CONF-970717--1

LA-UR- 97-2929

Approved for public release; distribution is unlimited.

Title: SOLVING NONLINEAR HEAT CONDUCTION PROBLEMS WITH MULTIGRID PRECONDITIONED NEWTON-KRYLOV METHODS

Author(s): WILLIAM J. RIDER, XHM, MS F663
DANA A KNOLL, T-15, MS K717

Submitted to: 3RD IMACS INTERNATIONAL SYMPOSIUM
JACKSON HOLE, WY
JULY 9-12, 1997

DISCLAIMER

Los Alamos
NATIONAL LABORATORY

LA-UR- 97-2929

Solving Nonlinear Heat Conduction Problems with Multigrid Preconditioned Newton-Krylov Methods *

William J. Rider† Dana A. Knoll‡

July 22, 1997

Abstract

Our objective is to investigate the utility of employing multigrid preconditioned Newton-Krylov methods for solving initial value problems. Multigrid based method promise better performance from the linear scaling associated with them. Our model problem is nonlinear heat conduction which can model idealized Marshak waves. Here we will investigate the efficiency of using a linear multigrid method to precondition a Krylov subspace method. In effect we will show that a fixed point nonlinear iterative method provides an effective preconditioner for the nonlinear problem.

1 Overview and Motivation

Nonlinear problems are ubiquitous in physics and their efficient solution is of great practical interest. In particular nonlinear initial value problems present a unique set of challenges especially with respect to the efficiency of the solution. It is our intention to investigate Newton-Krylov methods which have shown great promise in solving a wide class of nonlinear problems [2].

It is well known that the efficiency of the Newton-Krylov methods is critically dependent on the effectiveness of the preconditioner. Traditional preconditioning (typically ILU(n)) shows less than optimal scalability practically limiting time step size and mesh size [?]. Storage becomes an increasing issue with ILU(n) as the degree of fill-in increases. Our intention is to investigate the potential of multigrid preconditioning to alleviating this shortcoming. Furthermore, the basis of a simple nonlinear iteration such as a Picard iteration (based on a multigrid solver) can serve to precondition Newton's method implemented with a matrix-free Newton-Krylov algorithm.

*This work performed under the auspices of the U.S. Department of Energy by Los Alamos National Laboratory under Contract W-7405-ENG-36. Presented at the 3rd IMACS International Symposium on Iterative Methods in Scientific Computation, July 9–12, 1997, Jackson Hole, Wyoming.

†Los Alamos National Laboratory, Applied Theoretical and Computational Physics Division, Hydrodynamics Methods Group wjr@lanl.gov

‡Los Alamos National Laboratory, Theoretical Division, Plasma Theory Group, nol@lanl.gov

It has been shown that GMRES [8] has advantageous properties for Newton-Krylov [5]. The Krylov vectors are well behaved and the convergence is monotone. When GMRES is used with as the Krylov method, the issues regarding the scaling of work and storage are especially critical. This is due to the required storage of the Krylov vectors and the increase in work per iteration associated with the orthogonalization process in the Arnoldi algorithm. The multigrid algorithm in addition to its scalability is also more effective per iteration than other typical preconditioners and should reduce the raw number of linear iterations significantly (reducing storage needs greatly for large problems). We are interested in combining multigrid with Newton-Krylov in order to give better performance.

Below we will introduce the physics of our model problem the Marshak wave as well as the definition of the problem. We then will describe its discrete solution in terms of spatial and temporal differencing. Our multigrid method is then discussed and applied to the Marshak wave. Its performance is accessed in a several respects. We then introduce coupling the multigrid method to Newton Krylov. An effective strategy to implement this approach involves the use of a simple nonlinear iterative method such as the Picard iteration to precondition the inexact Newton's method. Finally the results and performance of the combined multigrid matrix-free Newton-Krylov method is shown. In particular a direct comparison is made of the Picard and inexact Newton-Krylov methods is provided.

2 The Marshak Wave Problem

The Marshak wave results from solving the IVP for nonlinear heat conduction (an approximation to radiation transport),

$$\frac{\partial T}{\partial t} = \nabla \cdot \left(D\left(T, Z\right) \nabla T^4 \right) \tag{1}$$

where $D\left(T, Z\right) = CT^n/Z^3$ and $n = 0$ to 3. T is the temperature, D is the thermal diffusivity (opacity) and Z is the mass number of the material. This equation is a valid approximation in an optically thick region that in nonrelativistic where the thermal energy dominates the radiation energy ($E_{\text{Rad}} \propto T^4$)[1]

The solution of this equation presents us with two basic problems: the nonlinearity associated with T and the jump discontinuities associated with changes of material. Furthermore the jumps can be quite large and in the problems used here approach ratios in effective diffusion coefficients of 10^7 to 10^{10}. This makes the associated linear algebra problem quite difficult necessitating an emphasis on the robustness of the methodology.

