ_{ij} se reescribe en términos de una métrica conforme $\tilde{\gamma}_{ij}$ como:

$$\gamma_{ij} = \psi^4 \tilde{\gamma}_{ij}, \quad (3.4)$$

donde el factor conforme ψ es tal que el determinante de $\tilde{\gamma}_{ij}$ sea igual a 1, es decir:

$$\psi = \gamma^{1/12}, \quad (3.5)$$

$$\tilde{\gamma}_{ij} = \psi^{-4} \gamma_{ij} = \gamma^{-1/3} \gamma_{ij}, \quad (3.6)$$

$$\tilde{\gamma} = 1, \quad (3.7)$$

con γ el determinante de γ_{ij} y $\tilde{\gamma}$ el determinante de $\tilde{\gamma}_{ij}$. Por otro lado, la curvatura extrínseca se separa en su traza $K := \text{tr}K$ y su parte de traza cero:

$$A_{ij} = K_{ij} - \frac{1}{3} \gamma_{ij} K. \quad (3.8)$$

En vez de las variables ADM γ_{ij} y K_{ij} , se utilizan entonces las variables:

$$\phi = \ln \psi = \frac{1}{12} \ln \gamma, \quad (3.9)$$

$$K = \gamma_{ij} K^{ij}, \quad (3.10)$$

$$\tilde{\gamma}_{ij} = e^{-4\phi} \gamma_{ij}, \quad (3.11)$$

$$\tilde{A}_{ij} = e^{-4\phi} A_{ij}. \quad (3.12)$$

Además, se introducen tres variables auxiliares llamadas “funciones de conexión conformes” y definidas como:

$$\tilde{\Gamma}^i = \tilde{\gamma}^{jk} \tilde{\Gamma}_{jk}^i = -\partial_j \tilde{\gamma}^{ij}, \quad (3.13)$$

donde $\tilde{\Gamma}_{jk}^i$ son los símbolos de Christoffel de la métrica conforme, y donde la segunda igualdad se deriva de la definición de los símbolos de Christoffel en el caso en que el determinante de $\tilde{\gamma}$ sea igual a 1 (lo que debe ser cierto por construcción). Llamamos a las 17 variables ϕ , K , $\tilde{\gamma}_{ij}$, \tilde{A}_{ij} y $\tilde{\Gamma}^i$ las “variables BSSN”.

Hasta ahora no hemos hecho otra cosa que redefinir variables e introducir tres variables auxiliares adicionales. Aún falta dar las ecuaciones de evolución de las nuevas variables. Por simplicidad, de aquí en adelante asumiremos que estamos en vacío y que el vector de corrimiento β^i es cero (el caso general puede obtenerse fácilmente). De la ecuación de evolución para la métrica espacial (2.14) encontramos:

$$\partial_t \tilde{\gamma}_{ij} = -2\alpha \tilde{A}_{ij} , \quad (3.14)$$

$$\partial_t \phi = -\frac{1}{6}\alpha K , \quad (3.15)$$

mientras que de la ecuación de evolución para la curvatura extrínseca (2.31) obtenemos:

$$\partial_t \tilde{A}_{ij} = e^{-4\phi} [-D_i D_j \alpha + \alpha R_{ij}]^{\text{TF}} + \alpha (K \tilde{A}_{ij} - 2 \tilde{A}_{ik} \tilde{A}^k_j) , \quad (3.16)$$

$$\partial_t K = -D_i D^i \alpha + \alpha \left(\tilde{A}_{ij} \tilde{A}^{ij} + \frac{1}{3} K^2 \right) , \quad (3.17)$$

donde TF denota la parte de traza cero de la expresión dentro de los corchetes. Es importante notar que en la ecuación de evolución para K se ha utilizado ya la constricción hamiltoniana para eliminar al escalar de curvatura:

$$R = K_{ij} K^{ij} - K^2 = \tilde{A}_{ij} \tilde{A}^{ij} - \frac{2}{3} K^2 . \quad (3.18)$$

Vemos entonces que ya hemos comenzado a jugar el juego de sumar múltiplos de las constricciones a las ecuaciones de evolución.

Falta aún la ecuación de evolución para las $\tilde{\Gamma}^i$. Esta ecuación puede obtenerse directamente de (3.13) y (2.14). Encontramos:

$$\partial_t \tilde{\Gamma}^i = -2 \left(\alpha \partial_j \tilde{A}^{ij} + \tilde{A}^{ij} \partial_j \alpha \right) , \quad (3.19)$$

Nótese que en las ecuaciones de evolución para \tilde{A}_{ij} y K aparecen derivadas covariantes de la función de lapso con respecto a la métrica física γ_{ij} . Estas pueden calcularse fácilmente utilizando el hecho de que los símbolos de Christoffel están relacionados por:

$$\Gamma^k_{ij} = \tilde{\Gamma}^k_{ij} + 2 \left(\delta_i^k \partial_j \phi + \delta_j^k \partial_i \phi - \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \partial_l \phi \right) , \quad (3.20)$$

donde $\tilde{\Gamma}^k_{ij}$ son los símbolos de Christoffel de la métrica conforme.

$$D^i D_i \alpha = e^{-4\phi} \left(\tilde{\gamma}^{ij} \partial_i \partial_j \alpha - \tilde{\Gamma}^k \partial_k \alpha + 2 \tilde{\gamma}^{ij} \partial_i \phi \partial_j \alpha \right) . \quad (3.21)$$

Además, en la ecuación para \tilde{A}_{ij} aparece el tensor de Ricci asociado a γ_{ij} que se separa de la siguiente forma

$$R_{ij} = \tilde{R}_{ij} + R_{ij}^\phi, \quad (3.22)$$

donde \tilde{R}_{ij} es el tensor de Ricci asociado a la métrica conforme $\tilde{\gamma}_{ij}$:

$$\begin{aligned} \tilde{R}_{ij} = & -\frac{1}{2}\tilde{\gamma}^{lm}\partial_l\partial_m\tilde{\gamma}_{ij} + \tilde{\gamma}_{k(i}\partial_{j)}\tilde{\Gamma}^k + \tilde{\Gamma}^k\tilde{\Gamma}_{(ij)k} \\ & + \tilde{\gamma}^{lm}\left(2\tilde{\Gamma}^k{}_{l(i}\tilde{\Gamma}_{j)km} + \tilde{\Gamma}^k{}_{im}\tilde{\Gamma}_{klj}\right). \end{aligned} \quad (3.23)$$

y donde R_{ij}^ϕ denota términos adicionales que dependen de ϕ :

$$R_{ij}^\phi = -2\tilde{D}_i\tilde{D}_j\phi - 2\tilde{\gamma}_{ij}\tilde{D}^k\tilde{D}_k\phi + 4\tilde{D}_i\phi\tilde{D}_j\phi - 4\tilde{\gamma}_{ij}\tilde{D}^k\phi\tilde{D}_k\phi, \quad (3.24)$$

con \tilde{D}_i la derivada covariante con respecto a la métrica conforme.

La motivación para el cambio de variables realizado arriba es múltiple. La transformación conforme y la separación de la traza de la curvatura extrínseca se lleva a cabo para tener mejor control sobre las condiciones de lapso, que como se menciono en la sección 2.7.1, en general están relacionadas con la traza de K_{ij} y se escogen de manera que se pueda controlar mejor la evolución de los elementos de volumen. Por otro lado, la introducción de las variables auxiliares $\tilde{\Gamma}^i$ se debe a que, cuando estas se consideran variables independientes, las segundas derivadas de la métrica conforme que aparecen en el lado derecho de la ecuación de evolución (3.16) (contenidas en el tensor de Ricci (3.23)) se reducen simplemente al operador de Laplace $\tilde{\gamma}^{lm}\partial_l\partial_m\tilde{\gamma}_{ij}$. Todos los demás términos que tienen segundas derivadas de $\tilde{\gamma}_{ij}$ pueden reescribirse en términos de primeras derivadas de $\tilde{\Gamma}^i$. Esto significa que las ecuaciones (3.14) y (3.16) tienen la estructura de una ecuación de onda con una fuente complicada.

Falta aún un elemento clave en la formulación BSSN que no hemos mencionado. En la práctica resulta que, pese a las motivaciones arriba mencionadas, si uno utiliza las ecuaciones de evolución dadas por (3.14), (3.15), (3.16), (3.17), y (3.19) en una simulación numérica, el sistema resulta ser violentamente inestable (mucho peor que ADM).

Para corregir este problema escribimos primero las constricciones de momento, que en las variables BSSN toman la forma:

$$\partial_j\tilde{A}^{ij} = -\tilde{\Gamma}^i{}_{jk}\tilde{A}^{jk} - 6\tilde{A}^{ij}\partial_j\phi + \frac{2}{3}\tilde{\gamma}^{ij}\partial_jK. \quad (3.25)$$

Podemos ahora utilizar esta ecuación para substituir a la divergencia de \tilde{A}^{ij} que aparece en la ecuación de evolución (3.19) para $\tilde{\Gamma}^i$, obteniendo:

$$\partial_t\tilde{\Gamma}^i = -2\tilde{A}^{ij}\partial_j\alpha + 2\alpha\left(\tilde{\Gamma}^i{}_{jk}\tilde{A}^{jk} + 6\tilde{A}^{ij}\partial_j\phi - \frac{2}{3}\tilde{\gamma}^{ij}\partial_jK\right). \quad (3.26)$$

El sistema de ecuaciones de evolución es ahora: (3.14), (3.15), (3.16), (3.17), y (3.26). Este nuevo sistema no solo ya no presenta la violenta inestabilidad antes mencionada, sino que además resulta ser mucho mejor comportado que el sistema ADM en todos los casos estudiados a la fecha. Que esto es así fue mostrado por primera vez de forma empírica (es decir, mediante comparaciones directas de simulaciones numéricas) por Baumgarte y Shapiro en [14], y explicado por Alcubierre y colaboradores en [4]. La explicación de porqué BSSN es mejor que ADM está relacionada con el concepto de hiperbolicidad que discutiremos en la siguiente sección.

3.3. El concepto de hiperbolicidad

En esta sección daremos una breve introducción al concepto de hiperbolicidad de sistemas de ecuaciones de evolución, que resulta clave en el análisis de las propiedades matemáticas de los sistemas de ecuaciones utilizados en diversas áreas de la física, y en particular en la relatividad numérica. Una descripción más detallada de la teoría de sistemas hiperbólicos puede encontrarse en [36].

Consideremos un sistema de ecuaciones de evolución en una dimensión espacial de la forma:

$$\partial_t u_i + \partial_x F_i = q_i \quad i \in \{1, \dots, N_u\}, \quad (3.27)$$

donde F_i y q_i son funciones arbitrarias, posiblemente no lineales, de las u 's pero no de sus derivadas. Este sistema también puede escribirse como:

$$\partial_t u_i + \sum_j M_{ij} \partial_x u_j = q_i \quad i \in \{1, \dots, N_u\}, \quad (3.28)$$

con $M_{ij} = \partial F_i / \partial u_j$ la llamada “matriz jacobiana”.

Es importante mencionar que la mayor parte de las ecuaciones diferenciales de evolución de la física se pueden escribir de esta forma. En el caso en el que haya derivadas de orden mayor, se puede simplemente definir variables auxiliares para obtener un sistema de primer orden (más abajo veremos como se hace esto para la ecuación de onda).

Sean ahora λ_i los eigenvalores de la matriz jacobiana M . El sistema de ecuaciones de evolución se denomina “hiperbólico” si todas las λ_i resultan ser funciones reales. Más aún, el sistema se denomina “fuertemente hiperbólico” si la matriz M tiene un conjunto completo de eigenvectores. En caso que todos los eigenvalores sean reales, pero no exista un conjunto completo de eigenvectores, el sistema se denomina “débilmente hiperbólico”.⁴

⁴La distinción importante entre sistemas fuertemente hiperbólicos y simétricamente hiperbólicos no existe en el caso de una sola dimensión espacial. En el caso de múltiples dimensiones, los sistemas simétricamente hiperbólicos son aquellos que pueden diagonalizarse de forma independiente de la dirección.

Asumamos ahora que el sistema es fuertemente hiperbólico. En ese caso, se definen los “eigencampos” w_i como:

$$\mathbf{u} = R \mathbf{w} \quad \Rightarrow \quad \mathbf{w} = R^{-1} \mathbf{u} , \quad (3.29)$$

donde R es la matriz de eigenvectores columna \mathbf{e}_i . Es posible mostrar que la matriz R es tal que

$$RMR^{-1} = \Lambda , \quad (3.30)$$

con $\Lambda = \text{diag}(\lambda_i)$. Es decir, la matriz R diagonaliza a la matriz jacobiana M (esto es básicamente un cambio de los vectores de la base).

Las ecuaciones de evolución para los eigencampos w_i resultan ser entonces

$$\partial_t w_i + \lambda_i \partial_x w_i = q'_i , \quad (3.31)$$

con q'_i funciones de las w 's pero no sus derivadas. El sistema se ha transformado entonces en una serie de ecuaciones de advección con velocidades características dadas por los eigenvalores λ_i . Dicho de otro modo, tenemos una serie de perturbaciones propagándose con velocidades λ_i .

La hiperbolicidad es de fundamental importancia en el estudio de las ecuaciones de evolución asociadas con un problema de Cauchy. En un sentido físico, la hiperbolicidad implica que el sistema de ecuaciones es causal y local, es decir, que la solución en un punto dado del espacio-tiempo depende solo de información contenida en una región compacta al pasado de ese punto, el llamado “cono característico” (o cono de luz en el caso de la relatividad). Matemáticamente, se puede mostrar que un sistema fuertemente hiperbólico está “bien planteado”, es decir, sus soluciones existen y son únicas (al menos localmente), y además las soluciones son estables en el sentido de que cambios pequeños en los datos iniciales corresponden a cambios pequeños en la solución.

El concepto de hiperbolicidad se puede extender fácilmente a tres dimensiones considerando ecuaciones del tipo:

$$\partial_t u_i + \partial_x F_i^x + \partial_y F_i^y + \partial_z F_i^z = q_i , \quad i \in \{1, \dots, N_u\} , \quad (3.32)$$

y analizando las tres matrices jacobianas $M_{ij}^k = \partial F_i^k / \partial u_j$ ($k = x, y, z$).

Para ver un ejemplo sencillo de hiperbolicidad, consideremos la ecuación de onda en una dimensión:

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \frac{\partial^2 \phi}{\partial x^2} = 0 , \quad (3.33)$$

donde ϕ es la función de onda y c la velocidad de onda. Esta ecuación es de segundo orden, pero podemos pasarla a primer orden si introducimos las variables auxiliares:

$$\Pi := \partial \phi / \partial t , \quad \Psi := \partial \phi / \partial x . \quad (3.34)$$

La ecuación de onda puede reescribirse entonces como el sistema:

$$\frac{\partial \phi}{\partial t} = \Pi, \quad \frac{\partial \Pi}{\partial t} - c^2 \frac{\partial \Psi}{\partial x} = 0, \quad \frac{\partial \Psi}{\partial t} - \frac{\partial \Pi}{\partial x} = 0, \quad (3.35)$$

donde la primera ecuación es la definición de Π , la segunda es la ecuación de onda propiamente dicha, y la tercera es el requerimiento de que las derivadas de ϕ conmuten. El sistema anterior claramente tiene ya la forma (3.27). Como la función ϕ no aparece en ninguna de las dos últimas ecuaciones, podemos analizar únicamente el sub-sistema $\{\Pi, \Psi\}$. La matriz jacobiana en este caso es de 2×2 , y resulta ser:

$$\mathbf{M} = \begin{pmatrix} 0 & -c^2 \\ -1 & 0 \end{pmatrix}. \quad (3.36)$$

Sus correspondientes eigenvalores son reales y están dados por $\lambda_{\pm} = \pm c$, es decir, las ondas pueden moverse a la izquierda y a la derecha con velocidad c . La matriz M tiene dos eigenvectores independientes, es decir un conjunto completo, por lo que el sistema es fuertemente hiperbólico. Los eigenvectores resultan ser $v_{\pm} = (\mp c, 1)$, de donde podemos encontrar fácilmente las eigenfunciones $\omega_{\pm} = \Pi \mp c\Psi$ (nótese que a la velocidad $+c$ corresponde la eigenfunción $\Pi - c\Psi$ y viceversa). Las ecuaciones de evolución para las eigenfunciones son claramente:

$$\frac{\partial \omega_{\pm}}{\partial t} \pm c \frac{\partial \omega_{\pm}}{\partial x} = 0. \quad (3.37)$$

Vemos que las eigenfunciones se propagan en direcciones opuestas de manera independiente.

