Lectures for Taller de Verano 1999 de FENOMEC
Numerical Analysis with applications in Theoretical Physics

Matthew W. Choptuik
Dept. of Physics & Astronomy
University of British Columbia
choptuik@physics.ubc.ca

July 1999

1 Basic Finite Difference Techniques for Time Dependent PDEs

1.1 Preliminaries

We can divide time-dependent PDEs into two broad classes:

1. **Initial-value Problems (Cauchy Problems)**, spatial domain has no boundaries (either infinite or "closed"—e.g. "periodic boundary conditions"
2. **Initial-Boundary-Value Problems**, spatial domain *finite*, need to specify boundary conditions

Note: Even if *physical* problem is really of Type 1, finite computational resources $\longrightarrow$ finite spatial domain $\longrightarrow$ approximate as Type 2; will hereafter loosely refer to either type as an IVP.

Working Definition: **Initial Value Problem**

- State of physical system arbitrarily (usually) specified at some initial time $t = t_0$.
- Solution exists for $t \geq t_0$; uniquely determined by equations of motion (EOM) and boundary conditions (BCs).

Issues in Finite Difference (FD) Approximation of IVPs

- Discretization (Derivation of FDA's)
- Solution of algebraic systems resulting from discretization
- Consistency
- Accuracy
- Stability
- Converegence
- Dispersion / Dissipation
- Treatment of Non-linearities
- Computational cost—expect $O(N)$ work ($N \equiv$ number of "grid points" (discrete events at which approximate solution is computed)

1.2 Types of IVPs (by example)

In the following three examples, u is always a function of one space and one time variable, i.e. $u \equiv u(x,t)$. Such a problem is often referred to as "1-d" by numericists, the time dimension being implicit in this nomenclature. I will also use the subscript notation for partial differentiation, e.g. $u_t \equiv \partial_t u$.

1.2.1 Wave/"Wave-Like" ("Hyperbolic"): The 1-d Wave Equation

$$
\begin{aligned}
u_{tt} &= c^2 u_{xx} \qquad c \in \mathbf{R}, \\
u(x,0) &= u_0(x) \\
u_t(x,0) &= v_0(x)
\end{aligned} \tag{1}
$$

1.2.2 Diffusion ("Parabolic"): The 1-d Diffusion Equation

$$
\begin{aligned}
u_t &= \sigma u_{xx} \qquad \sigma \in \mathbf{R}, \quad \sigma > 0. \\
u(x,0) &= u_0(x)
\end{aligned} \tag{2}
$$

1.2.3 Schrödinger: The 1-d Schrödinger Equation

$$
\begin{aligned}
i\psi_t &= -\frac{\hbar}{2m}\psi_{xx} + V(x,t)\psi \qquad \psi \in \mathbf{C} \\
\psi(x,0) &= \psi_0(x)
\end{aligned} \tag{3}
$$

Note: Although $\psi(x,t)$ is *complex* in this case, we can rewrite 3 as a *system* of 2 coupled scalar, real-valued equations.

1.3 Basic Concepts, Definitions and Techniques

We will be considering the finite-difference approximation (FDA) of PDEs, and as such, will generally we interested in the continuum limit, where the *mesh spacing*, or *grid spacing*, usually denoted h, tends to 0. Because any specific calculation must necessarily be performed at some specific, *finite* value of h, we will also be (extremely!) interested in the way that our discrete solution varies as a function of h. In fact, we will *always* view h as the basic "control" parameter of a typical FDA. Fundamentally, for sensibly constructed FDAs, we expect the error in the approximation to go to 0, as h goes to 0.

Let

$$Lu = f \tag{4}$$

denote a general *differential* system.

For simplicity and concreteness, you can think of $u = u(x,t)$ as a single function of one space variable and time, but the discussion in this section applies to cases in more independent variables ($u(x,y,t)$, $u(x,y,z,t)$ $\cdots$ etc.), as well as multiple *dependent* variables ($u = \mathbf{u} = [u_1, u_2, \cdots, u_n]$).

In (4), L is some differential operator (such as $\partial_{tt} - \partial_{xx}$) in our wave equation example), u is the unknown, and f is some specified function (frequently called a *source* function) of the independent variables.

Here and in Section 1.9 it will be convenient to adopt a notation where a superscript h on a symbol indicates that it is discrete, or associated with the FDA, rather than the continuum. (Note, however, that for simplicity of presentation, we will *not* adopt this notation in much of the development below). With this notation, we will generically denote an FDA of (4) by

$$L^h u^h = f^h \tag{5}$$

where u^h is the discrete solution, f^h is the specified function evaluated on the finite-difference mesh, and L^h is the finite-difference approximation of L.

1.3.1 Residual

Note that another way of writing our FDA is

$$L^h u^h - f^h = 0 \tag{6}$$

It is often useful to view FDAs in this form for the following reason. First, we have a canonical view of what it means to solve the FDA—"drive the left-hand side to 0". Furthermore, for iterative approaches to the solution of the FDA (which are common, since it may be too expensive to solve the algebraic equations directly), we are naturally lead to the concept of a *residual*. The residual is simply the level of "non-satisfaction" of our FDA (and, indeed, of any algebraic expression). Specifically, if $\tilde{u}^h$ is some approximation to the true solution of the FDA, u^h, then the residual, r^h, associated with $\tilde{u}^h$ is just

$$r^h \equiv L^h \tilde{u}^h - f^h \tag{7}$$

This leads to the view of a convergent, iterative process as being one which "drives the residual to 0".

1.3.2 Truncation Error

The *truncation error*, τ^h, of an FDA is defined by

$$\tau^h \equiv L^h u - f^h \tag{8}$$

where u satisfies the continuum PDE (4). We note that the *form* of the truncation error can always be computed (typically using Taylor series) from the finite difference approximation and the differential equations.

1.3.3 Convergence

Assuming that our FDA is characterized by a *single* discretization scale, h, we say that the approximation *converges* iff

$$u^h \to u \quad \text{as} \quad h \to 0. \tag{9}$$

Operationally (i.e. in practice), convergence is clearly our chief concern as numerical analysts, particularly if there is reason to suspect that the solutions of our PDEs are good models for real phenomena. We note that this is believed to be the case for many interesting problems in general relativistic astrophysics—the two black hole problem being an excellent example.

1.3.4 Consistency

Assuming that the FDA with truncation error τ^h is characterized by a single discretization scale, h, we say that the FDA is *consistent* if

$$\tau^h \to 0 \quad \text{as} \quad h \to 0. \tag{10}$$

Consistency is obviously a necessary condition for convergence.

1.3.5 Order of an FDA

Assuming that the FDA is characterized by a single discretization scale, h, we say that the FDA is *p-th order accurate* or simply *p-th order* if

$$\lim_{h \to 0} \tau^h = O(h^p) \qquad \text{for some integer } p \tag{11}$$

1.3.6 Solution Error

The solution error, e^h, associated with an FDA is defined by

$$e^h \equiv u - u^h \tag{12}$$

1.3.7 Relation Between Truncation Error and Solution Error

Is is common to tacitly assume that

$$\tau^h = O(h^p) \qquad \longrightarrow \qquad e^h = O(h^p)$$

This assumption is often warranted, but it is extremely instructive to consider *why* it is warranted and to investigate (following Richardson 1910 (!) [5]) in some detail the *nature* of the solution error. We will return to this issue in more detail in Section 1.9.

1.3.8 Deriving Finite Difference Formulae

The essence of finite-difference approximation of a PDE is the replacement of the continuum by a discrete lattice of grid points, and the replacement of derivatives/differential operators by finite-difference expressions. These finite-difference expressions (finite-difference quotients) approximate the derivatives of functions at grid points, using the grid values themselves. All of the operators and expressions we need can easily be worked out using Taylor series techniques. For example, let us consider the task of approximating the first derivative $u_x(x)$ of a function $u(x)$, given a discrete set of values $u_j \equiv u(jh)$ as shown in Figure 1.

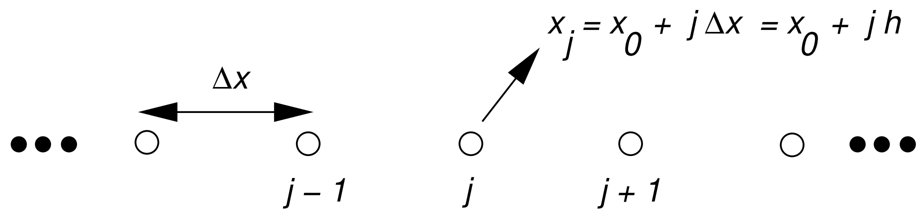

Figure 1. A one-dimensional, uniform finite difference mesh. Note that the spacing, $\triangle x = h$, between adjacent mesh points is *constant.* In the text we tacitly assume that the origin, x_0, of our coordinate system is $x_0 = 0$.

As it turns out, given the three values $u(x_j - h), u(x_j)$ and $u(x_j + h)$, which we will denote u_{j-1}, u_j, and u_{j+1} respectively, we can compute an $O(h^2)$ approximation to $u_x(x_j) \equiv (u_x)_j$ as follows. Taylor expanding, we have

$$
\begin{aligned}
u_{j-1} &= u_j - h(u_x)_j + \frac{1}{2}h^2(u_{xx})_j - \frac{1}{6}h^3(u_{xxx})_j + \frac{1}{24}h^4(u_{xxxx})_j + O(h^5) \\
u_j &= u_j \\
u_{j+1} &= u_j + h(u_x)_j + \frac{1}{2}h^2(u_{xx})_j + \frac{1}{6}h^3(u_{xxx})_j + \frac{1}{24}h^4(u_{xxxx})_j + O(h^5)
\end{aligned}
$$

We now seek a linear combination of u_{j-1}, u_j, and u_{j+1} which yields $(u_x)_j$ to $O(h^2)$ accuracy, i.e. we seek c_-, c_0 and c_+ such that

$$c_-\, u_{j-1} + c_0\, u_j + c_+\, u_{j+1} = (u_x)_j + O(h^2)$$

This results in a system of three linear equations for u_{j-1}, u_j, and u_{j+1}:

$$\begin{aligned} c_- + c_0 + c_+ &= 0 \\ -hc_- + hc_+ &= 1 \\ \frac{1}{2}h^2 c_- + \frac{1}{2}h^2 c_+ &= 0 \end{aligned}$$

which has the solution

$$\begin{aligned} c_- &= -\frac{1}{2h} \\ c_0 &= 0 \\ c_+ &= +\frac{1}{2h} \end{aligned}$$

Thus our $O(h^2)$ finite difference approximation for the first derivative is

$$\frac{u(x+h) - u(x-h)}{2h} = u_x(x) + O(h^2) \tag{13}$$

Note that it may not be obvious to you *a priori*, that the truncation error of this approximation *is* $O(h^2)$, since a naive consideration of the number of terms in the Taylor series expansion which can be eliminated using 2 values (namely $u(x+h)$ and $u(x-h)$) suggests that the error might be $O(h)$. The fact that the $O(h)$ term "drops out" is a consequence of the *symmetry*, or *centering* of the stencil, and is a common theme in such FDAs (which, naturally enough, are called centred difference approximations).

Using the same technique, we can easily generate the $O(h^2)$ expression for the *second* derivative, which uses the same difference stencil as the above approximation for the first derivative.

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = u_{xx}(x) + O(h^2) \tag{14}$$

Exercise: Compute the precise form of the $O(h^2)$ terms in expressions (13) and (14).

1.4 Sample Discretizations / FDAs

1.4.1 1-d Wave equation with fixed (Dirichlet) boundary conditions

$$
\begin{aligned}
u_{tt} &= u_{xx} \qquad (c=1) \qquad 0 \le x \le 1; \quad t \ge 0 \qquad (15)\\
u(x,0) &= u_0(x)\\
u_t(x,0) &= v_0(x)\\
u(0,t) &= u(1,t) = 0
\end{aligned}
$$

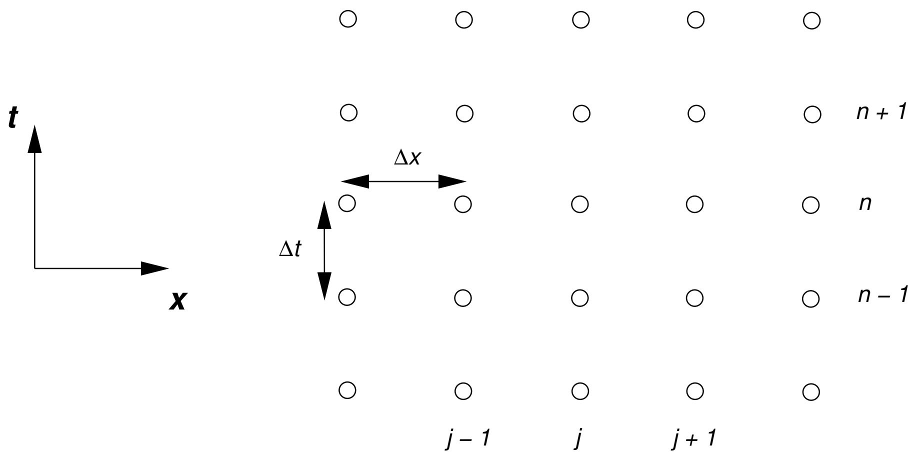

Figure 2. Portion of uniform finite-difference mesh (grid) for 1-d time-dependent problem. Note that the spacings in both the spatial *and* temporal directions are constant

We now introduce a discrete domain (uniform grid) (x_j, t^n)—part of which is shown in Figure 2.

$$
\begin{aligned}
t^n &\equiv n\,\triangle t\,, \qquad n = 0, 1, 2, \cdots\\
x_j &\equiv (j-1)\,\triangle x\,, \qquad j = 1, 2, \cdots J\\
u_j^n &\equiv u(n\,\triangle t\,, (j-1)\,\triangle x\,)\\
\triangle x &= (J-1)^{-1}\\
\triangle t &= \lambda\,\triangle x \qquad \lambda \equiv \text{“Courant number”}
\end{aligned}
$$

Note: When solving wave equations using FDAs, we will typically keep λ constant when we vary $\triangle x$. Thus, our FDA will always be characterized by *single* discretization scale, h.

$$\begin{aligned} \triangle x &\equiv h \\ \triangle t &\equiv \lambda h \end{aligned}$$

(Also note the `Fortran`-style indexing of the spatial grid index ($j = 1, 2, \cdots$) and the `C`-style indexing of the temporal one ($n = 0, 1, \cdots$). This is a particular convention which I, as a predominantly `Fortran` programmer, find convenient.)

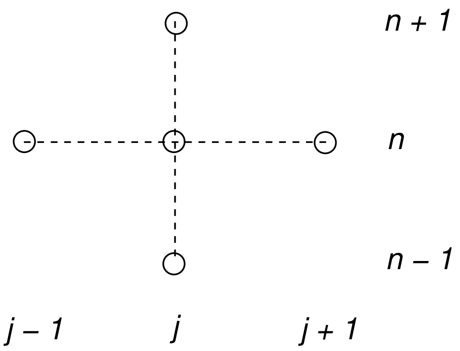

Figure 3. Stencil (molecule/star) for "standard" $O(h^2)$ approximation of (15).

FDA: "standard $O(h^2)$"

Discretized Interior equation:

$$
\begin{aligned}
(\triangle t)^{-2}\left(u_j^{n+1} - 2u_j^n + u_j^{n-1}\right) &= (u_{tt})_j^n + \frac{1}{12}\triangle t^2 (u_{tttt})_j^n + O(\triangle t^4) \\
&= (u_{tt})_j^n + O(h^2) \\
(\triangle x)^{-2}\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right) &= (u_{xx})_j^n + \frac{1}{12}\triangle x^2 (u_{xxxx})_j^n + O(\triangle x^4) \\
&= (u_{xx})_j^n + O(h^2)
\end{aligned}
$$

Putting these two together, we get the $O(h^2)$ approximation

$$
\frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{\triangle t^2} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\triangle x^2} \qquad j = 2, 3, \cdots, J-1 \tag{16}
$$

Note that a scheme such as (16) is often called a *three level scheme* since it couples *three "time levels"* of data (i.e. unknowns at three distinct, disrete times t^{n-1}, t^n, t^{n+1}.