2.1 Problem Statement

The specific problem that is studied here is shown in Figure 1. The problem consists of a low atomic mass number ($Z = 10$) path embedded in a high atomic mass number material ($Z = 200$). A high temperature is applied to one boundary and a Marshak wave will

[1]In a real sense this equation represents a drastic simplification of radiation transport. More complicated models include multiple temperatures, frequencies and are transported rather than diffused. The physics is discussed in more detail by Zeldovich and Raizer [?] and Mihalas and Mihalas [?].

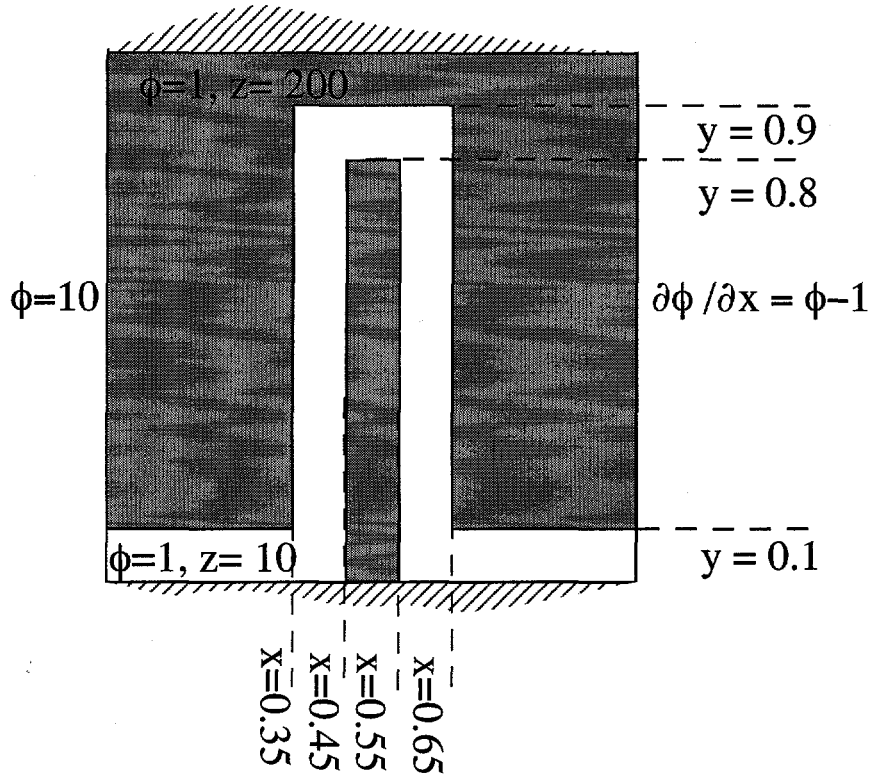


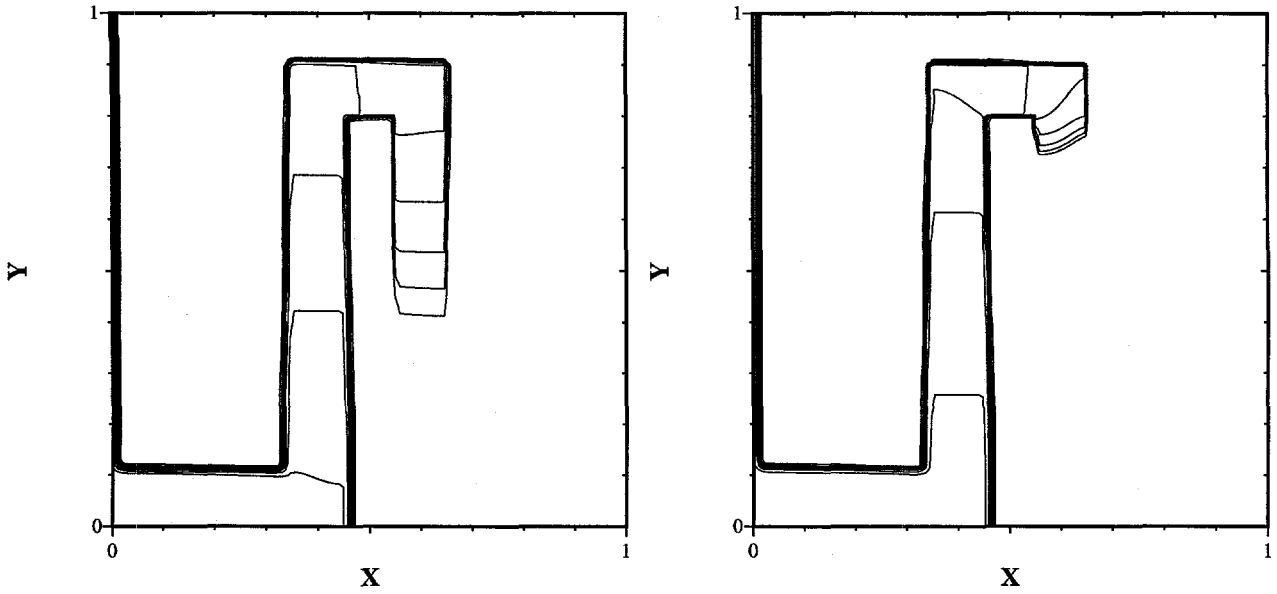
Figure 1: Problem geometry, initial and boundary conditions for the examples given in this paper.

progress quickly through the low Z material. For $D \propto T^3$, the boundary conditions and the material given, the jump in the diffusion coefficient can be as large as 10^{10} across a material interface (10^7 for $D \propto T^0$). Figure 2 shows two example solutions for different forms of the opacity's dependence on temperature. As the nonlinearity increases, the sharpness of the wave increases and as we will see later, the difficulty of obtaining the nonlinear solution. Next, the discretization of this problem is given.