3.4. Hiperbolicidad de las ecuaciones de evolución en 3+1

Vamos ahora a estudiar la hiperbolicidad de las ecuaciones de evolución de ADM, ecuaciones (2.14) y (2.31). Como solo estamos interesados en la idea básica, vamos a analizar un caso muy simple (las conclusiones no se modifican fundamentalmente en el caso general). En primer lugar, asumiremos que estamos en el vacío y que el vector de corrimiento β^i es cero. Las ecuaciones que vamos a utilizar son entonces:

$$\partial_t \gamma_{ij} = -2\alpha K_{ij}, \quad (3.38)$$

$$\partial_t K_{ij} = -D_i D_j \alpha + \alpha [R_{ij}^{(3)} - 2K_{ia} K_j^a + K_{ij} \text{tr} K], \quad (3.39)$$

Para simplificar aún más las cosas, consideraremos perturbaciones lineales de un espacio-tiempo plano. En este caso, la métrica espacial y la función de lapso se pueden

escribir como:

$$\alpha = 1 + a, \quad (3.40)$$

$$\gamma_{ij} = \delta_{ij} + h_{ij}, \quad (3.41)$$

con a y h_{ij} mucho menores que 1. A primer orden en cantidades pequeñas, las ecuaciones de evolución toman la forma simplificada (nótese que en estas circunstancias K_{ij} resulta ser una cantidad pequeña):

$$\partial_t h_{ij} = -2K_{ij}, \quad (3.42)$$

$$\partial_t K_{ij} = -\partial_i \partial_j a + R_{ij}^{(1)}, \quad (3.43)$$

donde el tensor de Ricci linearizado está dado por:

$$R_{ij}^{(1)} = -1/2 \left(\nabla_{\text{plano}}^2 h_{ij} - \partial_i \Gamma_j - \partial_j \Gamma_i \right), \quad (3.44)$$

y donde hemos definido:

$$\Gamma_i := \sum_k \partial_k h_{ik} - 1/2 \partial_i h, \quad (3.45)$$

con $h = \sum_i h_{ii}$.

Para continuar el análisis debemos ahora elegir nuestra condición de foliación. La opción más simple es tomar $a = 0$, pero eso equivale a la foliación geodésica y ya hemos mencionado que esa no es una buena elección. Tomaremos mejor el lapso armónico, que en esta aproximación corresponde a:

$$\partial_t a = -K, \quad (3.46)$$

donde ahora $K = \sum_i K_{ii}$.

El siguiente paso es reescribir estas ecuaciones como un sistema de primer orden. Para ello debemos introducir las variables auxiliares:

$$A_i := \partial_i a, \quad D_{ijk} := \frac{1}{2} \partial_i h_{jk}. \quad (3.47)$$

El sistema de ecuaciones toma ahora la forma:

$$\partial_t a = -K, \quad (3.48)$$

$$\partial_t h_{ij} = -2K_{ij}, \quad (3.49)$$

$$\partial_t A_i = -\partial_i K, \quad (3.50)$$

$$\partial_t D_{ijk} = -\partial_i K_{jk}, \quad (3.51)$$

$$\partial_t K_{ij} = -\partial_{(i} A_{j)} + \sum_k \left(2\partial_{(i} D_{kkj)} - \partial_{(i} D_{j)kk} - \partial_k D_{kij} \right). \quad (3.52)$$

Para hacer el análisis de hiperbolicidad notamos primero que las variables a y h_{ij} evolucionan solo con términos de fuente (no hay derivadas del lado derecho de sus ecuaciones de evolución). Además, no aparecen en las siguientes ecuaciones, por lo que pueden ignorarse al estudiar la hiperbolicidad. Nos concentraremos entonces en el sub-sistema $\{A_i, D_{ijk}, K_{ij}\}$. Como el sistema es tridimensional, en principio debemos construir las tres matrices jacobianas. Analizaremos únicamente la dirección x , las otras dos direcciones son enteramente análogas.

Si consideramos solo derivadas en la dirección x , notamos que únicamente debemos considerar las componentes A_x y D_{xjk} , pues las componentes A_q y D_{qjk} , para q distinta de x , se pueden tomar como fijas en este caso (podemos asumir que son cero sin afectar el análisis). Haciendo un poco de álgebra encontramos que el sistema a analizar toma la forma:

$$\partial_t A_x + \partial_x (K_{xx} + K_{yy} + K_{zz}) = 0 , \quad (3.53)$$

$$\partial_t K_{xx} + \partial_x A_x + \partial_x (D_{xyy} + D_{xzz}) = 0 , \quad (3.54)$$

$$\partial_t K_{yy} + \partial_x D_{xyy} = 0 , \quad (3.55)$$

$$\partial_t K_{zz} + \partial_x D_{xzz} = 0 , \quad (3.56)$$

$$\partial_t K_{xy} = 0 , \quad (3.57)$$

$$\partial_t K_{xz} = 0 , \quad (3.58)$$

$$\partial_t K_{yz} + \partial_x D_{xyz} = 0 , \quad (3.59)$$

$$\partial_t D_{xjk} + \partial_x K_{jk} = 0 . \quad (3.60)$$

Este es un sistema de trece variables:

$$u := \{A_x, K_{xx}, K_{yy}, K_{zz}, K_{xy}, K_{xz}, K_{yz}, D_{xxx}, D_{xyy}, D_{xzz}, D_{xyx}, D_{xxz}, D_{xyz}, \}$$

La matriz jacobiana resulta ser:

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (3.61)$$

La matriz anterior, de 13×13 elementos, tal vez parezca muy grande, pero la mayor parte de sus entradas son ceros, por lo que no es difícil encontrar sus eigenvalores. Estos resultan ser: $+1$ con multiplicidad 4, -1 con multiplicidad 4, y 0 con multiplicidad 5, para un total de 13. Como todos estos eigenvalores son reales, el sistema es hiperbólico.

Más interesantes son los eigenvectores. Asociados al eigenvalor $\lambda = -1$ tenemos los siguientes tres eigenvectores:

$$v_1^- = (1, -1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0), \quad (3.62)$$

$$v_2^- = (0, 0, 1, -1, 0, 0, 0, 0, -1, 1, 0, 0, 0, 0), \quad (3.63)$$

$$v_3^- = (0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, -1), \quad (3.64)$$

asociados al eigenvalor $\lambda = +1$ los tres eigenvectores:

$$v_1^+ = (1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0), \quad (3.65)$$

$$v_2^+ = (0, 0, 1, -1, 0, 0, 0, 0, 1, -1, 0, 0, 0, 0), \quad (3.66)$$

$$v_3^+ = (0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1), \quad (3.67)$$

y asociados al eigenvalor $\lambda = 0$ los tres eigenvalores:

$$v_1^0 = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0), \quad (3.68)$$

$$v_2^0 = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0), \quad (3.69)$$

$$v_3^0 = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0), \quad (3.70)$$

para un total de solo nueve eigenvectores. Es decir, no hay un conjunto completo de eigenvectores y el sistema es solo débilmente hiperbólico. Como hemos mencionado, un sistema débilmente hiperbólico no es bien comportando matemáticamente, por lo que deberíamos esperar problemas cuando utilizamos las ecuaciones de evolución de ADM para simulaciones numéricas.

Es posible resolver el problema de la hiperbolicidad débil de ADM añadiendo múltiplos de las constricciones en lugares adecuados. Para ver esto, primero escribimos la aproximación lineal a las constricciones que resulta ser:

$$\sum_j \partial_j f_j = 0, \quad (\text{hamiltoniana}) \quad (3.71)$$

$$\sum_j \partial_j K_{ij} - \partial_i K = 0, \quad (\text{momento}) \quad (3.72)$$

donde hemos definido $f_i := \sum_k \partial_k h_{ik} - \partial_i h = 2 \sum_k (D_{kki} - D_{ikk})$. No es difícil ver que si consideramos ahora solo derivadas en la dirección x (e ignoramos de nuevo las componentes D_{qjk} con q distinta de x), las constricciones toman la forma:

$$\partial_x (D_{xyy} + D_{xzz}) = 0, \quad (\text{hamiltoniana}) \quad (3.73)$$

$$\partial_x (K_{yy} + K_{zz}) = 0, \quad (\text{momentum x}) \quad (3.74)$$

$$\partial_x K_{xy} = 0, \quad (\text{momento y}) \quad (3.75)$$

$$\partial_x K_{xz} = 0, \quad (\text{momento z}) \quad (3.76)$$

Substituyendo estas constricciones en las ecuaciones de evolución, la matriz jacobiana se transforma en:

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (3.77)$$

Los eigenvalores resultan ser los mismos que antes, pero ahora si se encuentra un conjunto completo de eigenvectores. La lección aprendida es la siguiente: substituyendo las constricciones en las ecuaciones de evolución es posible transformar a las ecuaciones linearizadas de ADM en un sistema fuertemente hiperbólico.

En el caso general, cuando estamos en un campo no linearizado y con un vector de corrimiento distinto de cero se puede hacer un análisis similar. El problema es que en ese caso no es nada evidente como se deben substituir las constricciones en las ecuaciones de evolución para obtener sistemas fuertemente hiperbólicos. De hecho hay muchas formas distintas de hacerlo, y ya se han construido varias familias de formulaciones fuertemente hiperbólicas, entre las que se pueden mencionar las formulaciones de Bona-Masso [16, 17], Frittelli-Reula [28], Kidder-Scheel-Teukolsy [34]. Incluso es posible mostrar que la formulación BSSN que se introdujo en la sección anterior es fuertemente hiperbólica bajo ciertas suposiciones [46].

Hoy en día, el problema de encontrar criterios que digan que formulación hiperbólica es mejor que otra se ha convertido en un área de investigación de gran actualidad, y varios grupos trabajan activamente en ella en diversos lugares del mundo.

4. Aproximaciones en diferencias finitas

4.1. Introducción

Las teorías de campo juegan un papel fundamental en la física moderna. Desde la electrodinámica clásica de Maxwell, hasta las teorías cuánticas de campo, pasando por la ecuación de Schrödinger, la hidrodinámica, y desde luego la relatividad general, la noción de campo como una entidad física en si misma ha tenido implicaciones profundas en nuestra concepción del Universo. Los campos son funciones continuas del espacio y el tiempo, y la descripción matemática de sus leyes dinámicas se realiza en el contexto de las ecuaciones diferenciales parciales.

Las ecuaciones diferenciales parciales asociadas a teorías físicas son en general imposibles de resolver analíticamente excepto en casos muy idealizados. Esta dificultad puede tener diversos orígenes, desde la presencia de fronteras irregulares, a la existencia de términos no lineales en las ecuaciones mismas. Para resolver este tipo de ecuaciones en situaciones dinámicas generales resulta inevitable utilizar aproximaciones numéricas.

Existen muchas formas distintas de resolver ecuaciones diferenciales parciales de forma numérica. Los métodos más populares son tres: las “diferencias finitas” [45], los “elementos finitos” [43], y los “métodos espectrales”. En este curso nos limitaremos a estudiar las diferencias finitas por ser el método conceptualmente más simple y también por ser el más común en la relatividad numérica (aunque no el único, en particular los métodos espectrales han comenzado a ganar popularidad en los últimos años [18, 33, 35]).

4.2. Ideas fundamentales en las diferencias finitas

Cuando uno estudia un campo en un espacio-tiempo continuo, se ve en la necesidad de considerar un número infinito (y no contable) de variables desconocidas: el valor del campo en todo punto del espacio y a todo tiempo. Para encontrar el valor del campo utilizando aproximaciones numéricas primero se debe reducir el número de variables a una cantidad finita. Hay muchas formas de hacer esto. Los métodos espectrales, por ejemplo, expanden la solución como una combinación lineal finita de una base adecuada de funciones. Las variables a resolver son entonces los coeficientes de dicha expansión. Un enfoque distinto es tomado por las diferencias y los elementos finitos. En ambos casos se reduce el número de variables discretizando el dominio de dependencia de las funciones, aunque con distintas estrategias.

La idea básica de las diferencias finitas es substituir al espacio-tiempo continuo por un conjunto discreto de puntos. Este conjunto de puntos se conoce como la “malla” o “red” computacional. Las distancias en el espacio entre los puntos de esta red no tienen

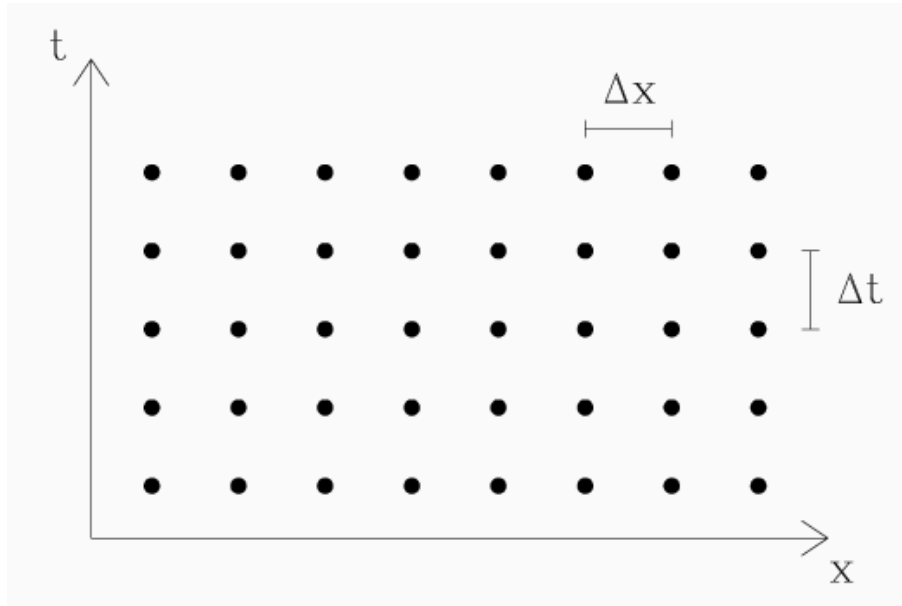


Figura 6: Discretización del espacio-tiempo utilizada en diferencias finitas.

porqué ser uniformes, pero en estas notas asumiremos que si lo son. El paso de tiempo entre dos niveles consecutivos se denomina Δt , y la distancia entre dos puntos adyacentes en el espacio Δx . La Figura 6 es una representación gráfica de la red computacional.

