

Fundamental Issues of Numerical Relativity

Matthew W. Choptuik
CIAR Cosmology and Gravity Program
Dept. of Physics & Astronomy, UBC

IMA Numerical Relativity Workshop
Minneapolis, MN
June 26 2002

Research supported by NSERC and CIAR

Current UBC Group

Kevin Lai

Luis Lehner

Scott Noble

Inaki Olabarrieta

Roman Petryk

Frans Pretorius

Bruno Rousseau

Jason Ventrella

Hugo Villegas

Collaborators

Eric Hirshcman (BYU)

Steve Liebling (LIU, Southampton)

Outline

- Target Physics
- Nature of the Gravitational Interaction
- Some Guiding Principles
- The 3+1 Approach and Update Freedom
- Coordinates, Good and Bad
- Boundary Conditions
- Discretization Issues
- Computational Demands and Model Problems
- Convergence, Stability and Dissipation
- Resolution and Adaptive Mesh Refinement
- Validation and Error Analysis
- Implementation Issues
- Other Issues

SOMETHING TO CONTEMPLATE ...



Relativists' Units

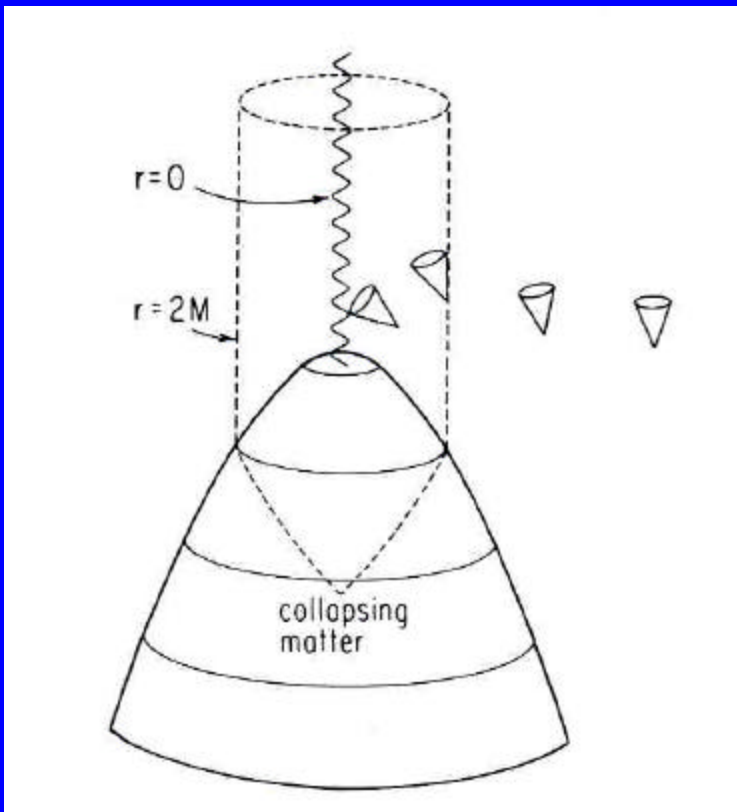
- Relativists like to choose units wherein Newton's gravitational constant, and the speed of light are both set to unity

$$G=1 \quad c=1$$

- Any physical quantity can then be expressed in terms of, e.g. a single mass, M , such as the (approximate) mass of a black hole being considered; for example

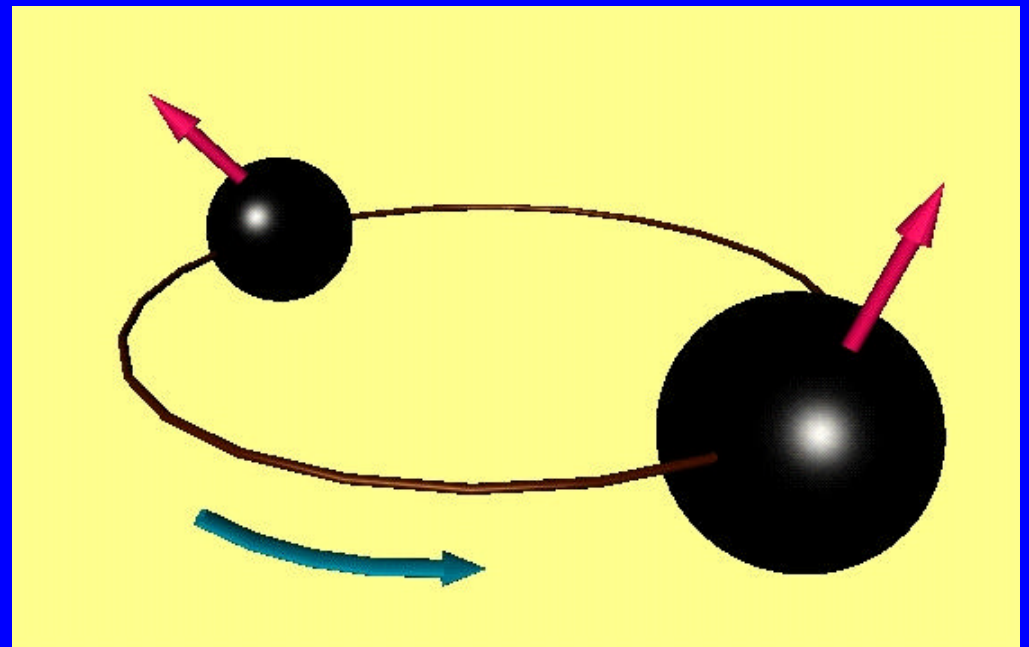
$$R_S = 2M \quad \rightarrow \quad R_S = 2 \frac{G}{c^2} M$$

Target Physics



Gravitational Collapse

Black Hole Binaries

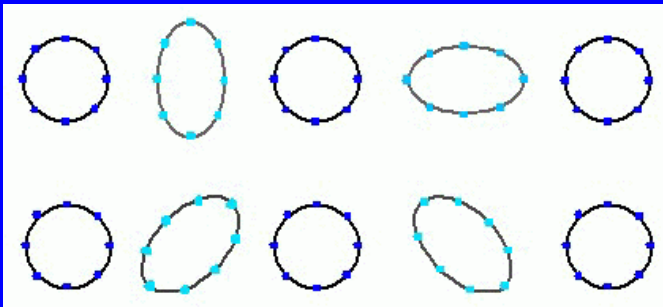


Gravitational Radiation

- **Gravitational waves:** “ripples” in the curvature of spacetime
- At least in weak field limit, very much analogous to electromagnetic radiation; propagate at speed of light, transverse, two polarizations, frequency set by dynamics of source

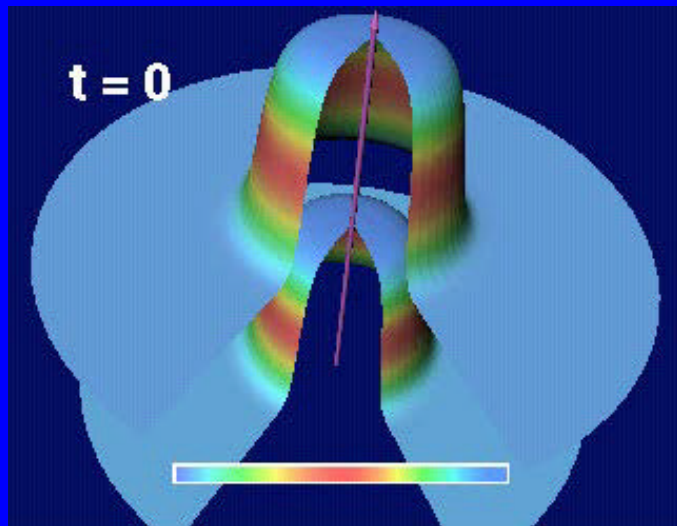


The Laser Interferometer Gravitational Wave Observatory (LIGO) installation near Hanford WA. Each interferometer arm is 4 km long. A similar instrument is located near Livingston LA (www.ligo.caltech.edu)



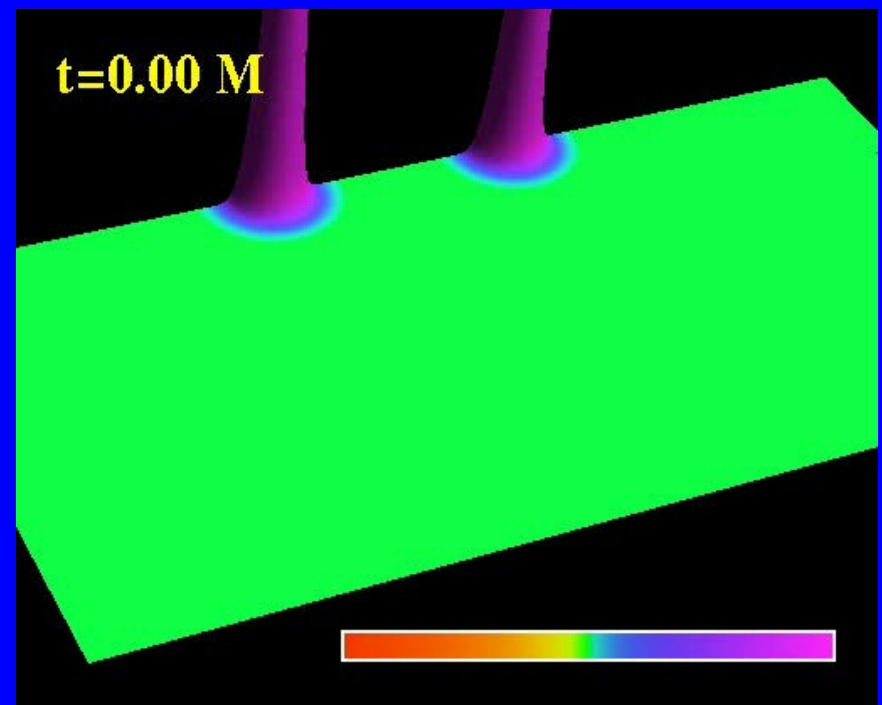
Cause periodic, quadrupolar variations in distance between freely falling objects (or induce strains in objects with interactions)