Discretized Boundary conditions:

$$
u_1^{n+1} = u_J^{n+1} = 0
$$

Discretized Initial conditions:
We need to specify *two* "time levels" of data (effectively $u(x,0)$ and $u_t(x,0)$), i.e. we must specify

$$
\begin{aligned}
u_j^0 \quad &, \quad j = 1, 2, \cdots, J \\
u_j^1 \quad &, \quad j = 1, 2, \cdots, J
\end{aligned}
$$

ensuring that the initial values are compatible with the boundary conditions.

Note that we can solve (16) *explicitly* for u_j^{n+1}:

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_j^{n-1} \right) \tag{17}$$

Also note that (17) is actually a *linear system* for the unknowns u_j^{n+1}, $j = 1, 2, \cdots, J$; in combination with the discrete boundary conditions we can write

$$\mathbf{A}\ \mathbf{u}^{n+1} = \mathbf{b} \tag{18}$$

where $\mathbf{A}$ is a *diagonal* $J \times J$ matrix and $\mathbf{u}^{n+1}$ and $\mathbf{b}$ are vectors of length J. Such a difference scheme for an IVP is called an *explicit* scheme.

1.4.2 1-d Diffusion equation with Dirichlet boundary conditions

$$
\begin{aligned}
u_t &= u_{xx} \qquad (\sigma = 1) \qquad 0 \le x \le 1; \quad t \ge 0 \\
u(x,0) &= u_0(x) \\
u(0,t) &= u_(1,t) = 0
\end{aligned}
\tag{19}
$$

We will use same discrete domain (grid) as for the 1-d wave equation.

FDA: Crank-Nicholson

This scheme illustrates a useful "rule of thumb": *Keep difference scheme "centred"*

- centred in time, centred in space
- minimizes truncation error for given h
- tends to minimize instabilities

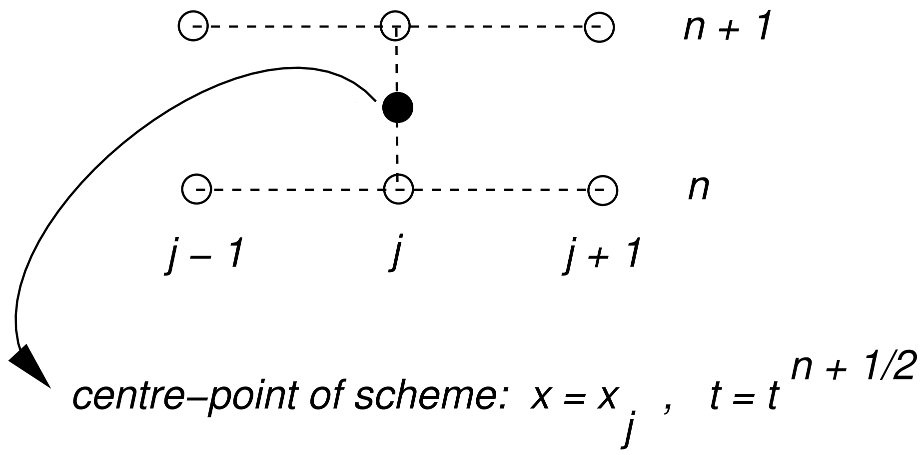

Figure 4. Stencil (molecule/star) for $O(h^2)$ Crank-Nicholson approximation of (19).

Discretization of time derivative:

$$
\begin{aligned}
\triangle t^{-1}\left(u_j^{n+1} - u_j^n\right) &= (u_t)_j^{n+\frac{1}{2}} + \frac{1}{24}\triangle t^{2}\,(u_{ttt})_j^{n+\frac{1}{2}} + O(\triangle t^{4}) \qquad (20)\\
&= (u_t)_j^{n+\frac{1}{2}} + O(\triangle t^{2})
\end{aligned}
$$

$O(h^2)$ second-derivative operator:

$$
\begin{aligned}
D_{xx}\,u_j^n &\equiv \triangle x^{-2}\left(u_{j+1}^n - 2u_j^n + u_{j+1}^n\right) \qquad (21)\\
D_{xx} &= \partial_{xx} + \frac{1}{12}\triangle x^{2}\,\partial_{xxxx} + O(\triangle x^{4}) \qquad (22)
\end{aligned}
$$

(Forward) Time-averaging operator, μ_t:

$$
\begin{aligned}
\mu_t\,u_j^n &\equiv \frac{1}{2}\left(u_j^{n+1} + u_j^{n-1}\right) = u_j^{n+\frac{1}{2}} + \frac{1}{8}\triangle t^{2}\,(u_{tt})_j^{n+\frac{1}{2}} + O(\triangle t^{4}) \qquad (23)\\
\mu_t &= \left[I + \frac{1}{8}\triangle t^{2}\,\partial_{tt} + O(\triangle t^{4})\right]_{t=t^{n+1/2}} \qquad (24)
\end{aligned}
$$

where I is the identity operator. Assuming that $\triangle t = O(\triangle x) = O(h)$, is is easy to show (*exercise*) that

$$
\mu_t\left[D_{xx}\,u_j^n\right] = (u_{xx})_j^{n+\frac{1}{2}} + O(h^2)
$$

Putting the above results together, we are lead to the $(O(h^2))$ Crank-Nicholson approximation of (19):

$$\frac{u_j^{n+1} - u_j^n}{\triangle t} = \mu_t \left[D_{xx}\, u_j^n \right] \tag{25}$$

Written out in full, this is

$$\frac{u_j^{n+1} - u_j^n}{\triangle t} = \frac{1}{2} \left[\frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{\triangle x^{\,2}} + \frac{u_{j+1}^{n} - 2u_j^{n} + u_{j-1}^{n}}{\triangle x^{\,2}} \right] \tag{26}$$

for $j = 2, 3, \cdots, J-1$.

We can rewrite (26) in the form

$$a_+ \, u_{j+1}^{n+1} + a_0 \, u_j^{n+1} + a_- \, u_{j-1}^{n+1} = b_j \qquad j = 2, 3, \cdots, J-1 \tag{27}$$

where

$$\begin{aligned}
a_+ &\equiv -\frac{1}{2} \triangle x^{\,-2} \\
a_0 &\equiv \triangle t^{\,-1} + \triangle x^{\,-2} \\
a_- &\equiv -\frac{1}{2} \triangle x^{\,-2} \\
b_j &\equiv \left(\triangle t^{\,-1} - \triangle x^{\,-2} \right) u_j^n + \frac{1}{2} \triangle x^{\,-2} \left(u_{j+1}^n + u_{j-1}^n \right)
\end{aligned}$$

which, along with the BCs ($u_1^{n+1} = u_J^{n+1} = 0$), is again a linear system of the form

$$\mathbf{A} \; \mathbf{u}^{n+1} = \mathbf{b}$$

for the "unknown vector" $\mathbf{u}^{n+1}$. This time, however, the matrix $\mathbf{A}$, is *not* diagonal, and the scheme is called *implicit*—i.e. the scheme *couples* unknowns at the *advanced* time level, $t = t^{n+1}$.

Note that $\mathbf{A}$ is a *tridiagonal* matrix: all elements A_{ij} for which $j \neq i + 1, i$ or $i - 1$ vanish. The solution of tridiagonal systems can be performed very efficiently using special purpose routines (such as `DGTSV` in `LAPACK` [6]): specifically, the operation count for solution of (26) is $O(J)$.

Also note that we can immediately write down the analogous scheme for the Schrödinger equation (3):

$$i\,\frac{\psi_j^{n+1}-\psi_j^n}{\triangle t} = -\frac{\hbar}{2m}\mu_t\left[D_{xx}\,\psi_j^n\right] + V(x_j)\,\mu_t\psi_j^n \tag{28}$$

In this case we get a *complex* tridiagonal system, which can also be solved in $O(J)$ time, using, for example, the `LAPACK` routine `ZGTSV`.

1.5 The 1-D Wave Equation in More Detail

Recall our "standard" $O(h^2)$ discretization:

$$
\begin{aligned}
u_j^{n+1} &= 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right), \qquad j = 2, 3, \cdots, J-1 \\
u_1^{n+1} &= u_J^{n+1} = 0
\end{aligned}
$$

As we have discussed, to initialize the scheme, we need to specify u_j^0 and u_j^1, which is equivalent (in the limit $h \to 0$) to specifying $u(x,0)$ and $u_t(x,0)$.

Before proceeding to a discussion of a "proper initialization", let us briefly digress and consider the continuum case, and, for the sake of presentation, assume that we are considering a true IVP on an unbounded domain; i.e. we wish to solve

$$
u_{tt} = u_{xx} \qquad -\infty < x < \infty \quad , \quad t \geq 0 \tag{29}
$$

As is well known, the general solution of (29) is the superposition of an arbitrary *left-moving* profile ($v = -c = -1$), and an arbitrary *right-moving* profile ($v = +c = +1$); i.e.

$$
u(x,t) = \ell(x+t) + r(x-t) \tag{30}
$$

where (see Figure 5)

ℓ : constant along "left-directed" characteristics

r : constant along "right-directed" characteristics

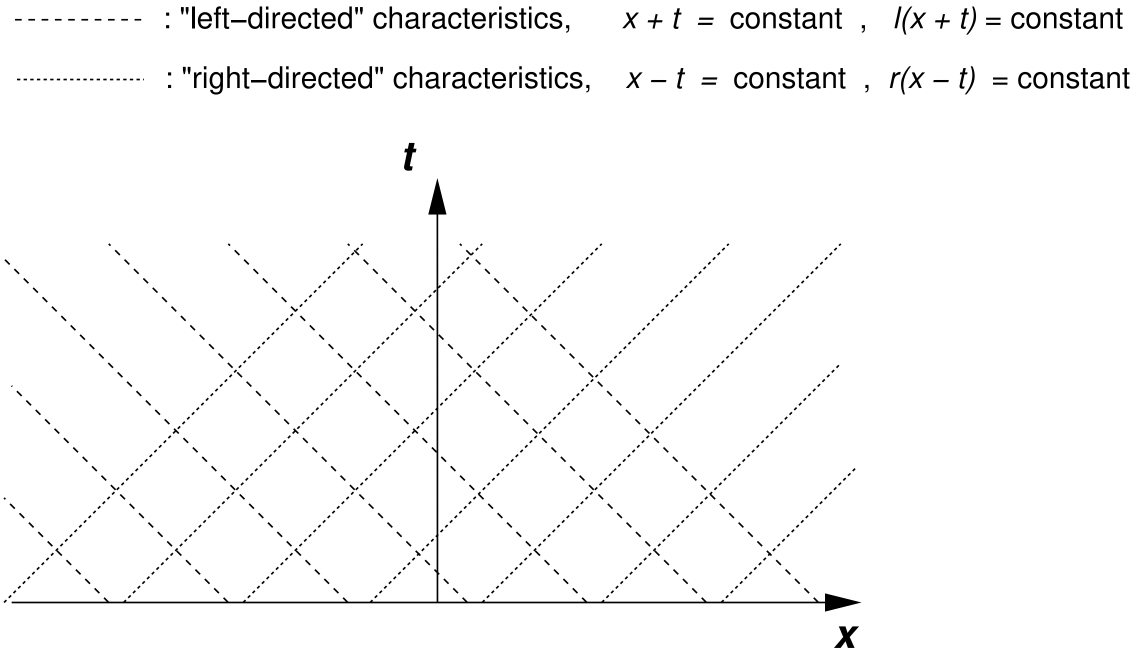

Figure 5. Characteristics of the wave equation: $u_{xx} = u_{tt}$. Signals (disturbances) travel along the characteristics (dashed and dotted lines.)

This observation provides us with an alternative way of specifying initial values, which is often quite convenient in practice. Rather than specifying $u(x, 0)$ and $u_t(x, 0)$ directly, we can specify *initial* left-moving and right-moving parts of the solution, $\ell(x)$ and $r(x)$. Specifically, we set

$$u(x,0) \;=\; \ell(x) + r(x) \tag{31}$$

$$u_t(x,0) \;=\; \ell'(x) - r'(x) \equiv \frac{d\ell}{dx}(x) - \frac{dr}{dx}(x) \tag{32}$$

Returning now to the solution of the finite-differenced version of the wave equation, it is clear that given the initial data (31–32), we can trivially initialize u_j^0 with *exact* values, but that we can only approximately initialize u_j^1.

The question then arises: *How accurately must we initialize the advanced values so as to ensure second order ($O(h^2)$) accuracy of the difference scheme?*

A brief, heuristic answer to this question (which can be more rigorously justified) is as follows. We have $\triangle x = O(h)$, $\triangle t = O(h)$ and the FDA is $O(h^2)$. Since the scheme is $O(h^2)$, we expect that

$$u_{\text{exact}}(x,t) - u_{\text{FD}}(x,t) = O(h^2)$$

for arbitrary, *fixed, FINITE* t. However, the number of time steps required to integrate to time t is $O(\triangle t^{-1}) = O(h^{-1})$. Thus, the per-time-step error must be $O(h^2)/O(h^{-1}) = O(h^3)$, and, therefore, we require

$$(u_{\text{FD}})^{\,1}_{\,j} = (u_{\text{exact}})^{\,1}_{\,j} + O(h^3)$$

We can readily accomplish this using (1) Taylor series and (2) the equation of motion to rewrite higher time derivatives in terms of spatial derivatives:

$$\begin{aligned} u^{1}_{j} &= u^{0}_{j} + \triangle t\ (u_t)^{\,0}_{\,j} + \frac{1}{2}\triangle t^{\,2}\,(u_{tt})^{\,0}_{\,j} + O(\triangle t^{\,3}) && (33)\\ &= u^{0}_{j} + \triangle t\ (u_t) + \frac{1}{2}\triangle t^{\,2}\,(u_{xx})^{\,0}_{\,j} + O(\triangle t^{\,3}) && (34) \end{aligned}$$

which, using results from above, can be written as

$$u^{1}_{j} = (\ell + r)_{\,j} + \triangle t\ (\ell' - r')_{\,j} + \frac{1}{2}\triangle t^{\,2}\,(\ell'' + r'')_{\,j} \tag{35}$$

1.6 Stability Analysis

One of the most frustrating—yet fascinating—features of FD solutions of time dependent problems, is that the discrete solutions often "blow up"—e.g. floating-point overflows are generated at some point in the evolution. Although "blow-ups" can sometimes be caused by legitimate (!) "bugs"—i.e. an incorrect implementation—at other times it is simply the *nature of the FD scheme* which causes problems. We are thus lead to consider the *stability* of solutions of difference equations (as well as their differential-equation progenitors).

Let us again consider the 1-d wave equation (15) and let us now remark that this is a *linear, non-dispersive* wave equation, a consequence of which is the fact that the "size" of the solution does *not* change with time:

$$\|u(x,t)\| \sim \|u(x,0)\|\,, \tag{36}$$

where $\|\cdot\|$ is an suitable norm, such as the L_2 norm:

$$\|u(x,t)\| \equiv \left(\int_0^1 u(x,t)^2\,dx\right)^{1/2}\,. \tag{37}$$

We will use the property captured by (36) as our working definition of stability. In particular, if you believe (36) is true for the wave equation, then you believe the wave equation is stable.

Fundamentally, if our FDA approximation *converges*, then we expect the same behaviour for the difference solution:

$$\|u^n_j\| \sim \|u^0_j\|\,. \tag{38}$$

Now, we construct our FD solution by *iterating in time*, generating

$$u_j^0,\ u_j^1,\ u_j^2,\ u_j^3,\ u_j^4,\ \cdots$$

in succession, using the FD equation

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right) .$$

As it turns out, we are *not* guaranteed that (38) holds for all values of $\lambda \equiv \triangle t \,/\, \triangle x$. In fact, for certain λ (all $\lambda > 1$, as we shall see), we have

$$\|u_j^n\| \gg \|u_j^0\| \,,$$

and for those λ, $\|u^n\|$ *diverges* from u, even (especially!) as $h \to 0$—that is, the difference scheme is *unstable.*

In fact, for many wave problems (including all linear problems), given that a FD scheme is *consistent* (i.e. so that $\hat{\tau} \to 0$ as $h \to 0$), stability is the necessary and sufficient condition for convergence (Lax's theorem).