Una vez establecida la malla computacional el siguiente paso es substituir las ecuaciones diferenciales por un sistema de ecuaciones algebraicas. Esto se logra aproximando los operadores diferenciales por diferencias finitas entre los valores de las funciones en puntos adyacentes de la malla. De esta forma se obtiene una ecuación algebraica en cada punto de la malla por cada ecuación diferencial. Estas ecuaciones algebraicas involucran los valores de las funciones en el punto considerado y en sus vecinos más cercanos. El sistema de ecuaciones algebraicas se puede resolver de manera sencilla, el precio que hemos pagado es que ahora tenemos muchísimas ecuaciones algebraicas, por lo que se requiere utilizar una computadora.

Para ver como se hace esto en la práctica, consideremos como modelo la ecuación de onda en una dimensión. Esta ecuación presenta muchas ventajas. En primer lugar se puede resolver de manera exacta y la solución exacta se puede utilizar para comparar con la solución numérica. Además, la mayoría de las ecuaciones fundamentales en las teorías de campo modernas se pueden ver como generalizaciones de diversos tipos de la ecuación de onda.

4.3. La ecuación de onda en una dimensión

La ecuación de onda en una dimensión (en un espacio plano) tiene la forma:

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0, \quad (4.1)$$

donde ϕ es la función de onda y c la velocidad de onda.

Para encontrar una aproximación en diferencias finitas a esta ecuación comenzamos por introducir la siguiente notación para los valores de ϕ en la red computacional:

$$\phi_m^n := \phi(n\Delta t, m\Delta x), \quad (4.2)$$

Podemos ahora aproximar los operadores diferenciales que aparecen en la ecuación (4.1) utilizando expansiones de Taylor de ϕ alrededor del punto (n, m) . Consideremos, por ejemplo, el valor de ϕ en los puntos $(n, m+1)$ y $(n, m-1)$:

$$\phi_{m+1}^n = \phi_m^n + \left(\frac{\partial \phi}{\partial x}\right) \Delta x + \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial x^2}\right) (\Delta x)^2 + \frac{1}{6} \left(\frac{\partial^3 \phi}{\partial x^3}\right) (\Delta x)^3 + \dots, \quad (4.3)$$

$$\phi_{m-1}^n = \phi_m^n - \left(\frac{\partial \phi}{\partial x}\right) \Delta x + \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial x^2}\right) (\Delta x)^2 - \frac{1}{6} \left(\frac{\partial^3 \phi}{\partial x^3}\right) (\Delta x)^3 + \dots, \quad (4.4)$$

donde todas las derivadas deben ser evaluadas en el punto $(t = n\Delta t, x = m\Delta x)$. A partir de estas expresiones es fácil ver que:

$$\left(\frac{\partial^2 \phi}{\partial x^2}\right) = \frac{\phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n}{(\Delta x)^2} + \frac{(\Delta x)^2}{12} \left(\frac{\partial^4 \phi}{\partial x^4}\right) + \dots. \quad (4.5)$$

Podemos entonces aproximar la segunda derivada como:

$$\left(\frac{\partial^2 \phi}{\partial x^2}\right) \simeq \frac{\phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n}{(\Delta x)^2}. \quad (4.6)$$

Que tan buena es esta aproximación depende, desde luego, del tamaño de la malla Δx . Si Δx es pequeño en el sentido de que la función ϕ varía muy poco en una región de ese tamaño, entonces la aproximación puede ser muy buena. El error involucrado en esta aproximación es llamado “error de truncación” y como vemos en este caso su parte dominante resulta ser proporcional a Δx^2 , por lo que se dice que esta aproximación es de segundo orden.

La segunda derivada de ϕ con respecto a t se puede aproximar exactamente de la misma manera. De esta forma obtenemos la siguiente aproximación a segundo orden de la ecuación de onda:

$$\frac{\phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n}{(\Delta x)^2} - \frac{1}{c^2} \frac{\phi_m^{n+1} - 2\phi_m^n + \phi_m^{n-1}}{(\Delta t)^2} = 0 . \quad (4.7)$$

Podemos reescribir esta ecuación de forma más compacta si introducimos el llamado “parámetro de Courant”: $\rho := c\Delta t/\Delta x$. La aproximación toma la forma final:

$$\rho^2 \left(\phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n \right) - \left(\phi_m^{n+1} - 2\phi_m^n + \phi_m^{n-1} \right) = 0 . \quad (4.8)$$

Esta ecuación tiene una propiedad muy importante: involucra solo un valor de la función de onda en el último nivel de tiempo, el valor ϕ_m^{n+1} . Podemos entonces despejar este valor en términos de valores en tiempos anteriores para obtener:

$$\phi_m^{n+1} = 2\phi_m^n - \phi_m^{n-1} + \rho^2 \left(\phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n \right) . \quad (4.9)$$

Por esta propiedad la aproximación anterior se conoce como “aproximación explícita”. Si conocemos los valores de la función ϕ en los niveles n y $n-1$, podemos usar la ecuación anterior para calcular directamente los valores de la función en el nuevo paso de tiempo $n+1$. El proceso puede luego iterarse tantas veces como se quiera.

Es evidente que todo lo que se necesita para comenzar la evolución es el conocimiento del valor de la función de onda en los primeros dos pasos de tiempo. Pero obtener estos datos es muy fácil de hacer. Como se trata de una ecuación de segundo orden, los datos iniciales incluyen:

$$f(x) := \phi(0, x) , \quad g(x) := \left. \frac{\partial \phi}{\partial t} \right|_{t=0} . \quad (4.10)$$

El conocimiento de $f(x)$ evidentemente nos da el primer nivel de tiempo:

$$\phi_m^0 = f(m \Delta x) . \quad (4.11)$$

Para el segundo nivel basta con aproximar la primera derivada en el tiempo en diferencias finitas. Una posible aproximación es:

$$g(m \Delta x) = \frac{\phi_m^1 - \phi_m^0}{\Delta t} , \quad (4.12)$$

de donde obtenemos:

$$\phi_m^1 = g(m \Delta x) \Delta t + \phi_m^0 . \quad (4.13)$$

La expresión anterior tiene un inconveniente importante. De la expansión en Taylor podemos ver fácilmente que el error de truncación para esta expresión es de orden Δt , por lo que la aproximación es solo de primer orden. Es claro que si comenzamos la evolución con un error de primer orden, el segundo orden de todo el esquema se pierde. Sin embargo, es fácil resolver este problema. Una aproximación a segundo orden de la derivada temporal resulta ser:

$$g(m \Delta x) = \frac{\phi_m^1 - \phi_m^{-1}}{2 \Delta t} . \quad (4.14)$$

El problema ahora es que esta expresión hace referencia al valor de la función ϕ_m^{-1} que también es desconocido. Pero ya tenemos otra ecuación que hace referencia a ϕ_m^1 y ϕ_m^{-1} : la aproximación a la ecuación de onda (4.8) evaluada en $n = 0$. Podemos entonces utilizar estas dos ecuaciones para eliminar ϕ_m^{-1} y resolver para ϕ_m^1 . De esta forma obtenemos la siguiente aproximación a segundo orden para el segundo nivel de tiempo:

$$\phi_m^1 = \phi_m^0 + \frac{\rho^2}{2} (\phi_{m+1}^0 - 2\phi_m^0 + \phi_{m-1}^0) + \Delta t g(m \Delta x) . \quad (4.15)$$

Las ecuaciones (4.11) y (4.15) nos dan ahora toda la información necesaria para comenzar la evolución.

Hay un punto importante que debe mencionarse aquí. Para reducir el número total de variables a un número finito, también es necesario tomar una región finita del espacio, conocida como el “dominio computacional”, con un número finito de puntos N . Resulta entonces crucial especificar las condiciones de frontera que deberán aplicarse a los extremos de la red computacional. Es claro que la aproximación a la ecuación de onda (4.8) no puede ser utilizada en las fronteras debido a que hace referencia a puntos fuera del dominio computacional. Existen diversas maneras conocidas de imponer condiciones de frontera para la ecuación de onda. Sin embargo, para sistemas más complejos como las ecuaciones de Einstein, la elección de condiciones de frontera adecuadas y consistentes es un problema no resuelto de gran actualidad [57, 58]. Para los objetivos de estas notas podemos olvidarnos del problema de las condiciones de frontera y asumir que tenemos un espacio periódico, de manera que la elección de las condiciones de frontera será simplemente:

$$\phi_0^n \equiv \phi_N^n . \quad (4.16)$$

Esta elección, además de ser sumamente simple, es equivalente a utilizar la aproximación interna en todos lados, por lo que nos permite concentrarnos en las propiedades del método interior únicamente, sin preocuparnos por los efectos que puedan introducir las fronteras.

4.4. Aproximaciones implícitas y moléculas computacionales

En la sección anterior se introdujeron las ideas básicas de las aproximaciones en diferencias finitas utilizando la ecuación de onda en una dimensión como ejemplo. La aproximación que encontramos, sin embargo, esta lejos de ser única. En principio, existe un número infinito de formas distintas de aproximar una misma ecuación diferencial utilizando diferencias finitas. Diferentes aproximaciones tienen distintas propiedades. En la siguiente sección se mencionarán cuales de estas propiedades pueden hacer una aproximación más útil que otra.

Para hacer las cosas más fáciles, introduciremos una notación compacta para las diferencias finitas. Definimos los operadores de “primeras diferencias centradas” como:

$$\delta_x \phi_m^n := \frac{1}{2} (\phi_{m+1}^n - \phi_{m-1}^n) , \quad (4.17)$$

$$\delta_t \phi_m^n := \frac{1}{2} (\phi_m^{n+1} - \phi_m^{n-1}) , \quad (4.18)$$

y los operadores de “segundas diferencias centradas” como:

$$\delta_x^2 \phi_m^n := \phi_{m+1}^n - 2\phi_m^n + \phi_{m-1}^n , \quad (4.19)$$

$$\delta_t^2 \phi_m^n := \phi_m^{n+1} - 2\phi_m^n + \phi_m^{n-1} . \quad (4.20)$$

(Cuidado: Con esta notación $(\delta_x)^2 \neq \delta_x^2$).

Habiendo definido estos operadores, podemos regresar a las aproximaciones que hicimos a los operadores diferenciales que aparecen en la ecuación de onda. Partiendo de nuevo de las series de Taylor, es posible mostrar que la segunda derivada espacial puede aproximarse de manera más general como:

$$\left(\frac{\partial^2 \phi}{\partial x^2} \right) \simeq \frac{1}{(\Delta x)^2} \delta_x^2 \left[\frac{\theta}{2} (\phi_m^{n+1} + \phi_m^{n-1}) + (1 - \theta) \phi_m^n \right] , \quad (4.21)$$

con θ un parámetro arbitrario. La expresión que teníamos antes, ecuación (4.6), puede recuperarse tomando $\theta = 0$. Esta nueva aproximación corresponde a tomar promedios, con un cierto peso, de operadores en diferencias finitas actuando en diferentes pasos de tiempo. En el caso particular en que $\theta = 1$, la contribución del paso de tiempo intermedio de hecho desaparece por completo.

Si utilizamos esta aproximación para la segunda derivada espacial, pero mantenemos la misma aproximación que antes para la derivada temporal, obtenemos la siguiente aproximación en diferencias finitas de la ecuación de onda:

$$\rho^2 \delta_x^2 \left[\frac{\theta}{2} (\phi_m^{n+1} + \phi_m^{n-1}) + (1 - \theta) \phi_m^n \right] - \delta_t^2 \phi_m^n = 0 . \quad (4.22)$$

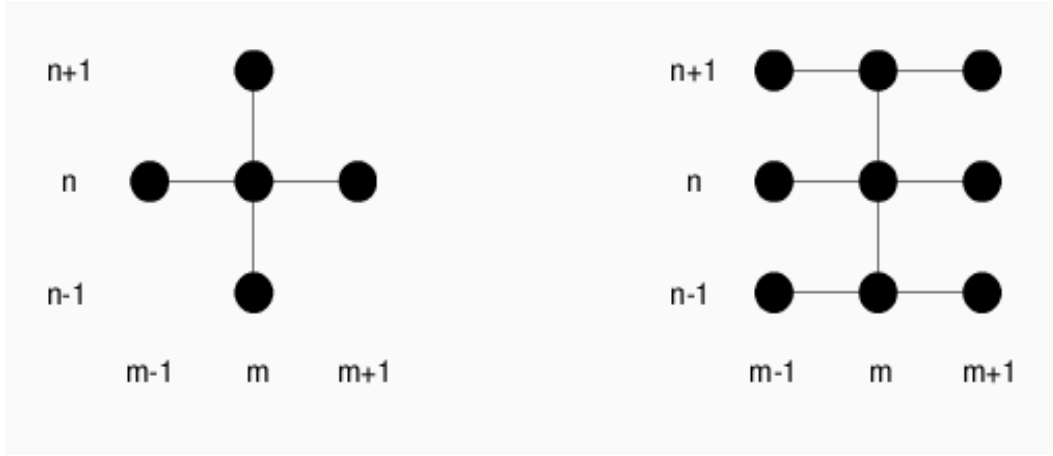


Figura 7: Moléculas computacionales.

Esta es una posible generalización de (4.8), pero claramente no la única, es claro que podemos jugar este juego de muchas maneras y obtener aproximaciones aún más generales, todas ellas válidas, y todas ellas a segundo orden (incluso es posible ingeniárselas para encontrar aproximaciones a cuarto orden o mayor). La aproximación dada por (4.22) tiene una nueva propiedad muy importante: hace referencia no a uno, sino a tres valores diferentes de ϕ en el último paso de tiempo. Esto significa que ahora no podemos resolver para ϕ en el último paso de manera explícita en términos de los valores en los dos pasos anteriores. Debido a esto, a la aproximación (4.22) se le conoce como “aproximación implícita”.

Cuando las ecuaciones para todos los puntos de la malla, incluyendo las fronteras, se consideran a la vez, es posible resolver el sistema completo invirtiendo una matriz no trivial, lo que desde luego toma más tiempo que el necesario para una aproximación explícita. Parecería entonces que no tiene ningún sentido considerar aproximaciones de tipo implícito. Sin embargo, en muchas ocasiones las aproximaciones implícitas resultan tener mejores propiedades que las explícitas, en particular relacionadas con la estabilidad del esquema numérico, concepto que se discutirá en la siguiente sección.

La diferencia entre una aproximación implícita y otra explícita puede verse gráficamente utilizando el concepto de “molécula computacional”, que no es otra cosa que un diagrama que muestra las relaciones entre los distintos puntos utilizados en una aproximación en diferencias finitas. La Figura 7 muestra las moléculas computacionales para los casos implícito y explícito que hemos considerado.

4.5. Consistencia, convergencia y estabilidad

En la última sección nos topamos con quizá uno de los conceptos más importantes de las diferencias finitas: Existe, en general, un número infinito de maneras posibles de aproximar una misma ecuación diferencial. Esto sigue siendo cierto incluso si uno se restringe a considerar aproximaciones con un mismo orden de error.

La multiplicidad de posibles aproximaciones nos lleva inmediatamente a la siguiente pregunta: ¿Cómo saber en que caso usar una cierta aproximación y no otra? Desgraciadamente, no existe una respuesta general a esta pregunta, por lo que muchas veces se dice que las diferencias finitas son un arte más que una ciencia. Sin embargo, si existen guías que nos permiten escoger entre distintas aproximaciones en ciertos casos. Estas guías tienen que ver con los conceptos de consistencia, convergencia y estabilidad.