Target Physics



Gravitational Collapse

Black Hole Binaries

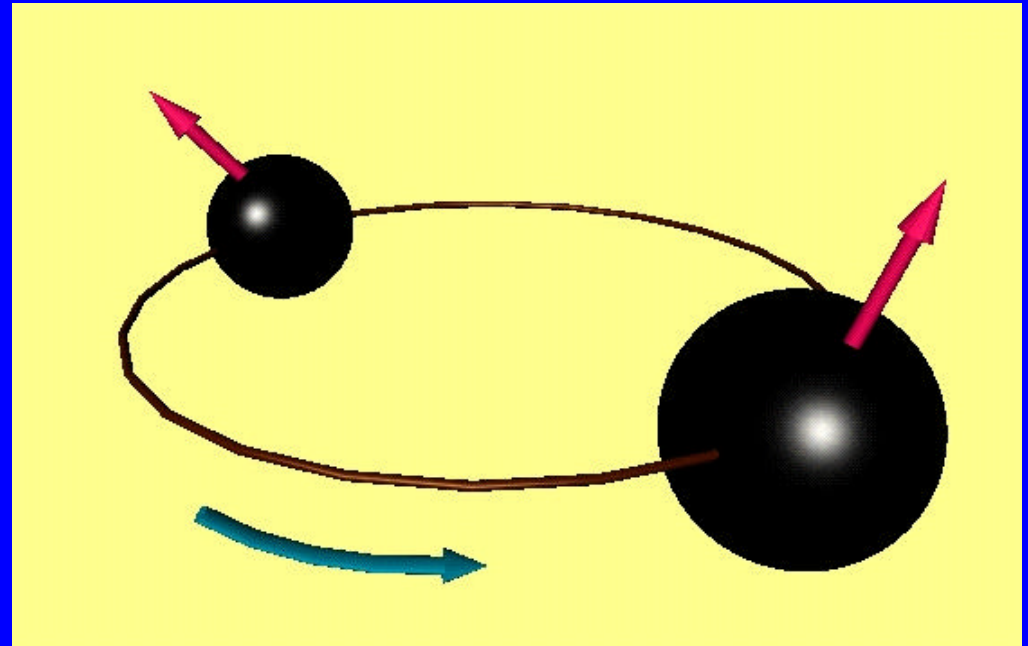


Assumptions / Restrictions (for this talk)

- Solution of Einstein field equations will proceed via “3+1” split; i.e. as a Cauchy problem
- Where specific EOM are discussed, will generally restrict to ADM form (clearly formulation of EOM is fundamental issue in NR, but has, and will be discussed extensively in other talks)
- Discussion of numerics will focus on finite difference techniques

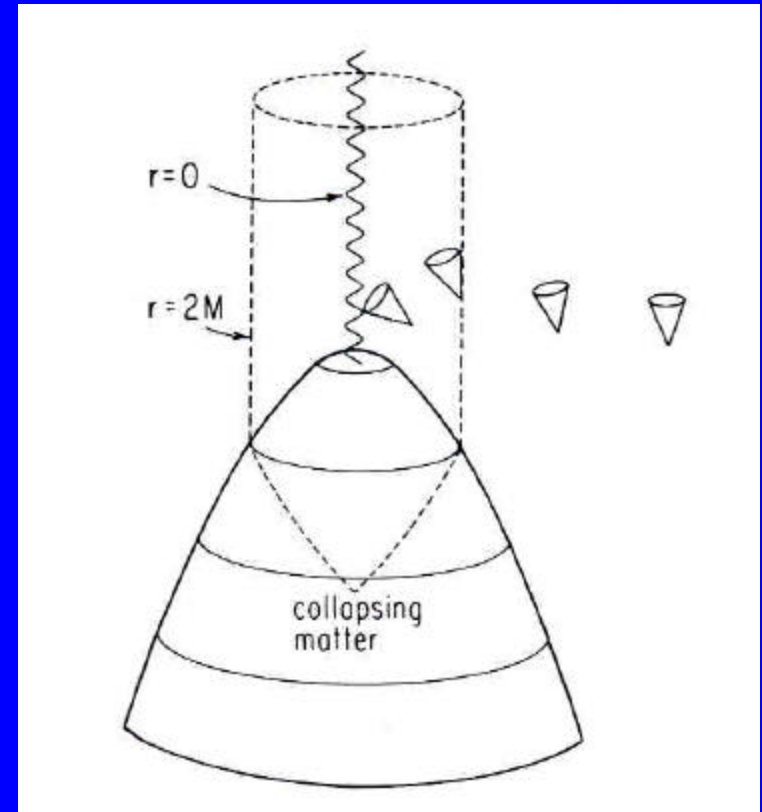
Nature of Relativistic Gravitation

- Like E & M, generic GR solution has both infinite-range “Coulomb” aspect, as well as radiation
- **Coulombic** character largely responsible for strong-field effects, singularity formation (gravity always attractive)
- **Radiative** component often quite weak, challenge is to hold strong field solution together while accurately representing and extracting radiation field



Black Holes and Physical Singularities

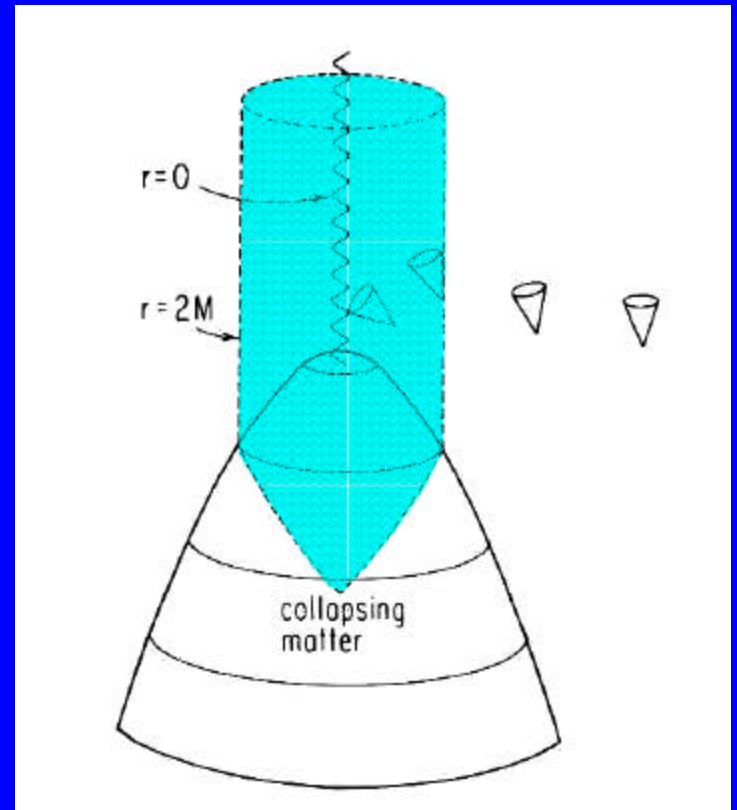
- **BLACK HOLE**: Region of spacetime from which no physical signal can escape
- During collapse of matter and/or radiation, BH forms when gravitational field becomes strong enough to “trap” light rays
- Surface of black hole is known as the *event horizon*
- *Singularities* (infinite, crushing gravitational forces) *inevitable inside black holes*



(From Wald, *General Relativity*, 1984)

Singularity Avoidance: Black Hole Excision

- To avoid singularity within black hole, *exclude interior of hole from computational domain (Unruh)*
- **Catch 22:** event horizon is *globally defined*, location unknown until complete spacetime geometry is in hand
- **Apparent horizon** functions as “instantaneous horizon” can be located at any instant of time
- Current conventional wisdom is that excision is crucial to our ability to accurately evolve black hole spacetimes for long time scales



Core Mathematical Problem

"They're just a set of PDEs, Larry"

- Solution of coupled, non-linear, time dependent PDEs in up to three space dimensions
 - Equations are principally of "evolution type", but elliptic equations **MUST** be solved at initial time, and possibly at subsequent times
- Boundary conditions are in principle set at infinity (asymptotic flatness)

The (Fundamentally) Good News

“But the fundamental differences between general relativity and classical field theories also bring major new computational issues.”