1.6.1 Heuristic Stability Analysis

Let us write a general time-dependent FDA in the form

$$\mathbf{u}^{n+1} = \mathbf{G}[\mathbf{u}^n], \tag{39}$$

where $\mathbf{G}$ is some *update operator* (linear in our example problem), and $\mathbf{u}$ is a column vector containing sufficient unknowns to write the problem in first-order-in-time form. For example, if we introduce a new, auxiliary set of unknowns, v^n_j, defined by

$$v^n_j = u^{n-1}_j ,$$

then we can rewrite the differenced-wave-equation (16) as

$$u^{n+1}_j = 2u^n_j - v^n_j + \lambda^2 \left(u^n_{j+1} - 2u^n_j + u^n_{j-1} \right) , \tag{40}$$

$$v^{n+1}_j = u^n_j , \tag{41}$$

so with

$$\mathbf{u}^n = [u^n_1, v^n_1, \; u^n_2, v^n_2, \; \cdots \; u^n_J, v^n_J] ,$$

(for example), (40-41) is clearly of the form (39). Equation (39) provides us with a compact way of describing the solution of the FDA. Given initial data, $\mathbf{u}^0$, the solution after n time-steps is

$$\mathbf{u}^n = \mathbf{G}^n \mathbf{u}^0 , \tag{42}$$

where $\mathbf{G}^n$ is the n-th power of the matrix $\mathbf{G}$.

Now, assume that $\mathbf{G}$ has a complete set of orthonormal eigenvectors

$$\mathbf{e}_k, \quad k = 1, 2, \cdots J\,,$$

and corresponding eigenvalues

$$\mu_k, \quad k = 1, 2, \cdots J\,,$$

so that

$$\mathbf{G}\,\mathbf{e}_k = \mu_k\,\mathbf{e}_k, \quad k = 1, 2, \cdots J\,.$$

We can then write the initial data as (spectral decomposition):

$$\mathbf{u}^0 = \sum_{k=1}^{J} c_k^0\,\mathbf{e}_k\,,$$

where the c_k^0 are coefficients. Using (42), the solution at time-step n is then

$$\begin{aligned} \mathbf{u}^n &= \mathbf{G}^n \left(\sum_{k=1}^{J} c_k^0\,\mathbf{e}_k \right) && (43) \\ &= \sum_{k=1}^{J} c_k^0\,(\mu_k)^n\,\mathbf{e}_k\,. && (44) \end{aligned}$$

Clearly, if the difference scheme is to be stable, we must have

$$|\mu_k| \le 1 \qquad k = 1, 2, \cdots J \qquad (45)$$

(Note: μ_k will be complex in general, so $|\mu|$ denotes complex modulus, $|\mu| \equiv \sqrt{\mu\mu^\star}$). Geometrically, then, the eigenvalues of the update matrix must lie on or within the unit circle (see Figure 6).

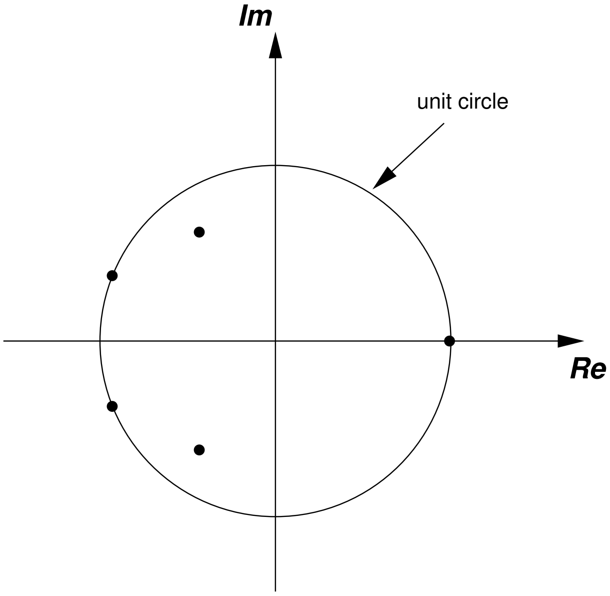

Figure 6. Schematic illustration of location in complex plane of eigenvalues of update matrix $\mathbf{G}$. In this case, all eigenvalues (dots) lie on or within the unit circle, indicating that the corresponding finite difference scheme is stable.

1.6.2 Von-Neumann (Fourier) Stability Analysis

Von-Neumann stability analysis is based on the ideas sketched above, but additionally assumes that the difference equation is linear with constant coefficients, and that the boundary conditions are periodic. We can then use Fourier analysis, which has the same benefits in the discrete domain—difference operators in real-space variable $x \longrightarrow$ algebraic operations in Fourier-space variable k—as it does in the continuum.

Schematically, instead of writing

$$\mathbf{u}^{n+1}(x) = \mathbf{G}[\mathbf{u}^n(x)]\,,$$

we consider the Fourier-domain equivalent:

$$\tilde{\mathbf{u}}^{n+1}(k) = \tilde{\mathbf{G}}[\tilde{\mathbf{u}}^n(k)]\,,$$

where k is the wave-number (Fourier-space variable) and $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{G}}$ are the Fourier-transforms of $\mathbf{u}$ and $\mathbf{G}$, respectively. Specifically, we define the Fourier-transformed grid function via

$$\tilde{\mathbf{u}}^n(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx}\, \mathbf{u}^n(x)\, dx\,. \tag{46}$$

For a general difference scheme, we will find that

$$\tilde{\mathbf{u}}^{n+1}(k) = \tilde{\mathbf{G}}(\xi)\, \tilde{\mathbf{u}}^n(k)\,,$$

where $\xi \equiv kh$, and we will have to show that $\tilde{\mathbf{G}}(\xi)$'s eigenvalues lie within or on the unit circle for all conceivable ξ. The appropriate range for ξ is

$$-\pi \le \xi \le \pi\,,$$

since the shortest wavelength representable on a uniform mesh with spacing h is $\lambda = 2h$ (Nyquist limit), corresponding to a maximum wave number $k = (2\pi)/\lambda = \pm\pi/h$.

Let us consider the application of the Von-Neumann stability analysis to our current model problem. We first define a (non-divided) difference operator D^2 as follows:

$$D^2\, u(x) = u(x+h) - 2u(x) + u(x-h)\,.$$

Then, suppressing the spatial grid index, we can write the first-order form of the difference equation (40-41) as

$$\begin{aligned} u^{n+1} &= 2u^{n} - v^{n} + \lambda^2\, D^2\, u^{n}\,, \\ v^{n+1} &= u^{n}\,, \end{aligned}$$

or

$$\begin{bmatrix} u \\ v \end{bmatrix}^{n+1} = \begin{bmatrix} 2+\lambda^2\, D^2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}^{n}. \tag{47}$$

In order to perform the Fourier transform, we need to know the action of D^2 in Fourier-space. Using the transform inverse to (46) we have

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx}\, \tilde{u}(k)\, dk\,,$$

so

$$\begin{aligned} D^2\, u(x) = u(x+h) - 2u(x) + u(x-h) &= \int_{-\infty}^{+\infty} \left(e^{ikh} - 2 + e^{-ikh}\right) e^{ikx}\, \tilde{u}(k)\, dk \\ &= \int_{-\infty}^{+\infty} \left(e^{i\xi} - 2 + e^{-i\xi}\right) e^{ikx}\, \tilde{u}(k)\, dk\,. \end{aligned}$$

Now consider the quantity $-4\sin^2(\xi/2)$:

$$\begin{aligned}
-4\sin^2\frac{\xi}{2} &= -4\left(\frac{e^{i\xi/2}-e^{-i\xi/2}}{2i}\right)^2 \\
&= \left(e^{i\xi/2}-e^{-i\xi/2}\right)^2 = e^{i\xi}-2+e^{-i\xi}\,,
\end{aligned}$$

so

$$D^2\,u(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(-4\sin^2\frac{\xi}{2}\right)e^{ikx}\,\tilde{u}(k)\,dk\,.$$

In summary, under Fourier transformation, we have

$$\begin{aligned}
\mathbf{u}(x) &\longrightarrow \tilde{\mathbf{u}}(k)\,, \\
D^2\mathbf{u}(x) &\longrightarrow -4\sin^2\frac{\xi}{2}\tilde{\mathbf{u}}(k)\,.
\end{aligned}$$

Using this result in the Fourier transform of (47), we see that we need to compute the eigenvalues of

$$\left[\begin{array}{cc} 2-4\lambda^2\sin^2(\xi/2) & -1 \\ 1 & 0 \end{array}\right],$$

and determine the conditions under which the eigenvalues lie on or within the unit circle. The characteristic equation (whose roots are the eigenvalues) is

$$\left|\begin{array}{cc} 2-4\lambda^2\sin^2(\xi/2)-\mu & -1 \\ 1 & -\mu\,, \end{array}\right| = 0$$

or

$$\mu^2 + \left(4\lambda^2 \sin^2 \frac{\xi}{2} - 2\right) \mu + 1 = 0\,.$$

This equation has roots

$$\mu(\xi) = \left(1 - 2\lambda^2 \sin^2 \frac{\xi}{2}\right) \pm \left(\left(1 - 2\lambda^2 \sin^2 \frac{\xi}{2}\right) - 1\right)^{1/2}\,.$$

We now need to find sufficient conditions for

$$|\mu(\xi)| \leq 1,$$

or equivalently

$$|\mu(\xi)|^2 \leq 1.$$

To this end, we note that we can write

$$\mu(\xi) = (1 - Q)\ \pm\ ((1 - Q)^2 - 1)^{1/2}\,,$$

where the quantity, Q

$$Q \equiv 2\lambda \sin^2 \frac{\xi}{2}\,,$$

is *real* and *non-negative* ($Q \geq 0$).

There are now two cases to consider:

1. $(1-Q)^2 - 1 \leq 0$,
2. $(1-Q)^2 - 1 > 0$.

In the first case, $((1-Q)^2-1)^{1/2}$ is purely imaginary, so we have

$$|\mu(\xi)|^2 = (1-Q)^2 + (1-(1-Q)^2) = 1 \,.$$

In the second case, $(1-Q)^2 - 1 > 0 \longrightarrow (1-Q)^2 > 1 \longrightarrow Q > 2$, and then we have

$$1 - Q - ((1-Q^2) - 1)^{1/2} \; < \; -1 \,,$$

so, in this case, our stability criterion will *always* be violated. We thus conclude that a necessary condition for Von-Neumann stability is

$$(1-Q)^2 - 1 \leq 0 \longrightarrow (1-Q)^2 \leq 1 \longrightarrow Q \leq 2 \,.$$

Since $Q \equiv 2\lambda \sin^2(\xi/2)$ and $\sin^2(\xi/2) \leq 1$, we must therefore have

$$\lambda \equiv \frac{\triangle t}{\triangle x} \leq 1 \,,$$

for stability of our scheme (16). This condition is often called the CFL condition—after Courant, Friedrichs and Lewy who derived it in 1928 (the ratio $\lambda = \triangle x \,/\, \triangle t$ is also frequently called the *Courant number*). In practical terms, we must limit the time-discretization scale , $\triangle t$, to values no larger than the space-discretization scale, $\triangle x$. Furthermore, this type of instability has a "physical" interpretation, often summarized by the statement *the numerical domain of dependence of an explicit difference scheme must contain the physical domain of dependence.*

1.7 Dispersion and Dissipation

Let us now consider an even simpler model "wave equation" than (1), the so-called *advection*, or *color* equation:

$$\begin{aligned} u_t &= a\,u_x \quad (a > 0) \qquad -\infty < x < \infty \quad , \quad t \geq 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{48}$$

which has the exact solution

$$u(x,t) = u_0(x + at) \tag{49}$$

Equation (48) is another example of a non-disspative, non-dispersive partial differential equation.

To remind ourselves what "non-dispersive" means, let us note that (48) admits "normal mode" solutions:

$$u(x,t) \sim e^{ik(x+at)} \equiv e^{i(kx+\omega t)}$$

where $\omega \equiv ka$; in general, of course, $\omega \equiv \omega(k)$ is known as the *dispersion relation*, and

$$\frac{d\omega}{dk} \equiv \text{ speed of propagation of mode with wave number } k$$

In the current case, we have

$$\frac{d\omega}{dk} = a = \text{constant}$$

which means that all modes propagate at the same speed, which is precisely what is meant by "non-dispersive". Further, if we consider resolving the general initial profile, $u_0(x)$ into "normal-mode" (Fourier) components, we find that the magnitudes of the components are preserved in time, i.e. equation (48) is also *non-dissipative*.

Ostensibly, we would like our finite-difference solutions to have the same properties—i.e. to be dissipationless and dispersionless, but, in general, this will not be (completely) possible. We will return to the issue of dissipation and dispersion in FDAs of wave problems below.

1.8 The Leap-Frog Scheme

First note that (48) is a good prototype for the general hyperbolic *system*:

$$\mathbf{u}_t = \mathbf{A}\mathbf{u}_x \tag{50}$$

where $\mathbf{u}$(x,t) is the n-component *solution vector*:

$$\mathbf{u}(x,t) = [u_1(x,t),\ u_2(x,t),\ \cdots\ u_n(x,t)] \tag{51}$$

and the $n \times n$ matrix $\mathbf{A}$ has distinct real eigenvalues

$$\lambda_1,\ \lambda_2,\ \cdots\ \lambda_n$$

so that, for example, there exists a similarity transformation $\mathbf{S}$ such that

$$\mathbf{S}\mathbf{A}\mathbf{S}^{-1} = \mathrm{diag}(\lambda_1,\ \lambda_2,\ \cdots\ \lambda_n)$$

The leap-frog scheme is a commonly used finite-difference approximation for hyperbolic systems. In the context of our simple scalar ($n = 1$) advection problem (48):

$$u_t = a\, u_x$$

an appropriate stencil is shown in Figure 7.

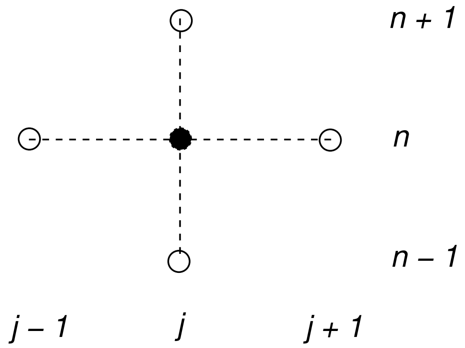

Figure 7. Stencil (molecule/star) for leap-frog scheme as applied to (48). Note that the central grid point has been filled in this figure to emphasize that the corresponding unknown, u^n_j, does not appear in the local discrete equation at that grid point (hence the term "leap-frog")

Applying the usual $O(h^2)$ approximations to ∂_x and ∂_t, our leap-frog (LF) scheme is

$$\frac{u^{n+1}_j - u^{n-1}_j}{2\,\triangle t} = a\,\frac{u^n_{j+1} - u^n_{j-1}}{2\,\triangle x} \tag{52}$$

or explicitly

$$u^{n+1}_j = u^{n-1}_j + a\lambda\left(u^n_{j+1} - u^n_{j-1}\right) \tag{53}$$

where

$$\lambda \equiv \frac{\triangle t}{\triangle x}$$

is the *Courant number* as previously.

Exercise: Perform a von Neumann stability analysis of (52) thus showing that $a\lambda \leq 1$ (which, you should note, is just the CFL condition) is necessary for stability.

Observe that the LF scheme (52) is a *three-level* scheme. As in our treatment of the wave equation, $u_{tt} = u_{xx}$ using the "standard scheme", we need to specify

$$u_j^0 \quad , \quad u_j^1 \qquad j = 1, 2, \cdots J$$

to "get the scheme going"—that is, we need to specify *two* numbers per spatial grid point. This should be contrasted to the continuum where we need to specify only *one* number per x_j, namely $u_0(x_j)$. Again, the initialization of the u_j^0 is trivial, given the (continuum) initial data $u_0(x)$, and, again, we need u_j^1 to $O(\triangle t^3) = O(h^3)$ accuracy for $O(h^2)$ global accuracy. Two possible approaches are as follows.