Consideremos una cierta aproximación en diferencias finitas a una ecuación diferencial. Cuando la malla se refina (es decir, cuando Δt y Δx se hacen cada vez más pequeños), uno esperaría que la aproximación fuera cada vez mejor en el sentido de que los errores de truncación se hacen cada vez más pequeños. Buscamos entonces que en el límite continuo nuestra aproximación se acerca a la ecuación diferencial original y no a otra. Cuando esto ocurre localmente se dice que nuestra aproximación es “consistente”. En general, esta propiedad es muy fácil de ver de la estructura de las aproximaciones en diferencias finitas, y puede comprobarse casi a ojo. Las excepciones importantes son situaciones en las que el sistema de coordenadas es singular, donde mostrar consistencia en el punto singular puede no ser trivial. Por ejemplo, es común que aproximaciones en diferencias finitas “estándar” fallen en el punto $r = 0$ cuando se utilizan coordenadas esféricas. La consistencia es claramente fundamental en una aproximación en diferencias finitas. Si falla, aunque sea en un solo punto, implica que no recuperaremos la solución correcta de la ecuación diferencial.

La consistencia es solo una propiedad local: una aproximación consistente se reduce *localmente* a la ecuación diferencial en el límite continuo. En la práctica, estamos realmente interesados en una propiedad más global. Lo que realmente buscamos es que la aproximación mejore *a un tiempo finito* T cuando refinamos la malla. Es decir, la diferencia entre la solución exacta y la solución numérica a un tiempo fijo T debe tender a cero en el límite continuo. Esta condición se conoce como “convergencia”.

La convergencia es claramente diferente a la consistencia: esquemas consistentes pueden fácilmente no ser convergentes. Esto no es difícil de entender si pensamos que en el límite cuando Δt se hace cero, un tiempo finito T se puede alcanzar solo después de un número infinito de pasos. Esto implica que incluso si el error en cada paso es infinitesimal, su integral total puede muy fácilmente ser finita. La solución numérica puede incluso divergir y el error de hecho resultar infinito. En general es muy difícil verificar analítica-

mente si un esquema de aproximación es convergente o no lo es. Numéricamente, por otro lado, es muy fácil ver si la solución aproximada converge a algo (es decir, no diverge). Lo difícil es saber si la solución numérica converge hacia la solución exacta y no a otra cosa.

Hay otra propiedad importante de las aproximaciones en diferencias finitas. Independientemente del comportamiento de la solución a la ecuación diferencial, debemos pedir que las *soluciones exactas de las ecuaciones en diferencias finitas* permanezcan acotadas para cualquier tiempo finito T y cualquier intervalo de tiempo Δt . Este requisito se conoce como “estabilidad”, e implica que ninguna componente de los datos iniciales debe amplificarse arbitrariamente. La estabilidad es una propiedad del sistema de ecuaciones en diferencias finitas, y es esencialmente el análogo discreto de la condición de que un sistema de ecuaciones diferenciales este bien planteado. Una aproximación inestable es inútil en la práctica.

Un resultado fundamental de la teoría de las aproximaciones en diferencias finitas es el teorema de Lax (para su demostración ver, por ejemplo, la referencia [45]):

TEOREMA: *Dado un problema de valores iniciales bien planteado matemáticamente y una aproximación en diferencias finitas a él que satisface la condición de consistencia, entonces la estabilidad es condición necesaria y suficiente para la convergencia.*

Este teorema es de gran importancia pues relaciona el objetivo final de toda aproximación en diferencias finitas, es decir, la convergencia a la solución exacta, con una propiedad que es mucho más fácil de probar: la estabilidad.

4.6. Estabilidad de von Neumann

Un método general para probar la estabilidad de los sistemas de ecuaciones en diferencias finitas se obtiene de la definición de estabilidad directamente. Comenzamos por escribir las ecuaciones en diferencias finitas como:

$$\mathbf{v}^{n+1} = \mathbf{B} \mathbf{v}^n, \quad (4.23)$$

donde \mathbf{v}^n es el vector solución en el nivel de tiempo n , y \mathbf{B} es una matriz (en general con muchos ceros). Es importante notar que toda aproximación en diferencias finitas, incluso aquellas que involucran más de dos niveles de tiempo (como las que introdujimos para la ecuación de onda en las secciones anteriores), puede escribirse de la forma (4.23) simplemente introduciendo variables auxiliares.

Como el vector \mathbf{v}^n se puede escribir como una combinación lineal de los eigenvectores de \mathbf{B} , el requerimiento de estabilidad se reduce a pedir que la matriz \mathbf{B} no amplifique ninguno de sus eigenvectores, es decir, que la magnitud del mayor de sus eigenvalores sea

menor o igual que 1. Dicho de otra forma, el “radio espectral” de \mathbf{B} debe ser menor o igual a 1.

El análisis de estabilidad basado en la idea que hemos expuesto es muy general, pero requiere del conocimiento de los coeficientes de \mathbf{B} en todo el espacio, incluida la frontera. Existe, sin embargo, un método muy popular de análisis de estabilidad que, aunque en principio solo nos proporciona condiciones necesarias para la estabilidad, en muchos casos resulta también dar condiciones suficientes. Este método, introducido originalmente por von Newmann, está basado en una descomposición en Fourier de la solución.

Para introducir este método, empezamos entonces por expandir la solución de (4.23) en una serie de Fourier:

$$\mathbf{v}^n(\mathbf{x}) = \sum_{\mathbf{k}} \tilde{\mathbf{v}}^n(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (4.24)$$

donde la suma es sobre todos los vectores de onda \mathbf{k} que pueden representarse en la malla.⁵ Si ahora sustituimos esto en la ecuación original (4.23) obtenemos:

$$\tilde{\mathbf{v}}^{n+1} = \mathbf{G}(\Delta x, \Delta t, \mathbf{k}) \tilde{\mathbf{v}}^n. \quad (4.25)$$

La matriz \mathbf{G} se conoce como “matriz de amplificación”. La condición de estabilidad ahora corresponde a pedir que ningún modo de Fourier se amplifique, es decir, el radio espectral de \mathbf{G} debe ser menor o igual que uno. Esta es la condición de estabilidad de von Newmann.

Es importante recalcar que para poder utilizar este criterio de estabilidad se han supuesto dos cosas: 1) Las condiciones de frontera son periódicas, pues de otra forma no se puede hacer la expansión en serie de Fourier, y 2) los elementos de la matriz \mathbf{B} son constantes, pues de otra forma no se pueden separar los diferentes modos de Fourier.

Como ejemplo del análisis de estabilidad de von Newmann vamos a utilizar la aproximación implícita a la ecuación de onda que derivamos antes (ecuación (4.22)):

$$\rho^2 \delta_x^2 \left[(\theta/2) (\phi_m^{n+1} + \phi_m^{n-1}) + (1 - \theta) \phi_m^n \right] - \delta_t^2 \phi_m^n = 0. \quad (4.26)$$

Consideraremos ahora un modo de Fourier de la forma:

$$\phi_m^n = \xi^n e^{imk\Delta x}. \quad (4.27)$$

Si sustituimos esto en la ecuación en diferencias finitas encontramos, después de un poco de álgebra, la siguiente ecuación cuadrática para ξ :

$$A\xi^2 + B\xi + C = 0, \quad (4.28)$$

⁵La menor longitud de onda que se puede representar en la malla es claramente $2\Delta x$, también conocida como la “longitud de onda de Niquist”. Esto implica que el valor máximo de cualquier componente del vector de onda es $\pi/\Delta x$.

con coeficientes dados por

$$A = \rho^2 \theta [\cos(k\Delta x) - 1] - 1, \quad (4.29)$$

$$B = 2\rho^2 (1 - \theta) [\cos(k\Delta x) - 1] + 2, \quad (4.30)$$

$$C = \rho^2 \theta [\cos(k\Delta x) - 1] - 1. \quad (4.31)$$

La dos raíces de la ecuación cuadrática son, claramente:

$$\xi_{\pm} = \frac{-B \pm (B^2 - 4AC)^{1/2}}{2A}, \quad (4.32)$$

y la solución general a la ecuación en diferencias finitas resulta ser:

$$\phi_m^n = \sum_k \left[Z_k^+ (\xi_+(k))^n + Z_k^- (\xi_-(k))^n \right] e^{imk\Delta x}, \quad (4.33)$$

donde Z_k^+ y Z_k^- son constantes arbitrarias.

Por otro lado, del hecho de que $A = C$ es fácil mostrar que:

$$|\xi_+ \xi_-| = \left| \frac{C}{A} \right| = 1. \quad (4.34)$$

Esta es una propiedad muy importante, implica que si el sistema es estable para toda k , es decir, si $|\xi_{\pm}(k)| \leq 1$, entonces necesariamente será también no disipativo (los modos de Fourier no solo no crecen, sino que tampoco se disipan). Para que el sistema sea estable debemos ahora pedir que:

$$\xi_+(k) = \xi_-(k) = 1. \quad (4.35)$$

Es fácil ver que esto ocurrirá siempre que:

$$B^2 - 4AC \leq 0. \quad (4.36)$$

Substituyendo los valores de los coeficientes A , B y C en esta expresión obtenemos la siguiente condición de estabilidad:

$$\rho^2 (1 - 2\theta) [1 - \cos(k\Delta x)] - 2 \leq 0. \quad (4.37)$$

Como queremos que esto se cumpla para toda k , debemos considerar el caso cuando el lado izquierdo alcanza su valor máximo. Si tomamos $\theta < 1/2$, esto ocurrirá para $k = \pi/\Delta x$, en cuyo caso la condición de estabilidad toma la forma simple:

$$\rho^2 \leq 1/(1 - 2\theta). \quad (4.38)$$

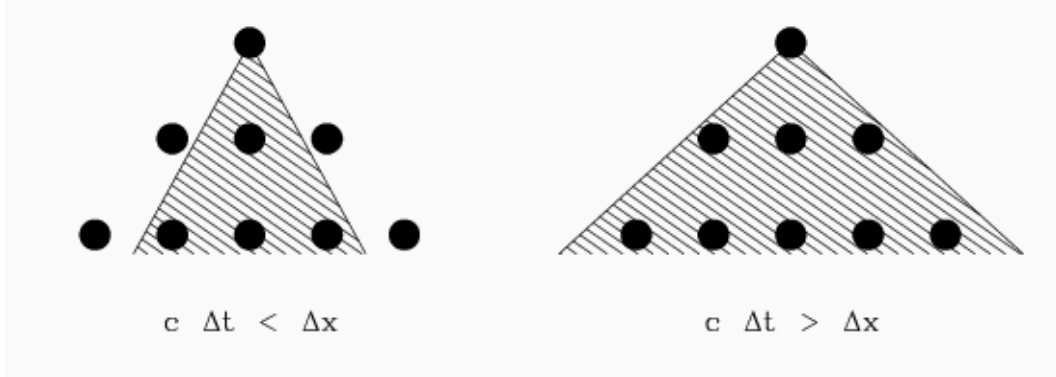


Figura 8: Condición de estabilidad CFL. Para $c\Delta t \leq \Delta x$, el dominio de dependencia numérico es mayor que el dominio de dependencia físico (región sombreada), y el sistema es estable. Para $c\Delta t > \Delta x$ la situación es la opuesta, y el sistema es inestable.

Para el esquema explícito se tiene $\theta = 0$, y la condición se reduce a la bien conocida condición de estabilidad de Courant-Friedrich-Lewy (CFL):

$$c\Delta t \leq \Delta x . \quad (4.39)$$

La condición CFL tiene una clara interpretación geométrica: el dominio de dependencia numérico debe ser mayor que el dominio de dependencia físico, y no al revés (ver Figura 8). Si esto no fuera así, resultaría imposible para la solución numérica converger a la solución exacta, pues al refinar la malla siempre habría información física relevante que quedaría fuera del dominio de dependencia numérico. Y, como hemos visto, el teorema de Lax implica que si no hay convergencia el sistema es inestable.

El argumento que acabamos de dar claramente solo se aplica a esquema explícitos. Esto se debe a que para un sistema implícito, el dominio de dependencia numérico es toda la malla. En este caso no hay un argumento geométrico simple que nos diga cual debe ser la condición de estabilidad.

Para llegar a la condición de estabilidad (4.38), supusimos que $\theta < 1/2$. Si, por otro lado, tomamos $\theta \geq 1/2$, debemos regresar a la condición general (4.37). Sin embargo, en este caso es fácil ver que la condición se satisface siempre. Esto significa que un sistema implícito con $\theta \geq 1/2$ es estable para todo valor de ρ , es decir, es “incondicionalmente estable”.

Esto nos lleva a quizá una de las lecciones más importantes de la teoría de las diferencias finitas: Los esquemas más simples no siempre tienen las mejores propiedades de

estabilidad.⁶

4.7. Ejemplos

Con el fin de ilustrar los diferentes conceptos que hemos introducido hasta ahora, vamos a considerar los resultados de experimentos numéricos utilizando la ecuación de onda y nuestra aproximación (4.22). En todas las simulaciones mostradas aquí se toma como región de integración $x \in [0, 100]$, y se utilizan condiciones de frontera periódicas. También tomamos la velocidad de onda c igual a 1.

Primero consideremos el caso cuando Δx y Δt son tales que

$$\Delta x = 1, \quad \Delta t = 1/2. \quad (4.40)$$

El parámetro de Courant resulta ser $\rho = 1/2$. La Figura 9 muestra resultados de dos simulaciones usando estos parámetros de malla. La gráfica de la izquierda corresponde a un esquema explícito, y la de la derecha a un sistema implícito con $\theta = 1/2$. En ambos casos los datos iniciales (línea punteada) corresponden a un paquete gaussiano moviéndose inicialmente a la derecha, y los resultados de la simulación numérica se muestran después de 120 pasos de tiempo.

De la figura vemos que ambas simulaciones son muy similares. En ambos casos se nota que el paquete inicial se ha dispersado, en contraste con la solución exacta para la cual el paquete se propaga manteniendo su forma original. La dispersión es un efecto numérico, ocasionado por el hecho de que, en la aproximación numérica, diferentes modos de Fourier viajan a diferentes velocidades.

La siguiente simulación corresponde a la misma situación, pero ahora tomando $\Delta x = \Delta t = 1$, correspondiente a un parámetro de Courant $\rho = 1$. La Figura 10 muestra los resultados de esta simulación. Nótese que como Δt es ahora el doble de grande, se requieren solo 60 pasos de tiempo para llegar a la misma etapa.

Lo primero que hay que notar es que para el esquema explícito la dispersión ha desaparecido. Esto es un resultado sorprendente, y solo ocurre para la ecuación de onda en una dimensión: En este caso es posible mostrar que para $\rho = 1$ todos los errores numéricos se cancelan y la solución numérica es de hecho exacta. El esquema implícito, sin embargo, sigue siendo dispersivo (incluso más que antes).

⁶Esto es aún más evidente para sistemas de ecuaciones de tipo “parabólico” (como la ecuación de calor), para los que los esquemas explícitos son prácticamente inútiles pues la condición CFL toma la forma $\Delta t < \Delta x^2$. El problema con esto es que significa que si reducimos Δt a la mitad, debemos reducir Δt a la cuarta parte, de manera que integrar hasta un tiempo finito T rápidamente se vuelve prohibitivo en tiempo de máquina.