- However, from the physicist’s perspective, GR is a classical field theory; i.e. it is non-quantum mechanical
- In addition, there are reasons to expect that solutions of most immediate astrophysical interest will be SMOOTH (in the gravitational field) away from singularities
 - No reason to expect shocks
 - No reason to expect turbulence

The (Fundamentally) Good News

- *NR is an ideal field for simulation work*
 - Once resolution is adequate, can expect to achieve near-optimal convergence; can (approximately) bound amount of computational effort required to solve specific problem
 - Relatively small numbers of calculations ($O(1)$) can be reasonably expected to reveal salient physics (no ensemble issues as in statistical mechanical, or quantum mechanical systems)
 - Relatively high payoff in terms of physics extracted per flop; lack of access to non-linear regime via observation, experiment, “analytic” techniques further leverages this payoff

The (Fundamentally) Good News

- *NR is an ideal field for simulation work*
 - We believe we understand the underlying physics very well for certain problems (most notably black hole collisions); i.e. modeling per se is not a key issue
 - Note, however, that coming up with good initial data for scenarios such as black hole collisions is a challenging task

A Principle One Can Take to the Bank (Brandt's "Golden Rule")

The amount of computational work should be proportional to the amount of real physical changes in the computed system. Stalling numerical processes must be wrong --- *Achi Brandt*

The rule of simulation that I would like to have is that the number of computer elements required to simulate a large physical system is only to be proportional to the space-time volume of the physical system. I don't want to have an explosion --- *Richard Feynman*

Optimization / Efficiency

- NR, and 3D NR in particular, is **NOT** a mature discipline
- In many instances, including binary inspiral, are looking for **ANY** reasonable algorithms which scale linearly with physical process

$$W \propto k P$$

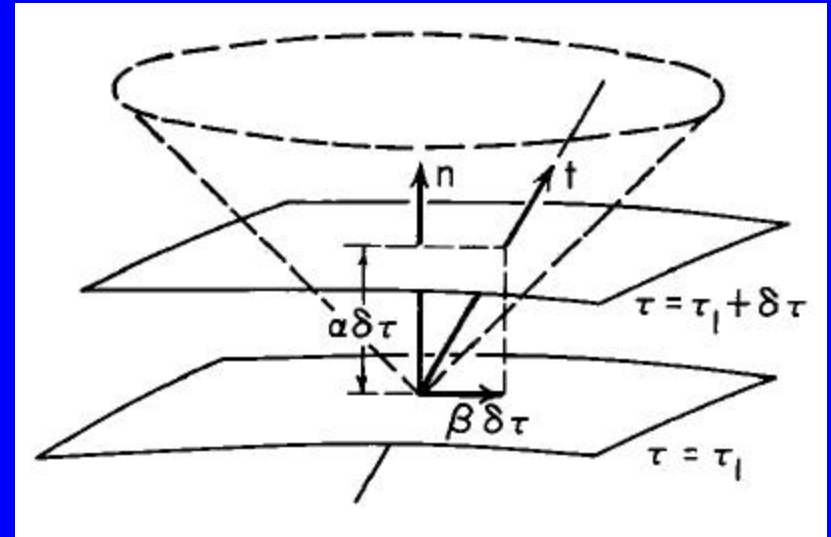
- Past experience suggests that once such methods are in hand, optimizations (both at the analytic and computational level) will evolve rapidly
- **CAVEAT:** Magnitude of computational problems means that optimizations **ARE** important, techniques which can reduce ***k*** significantly obviously worth pursuing

Flip Side of The Golden Rule (The No Free Lunch Principle)

- Required amount of computational work fundamentally set by **PHYSICS** being simulated
- Cleverness (specific formulation of EOM, coordinates, discretization, ...) without *a priori* information will only get one so far
- Those situations for which we most need NR are precisely those for which we have little *a priori* knowledge

The "3+1" split

- Slice spacetime into stack of 3-d *spacelike hypersurfaces*
- 4-geometry of spacetime (4-metric) becomes *time-evolution of 3-geometry* of initial hypersurface (dynamics of 3-metric)
- General covariance: *coordinates on spacetime arbitrary, don't affect physics*
- 4 degrees of coordinate-freedom ("gauge freedom")
- *Coordinates must be fixed: HOW?*



(J.W. York, Jr., 1979)

Dynamical Variables

$$g_{ij}(t, x^k), K_{ij}(t, x^k) \quad i, j, k = 1, 2, 3$$

Dynamical Degrees of Freedom

$$10 - 4 - 4 = 2$$

3+1 EOM (schematic)

- Fundamental variables: all functions of (x,y,z,t)

Dynamical: g_{ij} , K_{ij} Kinematical: \mathbf{a} , \mathbf{b}^i

- Evolution equations: (use 4 to 12)

$$\frac{\partial g_{ij}}{\partial t} = -2\mathbf{a}K_{ij} + D_i \mathbf{b}_j + D_j \mathbf{b}_i$$

$$\frac{\partial K_{ij}}{\partial t} = L_b K_{ij} - D_i D_j \mathbf{a} + \mathbf{a}(R_{ij} - 2K_{ik} K^k_j + K_{ij} K)$$

3+1 EOM (schematic)

- Constraint equations: (generally elliptic, use 0 to 4)

$$C_m[g_{ij}, K_{ij}] = 0 \quad m=0,1,2,3$$

- Coordinate conditions: (algebraic, elliptic, hyperbolic, need 4, can also be used to eliminate evolution equations)

$$F_m[\mathbf{a}, \mathbf{b}^i; g_{ij}, K_{ij}] = 0 \quad m=0,1,2,3$$

Update Possibilities

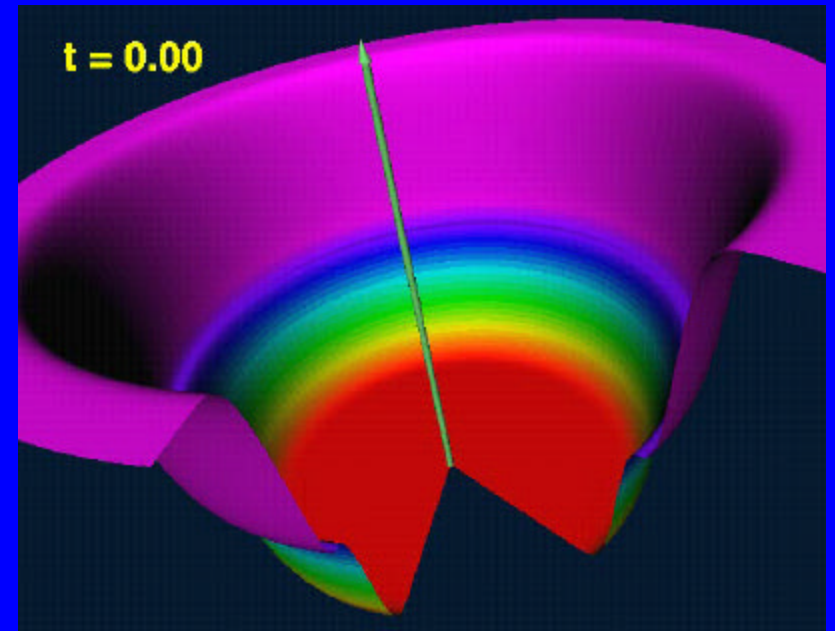
- Existence of constraints and coordinate conditions (“gauge freedom”) lead to huge variety of choices vis a vis use of “evolution equations”
- **Nomenclature:** fully constrained evolution, partially constrained evolution, free evolution etc.
- Situation further complicated by fact that can add multiples of constraints to evolution equations, thus changing the principal part of the evolution equations and impacting the well-posedness of the Cauchy problem

Update Possibilities

- Free evolution schemes currently preferred for 3-D work, apparently mainly due to difficulties in applying/solving constraint equations in generic circumstances
 - Note, however, that multi-grid technique provides means for solving coupled, non-linear, elliptic systems with $O(N)$ computational work, where N is the number of discrete unknowns
- Nature of update scheme has consequences for well-posedness, ability/ease to produce stable discretization, other issues ...

Coordinate Pathologies

- GR is a “background free” theory; arena in which physics plays out is THE dynamical quantity of interest
- No one tells us how to lay down coordinates; coordinate systems can become pathological dynamically, leading to actual or effective breakdown of scheme



$m(r, t)$

$$ds^2 = -\mathbf{a}^2(r, t) dt^2 + (1 - 2m(r, t)/r)^{-1} dr^2 + r^2 d\Omega^2$$

Building "Good" Coordinate Systems

- If possible, want one system to cover entire spacetime external to black holes
- System should be "well behaved" across black hole horizons
- Coordinates should expedite extraction of gravitational radiation, interpretation of physical result to extent possible
- Coordinates should result in computational economy if possible

Boundary Conditions (Outer)

- As is always the case when solving PDEs numerically, treatment of BC's is of paramount importance
- In "3+1" context, true boundary conditions are set at space-like infinity; in practice, if spatial computational domain is not compactified, will have to set approximate boundary conditions at some finite radius
- BC's imposed at finite radius must reflect a minimum of gravitational waves back into the interior of the computational domain; i.e. want outgoing, or absorbing boundary conditions, yet not all dynamical variables are purely radiative (i.e. have to worry about Coulombic pieces)

Boundary Conditions (Outer)

- Formulating good outer BCs has been notoriously difficult, particularly in 3-D context
- Much trial and error, some moderately successful *ad hoc* approaches have emerged; e.g. blending, which smoothly matches outgoing radiation conditions to Dirichlet conditions in a blending zone
- One area where emerging theory is particularly likely to have profound effect (Tiglio's, Lehner's talks)
- Also much room here for creative approaches (e.g. don't necessarily have to be using precise Einstein equations at distances greater than the radiation extraction radius)

Boundary Conditions (Inner)

- If excision is employed, additional interior boundaries are introduced
- In principle, if locus of excision boundary is luminal or superluminal, and if free evolution is employed, NO physical boundary conditions are required, “simply” impose evolution equations up to and including excision surfaces
- In practice, discretization methods may have to be modified (centred finite difference methods e.g.), has stability implications
- If update scheme involves elliptics, DO need interior boundary conditions; potential problems with kinematic quantities (lapse and shift)

Discretization Issues

- Finite difference methods have been dominant for historical reasons, as well as accessibility, flexibility
- Finite element methods have seen considerable success in context of initial value problem (PSU group), little impact on the evolution problem so far
- Spectral methods are being aggressively pursued by Cornell/Caltech group
- Within FD domain, what is the role, e.g., of advanced methods from computational fluid dynamics?