Taylor Series: The development here is parallel to that for the wave equation. We have

$$u_j^1 = u_j^0 + \triangle t \, (u_t)_j^0 + \frac{1}{2} \triangle t^2 \, (u_{tt})_j^0 + O(\triangle t^2)$$

also, from the equation of motion $u_t = a u_x$, we get

$$u_{tt} = (u_t)_t = (a u_x)_t = a \, (u_t)_x = a^2 u_{xx}.$$

so we have our desired initialization formula:

$$u_j^1 = u_j^0 + \triangle t \, (u_0')_j^0 + \frac{1}{2} \triangle t^2 \left(a^2 u_0''\right)_j^0 + O(\triangle t^3) \tag{54}$$

Self-Consistent Iterative Approach: The idea here is to initialize the u_j^1 from the u_j^0 and a version of the discrete equations of motion which introduces a "ficticious" half-time-level—see Figure 8.

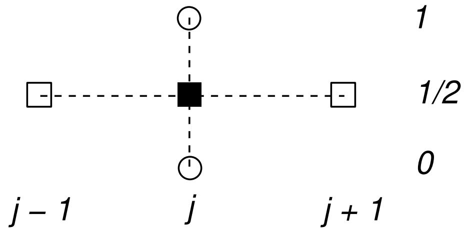

Figure 8. Stencil for initialization of leap-frog scheme for to (48). Note the introduction of the "fictitious" half-time level $t = t^{1/2}$ (squares).

Applying the leap-frog scheme on the stencil in the Figure, we have

$$\frac{u_j^1 - u_j^0}{\triangle t} = a\,\frac{u_{j+1}^{\frac{1}{2}} - u_{j-1}^{\frac{1}{2}}}{2\,\triangle x}$$

or, explicitly solving for u_j^1:

$$u_j^1 = u_j^0 + \frac{1}{2}\lambda\left(u_{j+1}^{\frac{1}{2}} - u_{j-1}^{\frac{1}{2}}\right)$$

It is a straightforward exercise to show that in order to retain $O(h^2)$ accuracy of the difference scheme, we need "fictitious-time" values, $u_j^{1/2}$ which are accurate to $O(h^2)$ (i.e. we can neglect terms which are of $O(h^2)$). In particular, if we *define* $u_j^{1/2}$, via

$$u_j^{\frac{1}{2}} = \frac{u_j^1 + u_j^0}{2}$$

which amounts to defining the half-time values via linear interpolation in the advanced and retarded unknowns, we will retain second-order accuracy.

We are thus led to the following initialization algorithm which is perhaps best expressed in pseudo-code (note, all loops over `j` are implicit:)

```
u[0,j] := u_0(x_j)
u[1,j] := u_0(x_j)
DO
   usave[j] := u[1,j]
   u[1/2,j] := (u[1,j] + u[0,j]) / 2

   u[1,j]   := u[0,j] + (lambda / 2) * (u[1/2,j+1] - u[1/2,j-1])

UNTIL  norm(usave[j] - u[1,j]) < epsilon
```

1.9 Error Analysis and Convergence Tests

As a side remark, we note that the discussion in this section applies to essentially *any* continuum problem which is solved using FDAs on a *uniform* mesh structure. In particular, the discussion applies to the treatment of ODEs and elliptic problems, where, in fact, convergence is often easier to achieve due to the fact that the FDAs are typically intrinsically stable (i.e. we have an easier time constructing stable FDAs for these types of problems). We also note that departures from non-uniformity in the mesh do not, in general, complete destroy the picture, but, rather, tend to distort it in ways which are beyond the scope of these notes. In any case, my colleagues have been known to paraphrase my whole approach to this subject as

Convergence!, Convergence!, Convergence!

1.9.1 Sample Analysis: The Advection Equation

We again consider the solution of the advection equation, but this time we impose periodic boundary conditions on our spatial domain, which we take to be $0 \le x \le 1$ with $x = 0$ and $x = 1$ identified (i.e. we solve the wave equation on $\mathbf{S}^1 \times \mathbf{R}$):

$$\begin{aligned} u_t &= a\,u_x \quad (a > 0) \qquad 0 \le x \le 1, \quad t \ge 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{55}$$

Note that the initial conditions $u_0(x)$ must be compatible with periodicity, i.e. we must specify periodic initial data.

Again, given the initial data, $u_0(x)$, we can immediately write down the full solution

$$u(x,t) = u_0(x + a\,t \bmod 1) \tag{56}$$

where mod is the usual modulus function which "wraps" $x + a\,t$, $t > 0$ onto the unit circle. As we shall see, because of the simplicity and solubility of this problem, one can perform a rather complete "analytic" (i.e. closed form) treatment of the convergence of simple FDAs of (55). The point of the exercise, however, is *not* to advocate parallel "analytic" treatments for more complicated problems. Rather, the key idea to be extracted from the following is that, in principle (always), and in practice (almost always, i.e. I've never seen a case where it *didn't* work, but then there's a lot of computations I haven't seen):

> *The error, e^h, of an FDA is no less computable than the solution, u^h itself.*

This has widespread ramifications, one of which is that there's really no excuse for publishing solutions of FDAs without error bars, or their equivalents!

Proceeding with our sample error analysis of the leap-frog scheme applied to the advection equation, we first introduce some difference operators for the usual $O(h^2)$ centred approximations of ∂_x and ∂_t:

$$D_x\, u_j^n \;\equiv\; \frac{u_{j+1}^n - u_{j-1}^n}{2\,\triangle x} \tag{57}$$

$$D_t\, u_j^n \;\equiv\; \frac{u_j^{n+1} - u_j^{n-1}}{2\,\triangle t} \tag{58}$$

We again take

$$\triangle x \;\equiv\; h \qquad \triangle t \;\equiv\; \lambda\,\triangle x \;=\; \lambda\, h$$

and will hold λ fixed as h varies, so that, as usual, our FDA is characterized by the single scale parameter, h.

The idea behind our error analysis is that we want to view the solution of the FDA as a *continuum* problem, and hence we will express both the difference operators and the FDA solution as asymptotic series (in h) of differential operators, and continuum functions, respectively. We have the following expansions for D_x and D_t:

$$D_x \;=\; \partial_x + \frac{1}{6}h^2\,\partial_{xxx} + O(h^4) \tag{59}$$

$$D_t \;=\; \partial_t + \frac{1}{6}\lambda^2 h^2\,\partial_{ttt} + O(h^4) \tag{60}$$

Now, in terms of the general, abstract formulation of 1.3 we have:

$$L\,u - f = 0 \qquad \Longleftrightarrow \qquad (\partial_t - a\,\partial_x)\,u = 0 \tag{61}$$

$$L^h u^h - f^h = 0 \qquad \Longleftrightarrow \qquad (D_t - a\,D_x)\,u^h = 0 \tag{62}$$

$$L^h u - f^h \equiv \tau^h \qquad \Longleftrightarrow \qquad (D_t - a\,D_x)\,u = \frac{1}{6}h^2\left(\lambda^2\partial_{ttt} - a\,\partial_{xxx}\right)u + \cdots \tag{63}$$

The Richardson ansatz: The key to our analysis is L.F. Richardson's old observation (*ansatz*) [5], that the solution, u^h, of *any* FDA which (1) uses a uniform mesh structure with scale parameter h, and (2) is completely centred, should have the following expansion in the limit $h \to 0$:

$$u^h(x,t) = u(x,t) + h^2 e_2(x,t) + h^4 e_4(x,t) + \cdots \qquad (64)$$

Here u is the continuum solution, while e_2, e_4, $\cdots$ are (continuum) *error functions* which *do not depend on h.* In a very real sense (64), is *the* key expression from which all error analysis of FDAs derives. We note that in the case that the FDA is *not* completely centred, we will have to modify the *ansatz.* In particular, for first order schemes (which are more common in relativistic astrophysics than one might expect!), we will have

$$u^h(x,t) = u(x,t) + h e_1(x,t) + h^2 e_x(x,t) + h^3 e_3(x,t) + \cdots \qquad (65)$$

Note that the Richardson *ansatz* (64) is completely compatible with the assertion discussed in 1.3.7, namely that

$$\tau^h = O(h^2) \quad \longrightarrow \quad e^h \equiv u - u^h = O(h^2) \qquad (66)$$

However, the Richardson form (64) contains much more information than "second-order truncation error should imply second-order solution error", telling us the precise form of the h dependence of u^h.

Given the Richardson expansion, we can now proceed with our error analysis. We start from the FDA, $L^h u^h - f^h = 0$, and replace both L^h and u^h with continuum expansions:

$$
\begin{aligned}
L^h u^h = 0 \quad &\longrightarrow \quad (D_t - a\,D_x)\left(u + h^2 e_2 + \cdots\right) = 0 \\
&\longrightarrow \quad \left(\partial_t + \frac{1}{6}\lambda^2 h^2 \partial_{ttt} - a\,\partial_x - \frac{1}{6} a h^2\,\partial_{xxx} + \cdots\right) \times \\
&\qquad\quad \left(u + h^2 e_2 + \cdots\right) = 0 \qquad (67)
\end{aligned}
$$

We now demand that terms in (67) vanish order-by-order in h. At $O(1)$ (zeroth-order), we have

$$(\partial_t - a\,\partial_x)\,u = 0 \qquad (68)$$

which is simply a statement of the *consistency* of the difference approximation. More interestingly, at $O(h^2)$ (second-order), we find

$$(\partial_t - a\,\partial_x)\,e_2 = \frac{1}{6}\left(a\partial_{xxx} - \lambda^2\partial_{ttt}\right)u \qquad (69)$$

which, assuming that we view u as a "known" function, is simply a PDE for the leading order error function, e_2. Moreover, the PDE governing e_2 is of *precisely* the same nature as the original PDE (48).

In fact, we can *solve* (69) for e_2. Given the "natural" initial conditions

$$e_2(x,0) = 0$$

(i.e. we initialize the FDA with the exact solution so that $u^h = u$ at $t = 0$), and defining $q(x + at)$:

$$q(x + at) \equiv \frac{1}{6} a\left(1 - \lambda^2 a^2\right)\partial_{xxx} u(x,t)$$

we have

$$e_2(x,t) = t\,q(x + at \bmod 1) \tag{70}$$

We note that, as is typical for leap-frog, we have *linear* growth of the finite difference error with time (to leading order in h). We also note that we can obviously push this analysis to higher order in h—what results, then, is an entire *hierarchy* of differential equations for u and the error functions e_2, e_4, e_6, $\cdots$. Indeed, it is useful to keep this view in mind:

> *When one solves an FDA of a PDE, one is* not *solving some system which is "simplified" realtive to the PDE, rather, one is solving a much* richer *system consisting of an (infinite) hierarchy of PDEs, one for each function appearing in the Richardson expansion (64).*

In the general case, of course, we will not be able to solve the PDE governing u, let alone that governing e_2—otherwise we wouldn't be considering the FDA in the first place! But it is precisely in this instance where the true power of Richardson's observation is evident. The key observation is that starting from (64), and computing FD solutions using the same initial data, but with differing values of h, we can learn a great deal about the error in our FD approximations. The whole game of investigating the manner in which a particular FDA converges or doesn't (i.e. looking at what happens as one varies h) is known as *convergence testing*. It is important to realize that there are no hard and fast rules for convergence testing; rather, one tends to tailor the tests to the specifics of the problem at hand, and, being largely an empirical approach, one gains experience and intuition as one works through more and more problems. That said, I emphasize again that the Richardson expansion, in some form or other, *always* underlies convergence analysis of FDAs.

A simple example of a convergence test, and the one I use most often in practice is the following. We compute three distinct FD solutions u^h, u^{2h}, u^{4h} at resolutions h, $2h$ and $4h$ respectively, but using the same initial data (as naturally expressed on the 3 distinct FD meshes). We also assume that the finite difference meshes "line up", i.e. that the $4h$ grid points are a subset of the $2h$ points which are a subset of the h points, so that, in particular, the $4h$ points constitute a common set of events (x_j, t^n) at which specific grid function values can be directly (i.e. no interpolation required) and meaningfully compared to one another. From the Richardson *ansatz* (64), we expect:

$$
\begin{aligned}
u^h &= u + h^2 e_2 + h^4 e_4 + \cdots \\
u^{2h} &= u + (2h)^2 e_2 + (2h)^4 e_4 + \cdots \\
u^{4h} &= u + (4h)^2 e_2 + (4h)^4 e_4 + \cdots
\end{aligned}
$$

We then compute a quantity $Q(t)$, which I will call a *convergence factor*, as follows:

$$
Q(t) \equiv \frac{\|u^{4h} - u^{2h}\|_x}{\|u^{2h} - u^h\|_x} \qquad (71)
$$

where $\|\cdot\|_x$ is any suitable discrete spatial norm, such as the ℓ_2 norm, $\|\cdot\|_2$:

$$
\|u^h\|_2 = \left(J^{-1} \sum_{j=1}^{J} \left(u_j^h\right)^2 \right)^{1/2} \qquad (72)
$$

and, for concreteness, the subtractions in (71) can be taken to involve the sets of mesh points which are common between u^{4h} and u^{2h}, and between u^{2h} and u^h. It is a simple exercise to show that, if our finite difference scheme is converging, then we should find:

$$
\lim_{h \to 0} Q(t) = 4. \qquad (73)
$$

In practice, one can use additional levels of discretization, $8h$, $16h$, etc. to extend this test to look for "trends" in $Q(t)$ and, in short, to convince oneself (and, with luck, others), that the FDA really *is* converging. Moreover, once convergence of an FDA has been established, then a point-wise subtraction of any two solutions computed at different resolutions, will immediately provide an estimate of the level of error in both. For example, if we have u^h and u^{2h}, then, again by the Richardson *ansatz* we have

$$\begin{aligned} u^{2h} - u^h &= \left(\left(u + (2h)^2 e_2 + \cdots\right) - \left(u + h^2 e_2 + \cdots\right)\right) \\ &= 3h^2 e_2 + O(h^4) \sim 3e^h \sim \frac{3}{4} e^{2h} \end{aligned} \tag{74}$$

Richardson extrapolation: Richardson's observation (64) also provides the basis for all the techniques of *Richardson extrapolation*, where solutions computed at different resolutions are linearly combined so as to *eliminate* leading order error terms, and hence provide more accurate solutions. As an example, given u^h and u^{2h} which satisfy (64), we can take the linear combination, $\bar{u}^h$:

$$\bar{u}^h \equiv \frac{4u^h - u^{2h}}{3} \tag{75}$$

which, by (64), is easily seen to be $O(h^4)$, i.e. *fourth*-order accurate!

$$\begin{aligned} \bar{u}^h &\equiv \frac{4u^h - u^{2h}}{3} = \frac{4\left(u + h^2 e_2 + h^4 e_4 + \cdots\right) - \left(u + 4h^2 e_2 + 16h^4 e_4 + \cdots\right)}{3} \\ &= -4h^4 e_4 + O(h^6) = O(h^4) \end{aligned} \tag{76}$$

When it works, Richardson extrapolation has an almost magical quality about it, but one generally has to start with fairly accurate (on the order of a few

%) solutions in order to see the dramatic improvement in accuracy suggested by (76). Partly because it is still a struggle to achieve that sort of accuracy (i.e. a few %) for *any* computation in many areas of numerical relativity/astrophysics, techniques based on Richardson extrapolation have not had a major impact in this context.

Independent Residual Evaluation A question which often arises in discussions of convergence testing is the following:

> "OK, you've established that u^h is converging as $h \to 0$, but how do you know you're converging to u, the solution of the continuum problem?"