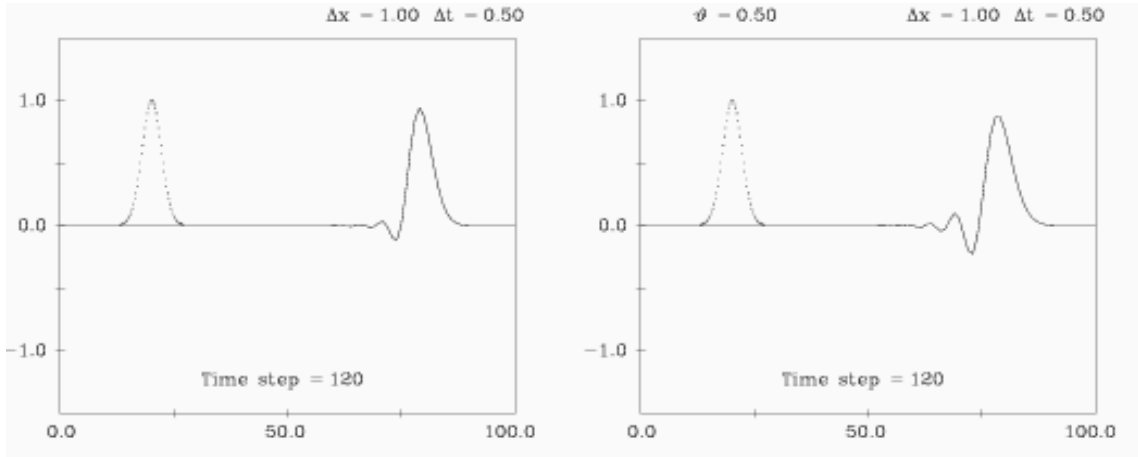


Figura 9: Simulación numérica para el caso $\rho = 1/2$, utilizando un sistema explícito (izquierda), y uno implícito con $\theta = 1/2$ (derecha).

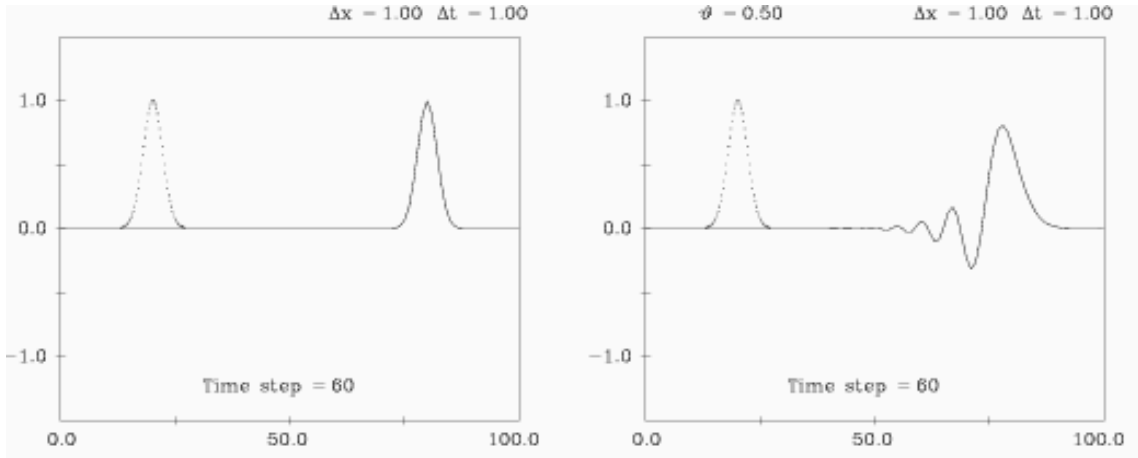


Figura 10: Simulación numérica para el caso $\rho = 1$.

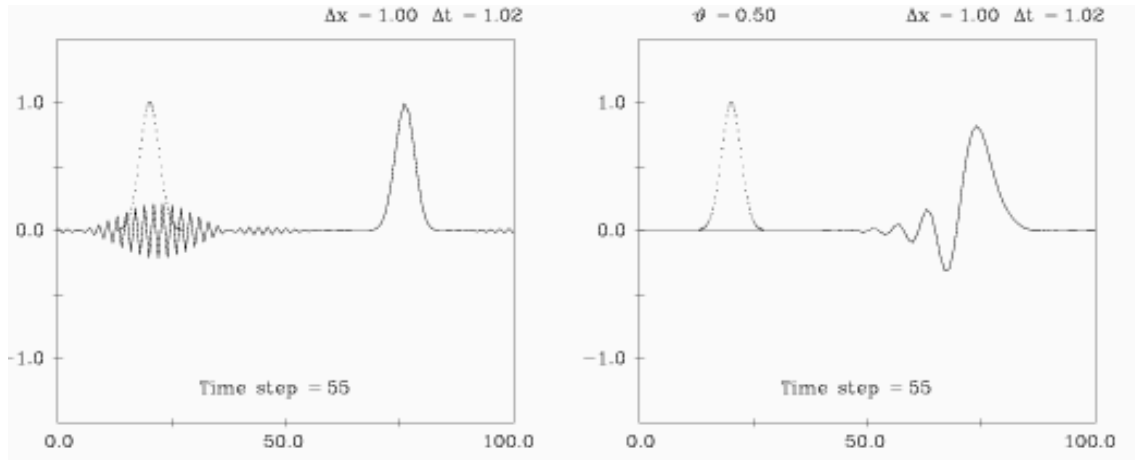


Figura 11: Simulación numérica para el caso $\rho = 1,02$.

Finalmente, en la Figura 11 se muestran resultados de una simulación con $\Delta x = 1$ y $\Delta t = 1.02$, para un parámetro de Courant de $\rho = 1,02$. Para el esquema explícito es evidente que una perturbación de alta frecuencia ha aparecido después de solo 55 pasos de tiempo. Si la simulación continua, esta perturbación crece exponencialmente y rápidamente domina la solución completa. Esta es una clásica inestabilidad numérica: una perturbación localizada, generalmente de alta frecuencia (aunque no siempre), que crece exponencialmente sin moverse. El esquema implícito, por otro lado, no muestra ninguna señal de una inestabilidad. De hecho, podemos continuar la integración en este caso y nunca encontrar ningún problema. La dispersión, sin embargo, si esta presente y eventualmente causará que la solución numérica difiera mucho de la solución analítica.

5. Ejemplos de aplicaciones

En esta sección consideraremos algunos ejemplos sencillos de simulaciones en relatividad numérica. Nos limitaremos a dos casos específicos: relatividad en 1+1 dimensiones (es decir, una sola dimensión espacial) y relatividad en simetría esférica. El caso de sistemas tridimensionales sin simetrías es considerablemente mas complicado y requeriría de un curso en si mismo.

5.1. Relatividad de juguete en 1+1 dimensiones

El primer ejemplo que consideraremos es el de relatividad en 1+1 dimensiones, es decir, una dimensión de tiempo y una de espacio. Ahora bien, la geometría diferencial nos dice que para un espacio de 2 dimensiones como el que estamos considerando (es decir, una superficie), el tensor de curvatura de Riemann tiene solo una componente independiente dada esencialmente por la curvatura gaussiana. Si asumimos, además, que estamos en el vacío, las ecuaciones de Einstein implican que dicha componente es cero, lo que nos lleva al bien conocido resultado de que en 1+1 dimensiones no hay gravedad en el vacío.

Podría uno entonces pensar que estudiar este sistema es totalmente trivial, y que no hay nada que aprender de él. Sin embargo, resulta ser que este es un sistema ideal para estudiar las propiedades de diferentes elecciones de norma. En efecto, aún cuando en una dimensión espacial estemos limitados al espacio de Minkowski, nada nos obliga a utilizar las coordenadas usuales, por lo que podemos tener evoluciones altamente no triviales del sistema de coordenadas que se reflejarán en la evolución de la métrica espacial y la curvatura extrínseca. Por simplicidad, en toda la discusión que sigue asumiremos que el vector de corrimiento es cero.

Comencemos por considerar la condición de foliación. Queda claro que en este caso la foliación maximal ($K = 0$) nos forzaría a considerar solamente hiper-superficies planas, es decir, nos obligaría a utilizar la foliación standard de Minkowski. Para obtener algo menos trivial es preferible utilizar la foliación de Bona-Masso (2.56):

$$\partial_t \alpha = -\alpha^2 f(\alpha) K , \quad (5.1)$$

donde en este caso la traza de la curvatura extrínseca esta dada simplemente por $K = K_x^x$.

Las ecuaciones ADM en el caso 1+1, junto con la condición de foliación, pueden entonces reescribirse a primer orden como:

$$\partial_t \alpha = -\alpha^2 f K , \quad (5.2)$$

$$\partial_t g = -2\alpha g K , \quad (5.3)$$

y

$$\partial_t D_\alpha + \partial_x (\alpha f K) = 0 , \quad (5.4)$$

$$\partial_t D_g + \partial_x (2\alpha K) = 0 , \quad (5.5)$$

$$\partial_t K + \partial_x (\alpha D_\alpha / g) = \alpha (K^2 - D_\alpha D_g / 2g) , \quad (5.6)$$

donde hemos definido $g := g_{xx}$, $D_\alpha := \partial_x \ln \alpha$ y $D_g := \partial_x \ln g$.

Resulta ser que la ecuación de evolución para K puede reescribirse como una ley de conservación de la siguiente forma:

$$\partial_t (g^{1/2} K) + \partial_x (\alpha D_\alpha / g^{1/2}) = 0 . \quad (5.7)$$

Si ahora definimos el vector $\vec{v} := (D_\alpha, D_g, \tilde{K})$, con $\tilde{K} := g^{1/2} K$, entonces las ecuaciones de evolución para las variables de primer orden pueden escribirse como un sistema conservativo de la forma:

$$\partial_t \vec{v} + \partial_x (\mathbf{M} \vec{v}) = 0 , \quad (5.8)$$

donde la matriz característica \mathbf{M} esta dada por:

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & \alpha f / g^{1/2} \\ 0 & 0 & 2\alpha / g^{1/2} \\ \alpha / g^{1/2} & 0 & 0 \end{pmatrix} . \quad (5.9)$$

Esta matriz tiene los siguientes eigenvalores

$$\lambda_0 = 0 , \quad \lambda_\pm = \pm \alpha (f/g)^{1/2} , \quad (5.10)$$

con correspondientes eigenvectores:

$$\vec{e}_0 = (0, 1, 0) , \quad \vec{e}_\pm = (f, 2, \pm f^{1/2}) . \quad (5.11)$$

Como los eigenvalores son reales para $f > 0$, y los eigenvectores son linealmente independientes, el sistema (5.8) es fuertemente hiperbólico. Los eigencampos están dados por:

$$\vec{\omega} = \mathbf{R}^{-1} \vec{v} , \quad (5.12)$$

con \mathbf{R} la matriz de eigenvectores columna. Encontramos (utilizando una elección adecuada de normalización):

$$\omega_0 = D_\alpha / f - D_g / 2 , \quad \omega_\pm = \tilde{K} \pm D_\alpha / f^{1/2} , \quad (5.13)$$

que puede invertirse fácilmente para dar:

$$\tilde{K} = \frac{(\omega_+ + \omega_-)}{2} , \quad (5.14)$$

$$D_\alpha = \frac{f^{1/2} (\omega_+ - \omega_-)}{2} , \quad (5.15)$$

$$D_g = \frac{(\omega_+ - \omega_-)}{f^{1/2}} - 2\omega_0 . \quad (5.16)$$

Claramente, el eigencampo ω_0 se “propaga” con velocidad cero, y los otros dos se propagan con la “velocidad de norma” $\lambda_\pm^f = \pm\alpha\sqrt{f/g}$. Dada una perturbación inicial, podemos esperar que en general esta de lugar a dos pulsos moviéndose en direcciones opuestas.

Es importante notar que con las eigenfunciones normalizadas como están arriba, sus ecuaciones de evolución también resultan conservativas, y toman la forma simple:

$$\partial_t \vec{\omega} + \partial_x (\mathbf{\Lambda} \vec{\omega}) = 0 , \quad (5.17)$$

con $\mathbf{\Lambda} = \text{diag} \{\lambda_i\}$. Si, por otro lado, las eigenfunciones son renormalizadas como $\omega'_i = F_i(\alpha, g) \omega_i$, entonces las ecuaciones de evolución para las ω'_i en general ya no serán conservativas y habrán fuentes no triviales. El punto crucial es que existe una normalización en la que las ecuaciones son puramente conservativas, aquella dada en (5.13).

Consideremos ahora algunos ejemplos numéricos. Para tener una evolución no trivial, tomamos datos iniciales correspondientes a una superficie curva en el espacio-tiempo de Minkowski dada en términos de las coordenadas usuales (t_M, x_M) como:

$$t_M = h(x_M) . \quad (5.18)$$

Si tomamos la coordenada espacial x de forma tal que inicialmente coincida con la coordenada de Minkowski x_M , no es difícil mostrar que los valores iniciales de la métrica y la curvatura extrínseca son:

$$g = 1 - h'^2 \Rightarrow D_g = -2h'h''/g , \quad (5.19)$$

$$K := K_x^x = -h''/g^{3/2} \Rightarrow \tilde{K} = -h''/g . \quad (5.20)$$

Para completar los datos iniciales tomamos como lapso inicial $\alpha = 1$.

En las simulaciones que discutimos a continuación, la función $h(x)$ tiene un perfil gaussiano:

$$h = A \exp \left[(x - x_c)^2 / \sigma^2 \right] . \quad (5.21)$$

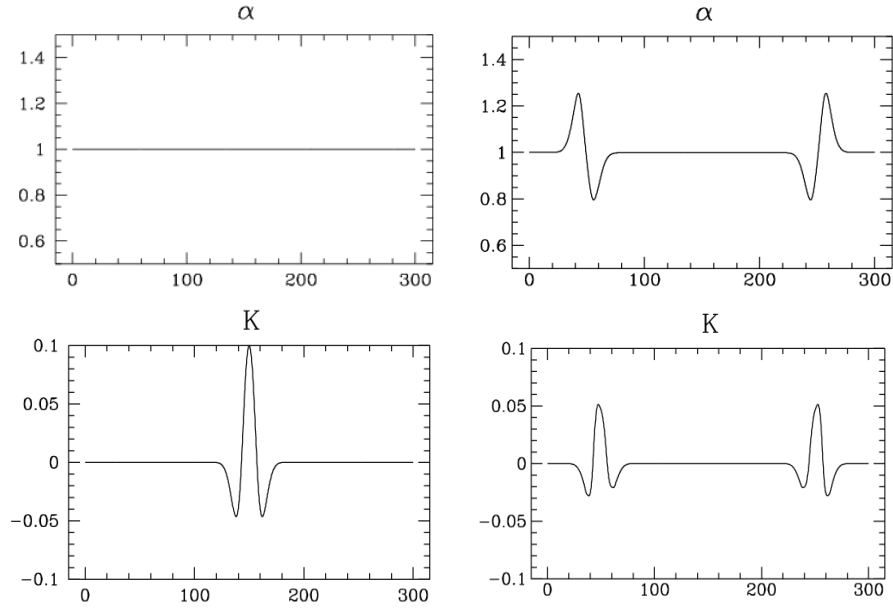


Figura 12: Evolución del lapso α y la traza de la curvatura extrínseca K para una simulación con $f = 1$. Columna izquierda: Datos iniciales. Columna derecha: Valores a $t = 100$.

Para concretar, tomaremos $A = 5$, $\sigma = 10$ y la gaussiana centrada en medio de la red computacional. Además, en todas las simulaciones mostradas a continuación usaremos un paso de tiempo y un intervalo espacial dados por $\{\Delta t = 0,125, \Delta x = 0,25\}$, y un dominio computacional en el intervalo $[0, 300]$.

Consideremos primero el caso $f = 1$, es decir, la foliación armónica. En la figura 12 se muestran los datos iniciales para α y K , así como sus valores a $t = 100$. De la figura podemos ver que el pulso inicial se separa en dos pulsos de menor tamaño que se mueven en direcciones opuestas con velocidad constante y conservando su forma. El resultado es muy similar al que se hubiera esperado de una simple ecuación de onda.