Computational Demands

- Again, am restricting attention to finite difference techniques

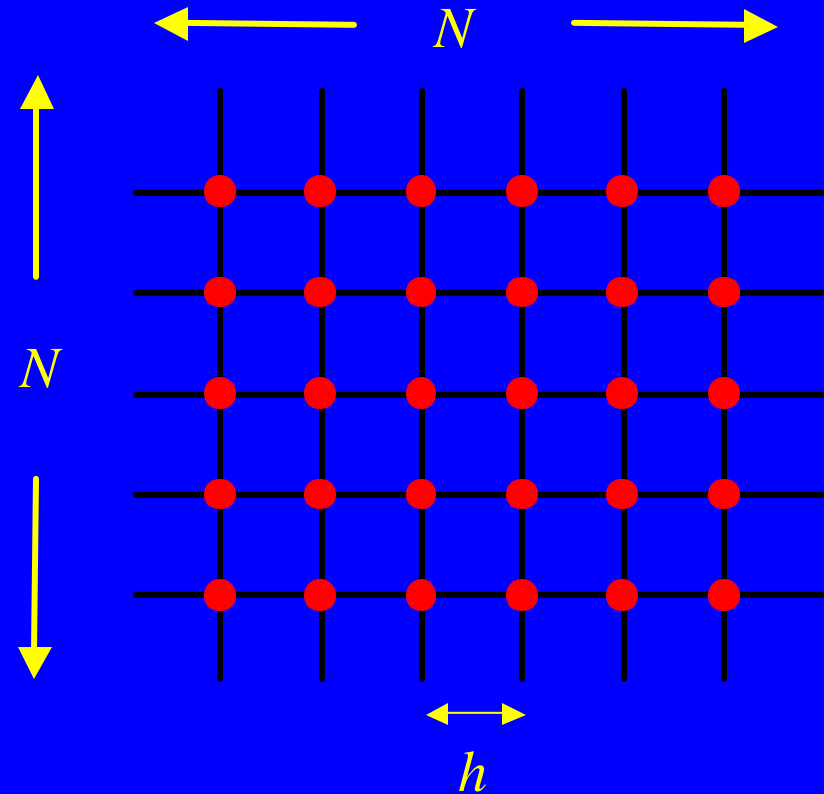
- D spatial dimensions:

$O(N^{D+1})$ CPU time

$O(N^D)$ Memory

- Minimum N for interesting problems:

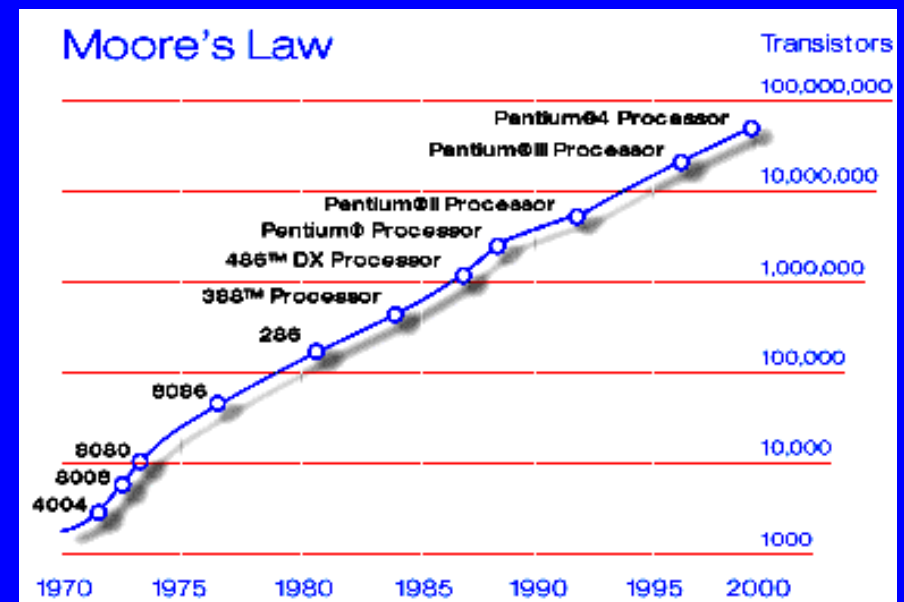
$$N \approx 1000$$



Computational Demands

- $D \rightarrow D+1$ thus requires of the order of 10 CPU-speed doublings, or about **15 years** of hardware evolution
- Estimate is pessimistic in some ways, but optimistic in others (e.g. increase in per-grid point complexity for higher D)

Crude estimates suggest calculations of general BH collisions require **many CPU days** on a **Teraflop/s, terabyte machine**



Computational Demands

- Simple minded analysis also ignores two key facts about simulation-based science
 1. When good solution method is not yet known, *want turn-around time on the order of hours* (at most) to make effective progress
 2. A few simulations generally won't cut it. Most interesting problems will have considerable parameter space to explore, which introduces additional effective dimensionality into the problem
- Lots of room for gains due to improved algorithms

Model Problems



- Reduce computational demands by imposing symmetry conditions which reduces effective D (spatial dimension)
 - $D=1$ **spherical symmetry** (symmetry under all rotations about a point)
 - $D=2$ **axisymmetry** (symmetry under rotations about an axis)
- Use for development of techniques for more realistic calculations
- *Models can reveal interesting physics which may not depend crucially on symmetry assumptions*

Model Problems

- Has been a tendency in field to move onto higher-D problems, before lower-D has been cleaned up
- E.g., very little work in axisymmetry at the current time, even though one can argue that only now do we routinely have access to enough computer power to do a reasonable job in 2-D
- Issue complicated by need/desire to simulate intrinsically 3-D phenomena, such as binary inspiral

Convergence and Stability

- Will assume that have well-posed system, and that physical solution does not blow up (implies excision for BH spacetimes)
- Have every reason to expect that some form of a Lax equivalence theorem applies, i.e. that convergence and stability are equivalent given consistency
- Not surprisingly, stability has emerged as THE KEY computational issue in NR
- Significantly complicated by issues related to formulation of equations of motion (particularly in context of free evolution schemes)

Convergence

- Assume finite difference scheme is characterized by unique discretization scale, h ; at least locally, can expect dynamical variables to admit Richardson expansion

$$u^h(x^m) = u(x^m) + h e_1(x^m) + h^2 e_2(x^m) + h^3 e_3(x^m) + \dots$$

- Provides basis for convergence testing, *intrinsic error estimation*, even increased accuracy via Richardson extrapolation when solutions are smooth enough
- NR has generally embraced the notion of convergence testing, particularly relative to other areas of computational astrophysics

Stability

- Schematic NR code

Read initial state

For NUM_STEPS do

For NUM_UPDATES & maybe until convergence

U[Grid Functions] -> Grid Functions

End for

End for

Write final state

- Key issue is spectrum of update operator(s) U ; heuristically, want eigenvalues within unit circle, but also want at least some of them as close as possible to unit circle to avoid undue dissipation of gravitational radiation

Stability and Dissipation

- One “well known” aspect of numerical analysis which NR has arguably not yet fully exploited
- *Historically, dissipative difference schemes played key role in analytic proofs of schemes’ stability (convergence)*
- For a wide variety of model problems, addition of explicit dissipation to otherwise non-dissipative schemes, has proven essential for successful numerical solution

Kreiss-Oliger Dissipation: Example

- Consider the simple “advection” equation

$$\partial_t u = \partial_x u$$

- Finite difference via $u_j^n \equiv u(n\Delta t, j\Delta x)$ and

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{1}{2} \left(\frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2\Delta x} + \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \right)$$