Here, the notion of an independent residual evaluation is very useful. The idea is as follows: we have our continuum PDE

$$Lu - f = 0 \tag{77}$$

and our FDA

$$L^h u^h - f^h = 0 \tag{78}$$

We have demonstrated that u^h is apparently converging by, for example, computing the convergence factor (71) and verifying that it tends to 4 as h tends to 0. However, we do not know if we have derived and/or implemented our discrete operator L^h correctly. Note that implicit in the "implementation" is the fact that, particularly for multi-dimensional and/or implicit and/or multi-component FDAs, considerable "work" (i.e. analysis and coding) may be involved in setting up and solving the algebraic equations for u^h. As a check that we *are* converging to u, we consider a *distinct* (i.e. independent) discretization of the PDE:

$$\tilde{L}^h \tilde{u}^h - f^h = 0 \tag{79}$$

The only thing we need from this FDA for the purposes of the independent residual test is the new FD operator $\tilde{L}^h$. As with L^h, we can expand $\tilde{L}^h$ in powers of the mesh spacing:

$$\tilde{L}^h = L + h^2 E_2 + h^4 E_4 + \cdots \tag{80}$$

where E_2, E_4, $\cdots$ are higher order (involve higher order derivatives than L) differential operators. We then simply apply the new operator $\tilde{L}^h$ to our FDA u^h and investigate what happens as $h \to 0$. If u^h *is* converging to the continuum solution, u, we will have

$$u^h = u + h^2 e_2 + O(h^4) \tag{81}$$

and we will compute

$$\begin{aligned} \tilde{L}^h u^h &= \left(L + h^2 E_2 + O(h^4)\right)\left(u + h^2 e_2 + O(h^4)\right) \\ &= Lu + h^2 (E_2\, u + L\, e_2) = O(h^2) \end{aligned} \tag{82}$$

i.e., $\tilde{L}^h u^h$ will be a residual-like quantity which converges quadratically as $h \to 0$. Conversely, if we have goofed in our derivation and/or implementation of $L^h u^h = f^h = 0$, but we still see convergence; i.e. we have, for example, $u^{2h} - u^h \to 0$ as $h \to 0$, then we must have something like

$$u^h = u + e_0 + h e_1 + h^2 e_2 + \cdots \tag{83}$$

where the crucial fact is that the error must have an $O(1)$ component, e_0. In this case, we will compute

$$
\begin{aligned}
\tilde{L}^h u^h &= \left(L + h^2 E_2 + O(h^4)\right)\left(u + e_0 + he_1 + h^2 e_2 + O(h^4)\right) \\
&= Lu + Le_0 + hLe_1 + O(h^2) = Le_0 + O(h)
\end{aligned}
\tag{84}
$$

and, unless we are *extraordinarily* lucky, and $L\,e_0$ vanishes, we will *not* observe the expected convergence, rather, we will see $\tilde{L}^h u^h - f^h$ tending to a *finite* ($O(1)$) value—a sure sign that something is wrong.

There is of course, the problem that we might have slipped up in our implementation of the "independent residual evaluator", $\tilde{L}^h$, in which case the results from our test will be ambigous at best! However, a key point here is that because $\tilde{L}^h$ is only used *a posteriori* on a computed solution (we never use it to compute $\tilde{u}^h$, for example) it is a relatively easy matter to ensure that $\tilde{L}^h$ has been implemented in an error-free fashion (perhaps using symbolic manipulation facilities). Furthermore, many of the restrictions commonly placed on the "real" discretization (such as stability and the ease of solution of the resulting algebraic equations) do not apply to $\tilde{L}^h$.

1.10 Dispersion and Dissipation in FDAs

We again consider the advection model problem, $u_t = a\, u_x$, but now discretize only in space (semi-discretization) using the usual $O(h^2)$ centred difference approximation:

$$u_t = a\, D_x\, u \equiv a\, \frac{u_{j+1} - u_{j-1}}{2\, \triangle x} \tag{85}$$

We now look for normal-mode solutions to (85) of the form

$$u = e^{ik\,(x+a'\,t)}$$

where the "discrete phase speed", a', is to be determined. Substitution of this *ansatz* in (85) yields

$$ika'u = \frac{a\,(2i\sin(k\,\triangle x\,))}{2\,\triangle x} u$$

or, solving for the discrete phase speed, a'

$$a' = a\frac{\sin(k\,\triangle x\,)}{k\,\triangle x} = a\frac{\sin\xi}{\xi}$$

where we have defined the dimensionless wave number, ξ:

$$\xi \equiv k\,\triangle x$$

In the *low frequency* limit, $\xi \to 0$, we have the expected result:

$$a' = a\frac{\sin\xi}{\xi} \to a$$

so that low frequency components propagate with the correct phase speed, a. However, in the *high frequency* limit, $\xi \to \pi$, we have

$$a' = a\frac{\sin\xi}{\xi} \to 0 \quad !!$$

i.e. the highest frequency components of the solution don't propagate at all! This is typical of FDAs of wave equations, particularly for relatively low-order schemes. The propagation of high frequency components of the difference solution is essentially completely wrong. Arguably then, there can be little harm in attenuating (dissipating) these components, and, in fact, since high frequency components are potentially troublesome (particularly *vis a vis* non-linearities and the treatment of boundaries), it is often *advantageous* to use a dissipative difference scheme.

Some FDAs are naturally dissipative (the Lax-Wendroff scheme, for example), while others, such as leap-frog, are not. In the case of a leap-frog-based scheme, the idea is to add dissipative terms to the method, but in such a way as to retain $O(h^2)$ accuracy of the scheme. Consider, for example, the leap-frog scheme as applied to the advection model problem:

$$u_j^{n+1} = u_j^{n-1} + a\lambda\left(u_{j+1}^n - u_{j-1}^n\right)$$

We add dissipation to the scheme by modifying it as follows:

$$\begin{aligned} u_j^{n+1} = u_j^{n-1} \ & + \ a\lambda\left(u_{j+1}^n - u_{j-1}^n\right) \\ & - \ \frac{\epsilon}{16}\left(u_{j+2}^{n-1} - 4u_{j+1}^{n-1} + 6u_j^{n-1} - 4u_{j-1}^{n-1} + u_{j-2}^{n-1}\right) \end{aligned} \tag{86}$$

where ϵ is an adjustable, non-negative parameter.

Note that

$$
\begin{aligned}
u^{n-1}_{j+2} - 4u^{n-1}_{j+1} + 6u^{n-1}_{j} - 4u^{n-1}_{j-1} + u^{n-1}_{j-2} &= \triangle x^{\,4} (u_{xxxx})^{n-1}_{j} + O(h^6) \\
&= \triangle x^{\,4} (u_{xxxx})^{n}_{j} + O(h^5) = O(h^4)
\end{aligned}
$$

so that the term which is added, does not change the leading order truncation error, which in the form we have written the equation, is $O(\triangle t^{\,3}) = O(h^3)$ (local/one-step truncation error).

A Von Neumann analysis of the modified scheme shows that, in addtion to the CFL condition $\lambda \le 1$, we must have $\epsilon < 1$ for stability, and, further, that the per-step amplification factor for a mode with wave number ξ is, to leading order

$$
1 - \epsilon \sin^4 \frac{\xi}{2}
$$

Thus the addition of the dissipation term is analagous to the use of an explicit "high frequency filter" (low-pass filter), which has a fairly sharp rollover as $\xi \to \pi$.

We note that an advantage to the use of explicit dissipation techniques (versus, for example, the use of an intrinsically dissipative scheme) is that the amount of dissipation can be controlled by tuning the dissipation parameter.

1.11 Lab Problem 1

Implement a `Fortran` or `C` program to solve the 1-d wave equation with fixed (Dirichlet) boundary conditions using the "standard" $O(h^2)$ discretization as discussed in Section . Some coding hints and routines which you can use to visualize your output will be available on-line.

1.12 Lab Problem 2

Perform an empirical convergence analysis of your implementation

2 A Very Brief Introduction to RNPL

2.1 Motivation

I quote directly from the Introduction of the RNPL User's Manual [23]

> *Writing a program to solve a system of partial differential equations takes a lot of work. It is not just a matter of implementing a clever solution scheme. A working program requires much more. It needs code for parameter fetching, initial data generation, input and output, memory management, and checkpointing as well as the actual routines for solving the equations.*
>
> *RNPL (Rapid Numerical Prototyping Language) was written to help solve scientists solve equations quickly by automatically taking care of everything except the inner-most parts of the solution routines. In many cases, RNPL can generate the entire program-updates and all.*
>
> *RNPL can be used in three basic ways: for producing complete programs, for producing skeleton programs, and for converting existing programs.*

2.2 A Simple Example: The 1-d Wave Equation

2.2.1 RNPL source file: wave_rnpl

```
#-------------------------------------------------
# RNPL program to solve 1-d wave equation
#
#   phi_tt = phi_xx
#
# with Dirichlet (fixed) boundary conditions.
#
# Initial data is a right-moving Gaussian pulse.
#-------------------------------------------------

parameter float xmin  := 0.0
parameter float xmax  := 1.0
parameter float amp   := 1.0
parameter float x0    := 0.5
parameter float sigma := 0.05

rect coordinates t,x

uniform rect grid g1 [1:Nx] {xmin:xmax}

float phi on g1 at -1,0,1 { out_gf := 1 }

operator D_LF(f,x,x) := (<0>f[1] - 2*<0>f[0] + <0>f[-1]) / (dx*dx)
operator D_LF(f,t,t) := (<1>f[0] - 2*<0>f[0] + <-1>f[0]) / (dt*dt)
operator D_LF(f,x)   := (<0>f[1] - <0>f[-1]) / (2*dx)
operator D_LF(f,t)   := (<1>f[0] - <-1>f[0]) / (2*dt)

evaluate residual phi {
     [1:1] := D_LF(phi,t);
  [2:Nx-1] := D_LF(phi,t,t) - D_LF(phi,x,x);
   [Nx:Nx] := D_LF(phi,t)
}
```

```
initialize phi {
    [1:Nx] := amp*exp(-(x-x0)^2/sigma^2)
}
initialize <-1>phi {
     [1:1] := 0;
  [2:Nx-1] := 0.5*(dt/dx)^2*(<0>phi[1] + <0>phi[-1]) +
              (1-(dt/dx)^2)*phi - dt*2*(x-x0)/sigma^2*phi;
   [Nx:Nx] := 0
}

looper iterative

auto update phi
```

2.2.2 Parameter file: `id0`

```
# parameters for wave
in_file := "in0.sdf"
out_file := "out0.sdf"
level := 0
output:=*-*

Nx0  := 128
iter := 256

tag    := ""
lambda := .5

xmin := 0.0
xmax := 1.0

ser  := 0
fout := 1

amp := 1.0
x0  := 0.5
sigma := 0.05
```

2.2.3 Building and running the application

```
% ls
Makefile    id0         wave_rnpl

% make
        /usr/local/bin/rnpl -l allf  wave_rnpl
        f77  -64 -O3   -YI,/usr/local/include  -c wave.f
        f77  -64 -O3   -YI,/usr/local/include  -c updates.f
        f77  -64 -O3   -YI,/usr/local/include  -c residuals.f
        f77  -64 -O3  -64 -L/usr/local/lib   -L/usr/local/lib  \
             wave.o updates.o residuals.o -lrnpl  -lvsso -lsv -o wave
        f77  -64 -O3   -YI,/usr/local/include  -c wave_init.f
        f77  -64 -O3   -YI,/usr/local/include  -c initializers.f
        f77  -64 -O3  -64 -L/usr/local/lib   -L/usr/local/lib  \
             wave_init.o updates.o residuals.o initializers.o -lrnpl  \
             -lvsso -lsv -o wave_init

% ls
Makefile         initializers.o   updates.f        wave_init*
gfuni0.inc       other_glbs.inc   updates.o        wave_init.f
globals.inc      residuals.f      wave*            wave_init.o
id0              residuals.o      wave.f           wave_rnpl
initializers.f   sys_param.inc    wave.o

% wave id0
 Can't open in0.sdf
 Calling initial data generator.
 WARNING: using default for parameter epsiter.
 WARNING: using default for parameter epsiterid.
 WARNING: using default for parameter maxstep.
 WARNING: using default for parameter maxstepid.
 WARNING: using default for parameter s_step.
 WARNING: using default for parameter start_t.
 WARNING: using default for parameter trace.
 WARNING: using default for parameter epsiter.
```

```
 WARNING: using default for parameter epsiterid.
 WARNING: using default for parameter maxstep.
 WARNING: using default for parameter maxstepid.
 WARNING: using default for parameter s_step.
 WARNING: using default for parameter start_t.
 WARNING: using default for parameter trace.
 Starting evolution.  step:            0 at t=  0.0
 step:            1 t=  3.9062500000E-03 steps=           1
 step:            2 t=  7.8125000000E-03 steps=           1
 step:            3 t=  1.1718750000E-02 steps=           1
                              .
                              .
                              .
 step:          253 t=  0.9882812500     steps=           1
 step:          254 t=  0.9921875000     steps=           1
 step:          255 t=  0.9960937500     steps=           1
 step:          256 t=   1.000000000     steps=           1

% ls *sdf
in0.sdf     out0.sdf    phi_0.sdf
```

3 An Introduction to Critical Behaviour in Gravitational Collapse

3.1 Overview and Synopsis

One development in relativity over the past few years has been the realization that the Einstein equations, with or without coupling to matter, admit special solutions which "sit at the threshold of black hole formation". These solutions are *strong-field*, and, unless one is very clever about the choice of one's coordinates, *non-trivially dynamic*. Thus, although they could have been discovered any time in the past 30 years or so, they remained largely hidden until we started to do extensive studies of the actual dynamics of "generic" initial data near the black hole threshold. These studies involved parametrized families of data, where the family parameter p could be tuned to control the amount of non-linearity in the generated spacetime; in particular, if p exceeded some threshold value $p^\star$ (which, of course, would generally be different for every family), a black hole would be formed. These threshold solutions are, by and large, *unstable* by construction, but, interestingly, generic collapse models seem to generically admit *minimally* unstable threshold solutions—that is, solutions with precisely one unstable (growing) mode in perturbation theory. This view of a critical (i.e. black-hole threshold) solution as an "intermediate attractor" with one unstable mode, has been very successful in explaining the observed scaling laws, such as

$$M_{\rm BH} \propto |p - p^\star|^\gamma$$

which were first observed in purely empirical studies of black hole critical behaviour (i.e. from "full blown" solution of the partial-differential equations of motion).

For a variety of reasons, much of what we know phenomenologically about black hole critical phenomena derives from such empirical studies in *spherical symmetry*. My collaborators and I, and others, are currently looking at critical phenomena in less symmetric cases (axisymmetry, as well as the generic "3-D" case), but discussion of those studies is clearly beyond the scope of these notes. More importantly, there is good reason to believe that many of the spherically symmetric critical solutions we know about are *stable* to departures from spherical symmetry—and thus should continue to be relevant even for generic critical collapse. Finally, calculations in spherical symmetry are sufficiently simple, and computer resources are sufficiently abundant, that the computations in [9], for example, could routinely be reproduced by a sufficiently inspired undergraduate with a laptop. This is important since, at least for me, half the fun of this whole game is setting up the equations of motion, "letting 'er rip" (in the immortal words of Stu Shapiro) and seeing (in real time if possible) what happens. In any case, in these notes we will restrict our attention to the spherically symmetric case.

3.2 Spherically Symmetric GR: Polar/Areal Coordinates

Many, many authors have worked in spherically symmetric numerical relativity. My particular approach (see [7, 8, 13, 11] for example), is based on the "usual" 3+1 decomposition, and closely parallels Piran's work in cylindrical symmetry [10]. Adopting the usual polar-spherical coordinates t, r, θ, φ, and working in relativists' units ($G = c = 1$), the *general* time-dependent spherically-symmetric metric can be written:

$$ds^2 = \left(-\alpha^2 + a^2\beta^2\right) dt^2 + 2a^2\beta dt dr + a^2 dr^2 + r^2 b^2 d\Omega^2 \tag{87}$$

where α, β, a and b are all functions of r and t, and $d\Omega^2$, defined by

$$d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\varphi^2$$

is the metric on the unit 2-sphere. The corresponding extrinsic curvature tensor is

$$K^i{}_j = \mathrm{diag}\left(K^r{}_r, K^\theta{}_\theta, K^\theta{}_\theta\right) \tag{88}$$

As usual, the lapse, α, and shift-vector-component, β, embody the coordinate invariance of general relativity, and must be prescribed, in some fashion or another, when any specific calculation is carried out. That is

Choosing a coordinate system $\Longleftrightarrow$ *Specifying the lapse and shift*

Not surprisingly, a particularly simple set of Einstein equations (and, when matter is coupled in, particularly simple sets of matter EOM) results from the choice of coordinates which, in some sense, most naturally generalize the Schwarzschild coordinates, which appear in the familiar form of the *static*, spherically-symmetric, vacuum line-element (i.e. the Schwarzschild solution):

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{89}$$

Note that r provides a direct measure of proper surface area: i.e. at any time, the proper area of an r = constant 2-sphere is simply $4\pi r^2$. Also note that the 4-metric is diagonal, i.e. the shift vector vanishes in (89). One way of generalizing these coordinates to the time dependent case, then, is to require that

1. The radial coordinate remain *areal*, i.e., in terms of the general form (87), that $b(r,t) \equiv 1$, so that the proper area of an r = constant 2-sphere remains $4\pi r^2$.
2. The 4-metric remain diagonal, i.e. that the shift vector continue to vanish.