Cuando f es una constante diferente de 1, la situación ya no es tan simple. La perturbación inicial se separa en dos pulsos, pero esos pulsos no conservan su forma al propagarse, y de hecho después de un tiempo dejan de ser suaves (desarrollan derivadas infinitas). Como ejemplo de esto, en la figura 13 se muestra el resultado de una simulación con $f = 1,69$ (correspondiente a una velocidad de norma $v \sim \sqrt{f} = 1,3$). Para el tiempo $t = 75$ podemos observar que los pulsos correspondientes al lapso α han desarrollado gradientes muy grandes al frente y en la parte posterior, mientras que los pulsos correspondientes a la curvatura extrínseca K han desarrollado picos muy angostos. La situación es similar

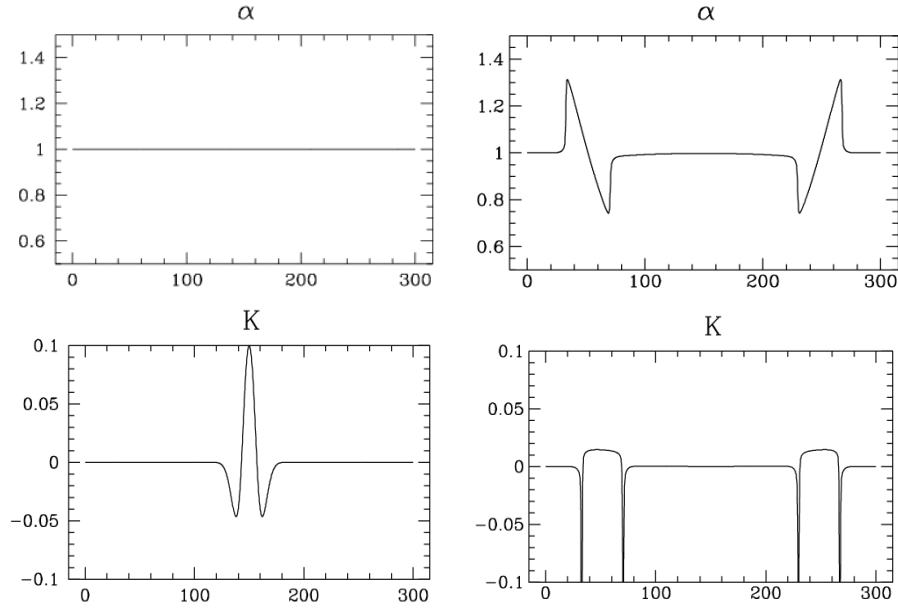


Figura 13: Evolución del lapso α y la traza de la curvatura extrínseca K para una simulación con $f = 1,69$. Columna izquierda: Datos iniciales. Columna derecha: Valores a $t = 75$.

para valores de f menores que uno (pero mayores que cero), salvo que en este caso los gradientes altos ocurren en la parte central de cada pulso y no en sus extremos.

Lo que ocurre para valores constantes de f diferentes de 1 es un fenómeno conocido como “choques de norma”, en el que la evolución del sistema de coordenadas da lugar a una singularidad de coordenadas en un tiempo finito, asociada al cruce de las líneas características. Este es un ejemplo claro de como datos iniciales perfectamente razonables, y una evolución con ecuaciones bien planteadas, pueden de todas formas dar lugar a situaciones patológicas. Los choques de norma son una propiedad de la familia de foliaciones de Bona-Masso y no son de ninguna manera un artefacto numérico. Una discusión sobre su origen esta fuera del alcance de este curso, pero los lectores interesados pueden consultar las referencias [2, 7, 8, 3].

5.2. Relatividad en simetría esférica

Como segundo ejemplo consideraremos el caso de espacio-tiempos con simetría esférica. En este caso la gravitación ya no es trivial, después de todo un gran número de sistemas astrofísicos interesantes (estrellas, agujeros negros, etc.) tienen, en muy buena aproxi-

mación, simetría esférica. Lo que si es importante mencionar es que, debido al carácter transversal de las ondas gravitacionales, en simetría esférica no se emite radiación gravitacional. De hecho, sucede igual en la teoría electromagnética, donde tampoco hay ondas electromagnéticas esféricas (monopolares).

Un fenómeno particularmente interesante que puede estudiarse en simetría esférica es el colapso gravitacional de diverso tipos de materia, y la formación de agujeros negros. Este tipo de estudios llevaron, a principios de los años noventa, al descubrimiento por parte de Choptuik de los fenómenos críticos asociados a la formación de agujeros negros [20].

Para comenzar nuestro análisis escribimos la métrica espacial en simetría esférica como:

$$dl^2 = A(r, t)dr^2 + r^2 B(r, t)d\Omega^2, \quad (5.22)$$

con A y B funciones positivas y $d\Omega^2$ el elemento de ángulo sólido: $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$. Nótese que ya hemos factorizado el término r^2 de la métrica angular, lo que facilita la regularización de las ecuaciones cerca del origen.

Como nos interesa escribir las ecuaciones de evolución a primer orden, también introducimos las siguientes variables auxiliares:

$$D_A := \partial_r \ln A, \quad D_B := \partial_r \ln B. \quad (5.23)$$

También trabajaremos con las componentes mixtas de la curvatura extrínseca: $K_A := K_r^r$, $K_B := K_\theta^\theta = K_\phi^\phi$.

Con esta notación, las ecuaciones de evolución de ADM toman la forma (para el caso de vacío y vector de corrimiento igual a cero):

$$\partial_t A = -2\alpha A K_A, \quad (5.24)$$

$$\partial_t B = -2\alpha B K_B, \quad (5.25)$$

$$\partial_t D_A = -2\alpha [K_A D_\alpha + \partial_r K_A], \quad (5.26)$$

$$\partial_t D_B = -2\alpha [K_B D_\alpha + \partial_r K_B], \quad (5.27)$$

$$\begin{aligned} \partial_t K_A = & -\frac{\alpha}{A} \left[\partial_r (D_\alpha + D_B) + D_\alpha^2 - \frac{D_\alpha D_A}{2} \right. \\ & + \frac{D_B^2}{2} - \frac{D_A D_B}{2} - A K_A (K_A + 2K_B) \\ & \left. - \frac{1}{r} (D_A - 2D_B) \right], \end{aligned} \quad (5.28)$$

$$\begin{aligned} \partial_t K_B = & -\frac{\alpha}{2A} \left[\partial_r D_B + D_\alpha D_B + D_B^2 - \frac{D_A D_B}{2} \right. \\ & \left. - \frac{1}{r} (D_A - 2D_\alpha - 4D_B) - \frac{2(A - B)}{r^2 B} \right] \\ & + \alpha K_B (K_A + 2K_B), \end{aligned} \quad (5.29)$$

donde, de manera similar a la sección anterior, hemos definido $D_\alpha := \partial_r \ln \alpha$.

Para las constricciones hamiltoniana y de momento encontramos, respectivamente:

$$\begin{aligned} \partial_r D_B &= \frac{1}{r^2 B} (A - B) + AK_B (2K_A + K_B) \\ &+ \frac{1}{r} (D_A - 3D_B) + \frac{D_A D_B}{2} - \frac{3D_B^2}{4}, \end{aligned} \quad (5.30)$$

$$\partial_r K_B = (K_A - K_B) \left[\frac{1}{r} + \frac{D_B}{2} \right], \quad (5.31)$$

Hace falta ahora mencionar un problema específico al uso de las coordenadas esféricas (algo similar ocurre en coordenadas cilíndricas). Como podemos ver de las ecuaciones anteriores, en el origen $r = 0$ varios términos se vuelven singulares. Para resolver el problema hacen falta dos ingredientes. El primer ingrediente consiste en fijar las condiciones de regularidad de las diferentes variables dinámicas, es decir, su comportamiento cerca del origen. No es difícil ver que el exigir que todas las variables estén bien definidas en el origen implica el siguiente comportamiento para r pequeña:

$$\alpha \sim \alpha_0 + \mathcal{O}(r^2), \quad (5.32)$$

$$A \sim A^0 + \mathcal{O}(r^2), \quad (5.33)$$

$$B \sim B^0 + \mathcal{O}(r^2), \quad (5.34)$$

$$D_\alpha \sim \mathcal{O}(r), \quad (5.35)$$

$$D_A \sim \mathcal{O}(r), \quad (5.36)$$

$$D_B \sim \mathcal{O}(r), \quad (5.37)$$

$$K_A \sim K_A^0 + \mathcal{O}(r^2), \quad (5.38)$$

$$K_B \sim K_B^0 + \mathcal{O}(r^2), \quad (5.39)$$

con $\{\alpha_0, A^0, B^0, K_A^0, K_B^0\}$ funciones del tiempo pero no de r . Estas condiciones de simetría son de hecho muy sencillas de imponer en una simulación numérica. Por ejemplo, se puede utilizar una malla que “salte” sobre el origen, comenzando en $r = \Delta r/2$, y después se pueden obtener datos en el punto ficticio de frontera en $r = -\Delta r/2$ demandando que $\{\alpha, A, B, K_A, K_B\}$ sean funciones pares en $r = 0$ y que $\{D_A, D_B\}$ sean impares. Como $\{D_\alpha, D_A, D_B\}$ son proporcionales a r cerca del origen, vemos que términos del tipo $D_{\{\alpha, A, B\}}/r$ en las ecuaciones ADM son regulares y no causan ningún problema (recuérdese que nunca evaluamos exactamente en $r = 0$).

Sin embargo, existe todavía un problema mas serio. Podemos ver que tanto en la constricción hamiltoniana como en la ecuación de evolución para K_B hay un término que tiene la forma $(A - B)/r^2$, mientras que en la constricción de momento hay un término

del tipo $(K_A - K_B)/r$. Dado el comportamiento de las variables cerca de $r = 0$, estos términos parecen divergir en el origen. La razón por la que esto no ocurre en realidad es que, cerca del origen, debemos pedir las condiciones de regularidad extra

$$A - B \sim \mathcal{O}(r^2) , \quad K_A - K_B \sim \mathcal{O}(r^2) , \quad (5.40)$$

es decir, $A^0 = B^0$ y $K_A^0 = K_B^0$. Estas condiciones extra provienen simplemente del hecho de que el espacio-tiempo debe permanecer localmente plano en $r = 0$.

Imponer todas las condiciones de regularidad en $r = 0$ numéricamente no es trivial. Esto se debe a que tenemos mas condiciones de regularidad que variables, por lo que el sistema esta sobre-determinado. En el caso analítico resulta fácil ver que si todas las condiciones se satisfacen inicialmente, entonces se mantendrán satisfechas a todo tiempo. Numéricamente, sin embargo, esto no es cierto debido a los errores de truncación. Esto significa que rápidamente alguna de las condiciones deja de satisfacerse y la evolución se vuelve irregular en el origen (los errores crecen sin control y el código falla). Existen métodos generales para regularizar las ecuaciones en simetría esférica, pero describirlos nos llevaría demasiado espacio en estas notas, los lectores interesados pueden consultar [6].

Es importante mencionar, sin embargo, que hay una manera muy sencilla, y por lo tanto muy comúnmente utilizada, de evitar el problema de la falta de regularidad. Este método consiste en restringir la elección de la foliación y utilizar la llamada “norma de área”, en la que se obliga a que la variable métrica angular sea $B = 1$. Esto implica que $g_{\theta\theta} = r^2$, por lo que el área de esferas es siempre $4\pi r^2$. Es por esto que en este caso a r se le llama el “radio de área”. Si uno toma esta elección, es fácil ver que imponiendo la condición de frontera $A(r = 0) = 1$ uno puede resolver para $A(r)$ simplemente integrando la constricción hamiltoniana hacia afuera con $B = 1$ y $D_B = 0$ (e ignorando las ecuaciones de evolución). En este caso el término $(A - B)/r^2$ ya no causa ningún problema. El forzar a que se tenga $B = 1$ todo el tiempo implica que $K_B = \partial_t K_B = 0$, lo que lleva a una ecuación diferencial para α en términos de r que se debe resolver a cada paso, dando lugar a la llamada “foliación polar”. La norma de área tiene ciertas desventajas, sin embargo. En primer lugar, solo puede utilizarse en situaciones en las que el área de esferas es una función creciente del radio, es decir, no permite considerar geometrías de tipo “bolsa de oro”.⁷ Además, tiene la desventaja de que no permite tampoco penetrar al interior de horizontes aparentes asociados a agujeros negros. Esto se debe a que dentro de un horizonte aparente es imposible mantener el área de una esfera constante sin utilizar un vector de corrimiento no trivial.

⁷Una geometría tipo “bolsa de oro” corresponde a un espacio tal que, al alejarnos de un cierto punto el área de las esferas primero aumenta, luego disminuye, y finalmente aumenta de nuevo, de la misma manera que ocurre con la circunferencia de círculos si nos movemos en el interior de una bolsa que ha sido atada con una cuerda.

Otro punto de importancia es que las ecuaciones ADM en simetría esférica dadas anteriormente no resultan ser fuertemente hiperbólicas cuando uno utiliza condiciones de foliación comunes (maximal, Bona-Masso, etc.). Este problema no es difícil de resolver, y el lector interesado puede consultar [6] donde se da un ejemplo sencillo de como construir un sistema fuertemente hiperbólico en simetría esférica utilizando las constricciones (la solución no es única, y no queda claro cual elección dentro de la infinidad posible es la mejor).

Para terminar, consideraremos como ejemplo de una simulación numérica el caso de un simple agujero negro de Schwarzschild. Uno podría pensar que, ya que el espacio-tiempo de Schwarzschild es estático, en la simulación numérica no habrá ninguna evolución. Esto no es correcto por dos razones: En primer lugar, el espacio-tiempo de Schwarzschild solo es estático fuera del horizonte del agujero negro, dentro del horizonte la solución es dinámica y la singularidad se alcanza en un tiempo propio finito. En segundo lugar, como hemos visto ya en la sección anterior, incluso en espacio-tiempos estáticos es interesante estudiar la evolución artificial inducida por una elección de norma no trivial. Estudiar el caso de Schwarzschild permite adquirir experiencia que luego puede utilizarse en situaciones realmente dinámicas, como el colapso gravitacional o la colisión de objetos compactos, donde uno espera encontrar agujeros negros durante la evolución incluso si no había ninguno presente inicialmente.

Comenzamos primero por la elección de los datos iniciales. Es claro que tomar simplemente la métrica de Schwarzschild evaluada en $t = 0$ no es una buena elección debido a que esta métrica es singular en el horizonte. Una mejor elección es utilizar la métrica en coordenadas isotrópicas que se obtienen a partir de las coordenadas estandar de Schwarzschild mediante el cambio en la coordenada radial $R = r(1 + M/2r)^2$, con R la coordenada radial de Schwarzschild y r la coordenada radial isotrópica (ver por ejemplo [42]). De hecho, ya obtuvimos la solución para la métrica espacial en este caso cuando discutimos la descomposición conforme de York-Lichnerowicz para resolver las constricciones en las secciones 2.5 y 2.6 de estas notas. Dicha solución estaba dada por la ecuación (2.48) que repetimos a continuación:

$$dl^2 = \psi^4 \left(dr^2 + r^2 d\Omega^2 \right) , \quad (5.41)$$

donde el factor conforme ψ es simplemente:

$$\psi = 1 + M/2r , \quad (5.42)$$

con M la masa del agujero negro. Esta es la elección mas común de datos iniciales para Schwarzschild, aunque no la única. Otra elección común son las llamadas “coordenadas de Kerr-Schild” que no discutiremos aquí (pero el lector puede referirse a [41]).