Kreiss-Oliger Dissipation: Example

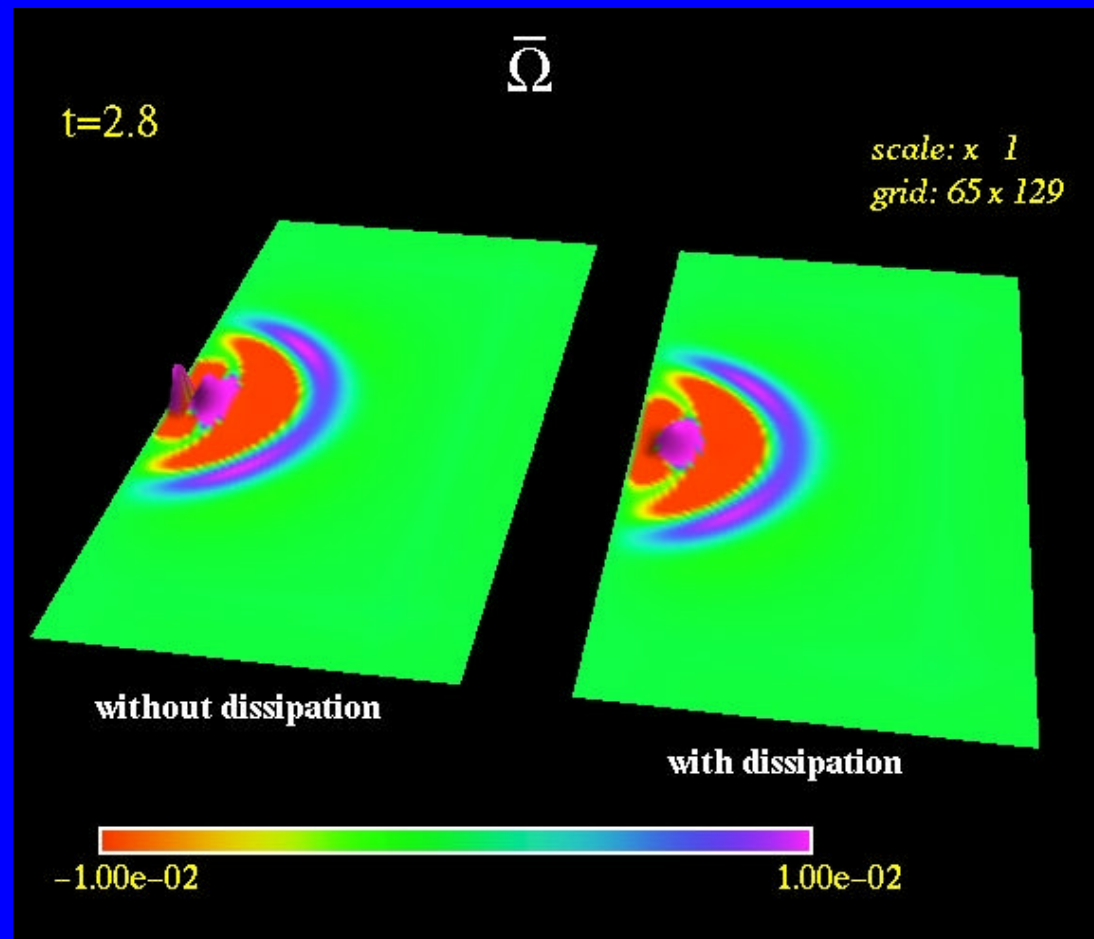
- Add “Kreiss-Oliger” dissipation via

$$u_j^n \rightarrow \left(1 - \frac{\mathbf{e}}{16} \Delta x^4 D^4 \right) u_j^n \quad (\text{LHS})$$

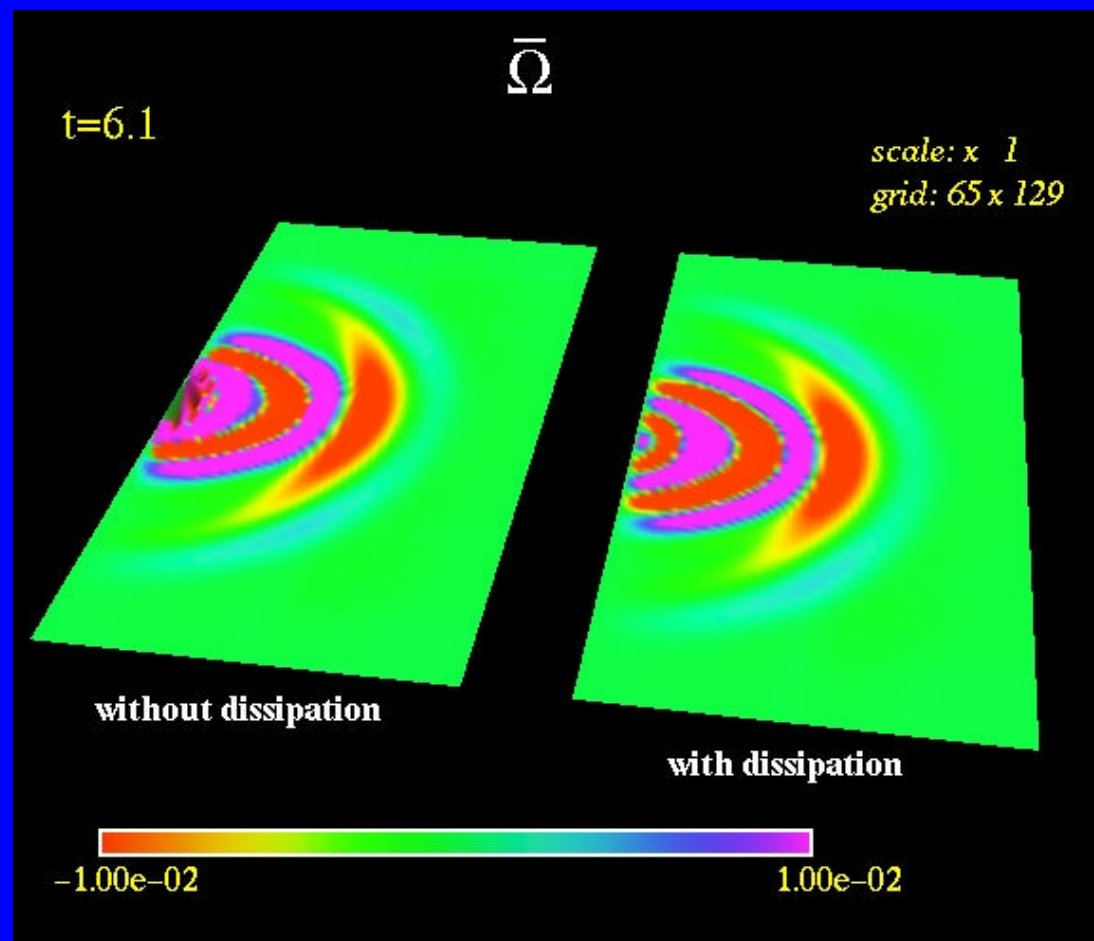
- Where $0 < \mathbf{e} < 1$ and

$$D^4 u_j^n \equiv \frac{u_{j+2}^n - 4u_{j+1}^n + 6u_j^n - 4u_{j-1}^n + u_{j-2}^n}{\Delta x^4}$$

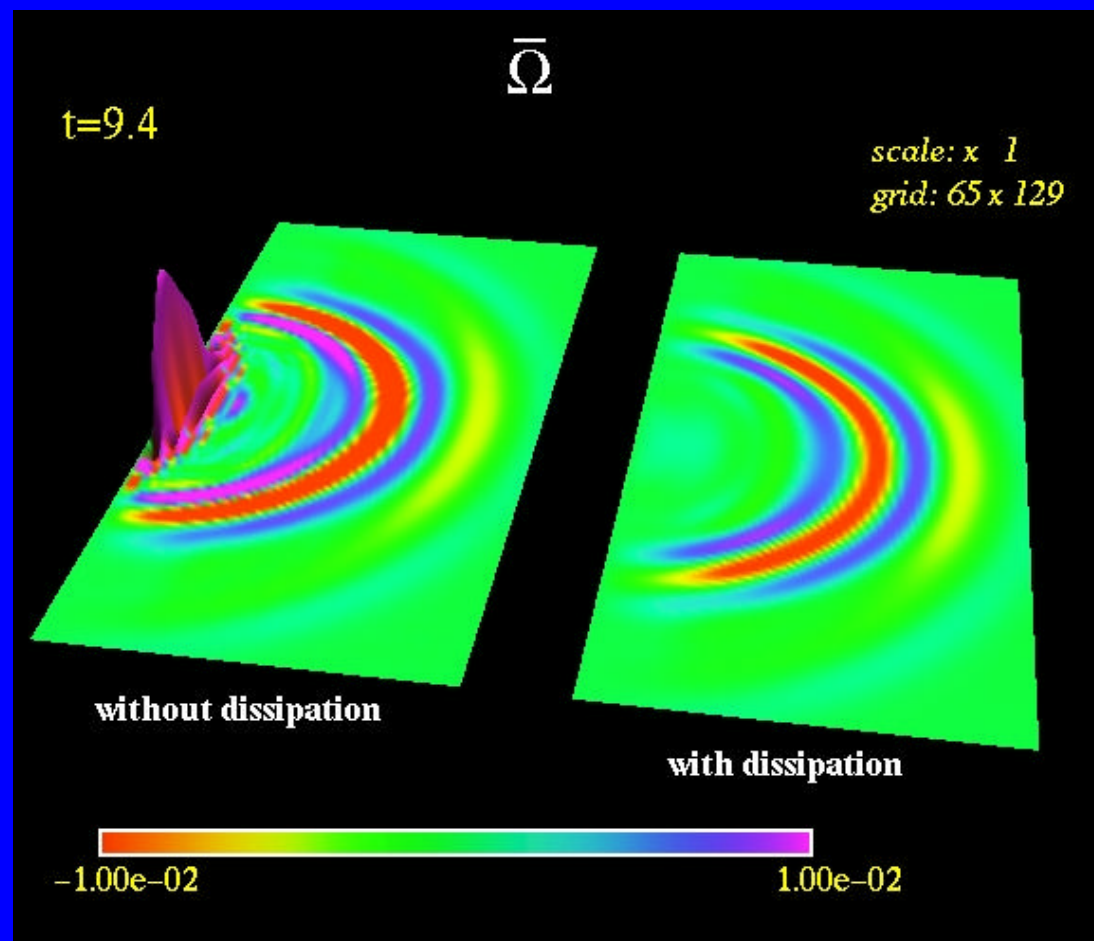
Effect of Dissipation
65 x 129 Grid
(Axisymmetric collapse of scalar field)