We thus have the desired form of our line element:

$$ds^2 = -\alpha^2(r,t)dt^2 + a^2(r,t)dr^2 + r^2 d\Omega^2 \tag{90}$$

and have essentially exhausted our coordinate freedom (sometimes called "gauge freedom"). The residual freedom is arbitrary reparametrizations of the time coordinate, $t \to \tilde{t} = \tilde{t}(t)$, which can be completely exhausted, for example, by requiring that $\alpha(r,t) \to 1$ as $r \to \infty$. This corresponds to setting up our coordinates so that coordinate time and proper time coincide at spatial infinity, which turns out to be a convenient choice for the numerical analysis.

It can be further shown that requiring the metric to have the form (90) implies that, K, the trace of the extrinsic curvature

$$K \equiv K^i{}_i = K^r{}_r + 2K^\theta{}_\theta$$

is equal to the radial component, $K^r{}_r$, of the extrinsic curvature, i.e. that

$$K^\theta{}_\theta = 0.$$

The slicing condition, then, can be expressed as a constraint on K:

$$K = K^r{}_r$$

and can be extended to more generic (less-symmetric) spacetimes [15]. It is known as *polar* slicing, and hence, one often speaks of as polar/areal coordinates in reference to the form (90).

We note that we have "killed three birds with two stones", in the sense that we have eliminated *three* geometric variables (β, b and $K^\theta{}_\theta$) using *two* coordinate conditions. This is an unusual occurence, dependent both on the high-symmetry of the fundamental *ansatz* (i.e. spherical symmetry), and on the especially simple form of the metric. Nonetheless, we are, of course, free to use such simplifications to our advantage. In particular, in polar/areal coordinates we are left with only the following geometric variables:

1. A single 3-metric function, a, which we can "solve for" using the Hamiltonian constraint, *or* via an evolution equation which is essentially just the definition of $K^r{}_r$.

2. The lapse function, α, which is constrained by the polar slicing condition, which follows from the demand that $K^\theta{}_\theta(r,t) = 0$ and $\partial_t K^\theta{}_\theta(r,t) = 0$.

3. The single non-trivial extrinsic curvature component, $K^r{}_r$, which can be updated using the momentum constraint, or using the evolution equation $\partial_t K^r{}_r = \cdots$ which follows from the Einstein field equations. Moreover, the momentum contraint is *algebraic* in this instance, so we can effectively eliminate $K^r{}_r$ from the scheme as well, replacing it with some function of the matter variables as dictated by the constraint.

3.2.1 Stress Energy Components and Geometric EOM

Before proceeding to a discussion of specific matter models for spherically-symmetric collapse, it is useful to write down, in as general fashion as possible, the various Einstein equations and coordinate-conditions which arise in polar/areal coordinates. To this end (and again, in the 3+1 spirit) we introduce the following components of the stress-energy tensor, $T^{\mu\nu}$

$$\begin{aligned} \rho &\equiv n_\mu n_\nu T^{\mu\nu} \\ j_i &\equiv -n_\mu T^\mu{}_i \\ S_{ij} &\equiv T_{ij} \end{aligned} \tag{91}$$

Here, $n_\mu = g_{\mu\nu} n^\nu$ where n^μ is the future-directed, unit timelike normal to the hypersurfaces $t =$ constant. In any 3+1 coordinate system we have

$$n_\mu = [-\alpha, 0, 0, 0] \tag{92}$$

so we have

$$\rho \equiv n_\mu n_\nu T^{\mu\nu} = \alpha^2 T^{tt} \tag{93}$$

and $j_i = [j_r, 0, 0]$, where

$$j_r = -n_\mu T^\mu{}_r = \alpha T^0{}_r \tag{94}$$

With these definitions, it turns out that the general form of the Hamiltonian constraint, which will be used to update a, is

$$\frac{a'}{a} + \frac{a^2 - 1}{2r} - 4\pi r a^2 \rho = 0 \tag{95}$$

while the general form of the slicing condition, which will be used to update α, is

$$\frac{\alpha'}{\alpha} - \frac{a^2 - 1}{2r} - 4\pi r a^2 S^r{}_r = 0 \tag{96}$$

Here and below, a prime will denote partial differentiation with respect to r. Also, note that j_i and S_{ij} are 3-tensors, so that, for example

$$S^i{}_j = \gamma^{ik} S_{kj} \tag{97}$$

where the 3-metric γ_{ij} is

$$\gamma_{ij} = \mathrm{diag}\left(a^2, r^2, r^2 \sin^2\theta\right) \tag{98}$$

with associated inverse γ^{ij}:

$$\gamma^{ij} = \mathrm{diag}\left(a^{-2}, r^{-2}, r^{-2} \sin^{-2}\theta\right) \tag{99}$$

We will (loosely) refer to ρ, j_r and S_{ij} as the "energy density", "momentum denstity", and "stresses", respectively.

Finally, we observe that it is often useful to introduce the so-called "mass aspect" function, $m(r,t)$, in analogy with the Schwarzschild form (89). Specifically we define $m(r,t)$ through

$$a^2(r,t) = \left(1 - \frac{2m(r,t)}{r}\right)^{-1} \tag{100}$$

It can be shown that, in a vacuum region, $m(r,t)$ measures the amount of (gravitating) mass contained within radius r at time t, and, in fact, has an intrinsic geometric interpretation (i.e. is "gauge independent"). Similarly, the quantity $m'(r,t) = dm/dr$ also has a geometric interpretation—and both m and m' are useful diagnostic quantities to monitor during the course of a dynamical evolution. Interestingly enough (although one should not dwell on the significance of the result, since it *does* depend on our specific choice of coordinate system), the Hamiltonian constraint can be re-written in the very suggestive form

$$\frac{dm}{dr} = 4\pi r^2 \rho \tag{101}$$

As is well known, *vacuum* general relativity is trivial in spherical symmetry in the sense that the only solution is static Schwarzschild, given, for example, by (89). In order to investigate *dynamics* in spherically symmetric GR, then, it is necessary to introduce one or more matter fields, and in the following we will consider two specific choices which have been well studied in the context of black-hole critical phenomena.

3.3 Minimally Coupled Massless Klein Gordon Field In this case, the matter field is a single, real, minimally coupled, massless scalar field, $\phi(r,t)$, with Lagrangian scalar

$$L_\phi = -\frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi \tag{102}$$

and stress-energy tensor

$$\begin{aligned} T_{\mu\nu} &\equiv -2\frac{\partial L_\phi}{\partial g^{\mu\nu}} + g_{\mu\nu}L_\phi & (103) \\ &= \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}\partial_\sigma\phi\partial^\sigma\phi & (104) \end{aligned}$$

The scalar field equation of motion is

$$\nabla_\mu\nabla^\mu\phi = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\sigma}\partial_\sigma\phi\right) = 0 \tag{105}$$

and can be derived either from the variation of the total action with respect to ϕ, or via conservation of the stress-tensor

$$\nabla_\mu T^{\mu\nu} = 0. \tag{106}$$

For numerical purposes (and consistent with the 3+1 philosophy), it is convenient to re-write the scalar field EOM in terms of "Hamiltonian" variables. To that end we define auxiliary variables

$$\begin{aligned} \Phi(r,t) &\equiv \frac{\partial\phi}{\partial r} & (107) \\ \Pi(r,t) &\equiv \frac{a}{\alpha}\frac{\partial\phi}{\partial t} & (108) \end{aligned}$$

Keeping in mind that we are restricting attention to polar/areal coordinates, it is easy to check that the following set of equations are equivalent to (105)

$$\frac{\partial \Phi}{\partial t} = \frac{\partial}{\partial r}\left(\frac{\alpha}{a}\Pi\right) \tag{109}$$

$$\frac{\partial \Pi}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\alpha}{a}\Phi\right) = 3\frac{\partial}{\partial r^3}\left(r^2\frac{\alpha}{a}\Phi\right) \tag{110}$$

Using (91-94) we find the following expressions for the stress-energy components we need:

$$\rho = \frac{\Phi^2 + \Pi^2}{2a^2}$$

$$S^r{}_r = \frac{\Phi^2 + \Pi^2}{2a^2} = \rho$$

$$j_r = -\frac{\Phi\Pi}{a}$$

Substitution of the above in the general (for polar/areal coordinates!) expressions for the Hamiltonian constraint (95) and slicing condition (96) we have

$$\frac{a'}{a} + \frac{a^2-1}{2r} - 2\pi r\left(\Phi^2 + \Pi^2\right) = 0 \tag{111}$$

$$\frac{\alpha'}{\alpha} - \frac{a^2-1}{2r} - 2\pi r\left(\Phi^2 + \Pi^2\right) = 0 \tag{112}$$

Equations (109-112) constitute a complete set of equations of motion for the spherically-symmetric EMKG (Einstein-Massless-Klein-Gordon) problem. However, before we can proceed with an FD implementation, we need to discuss a few more "details", including regularity at $r = 0$, and initial conditions. To this end, it is useful to first consider the "weak field" limit of the EMKG model, i.e. the limit in which the scalar field's gravitationally-induced backreaction on itself is negligible.

3.3.1 The Weak Field Limit

Here we have

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2 \tag{113}$$

that is

$$\alpha(r,t) \equiv 1 \qquad a(r,t) \equiv 1$$

The scalar field equation of motion(105) becomes

$$\phi_{tt} = \frac{1}{r^2}\left(r^2\phi_r\right)_r \tag{114}$$

which can be rewritten as the ordinary, one dimensional wave equation for the quantity $r\phi$

$$(r\phi)_{tt} = (r\phi)_{rr} \tag{115}$$

We can thus immediately write down, at least schematically, the general solution of our problem—as was the case in Section 1.5, it is simply a superposition of arbitrary in-going ("left-moving") and out-going ("right-moving") profiles:

$$(r\phi)(r,t) = f(t+t) + g(t-r) \tag{116}$$

As before, then, we can specify the initial data in terms of an initially in-going profile, $f(r)$ and an initially outgoing profile $g(r)$:

$$
\begin{aligned}
(r\phi)(r,0) &= f(r) + g(r) && (117)\\
(r\phi)(r,0)_t &= \frac{df}{dr}(r) - \frac{dg}{dr}(r) && (118)
\end{aligned}
$$

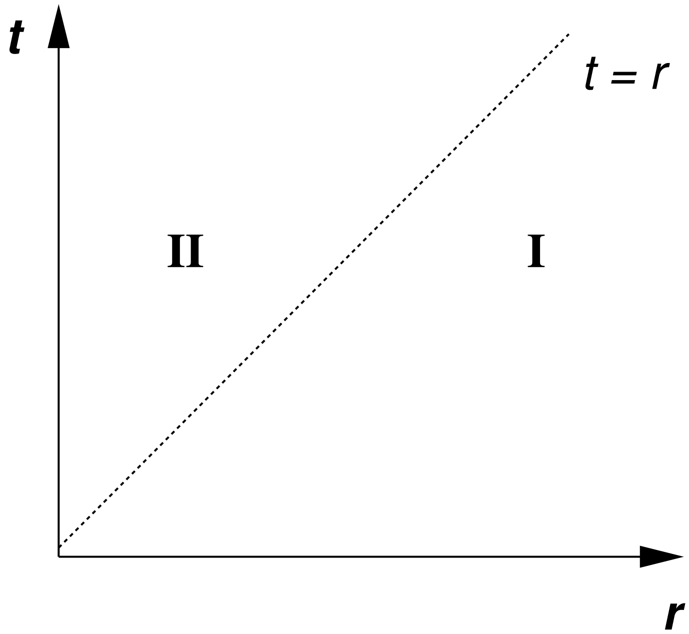

Figure 9. Illustration of solution domain for spherically symmetric wave equation.

We now consider the complete solution of the linear problem in more detail. First note that the solution domain is

$$r \geq 0 \qquad t \geq 0$$

(see Figure 9). We consider the two regions shown in the Figure separately. Note that the regions meet along $t = r$.

Region I: Here we have $t \leq r$ and, in terms of our in-going and out-going profiles f and g we have

$$(r\phi)\,(r,t) = f(t+r) + g(t-r) \tag{119}$$

Region II: Here we have $t > r$ and we note that only data which was *in-going* at $t = 0$ can have influence in this region. However, our solution must still be of the form

$$(r\phi)\,(r,t) = f(t+r) + h(t-r) \tag{120}$$

but we need to determine what the function h is. Rewriting the last expression, we have

$$\phi(r,t) = \frac{f(t+r)}{r} + \frac{h(t-r)}{r} \tag{121}$$

Let us now examine the behaviour near $r = 0$, assuming that $\phi(0,t)$, as well as f and h, are smooth; then as $r \to 0$ we have

$$\begin{aligned} f(t+r) &= f(t) + r\,f'(t) + \frac{1}{2}r^2\,f''(t) + \cdots && (122)\\ h(t-r) &= h(t) - r\,g'(t) + \frac{1}{2}r^2\,h''(t) + \cdots && (123) \end{aligned}$$

so

$$\lim_{r\to 0}\phi(t,r) = \frac{f(t)+h(t)}{r} \;+\; f'(t) - h'(t) \;+\; \frac{1}{2}r\left(f''(t) + h''(t)\right) + \cdots \tag{124}$$

Clearly, for the solution to be regular at $r = 0$ we must have

$$h(t) = -f(t) \;\longrightarrow\; h^{(n)} = -f^{(n)}$$

Then

$$\lim_{r \to 0} \phi(t, r) = 2f'(t) \,+\, \frac{1}{3} r^2 f'''(t) \,+\, \cdots \tag{125}$$

From this last result we can also see that we have

$$\phi_r(0, t) = 0 \tag{126}$$

which, in fact, is our desired *regularity* condition which will need to be enforced in some fashion in an FDA treatment. (Note that you will sometimes hear/see (126) referred to as a "boundary condition", since $r = 0$ is a "boundary" of the solution domain. It is important to keep in mind that $r = 0$ is *not* on a boundary—it is in the *interior* of our spherical domain! Thus, demanding that our solution be *regular* at $r = 0$ is *the* natural and physical thing to do; i.e. don't make $r = 0$ any more special than it has to be!)

Although our current problem is most naturally posed on the unbounded spatial domain, $r \geq 0$, if we wish to solve it via an FDA, we will have to introduce some arbitrary outer radius, $r_{\max}$, at which we "cut the computation off". In this case we will need boundary conditions on our dynamical variables, Φ and Π, which can be derived from the demand that there be *no in-coming waves* at large r (at $r = r_{\max}$ in particular). That is, following Sommerfeld, we demand that

$$\lim_{r\to\infty} r\phi(r,t) = r\phi(t-r) \tag{127}$$

In terms of the auxiliary variables, Φ and Π defined in (107-108)—which in the current case are just $\Phi \equiv \phi_r$ and $\Pi \equiv \phi_t$—we find from (127):

$$\begin{aligned}
\lim_{r\to\infty}\ \Phi_t + \Phi_r + \frac{\Phi}{r} &= O(r^{-3}) \approx 0 && (128)\\
\lim_{r\to\infty}\ \Pi_t + \Pi_r + \frac{\Pi}{r} &= 0 && (129)
\end{aligned}$$

Conditions such as these are often called *out-going radiation conditions* or *Sommerfeld conditions*. Although we cannot dwell on this issue in much detail, it should be pointed out that some care must be taken in differencing such boundary conditions—in conjunction with the interior differencing scheme—in order that the overall evolution be stable.