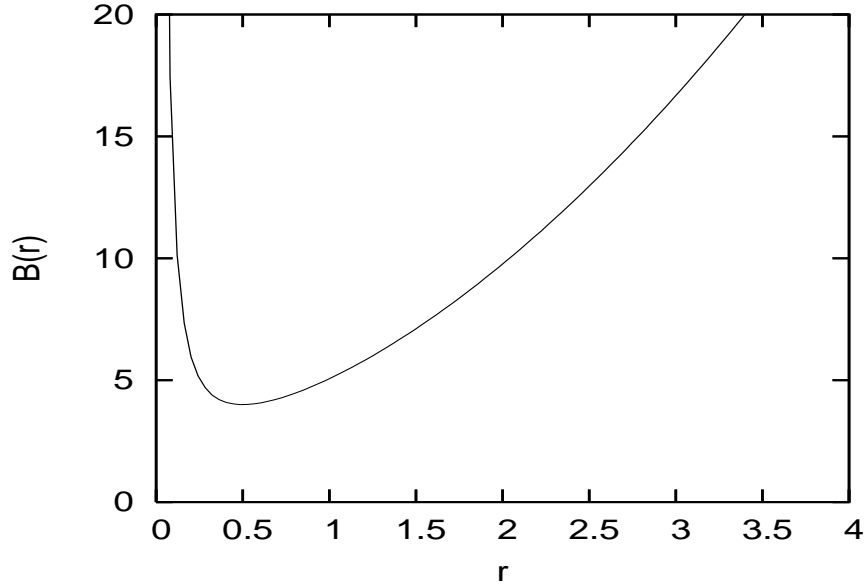


Figura 14: Métrica angular $B = r^2\psi^4$ para un agujero negro de Schwarzschild con $M = 1$ en coordenadas isotrópicas. La garganta del túnel de Einstein-Rosen esta en $r = 1/2$.

Es importante notar que, aunque la métrica anterior es singular en $r = 0$, esta singularidad no corresponde con la singularidad física en el espacio-tiempo de Schwarzschild. La singularidad física esta en el futuro y no esta contenida en nuestros datos iniciales. Recordemos que en coordenadas isotrópicas un agujero negro de Schwarzschild se representa como un agujero de gusano (túnel de Einstein-Rosen) entre dos universos distintos, lo que es fácil ver si se grafica la métrica angular $g_{\theta\theta} = r^2\psi^4$ que resulta ser creciente de $r = M/2$ al infinito, y decreciente (desde infinito) de $r = 0$ a $r = M/2$ (ver Figura 14). En estas coordenadas $r = M/2$ marca la garganta del túnel (la esfera de área mínima, correspondiente al horizonte del agujero negro), y $r = 0$ corresponde al infinito asintótico en el otro universo. Es decir, $r = 0$ es una singularidad de coordenadas introducida por haber compactificado todo el infinito asintótico del “otro” universo a un solo punto.

En situaciones como esta, donde se tiene un factor conforme singular en los datos iniciales, es común extraer analíticamente este factor de la evolución para que las variables dinámicas sean regulares. Definimos entonces nuevas variables:

$$\tilde{A} := A/\psi^4, \quad (5.43)$$

$$\tilde{B} := B/\psi^4, \quad (5.44)$$

$$\tilde{D}_A := D_A - 4 \partial_r \ln \psi, \quad (5.45)$$

$$\tilde{D}_B := D_B - 4 \partial_r \ln \psi , \quad (5.46)$$

(las variables K_A y K_B no se reescalan). Ahora reescribimos las ecuaciones ADM, o la formulación fuertemente hiperbólica equivalente, en términos de estas nuevas variables. Esta técnica de evolución en términos de variables reescaladas regulares y un factor conforme singular estático se conoce como “evolución de puntura”, y al punto $r = 0$ se le llama la puntura (ver por ejemplo [19, 9]).

La evolución de puntura no es la única técnica utilizada para tratar con las singularidades que se encuentran en el interior de los agujeros negros. Otra técnica muy importante esta basada en la idea de que el interior del agujero negro, donde se encuentra la singularidad, esta causalmente desconectado del exterior. Esto implica que uno puede simplemente cortar esa región del dominio computacional sin afectar la evolución exterior. Esta idea, conocida como “escisión de singularidades” fue sugerida desde los años 80 (Thornburg atribuye la idea original a Unruh en [60, 61]). Para simulaciones con simetría esférica la escisión de singularidades ha sido muy exitosa [51, 11, 48, 40, 31, 47, 49, 35], pero en el caso de simulaciones tridimensionales sin simetría ha resultado ser muy compleja computacionalmente. Sin embargo, aun se le considera como la técnica mas adecuada para tratar con el interior de los agujeros negros y hay mucho trabajo en torno a ella [51, 11, 5, 10, 53].

En términos de las nuevas variables, los datos iniciales para Schwarzschild corresponden simplemente a:

$$\tilde{A} = \tilde{B} = 1 , \quad \tilde{D}_A = \tilde{D}_B = 0 . \quad (5.47)$$

Como la métrica de Schwarzschild es estática, y el vector de corrimiento es cero en coordenadas isotrópicas, la curvatura extrínseca resulta ser:

$$K_A = K_B = 0 . \quad (5.48)$$

Quedan por elegir la condiciones de norma. Para el vector de corrimiento elegimos simplemente mantenerlo en cero, mientras que para el lapso elegimos la condición de foliación maximal (2.55), que en este caso se reduce a

$$\frac{1}{\tilde{A}\psi^4} \left[\partial_r^2 \alpha + \left(\frac{2}{r} + \tilde{D}_B - \frac{\tilde{D}_A}{2} + 2\partial_r \ln \psi \right) \partial_r \alpha \right] = K_A^2 + 2K_B^2 . \quad (5.49)$$

Esta ecuación debe integrarse numéricamente a cada paso de tiempo. Como se trata de una ecuación diferencial de segundo orden en r , es necesario dar dos condiciones de frontera. La condición de frontera que se pide en infinito (o en la práctica, en la frontera exterior de la red computacional) es simplemente

$$\partial_r \alpha|_{r \rightarrow \infty} = \frac{1 - \alpha}{r} , \quad (5.50)$$

que es una condición de frontera tipo “Robin” que exige que para $r \rightarrow \infty$ se tenga $\alpha = 1 + \mathcal{O}(r^{-1})$, es decir, el lapso se acerca a 1 (su valor para espacio plano) como $1/r$.

La otra condición de frontera debe darse en el interior, y hay básicamente tres elecciones de interés. La primera posibilidad es pedir $\alpha(r=0) = -1$. En este caso es posible mostrar que hay una solución analítica de la condición maximal dada por:

$$\alpha = \frac{1 - M/2r}{1 + M/2r} . \quad (5.51)$$

Este lapso es anti-simétrico respecto a la garganta del agujero de gusano en $r = M/2$, y es precisamente el lapso que nos da una solución estática, es decir, el lapso standard de Schwarzschild pero escrito en términos de la coordenada radial isotrópica. Esta elección es poco interesante pues, además de dar una situación completamente estática, no penetra el horizonte del agujero negro.

Otra posibilidad es pedir:

$$\partial_r \alpha|_{r=0} = 0 , \quad (5.52)$$

Esta elección ya no coincide con el lapso de Schwarzschild y nos da una situación dinámica (aunque la dinámica sea solamente producto de la elección de la norma). Además, en este caso si se penetra en el interior del horizonte del agujero negro. Esta es la elección que utilizaremos en las simulaciones numéricas que se presentan abajo.

Una tercera posibilidad consiste en pedir que el lapso sea simétrico respecto a la garganta en $r = M/2$. Esta elección también resulta en una evolución dinámica y penetra el horizonte. En el caso de una evolución de puntura como la que estamos considerando, la garganta no es la frontera natural de nuestra malla computacional por lo que el lapso simétrico es difícil de utilizar. Sin embargo, siempre es posible cambiar las cosas de manera que la malla comience precisamente en $r = M/2$ (utilizando por ejemplo la coordenada radial logarítmica $\eta = \ln(2r/M)$), de manera que el lapso simétrico sea una elección natural. Esta elección es muy común en la práctica y corresponde a mantener la isometría entre ambos universos.

Es interesante notar que la foliación maximal de Schwarzschild puede estudiarse de manera analítica en el caso general, y las diferencias entre las tres condiciones de frontera mencionadas pueden entenderse en detalle (ver [26, 15, 44]).

Con los ingredientes que hemos descrito estamos ahora listos para realizar la simulación numérica. Para la simulación mostrada aquí se he utilizado una malla con 1000 puntos en la dirección radial y con un intervalo de $\Delta r = 0,01$. El paso de tiempo utilizado es de $\Delta t = \Delta r/2 = 0,005$, y se han hecho un total de 4000 pasos de tiempo, para llegar a un tiempo final $t = 10$. Los datos iniciales corresponden a un agujero negro de masa $M = 1$,

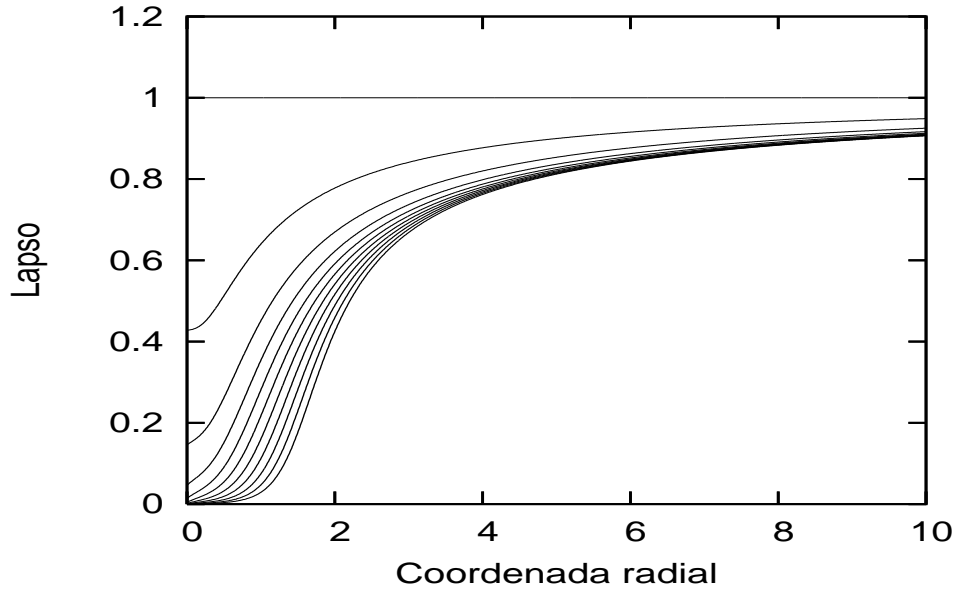


Figura 15: Evolución de la función de lapso α . El valor de α es mostrado cada $t = 1$. El colapso del lapso es evidente.

por lo que en estas unidades la simulación alcanza un tiempo igual a $10M$. Los resultados de la simulación se muestran en las figuras 15 y 16, donde se grafica la evolución del lapso α y la métrica radial A cada $t = 1$.

Consideremos primero la evolución del lapso mostrada en la figura 15. De la figura podemos ver como el lapso rápidamente evoluciona hacia cero en las regiones cercanas a la puntura ($r = 0$). Este fenómeno es el bien conocido “colapso del lapso”, y ocurre debido a que la condición de foliación maximal no permite que los elementos de volumen cambien, y cerca de la singularidad de un agujero negro esto la única manera de logra esto es congelando la evolución temporal. En las regiones centrales el lapso de hecho se acerca a cero exponencialmente.

En la figura 16 se muestra la evolución de la métrica radial A . Vemos como la métrica radial crece en la región cercana al horizonte del agujero negro. Este fenómeno también es bien conocido y se conoce como “estiramiento de la foliación”, o “estiramiento de la malla”.⁸ Lo que ocurre en este caso es una combinación de dos efectos. En primer

⁸Históricamente el nombre “estiramiento de la malla” ha sido comúnmente utilizado. Sin embargo este nombre es inadecuado ya que el fenómeno no tiene ninguna relación con la existencia de una malla numérica. Debido a esto más recientemente se ha preferido el nombre “estiramiento de la foliación”.

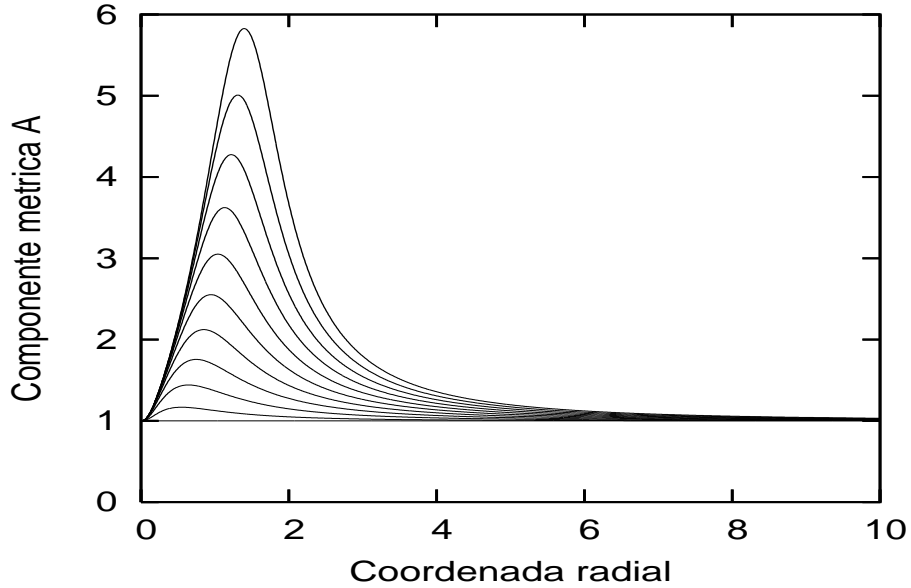


Figura 16: Evolución de la métrica radial A . El valor de A es mostrado cada $t = 1$. El estiramiento de la foliación es evidente.

lugar, como el lapso ha colapsado en las regiones centrales el tiempo ha dejado de avanzar ahí mientras que sigue avanzando afuera, por lo que las hiper-superficies se estiran. Además, incluso si el lapso se mantuviera igual a 1 durante toda la evolución, deberíamos esperar un crecimiento en la métrica radial debido a que observadores normales a diferentes distancias del agujero negro caen hacia este con distintas aceleraciones, por lo que la distancia entre ellos aumenta (recordemos que como el vector de corrimiento es cero, nuestras coordenadas están atadas a estos observadores).