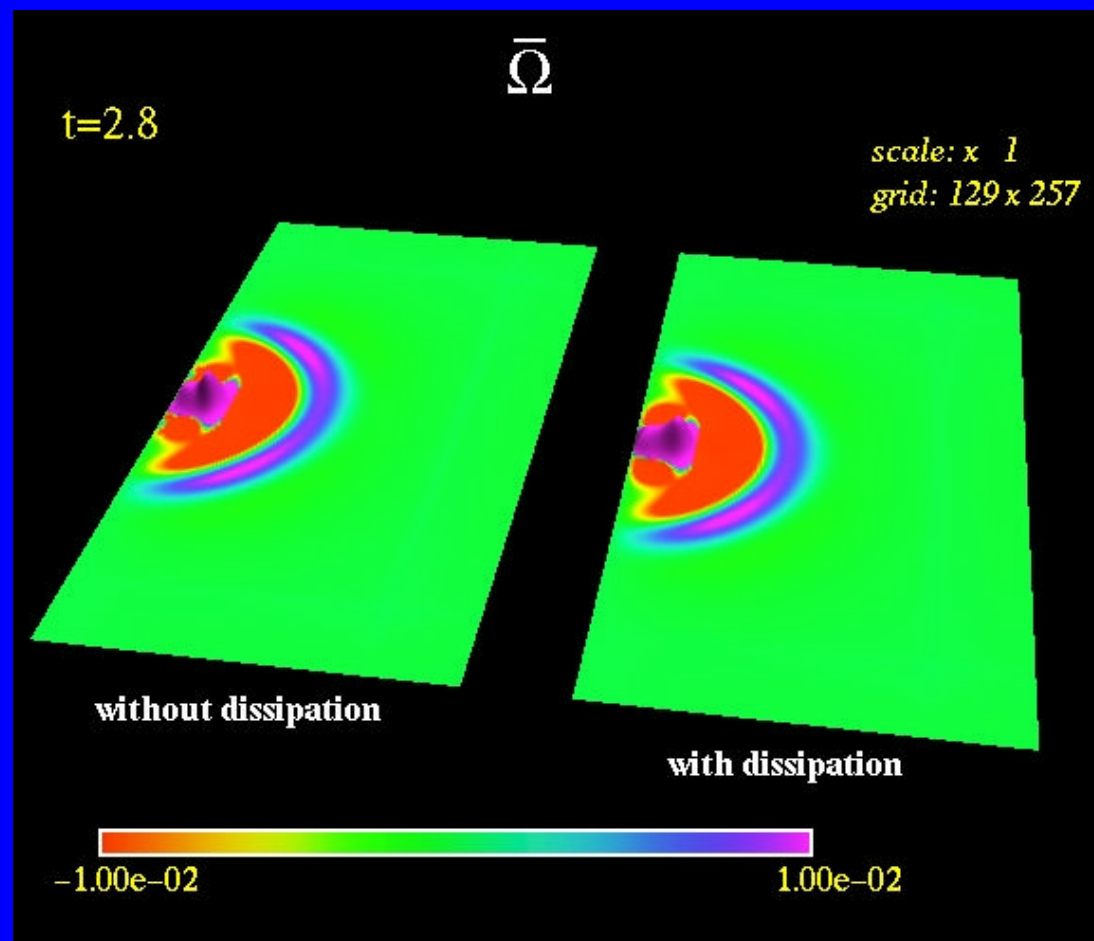
Effect of Dissipation 65 x 129 Grid



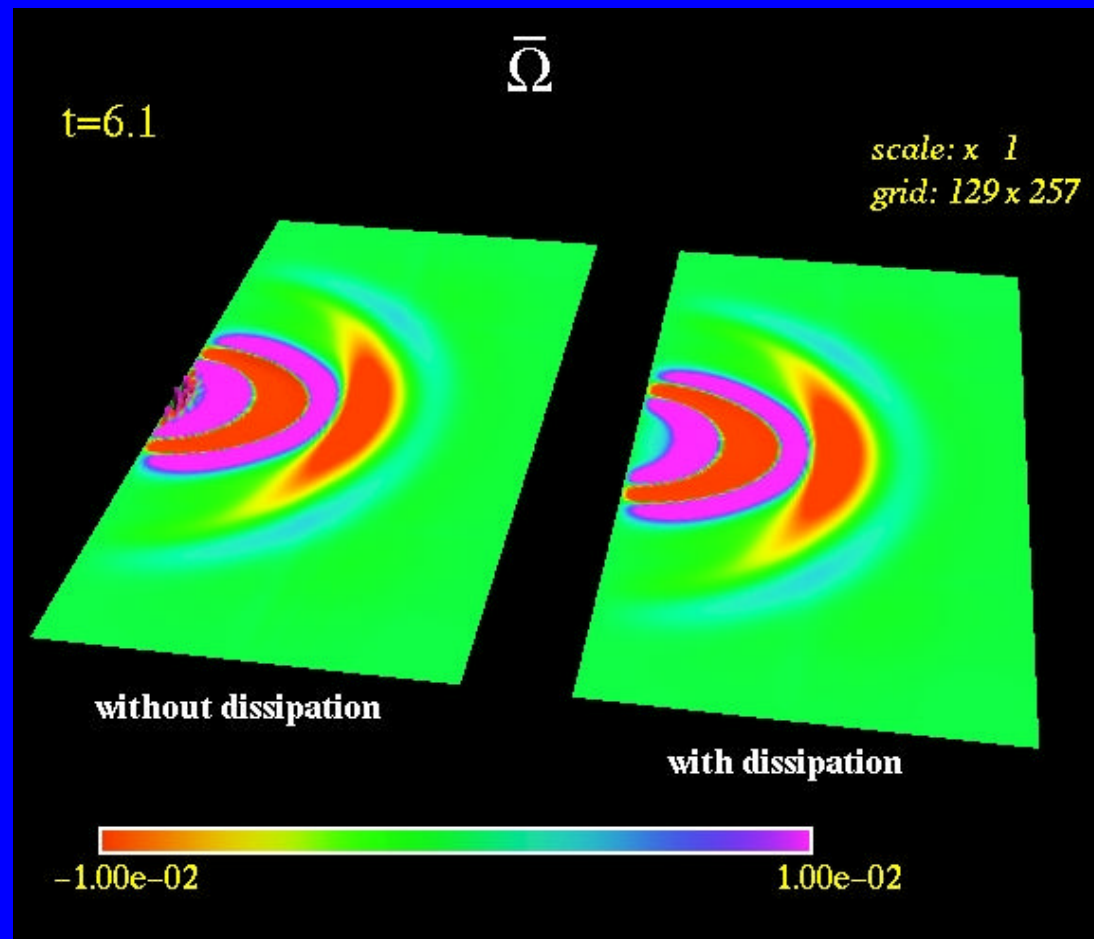
Effect of Dissipation 65 x 129 Grid



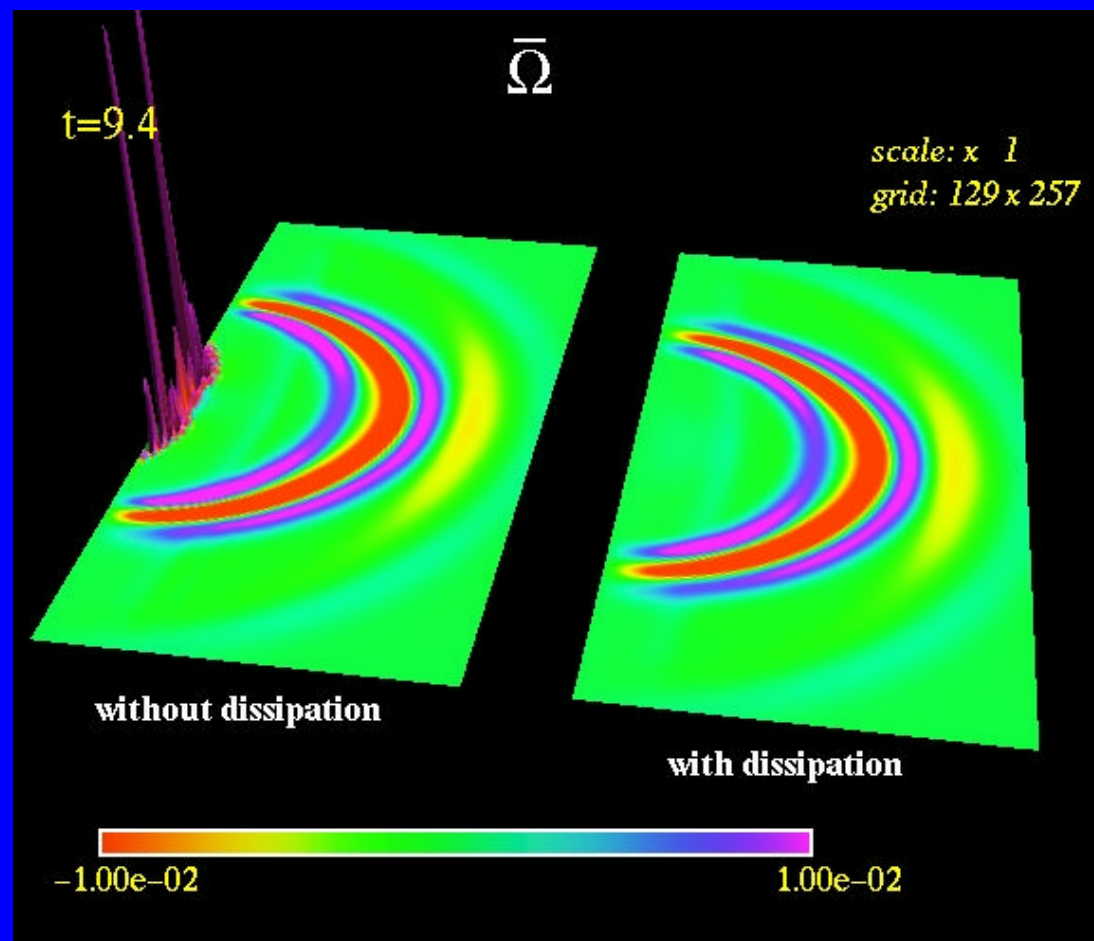
Effect of Dissipation 129 x 257 grid



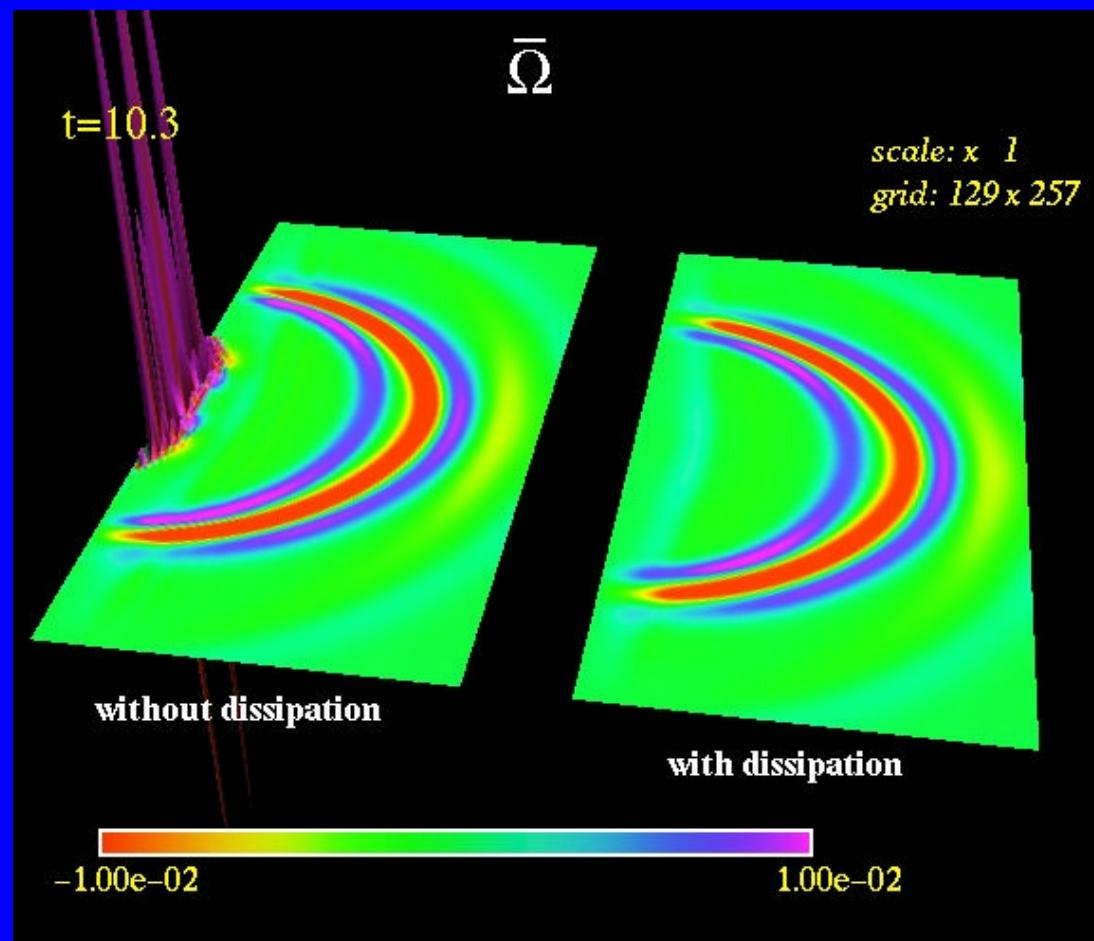
Effect of Dissipation 129 x 257 grid



Effect of Dissipation 129 x 257 grid



Effect of Dissipation 127 x 259 grid

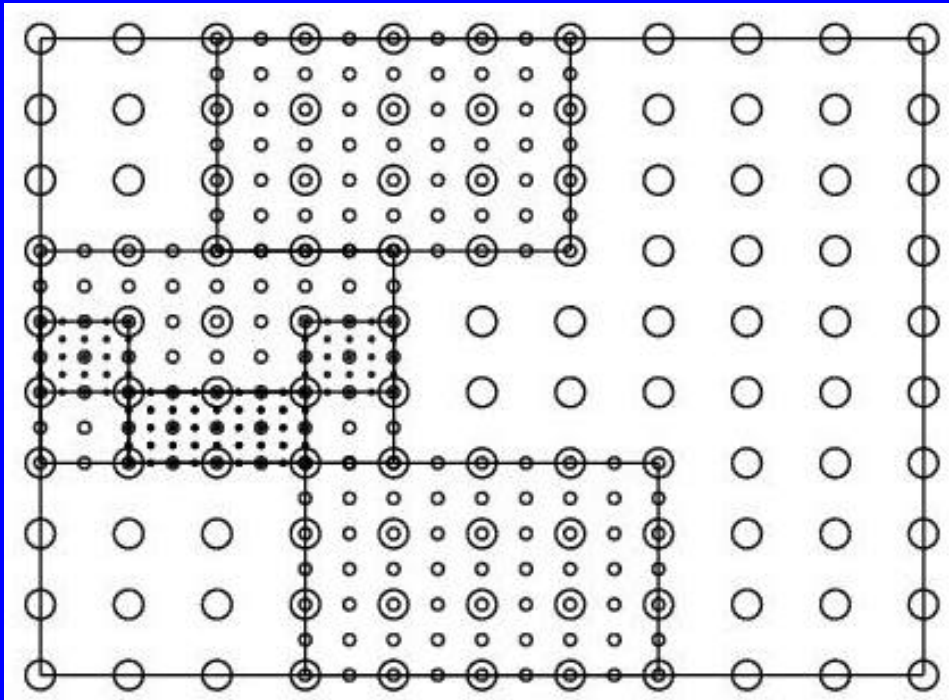


Resolution

- After stability, adequate resolution is arguably most important computational issue in numerical relativity