3.3.2 Strong Field Limit: Regularity, Boundary and Initial Conditions

Returning now to the strong field case, where the metric is

$$ds^2 = -\alpha^2 dt^2 + a^2 dr^2 + r^2 d\Omega^2$$

we first note that demanding that spacetime be *locally flat* near $r = 0$ gives us the following condition:

$$a(0,t) = 1 \tag{130}$$

Furthermore, additional considerations of regularity at $r = 0$ (see [15] for full details), yield the following conditions.

$$\begin{aligned} a_r(0,t) &= 0 && (131) \\ \alpha_r(0,t) &= 0 && (132) \end{aligned}$$

Given the "initial condition" (130), the Hamiltonian constraint

$$\frac{a'}{a} + \frac{a^2-1}{2r} - 2\pi r \left(\Phi^2 + \Pi^2\right) = 0$$

can be integrated radially outwards, and no additional condition is required at $r = r_{\max}$. Likewise, the slicing condition

$$\frac{\alpha'}{\alpha} - \frac{a^2-1}{2r} - 2\pi r \left(\Phi^2 + \Pi^2\right) = 0$$

can be integrated radially outwards for α, and, due to the linearity and homogeneity of the solution, we can choose the central value $\alpha(0,t)$ *arbitrarily*, then rescale the solution *a posteriori* via

$$\alpha(r,t) \longrightarrow \kappa\, \alpha(r,t)$$

with the constant κ chosen so that we end up with

$$\alpha(r_{\max},t) = \frac{1}{a(r_{\max},t)} \tag{133}$$

which corresponds to "proper-time at infinity" normalization of our time coordinate.

For the scalar field variables, Φ and Π, the regularity condition $\phi_r(0,t) = 0$ translates into

$$\begin{aligned} \Phi(0,t) &= 0 && (134) \\ \Pi_r(0,t) &= 0 && (135) \end{aligned}$$

If we demand that the gravitational field is always very weak at $r = r_{\max}$—so that, in particular, $\alpha(r_{\max},t) \sim a(r_{\max},t) \sim 1$—the outgoing radiation conditions (128-129) derived for the flat-space case will work well.

In contrast to less symmetric situations (and, again, partly because of our specific choice of coordinates), the "initial value problem" for our EMKG model is quite simple. We can take as our freely specifiable data initial profiles for Φ and Π:

$$\Phi(r,0) = \Phi_0(r) \tag{136}$$

$$\Pi(r,0) = \Pi_0(r) \tag{137}$$

after which the initial values, $\alpha(r,0)$ and $a(r,0)$, of the metric functions are fixed by the slicing and Hamiltonian constraints. In practice, if we are interested in investigating collapse and black hole formation, it is often convenient to specify a pulse-like profile, $\phi_0(r)$, for $\phi(r,0)$, assume that the pulse is propagating in flat space-time at $t=0$, and, finally, that it is purely in-going. Typically then, in terms of the initial pulse profile, $\phi_0(r)$, we will have

$$\Phi(r,0) = \phi_0' - \frac{\phi_0}{r} \tag{138}$$

$$\Pi(r,0) = -\phi_0' \tag{139}$$

3.4 $SU(2)$ Yang-Mills

A complete derivation of the equations of motion for this case is beyond the scope of these notes, and the interested reader is referred to [14] and references therein for a fairly complete description of spherically-symmetric Einstein-$SU(2)$ Yang Mills. The specific model considered here was first studied in the critical-phenomena context in [11], and is notable for exhibiting both of the general types of critical behaviour which have been seen to date.

The most immediate route to the field equations is to posit the existence of a spherically symmetric "Yang-Mills-field", $W(r,t)$ with an (effective) Lagrangian scalar:

$$L_W = -\left(\frac{\nabla^\mu W\,\nabla_\mu W}{r^2} + \frac{1}{2}\frac{\left(1-W^2\right)^2}{r^4}\right) \tag{140}$$

$W(r,t)$ is actually a specific component of the spherically-symmetric Yang-Mills connection, and the non-trivial fact that a single field can capture all the $SU(2)$ dynamics, follows from a specific *ansatz* (no electric charge density) as well as a special gauge choices. We again adopt polar/areal coordinates so that the metric is

$$ds^2 = -\alpha^2 dt^2 + a^2 dr^2 + r^s d\Omega^2$$

In the (massless, non self-interacting) scalar case, if $\phi(r,t)$ is a solution of the field equations, then so is $\phi(r,t)+\kappa$, for arbitrary constant κ.

In the YM case, the presence of the "self-interaction term"

$$\frac{1}{2}\frac{\left(1-W^2\right)^2}{r^4}$$

in the Lagrangian means that W must be in specific "vacuum" states at $r = 0$ and $r = \infty$ in order to have both a regular origin, and a finite-energy (i.e. asymptotically flat) configuration. Specifically we must have:

$$\lim_{r\to 0} W(r,t) = \pm 1 + r^2 W_2(t) + O(r^4) \tag{141}$$

$$\lim_{r\to\infty} W(r,t) = \pm 1 \tag{142}$$

The existence of these two distinct vacuua $W = +1$ and $W = -1$ is responsible for some of the interesting phenomenology in the model. In particular, as discovered about a decade ago by Bartnik and Mckinnon [12], it turns out that one can find static "kink" or "soliton"-like solutions to the EYM equations. These static solutions, $W_n(r)$, can be labelled by the positive integer n which counts the number of zero-crossings in the profile of W. The "lowest lying" member of the family, $W_1(r)$ is shown in Figure 10.

Soon after the discovery of the Bartnik-Mckinnon solutions, it was found that *all* of the $W_n(r)$ are *unstable* to perturbations, and in fact, W_n has precisely n unstable modes (within the context of the original *ansatz*). In particular, this means that the solution displayed in Figure 10 has *one* unstable mode, and thus hints that it may be relevant for critical collapse. This indeed turns out to be the case.

Following a development which precisely parallels the EMKG case, we again introduce auxiliary variables Φ_W and Π_W:

$$\Phi_w(r,t) \equiv \partial_r W(r,t) \tag{143}$$

$$\Pi_w(r,t) \equiv \frac{a}{\alpha}\partial_t W(r,t) \tag{144}$$

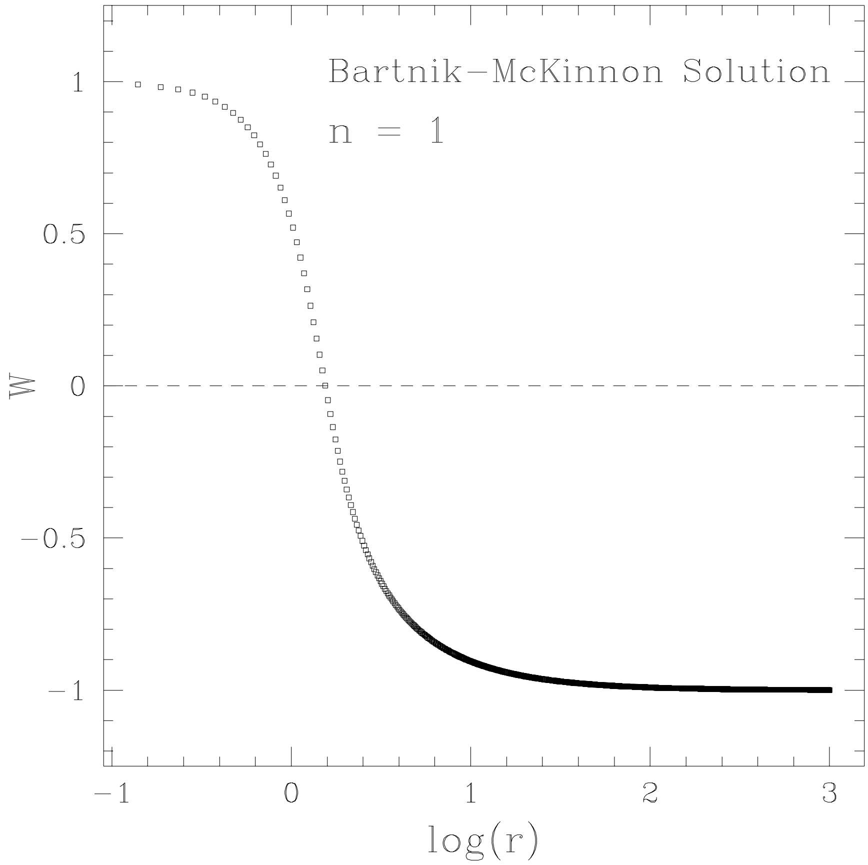

Figure 10. The static, $n = 1$ Bartnik-Mckinnon solution of the EYM equations in spherical symmetry.

We can then derive the following equations of motion for the EYM model [11]:

$$\frac{\partial \Phi_w}{\partial t} = \frac{\partial}{\partial r}\left(\frac{\alpha}{a}\Pi_w\right) \tag{145}$$

$$\frac{\partial \Pi_w}{\partial t} = \frac{\partial}{\partial r}\left(\frac{\alpha}{a}\Phi_w\right) + \alpha a \frac{W\left(1-W^2\right)}{r^2} \tag{146}$$

$$\frac{a'}{a} + \frac{a^2-1}{2r} - \frac{\Phi_w^2+\Pi_w^2}{r} - a^2\frac{(1-W^2)^2}{2r^3} = 0 \tag{147}$$

$$\frac{\alpha'}{\alpha} - \frac{a^2-1}{2r} - \frac{\Phi_w^2+\Pi_w^2}{r} + a^2\frac{(1-W^2)^2}{2r^3} = 0 \tag{148}$$

$$W(r,t) = \pm 1 + \int_0^r \Phi_w(\tilde{r},t)\, d\tilde{r} \tag{149}$$

The regularity and boundary conditions for the geometric variables in this case are identical to those for the EMKG model. For the matter variables, we have

$$\Phi_w(0,t) = \Pi_w(0,t) \tag{150}$$

and, by convention, we will typically take

$$W(0,t) = +1 \tag{151}$$

Asymptotic outgoing radiation conditions are even easier to impose here than for the scalar field: the self-interaction term falls off extremely rapidly, and then W propagates like a free field in *cartesian*, not spherical coordinates, i.e. we have

$$\lim_{r\to\infty} W(r,t) = W(r-t) \tag{152}$$

from which it immediately follows that

$$\begin{aligned} \lim_{r\to\infty} \; \partial_t \Phi_w(r,t) + \partial_r \Phi_w(r,t) &= 0 && (153) \\ \lim_{r\to\infty} \; \partial_t \Pi_w(r,t) + \partial_r \Pi_w(r,t) &= 0 && (154) \end{aligned}$$

Similarly, it is straightforward to specify in-coming initial data, by for example, giving an initial profile, $W_0(r)$, for $W(r,0)$, then setting

$$\Phi_w(r,0) = W_0'(r) \tag{155}$$

$$\Pi_w(r,0) = -W_0'(r) \tag{156}$$

Again, once the matter variables have been given in this fashion, the initial values of the metric components α and a are fixed from (147-148).

3.5 Parameter Space Surveys and the Black Hole Threshold

We now return attention to the EMKG model, and discuss the issues of parametrized families of solutions and the process of tuning to the black hole threshold. As we have seen, we can take the initial profiles $\Phi(r,0)$ and $\Pi(r,0)$ to be our initial data. We now want to consider *families* of initial data which will generate *families* of spacetimes describing the (non-linear) propagation (i.e. implosion and explosion) of self-gravitating, spherically symmetric scalar waves. Thus, our initial data will take the form

$$\Phi(r,0;p) \qquad \Pi(r,0;p)$$

For example, we could generate such data by taking the following form for the initial profile of ϕ

$$\phi(r,0) = \phi_0\, r^3 \, \exp\left(- \left(\frac{r - r_0}{\Delta} \right)^q \right) \tag{157}$$

and demanding that the waves be (almost) purely ingoing at $t = 0$. Note that in this example of a "Gaussian" initial pulse, any of ϕ_0, r_0, Δ or q could be used as the family parameter—for concreteness we will take the overall amplitude factor, ϕ_0, to be the control parameter.

As we saw in Section 3.3.1, in the weak-field limit—which in the current case is $\phi_0 \to 0$—we can solve the dynamics exactly. An initially imploding pulse of scalar radiation collapses towards the origin (roughly growing in amplitude as r^{-1} as it does so), passes through $r = 0$, and then emerges with the same shape, but inverted (i.e. $\phi \to -\phi$). The key point is that, in the long-time limit, all of the scalar radiation has dispersed to large radii, leaving (almost) completely flat spacetime in the interior.

As we increase the amplitude parameter ϕ_0, the ensuing gravitationally-induced self-interaction of the scalar field gets stronger at stronger (particularly during the time of "maximum concentration near the origin"), and eventually we reach a point where a black hole forms in the calculation.

Now, it should be noted that polar/areal coordinates share with the usual Schwarzschild coordinates the property that they *cannot* penetrate apparent horizons (instantaneously marginally-trapped surfaces—in our case, surfaces of constant r and t on which the divergence of outgoing null geodesics vanishes), and thus, for all practical purposes, they *cannot* penetrate event horizons. This is simply the statement that one never actually "sees" a black hole form in these coordinates—instead, when black hole formation is imminent, the lapse function quickly "collapses" in the central region, and the coordinate system starts to become singular near what *would be* the radius of the black hole. Nonetheless, precisely because the behaviour is qualitatively different for the case where a black hole *would* form, we have no problem either in detecting black hole formation, nor in getting a fairly accurate estimate of the final black hole mass in polar/areal coordinates. We can, for example, simply monitor the quantity

$$\frac{2m(r,t)}{r}$$

where $m(r,t)$ is the mass aspect function defined by (100). In a collapse calculation which produces a black hole, we will see, at some point in the evolution, that $2m/r$ rapidly asymptotes to 1 at some specific radius, $r = R_{\mathrm{BH}}$. R_{BH} is precisely the radius of the black hole and, in addition, the mass of the hole is then immediately given by the expected expression

$$M_{\mathrm{BH}} = \frac{R_{\mathrm{BH}}}{2}$$

A key point here is that, once one has an code to solve a FD version of the EMKG model, it is *not* a difficult task to construct families of solutions $\mathcal{S}[p]$ which "interpolate" between complete dispersal on one hand, and black hole formation on the other—just about any family one can think of will do the trick. Generically, one finds that black hole formation sets in at some specifc value $p^\star$, and, once an initial "bracket" $[p_{\mathrm{lo}}, p_{\mathrm{hi}}]$ has been found (so that a black hole forms for $p = p_{\mathrm{hi}}$, but not for $p = p_{\mathrm{lo}}$), it is an easy matter to use a bisection search, for example, to locate the critical value $p^\star$ to essentially arbitrary precision.

Once one has started thinking about interpolating families, and the black hole threshold, a particularly interesting question arises [16]:

> *What is the nature of the function $M_{\mathrm{BH}}(p)$ near and at the black hole threshold?*

In particular, for *generic* families, does black hole formation "turn on" at *finite* mass, or *infinitesimal* mass. With the benefit of hindsight, and viewing the mass as an order parameter we will refer to the two possibilities as Type I (finite mass) and Type II (infinitesimal mass) respectively. (The Type I/Type II terminology is motivated by loose analogy with statistical mechanical first and second order phase transitions where order parameters tend to be discontinuous/continuous respectively). Although, we now have examples of *both* types of transitions, it is the Type II critical solutions which have attracted most attention, not least since they provide a mechansim for producing arbitrarily small black holes, and arbitrarily large curvatures which are visible by distant observers.

3.6 Type II Critical Phenomena

Here I will simply summarize some key empirical results, to go along with the videos which I will show, referring the interested reader to the literature (especially Gundlach's review article [17]) for more details.