De las figuras, es claro que con las condiciones de norma que hemos utilizado, el espacio-tiempo de Schwarzschild no se ve estático. Esto puede verse aun más dramáticamente si se estudia la posición del horizonte de eventos durante la evolución. La figura 17 muestra el radio del horizonte del agujero negro r_h en función del tiempo. A $t = 0$ el horizonte se encuentra en $r = 0,5$, como corresponde a las coordenadas isotrópicas (para $M = 1$), pero durante la evolución el horizonte se mueve hacia afuera, y a $t = 10$ se encuentra ya en $r \sim 1,8$. El crecimiento aparente del horizonte no es real, es simplemente un efecto coordenado debido al hecho de que nuestra coordenada radial r esta atada a los observadores normales, y estos están cayendo dentro del agujero negro. Para convencernos de que el crecimiento no es real, en la figura 18 se muestra el área a_h del horizonte en función del tiempo para esta simulación. Vemos que el área permanece constante en

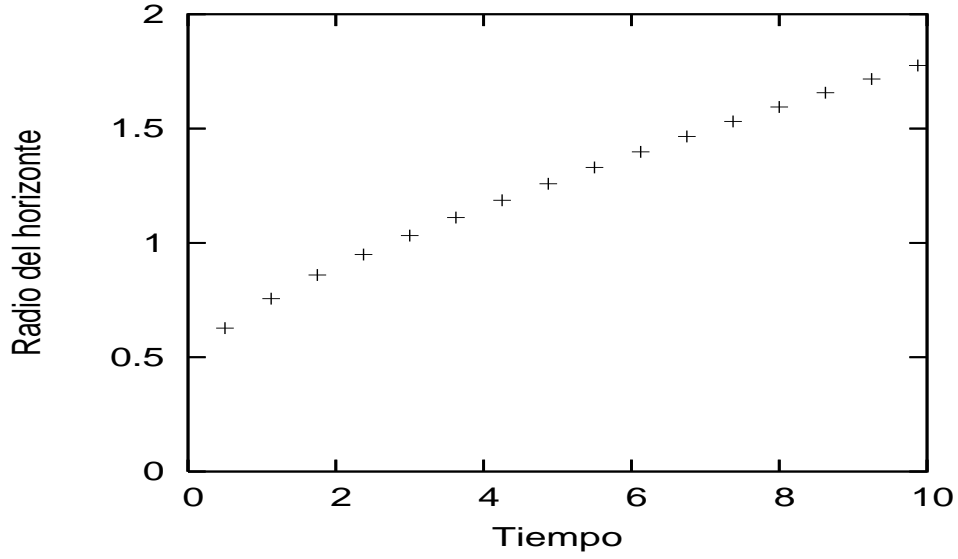


Figura 17: Radio del horizonte en función del tiempo. A $t = 0$ el horizonte se encuentra en $r = 0,5$, y para $t = 10$ su radio ha aumentado a $r = 1,8$.

$a_h \simeq 50,26$ durante toda la evolución, valor que corresponde con el área del horizonte de un agujero negro de Schwarzschild de masa unitaria $a_h = 16\pi = 50,26548246$.

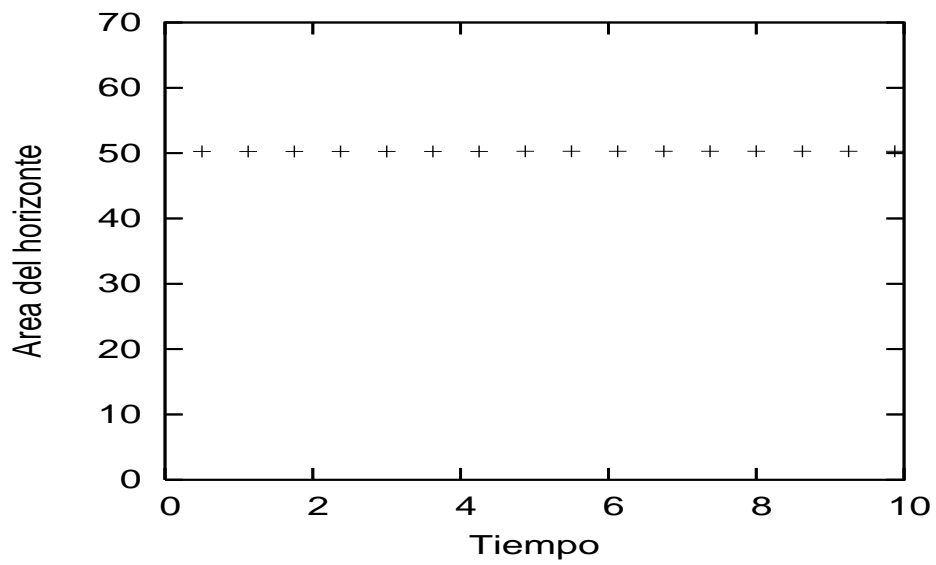


Figura 18: El área del horizonte se mantiene constante durante la evolución, con un valor cercano a la predicción teórica de $a_h = 16\pi \sim 50,2$.

Referencias

- [1] A. Abrahams, D. Bernstein, D. Hobill, E. Seidel, and L. Smarr. Numerically generated black hole spacetimes: Interaction with gravitational waves. *Phys. Rev. D*, 45:3544–3558, 1992.
- [2] M. Alcubierre. The appearance of coordinate shocks in hyperbolic formulations of general relativity. *Phys. Rev. D*, 55:5981–5991, 1997.
- [3] M. Alcubierre. Are gauge shocks really shocks? 2005. gr-qc/0503030.
- [4] M. Alcubierre, G. Allen, B. Brügmann, E. Seidel, and W.-M. Suen. Towards an understanding of the stability properties of the 3+1 evolution equations in general relativity. *Phys. Rev. D*, 62:124011, 2000.
- [5] M. Alcubierre and B. Brügmann. Simple excision of a black hole in 3+1 numerical relativity. *Phys. Rev. D*, 63:104006, 2001.
- [6] M. Alcubierre and J.A. González. Regularization of spherically symmetric evolution codes in numerical relativity. *Comp. Phys. Comm.*, 167:76–84, 2005. gr-qc/0401113.
- [7] M. Alcubierre and J. Massó. Pathologies of hyperbolic gauges in general relativity and other field theories. *Phys. Rev. D*, 57(8):4511–4515, 1998.
- [8] Miguel Alcubierre. Hyperbolic slicings of spacetime: singularity avoidance and gauge shocks. *Class. Quantum Grav.*, 20(4):607–624, 2003.
- [9] Miguel Alcubierre, Bernd Brügmann, Peter Diener, Michael Koppitz, Denis Pollney, Edward Seidel, and Ryoji Takahashi. Gauge conditions for long-term numerical black hole evolutions without excision. *Phys. Rev. D*, 67:084023, 2003.
- [10] Miguel Alcubierre, Bernd Brügmann, Denis Pollney, Edward Seidel, and Ryoji Takahashi. Black hole excision for dynamic black holes. *Phys. Rev. D*, 64:061501(R), 2001.
- [11] P. Anninos, G. Daues, J. Massó, E. Seidel, and W.-M. Suen. Horizon boundary conditions for black hole spacetimes. *Phys. Rev. D*, 51(10):5562–5578, 1995.
- [12] P. Anninos, D. Hobill, E. Seidel, L. Smarr, and W.-M. Suen. Collision of two black holes. *Phys. Rev. Lett.*, 71(18):2851–2854, 1993.

- [13] R. Arnowitt, S. Deser, and C. W. Misner. The dynamics of general relativity. In L. Witten, editor, *Gravitation: An Introduction to Current Research*, pages 227–265. John Wiley, New York, 1962.
- [14] T. W. Baumgarte and S. L. Shapiro. Numerical integration of Einstein’s field equations. *Phys. Rev. D*, 59:024007, 1999.
- [15] R. Beig and N. Ó Murchadha. Late time behavior of the maximal slicing of the schwarzschild black hole. *Phys. Rev. D*, 57(8):4728–4737, 1998. gr-qc/9706046.
- [16] C. Bona, J. Massó, E. Seidel, and J. Stela. New Formalism for Numerical Relativity. *Phys. Rev. Lett.*, 75:600–603, 1995.
- [17] C. Bona, J. Massó, E. Seidel, and J. Stela. First order hyperbolic formalism for numerical relativity. *Phys. Rev. D*, 56:3405–3415, 1997.
- [18] S. Bonazzola and J.-A. Marck. Pseudo-spectral methods applied to gravitational collapse. In C. Evans, L. Finn, and D. Hobill, editors, *Frontiers in Numerical Relativity*, pages 239–253. Cambridge University Press, Cambridge, England, 1989.
- [19] B. Brügmann. Binary black hole mergers in 3D numerical relativity. *Int. J. Mod. Phys. D*, 8:85, 1999.
- [20] M. W. Choptuik. Universality and scaling in gravitational collapse of massless scalar field. *Phys. Rev. Lett.*, 70:9, 1993.
- [21] Gregory B. Cook. Initial data for numerical relativity. *Living Rev. Rel.*, 3:5, 2000. <http://relativity.livingreviews.org/Articles/lrr-2000-5/index.html>.
- [22] A. Einstein. Die feldgleichungen der gravitation. *Preuss. Akad. Wiss. Berlin, Sitzber*, pages 844–847, 1915.
- [23] A. Einstein. Zur allgemeinen relativitätstheorie. *Preuss. Akad. Wiss. Berlin, Sitzber*, pages 778–786, 1915.
- [24] K. Eppley. *The Numerical Evolution of the Collision of Two Black Holes*. PhD thesis, Princeton University, Princeton, New Jersey, 1975.
- [25] K. Eppley. Evolution of time-symmetric gravitational waves: Initial data and apparent horizons. *Phys. Rev. D*, 16(6):1609–1614, 1977.

- [26] Frank Estabrook, Hugo Wahlquist, Steve Christensen, Bryce DeWitt, Larry Smarr, and Elaine Tsiang. Maximally slicing a black hole. *Phys. Rev. D*, 7(10):2814–2817, 1973.
- [27] H. Friedrich. The asymptotic characteristic initial value problem for einstein’s vacuum field equations as an initial value problem for a first order quasi-linear symmetric hyperbolic system. *Proc. Roy. Soc. London, A* 378:401–421, 1981.
- [28] S. Frittelli and O. Reula. First-order symmetric-hyperbolic Einstein equations with arbitrary fixed gauge. *Phys. Rev. Lett.*, 76:4667–4670, 1996.
- [29] GEO600 - <http://www.geo600.uni-hannover.de/>.
- [30] C. Gundlach. Critical phenomena in gravitational collapse. *Living Rev. Rel.*, 2:4, 1999.
- [31] C. Gundlach and P. Walker. Causal differencing of flux-conservative equations applied to black hole spacetimes. *Class. Quantum Grav.*, 16:991–1010, 1999. gr-qc/9809021.
- [32] S. G. Hahn and R. W. Lindquist. The two body problem in geometrodynamics. *Ann. Phys.*, 29:304–331, 1964.
- [33] L. E. Kidder and L. S. Finn. Spectral methods for numerical relativity. the initial data problem. *Phys. Rev. D*, 62:084026, 2000.
- [34] L. E. Kidder, M. A. Scheel, and S. A. Teukolsky. Extending the lifetime of 3D black hole computations with a new hyperbolic system of evolution equations. *Phys. Rev. D*, 64:064017, 2001.
- [35] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson, and G. B. Cook. Black hole evolution by spectral methods. *Phys. Rev. D*, 62:084032, 2000.
- [36] Heinz-Otto Kreiss and J. Lorenz. *Initial-Boundary Value Problems and the Navier-Stokes Equations*. Academic Press, New York, 1989.
- [37] Luis Lehner. Numerical relativity: A review. *Class. Quantum Grav.*, 18:R25–R86, 2001.
- [38] A. Lichnerowicz. L’intégration des équations de la gravitation relativiste et la problème des n corps. *J. Math Pures et Appl.*, 23:37, 1944.
- [39] LIGO - <http://www.ligo.caltech.edu/>.

- [40] R. L. Marsa and M. W. Choptuik. Black hole–scalar field interactions in spherical symmetry. *Phys Rev D*, 54:4929–4943, 1996.
- [41] Richard A. Matzner, Mijan F. Huq, and Deirdre Shoemaker. Initial data and coordinates for multiple black hole systems. *Phys. Rev. D*, 59:024015, 1999.
- [42] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.
- [43] A. R. Mitchell. *The finite element method in partial differential equations*. J. Wiley and Sons, U.S.A., 1977.
- [44] Bernd Reimann and Bernd Brügmann. Maximal slicing for puncture evolutions of Schwarzschild and Reissner - Nordström black holes. *Phys. Rev. D*, 69:044006, 2004.
- [45] R. D. Richtmyer and K.W. Morton. *Difference Methods for Initial Value Problems*. Interscience Publishers, New York, 1967.
- [46] O. Sarbach, G. Calabrese, J. Pullin, and M. Tiglio. Hyperbolicity of the BSSN system of Einstein evolution equations. *Phys. Rev. D*, 66:064002, 2002.
- [47] M. Scheel, T. Baumgarte, G. Cook, S. L. Shapiro, and S. Teukolsky. Numerical evolution of black holes with a hyperbolic formulation of general relativity. *Phys. Rev. D*, 56:6320–6335, 1997.
- [48] M. A. Scheel, S. L. Shapiro, and S. A. Teukolsky. Collapse to black holes in Brans-Dicke theory: I. horizon boundary conditions for dynamical spacetimes. *Phys. Rev. D*, 51(8):4208–4235, 1995.
- [49] Mark A. Scheel, Thomas W. Baumgarte, Gregory B. Cook, Stuart L. Shapiro, and Saul A. Teukolsky. Treating instabilities in a hyperbolic formulation of Einstein’s equations. *Phys. Rev. D*, 58:044020, 1998.
- [50] B. Schutz. *A First Course in General Relativity*. Cambridge University Press, 1985.
- [51] E. Seidel and W.-M. Suen. Towards a singularity-proof scheme in numerical relativity. *Phys. Rev. Lett.*, 69(13):1845–1848, 1992.
- [52] M. Shibata and Takashi Nakamura. Evolution of three-dimensional gravitational waves: Harmonic slicing case. *Phys. Rev. D*, 52:5428, 1995.

- [53] Deirdre Shoemaker, Kenneth L. Smith, Ulrich Sperhake, Pablo Laguna, Erik Schnetter, and David Fiske. Moving black holes via singularity excision. *Class. Quantum Grav.*, 20:3729–3744, 2003. gr-qc/0301111.
- [54] L. Smarr, A. Čadež, B. DeWitt, and K. Eppley. Collision of two black holes: Theoretical framework. *Phys. Rev. D*, 14(10):2443–2452, 1976.
- [55] L. Smarr and J. York. Radiation gauge in general relativity. *Phys. Rev. D*, 17:1945, 1978.
- [56] Larry Smarr and James W. York, Jr. Kinematical conditions in the construction of spacetime. *Phys. Rev. D*, 17(10):2529–2552, 15 May 1978.
- [57] B. Szilágyi, Roberto Gómez, N. T. Bishop, and Jeffrey Winicour. Cauchy boundaries in linearized gravitational theory. *Phys. Rev. D*, 62:104006, 2000.
- [58] B. Szilagyi, B. Schmidt, and Jeffrey Winicour. Boundary conditions in linearized harmonic gravity. *Phys. Rev. D*, 65:064015, 2002.
- [59] TAMA - <http://tamago.mtk.nao.ac.jp/>.
- [60] Jonathan Thornburg. Coordinates and boundary conditions for the general relativistic initial data problem. *Class. Quantum Grav.*, 4(5):1119–1131, September 1987.
- [61] Jonathan Thornburg. *Numerical Relativity in Black Hole Spacetimes*. PhD thesis, University of British Columbia, Vancouver, British Columbia, 1993.
- [62] VIRGO - <http://www.virgo.infn.it/>.
- [63] R. M. Wald. *General Relativity*. The University of Chicago Press, Chicago, 1984.
- [64] John A. Wheeler. *A journey into gravity and spacetime*. Scientific American Library, distributed by W. H. Freeman, New York, U.S.A., 1990.
- [65] Jeffrey Winicour. Characteristic evolution and matching. *Living Rev. Rel.*, 1:5, 1998. [Online article].
- [66] J. W. York. Conformal ‘thin-sandwich’ data for the initial-value problem of general relativity. *Phys. Rev. Lett.*, 82:1350–1353, 1999.
- [67] James W. York. Kinematics and dynamics of general relativity. In Larry L. Smarr, editor, *Sources of Gravitational Radiation*, pages 83–126. Cambridge University Press, Cambridge, UK, 1979.