- Most problems of astrophysical interest exhibit enormous dynamic range, for two inspiraling black holes of equal mass, M , emitted radiation wavelength is of order $100M$, outer boundary needs to be minimum of one wavelength from strong-field regime
- Some form of adaptivity crucial for efficient solution, especially if there are any “multi-scale surprises” in the strong-field merger dynamics

Adaptive Mesh Refinement

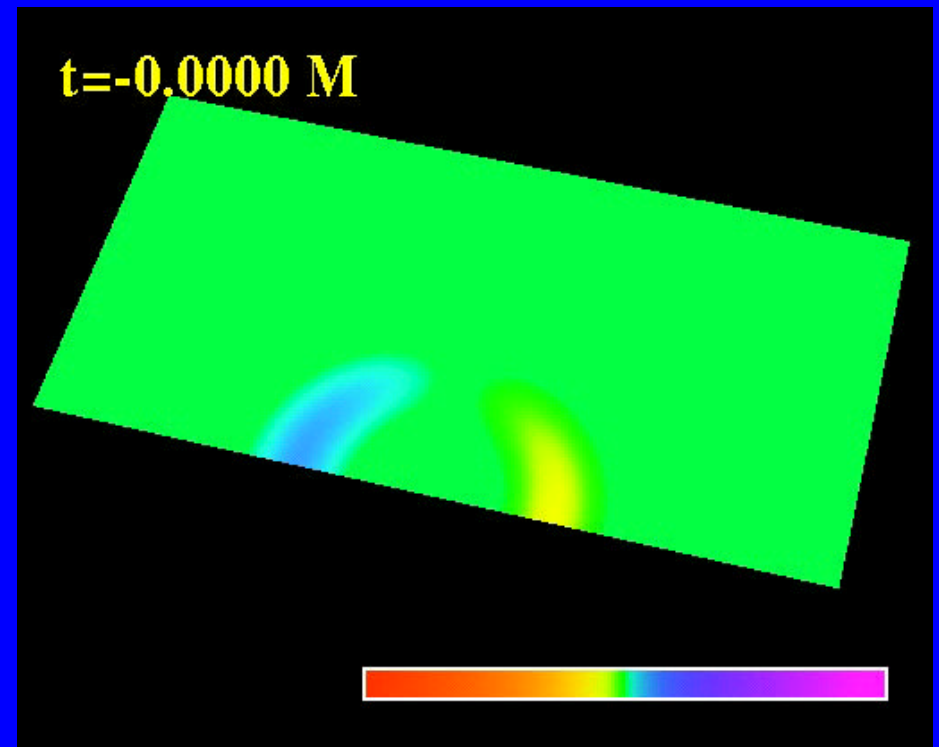
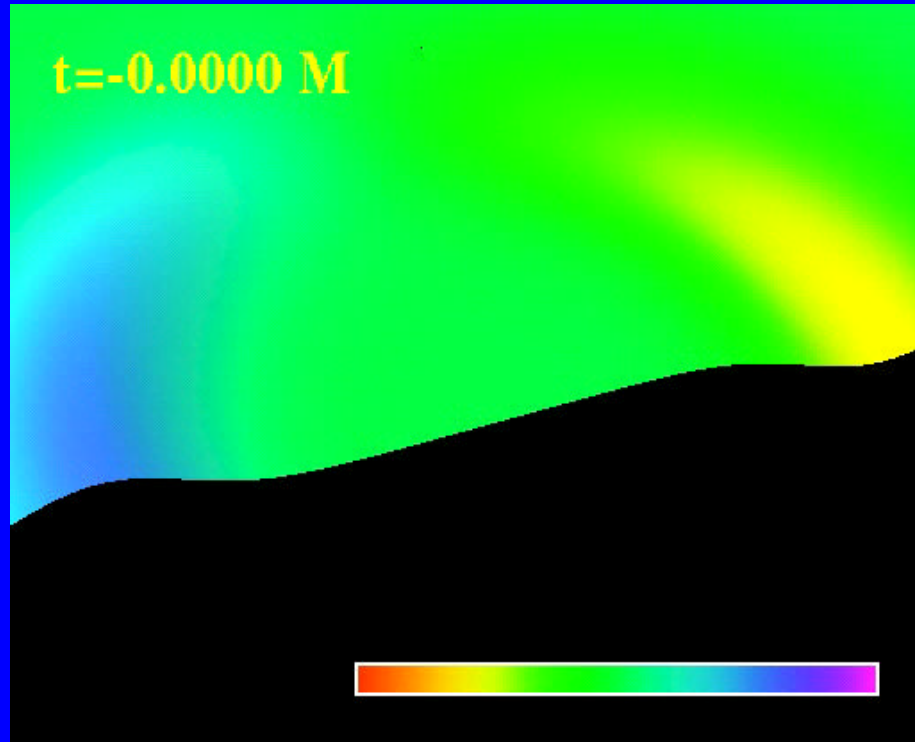
(Pretorius, Choptuik, Hirschmann and Liebling;
following Berger & Oliger, JCP, 1984)