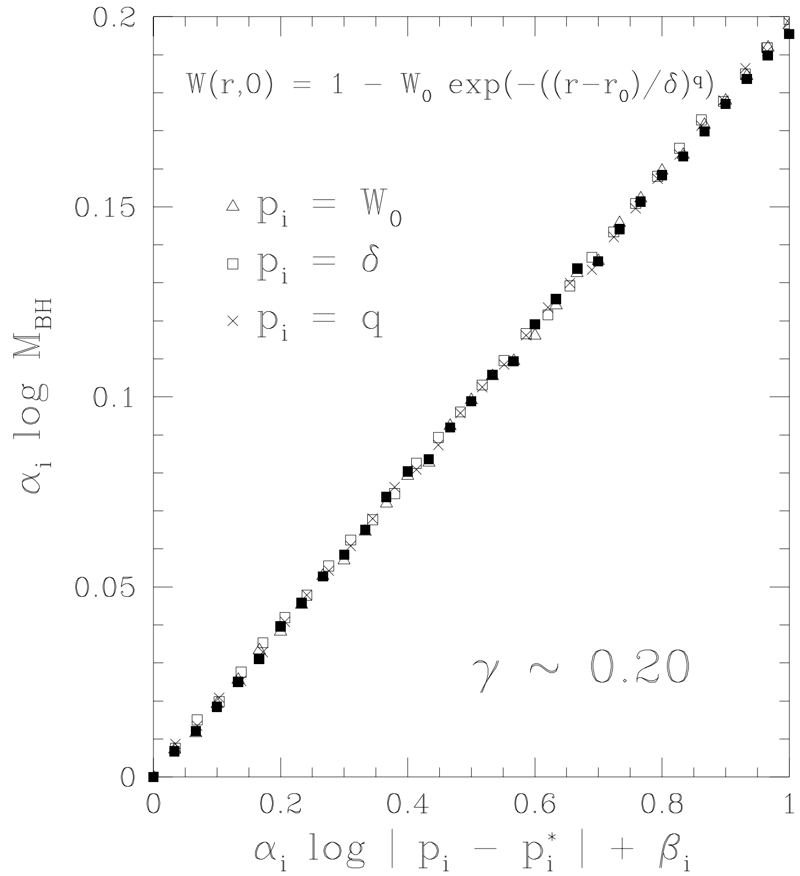

Figure 11. Illustration of critical mass-scaling for $SU(2)$Einstein Yang-Mills collapse. Each marker type corresponds to a different family of super-critical computations. For each family, constants α_i and β_i are chosen to unit-normalize the x-range and place the first data point (smallest black hole) at the origin. For all plots, the lesat squares fit for the slope, γ, is 0.20, with an estimated uncertainty of a few percent.

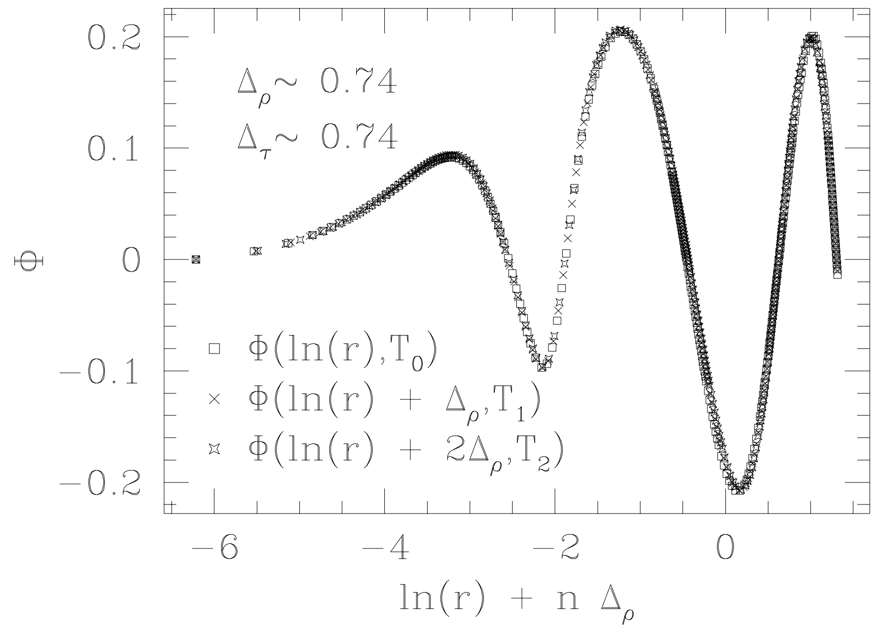

Figure 12. Illustration of scale-periodicity (discrete self-similarity) of the Type II solution in EYM collapse. This plot shows the superposition of a near-critical profile of Φ_w (at a particular time), with the first two echoes which subsequently develop. The echoing exponent Δ measures the amount by which the solution is re-scaled after each echo; i.e. here, the echoing factor is $e^{\Delta} \approx e^{0.74} \approx 2$.

By definition, Type II critical behaviour is characterized by infinitesimally-massed black holes at threshold. We see Type II behaviour in both the EMKG and EYM model, as well as in sufficiently relativistic collpase of perfect fluids [19, 20], and in the vacuum, *axisymmetric* collapse of gravitational radiation [18]. In all cases, power-law scaling of the black hole mass as a function of family parameter is observed:

$$M_{\mathrm{BH}} = c_f |p - p^\star|^\gamma \tag{158}$$

where c_f is a family-dependent normalization factor, but γ is a universal (within a single collapse model) scaling exponent. I will not dwell on any of the numerology here, suffice it to say that, despite some early suspicions to the contrary, γ turns out to be fairly sensitive on the specific type of collapse being studied (in particular, in the case of perfect fluid collapse, γ can range from values near 0 to values near 1, depending on the equation of state.) Typical empirical data illustrating mass scaling (for the case of EYM collapse) is shown in Figure 11.

Another feature of Type II near-critical evolution which emerges from the empirical studies is *self-similarity* of the critical solutions. In fact, for many cases (EMKG, EYM, axisymmetric gravitational radiation), the self-similarity is *discrete* rather that *continuous*. Heuristically, the self-similarity can be "explained" as follows: Roughly speaking, the critical phenomena which arise at black hole threshold result from *competition* in the system between the "matter kinetic energy" which wants to disperse to infinity, and the "gravitational potential energy", which wants to trap the matter-energy in a black hole. As we tune towards threshold, we find that we can "balance" the competitive effects, so that the strong-field dynamics persists to smaller and smaller spatial scales (i.e. closer and closer to $r = 0$), and unfold on faster and faster temporal scales. Precisely at threshold, the strong-field dynamics propagates down to an arbitrarily small scale, resulting in singularity formation at $r = 0$ at some *finite* (coordinate) time $t^\star$. Once we have located such a critical solution (in polar/areal coordinates for example), we can introduce new coordinates which are specially adapted to the scaling symmetry of the critical solution. In particular, if we first relabel our $t =$ constant slices:

$$t \;\rightarrow\; T_0(t) = \int_0^t \alpha(0, \tilde{t}) d\tilde{t}$$

then the new coordinate T, defined by

$$T = T_0^\star - T_0 \tag{159}$$

is the natural temporal analogue of the areal coordinate, r.

If we then transform into *logarithmic* coordinates, τ and ρ defined by

$$\begin{aligned}
\tau &\equiv \ln\left(T_0^\star - T_0\right) = \ln T && (160)\\
\rho &\equiv \ln r && (161)
\end{aligned}$$

we can more precisely express the property of discrete self-similarity. Schematically denoting the critical solution as $Z^\star(\rho,\tau)$ (so that Z represents any and all "scale invariant" variables, such as ϕ, $r\phi_r$, m/r, dm/dr, $\cdots$), discrete self-similarity means that we have

$$Z^\star(\rho,\tau) \sim Z^\star\left(\rho \pm n\Delta, \tau \pm n\Delta\right) \qquad (162)$$

Here Δ is another *universal* (i.e. does not depend on which family generates the initial solution) exponent, which measures the amount of spatial and temporal re-scaling which occurs per echo.

Some Type II critical solutions—most notably those which arise in perfect fluid collapse—exhibit the more familiar *continuous* self-similarity (CSS). It remains an open question precisely *why* some models exhibit a DSS critical solution rather than a CSS one, particularly since there are examples where the tuning of model parameters results in a transition from the CSS to DSS case [21].

3.7 Type I Critical Phenomena

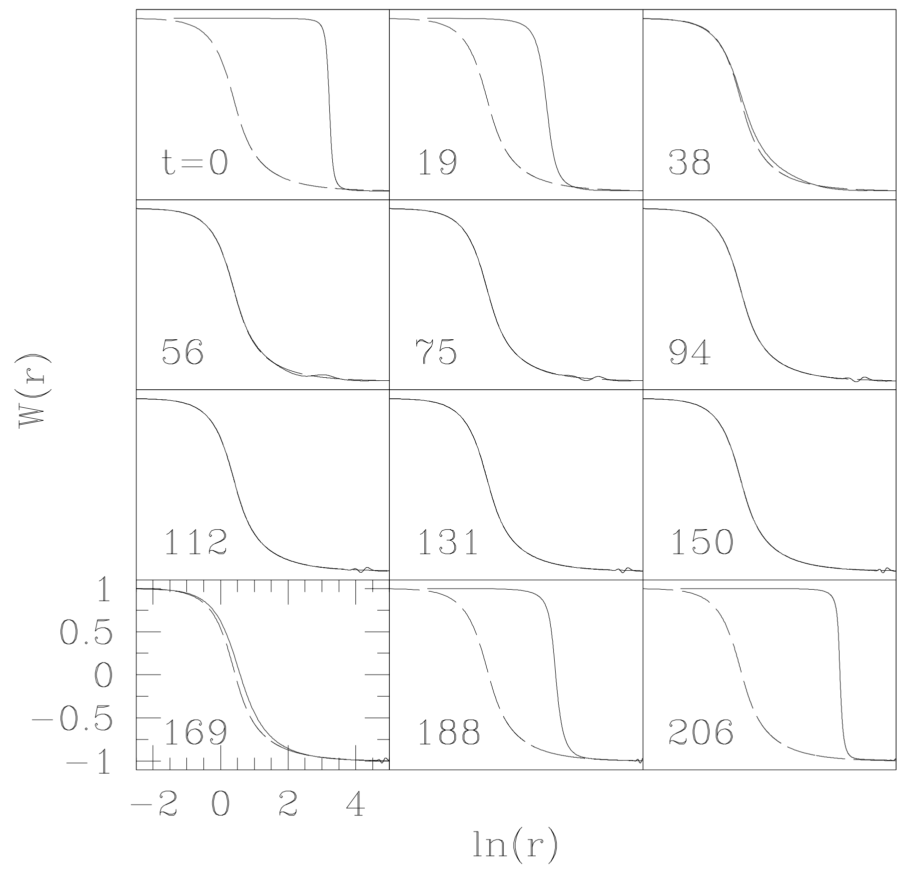

Figure 13. Marginally sub-critical Type I evolution in EYM collapse. Here, we plot the dynamical evolution of $W(r,t)$ (solid line) and superimpose the static Bartnik-Mckinnon configuration $W_1(r)$ (dashed line). Intially, the evolution is nearly linear and almost purely ingoing. When the pulse arrives at the center, it sheds off YM radiation, approaches $W_1(r)$ and stays near it for some time, and then disperses to infinity.

Whereas Type II critical behaviour is characterized by infinitesimal black hole mass at threshold, Type I transitions exhibit a "mass gap" at $p = p^\star$, so that for marginally super-critical evolutions, a black hole with *finite* mass is formed. To date, this type of transition has been observed in two models: EYM collapse [11] and the collapse of a *massive* Klein-Gordon field [22]. As with Type II collapse, the black-hole threshold is again characterized by a (locally) unique solution of the coupled field equations, but this time the solution is either *static* or *periodic*, rather than being self-similar. Empirically, a natural scaling law which emerges is the "lifetime" of the near-critical solution as a function of family parameter. We find that the lifetime, τ, scales like

$$\tau \sim -\sigma \ln |p - p^{\star}| \tag{163}$$

where σ is yet another "universal" scaling exponent. Figure 13. shows a typical near-critical evolution in the EYM model. Note how the *static* $n = 1$ Bartnik-Mckinnon solution, $W_1(r)$, acts as the "intermediate attractor" in this case. The evolution shown in the figure is slightly *sub*-critical; the "movie" for a slightly *super*-critical case would appear virtually identical, except that in the final stages, almost the entire central configuration would collapse, forming a finite-mass black hole.

3.8 Lab Problem 3

Download and compile the RNPL code `eym`. Use it to investigate Type I critical behaviour in the collapse of an $SU(2)$ Yang-Mills field in spherical symmetry. Additional information concerning this exercise will be available on-line. If time permits, download and compile the adaptive code `eymcnad` and use it to investigate the Type II behaviour in the model.

References

[1] Mitchell, A. R., and D. F. Griffiths, **The Finite Difference Method in Partial Differential Equations**, New York: Wiley (1980)

[2] Richtmeyer, R. D., and Morton, K. W., **Difference Methods for Initial-Value Problems**, New York: Interscience (1967)

[3] H.-O. Kreiss and J. Oliger, **Methods for the Approximate Solution of Time Dependent Problems**, GARP Publications Series No. 10, (1973)

[4] Gustatsson, B., H. Kreiss and J. Oliger, **Time-Dependent Problems and Difference Methods**, New York: Wiley (1995)

[5] Richardson, L. F., "The Approximate Arithmetical Solution by Finite Differences of Physical Problems involving Differential Equations, with an Application to the Stresses in a Masonry Dam", *Phil. Trans. Roy. Soc.*, **210**, (1910) 307–357.

[6] Anderson, E. em et al "Lapack Users' Guide - Release 2.0", (1994) `http://www.netlib.org/lapack/lug/lapack_lug.html`

[7] Choptuik, M. W., "A Study of Numerical Techniques for Radiative Problems in General Relativity", PhD Thesis, The University of British Columbia, (1986)

[8] Choptuik, M. W., "Consistency of Finite Difference Solutions of Einstein's Equations", *Phys. Rev.* **D 44** (1991), 3124–3135

[9] Choptuik, M. W., "Universality and Scaling in Gravitational Collapse of a Massless Scalar Field", *Phys. Rev. Lett.,* **70** (1993), 9–12

[10] Piran, T., "Numerical Codes for Cylindrical General Relativistic Systems", *J. Comp. Phys.*, **35** (1980), 254–283

[11] Choptuik, M. W., T. Chmaj and P. Bizon, "Critical Behaviour in Gravitational Collapse of a Yang-Mills Field", *Phys. Rev. Lett.,* **77** (1996), 424–427

[12] Bartnik, R. and J. Mckinnon, "Particle-like Solutions of the Einstein Yang-Mills Equations", *Phys. Rev. Lett.,* **61** (1988), 141–144

[13] Marsa, R. L. and M. W. Choptuik, "Black Hole–Scalar Field Interactions in Spherical Symmetry", *Phys. Rev.,* **D 54** (1996), 4929–4943

[14] Choptuik, M. W., E. W. Hirschmann and R. L. Marsa, "New Critical Behavior in Einstein-Yang-Mills Collapse", `gr-qc/9903081` (1999), submitted to *Phys. Rev. D*

[15] Bardeen, J. M., and Piran, T., "General Relativistic Axisymmetric Rotating Systems: Coordinates and Equations", *Phys. Rep.* **96**, (1983) 205-250.

[16] Christodoulou, D. (personal communication) (1987)

[17] Gundlach, C., "Critical Phenomena in Gravitational Collapse", *Adv. Theor. Math. Phys.* **2**, (1997) 1-49.

[18] Abrahams, A. M. and C. R. Evans, "Critical Behavior and Scaling in Vacuum Axisymmetric Gravitational Collapse", *Phys. Rev. Lett.*, **70** (1993) 2980-2983

[19] Evans, C. R. and J. S. Coleman, "Critical Phenomena and Self-similarity in the Gravitational Collapse of Radiation Fluid", *Phys. Rev. Lett.*, **72**, 1782-1785 (1994)

[20] Neilsen, D. W. and M.W. Choptuik, "Critical Phenomena in Perfect Fluids", `gr-qc/9812053` (1998)

[21] Liebling, S. L. and M. W. Choptuik, "Black Hole Criticality in the Brans-Dicke Model", *Phys. Rev. Lett.,* **77** (1996), 1424–1427.

[22] Brady, P. R., C. M. Chambers and S. M. C. V. Goncalves, "Phases of Massive Scalar Field Collapse", *Phys. Rev.,* **D56** (1997) 6057-6061.

[23] Marsa, R. L. and M. W.Choptuik, "The RNPL Reference Manual" & "The RNPL User's Guide"
`http://godel.ph.utexas.edu/Members/marsa/rnpl/refman/refman.html`

`http://godel.ph.utexas.edu/Members/marsa/rnpl/users_guide/users_guide.htm`