- Employ hierarchy of locally uniform finite difference meshes
- Remesh dynamically using truncation error estimates
- Refine in time as well as space, take coarse time steps first in order to provide boundary conditions on finer meshes via interpolation
- Incorporation of elliptics complicates matters, but not fatally so

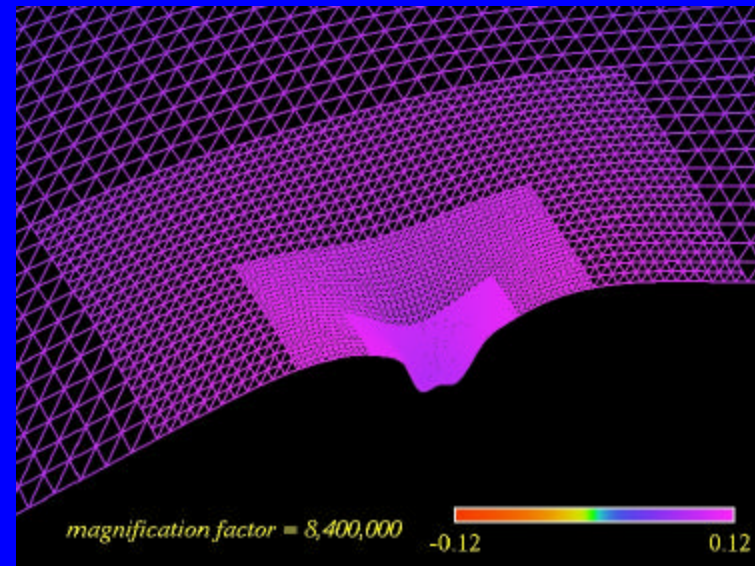
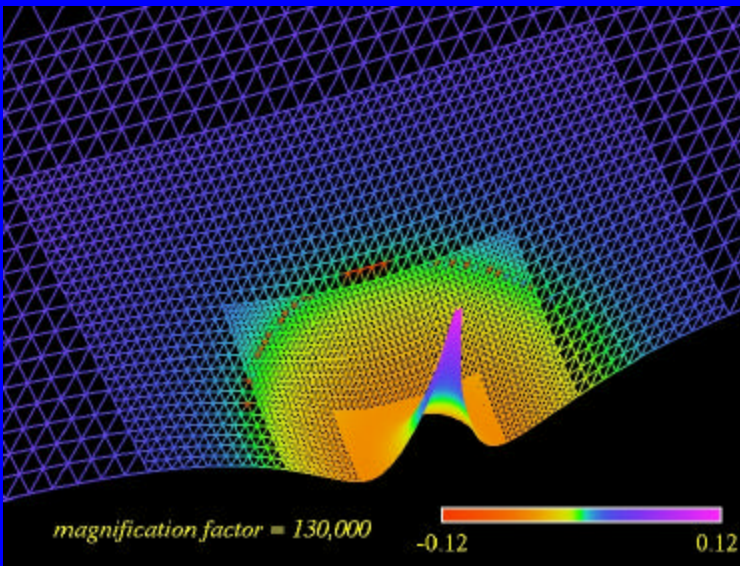
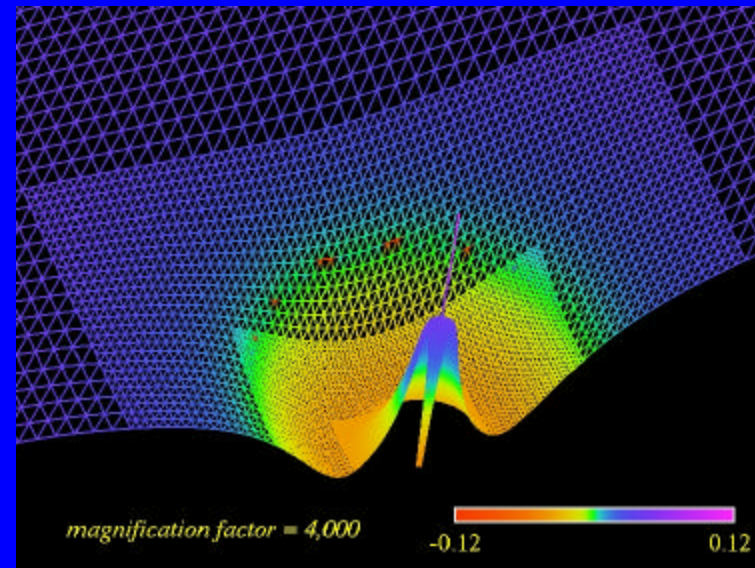
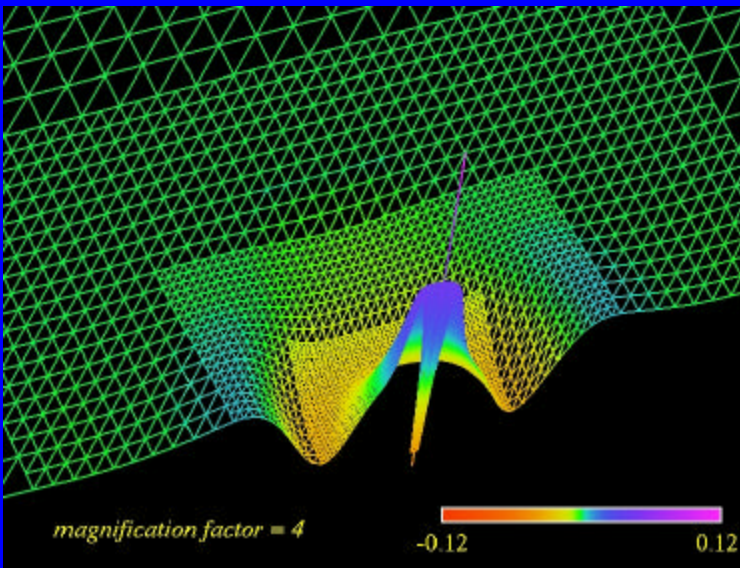
Adaptive Mesh Refinement

(near-critical axisymmetric collapse of scalar field)



Adaptive Mesh Refinement

(near-critical axisymmetric collapse of scalar field)



Validation / Error Analysis

- Known exact solutions tend to be of limited use:
 - Highly specialized/symmetric
 - Typically don't exercise all terms in EOM
 - Specific form highly coordinate dependent, expressing known solution in arbitrary coordinate system may not be significantly easier than the problem at hand
 - Non-linearity makes attempt to use deviation from exact solution to calibrate level of accuracy in solution of interest suspect

Validation / Error Analysis

- Intrinsic convergence tests, and “independent residual evaluation” in particular, are of general utility:
- Consider continuum system:

$$Lu = 0$$

- Corresponding discrete system:

$$L^h u^h = 0$$

- Truncation error (assume second-order discretization)

$$\mathbf{t}^h \equiv L^h u = h^2 \mathbf{t}_2 + \dots$$

Validation / Error Analysis

- Consider alternate/distinct/independent discretization:

$$\tilde{L}^h \tilde{u}^h = 0$$

- Corresponding truncation error (again assume second order discretization):

$$\boldsymbol{\epsilon}^h \equiv \tilde{L}^h u = h^2 \boldsymbol{\epsilon}_2 + \dots$$

- Use primary code to solve for u^h , then verify that

$$\tilde{L}^h u^h = O(h^2)$$

Validation / Error Analysis

- **NOTE:** Second (independent) discretization is only used in an applicative capacity, don't have to worry about stability, efficiency of solution of $\tilde{L}^h \tilde{u}^h = 0$
- If test fails, may not have a lot of information about what went wrong
 - Want to be as certain as possible that independent residual evaluator is correctly implemented
 - Symbolic algebra techniques can help here
- When test succeeds, tend to have a lot of confidence in correctness of solution

Other Issues

- Complexity of EOM
 - Symbolic manipulation frequently employed
- Analysis of results
 - Tensor components of metric, extrinsic curvature largely meaningless in and of themselves
 - Invariants (scalars) can be computed, but functional dependence on particular coordinate system is still problematic
- Parallelization
 - Field in relatively good shape due to “Grand Challenge” efforts and follow up projects; most notably CACTUS

Other Issues

- Generation of realistic initial data for, e.g., binary black hole inspiral which can be evolved to final state with a reasonable amount of computer time
- When coupling to matter, inherit all of the problems associated with solution of associated PDEs (e.g. hydrodynamics); new difficulties associated with coupling often arise