2.2 Discrete Representation

For the nonlinear diffusion operator the discrete form chosen is a five point Laplacian,

$$N(T) \approx \nabla \cdot D(x,y) \nabla T^4 \approx \tag{2}$$
$$\frac{1}{\Delta x}\left[D_{i+\frac{1}{2},j}\frac{T^4_{i+1,j}-T^4_{i,j}}{\Delta x} - D_{i-\frac{1}{2},j}\frac{T^4_{i,j}-T^4_{i-1,j}}{\Delta x}\right]$$
$$+\frac{1}{\Delta y}\left[D_{i,j+\frac{1}{2}}\frac{T^4_{i,j+1}-T^4_{i,j}}{\Delta y} - D_{i,j-\frac{1}{2}}\frac{T^4_{i,j}-T^4_{i,j-1}}{\Delta y}\right],$$



(a) Solution at $t = 1.0$, $D \propto T^0$ 128^2 grid.

(b) Solution at $t = 1.5 \times 10^{-3}$, $D \propto T^3$ 128^2 grid.

Figure 2: Two example solutions showing the general characteristics of the solution to the problem given in Figure 1

For use with multigrid a fully linearized form must be used,

$$L(T) \approx \nabla \cdot \bar{D}(x,y) \nabla T \approx \tag{3}$$
$$\frac{1}{\Delta x}\left[\bar{D}_{i+\frac{1}{2},j}\frac{T_{i+1,j}-T_{i,j}}{\Delta x} - \bar{D}_{i-\frac{1}{2},j}\frac{T_{i,j}-T_{i-1,j}}{\Delta x}\right]$$
$$+\frac{1}{\Delta y}\left[\bar{D}_{i,j+\frac{1}{2}}\frac{T_{i,j+1}-T_{i,j}}{\Delta y} - \bar{D}_{i,j-\frac{1}{2}}\frac{T_{i,j}-T_{i,j-1}}{\Delta y}\right],$$

The effective diffusion coefficient is split into three parts: a material dependent term.

$$D_m = C/Z^3,$$

which uses a harmonic mean, $D = 2D_1D_2/(D_1+D_2)$. A temperature dependent term

$$D_T = T^n$$

which is developed assuming a linear variation in temperature between grid points and determining an integral average $D(T_1,T_2) = \left(T_2^{n+1} - T_1^{n+1}\right)/\left((n+1)(T_2-T_1)\right)$. The overall diffusion coefficient is $D = D_m D_T$. A radiation term for the fully linearized equation

$$D_4(T_1,T_2) = (T_1+T_2)\left(T_1^2+T_2^2\right)$$

thus

$$D_4(T_1,T_2)(T_2-T_1) = T_2^4 - T_1^4$$

The overall linearized diffusion coefficient is $\bar{D} = DD_4$.

The form of the operator and the diffusion coefficient can have a profound impact on the linear algebra. Irrespective of the form of the diffusion coefficient, the form of L is symmetric positive definite. On the other hand, the Jacobian of N is nonsymmetric and as the temperature front becomes sharper, can become indefinite. It should be obvious that working with L is advantageous.

An explicit scheme can be defined by the forward Euler's method

$$T^{n+1} = T^n + \Delta t N\left(T^n\right),$$

or improved Euler's method

$$T^{n+\frac{1}{2}} = T^n + \frac{1}{2}\Delta t N\left(T^n\right),$$

$$T^{n+1} = T^n + \Delta t N\left(T^{n+\frac{1}{2}}\right),$$

with 1 or 2 function evaluations per grid point per time step. The integration method is stable if

$$\frac{D\Delta t}{4h^2} < 1.$$

Unfortunately this is extremely restrictive for this problem making an explicit integration impractical.

The standard approach is a semi-implicit method. The semi-implicit method is defined by an analytical linearization of (1) and a backward Euler's method,

$$\delta T - \Delta t L\left(\delta T\right) = \Delta t N\left(T^n\right), \tag{4a}$$

$$T^{n+1} = T^n + \delta T. \tag{4b}$$

This method is unconditionally stable (linearly). While it is stable, it is nonlinearly inaccurate for large time steps, $\Delta t \gg \Delta t_{\text{explicit}}$. Amount of work proportional to the "work units" used in solving the multigrid problem. Roughly speaking the solution must be at least 20 times greater than the explicit stability limit if the implicit solution takes 5 iterations/V-cycles with the MGCG algorithm. At time steps of this size or larger the accuracy of the solution is an issue especially when one considers the effect of the lack of convergence in the nonlinear sense.

Picard (fixed-point) iteration uses the semi-implicit method in a iterative sense. The time step advancement now becomes an iteration,

$$\delta T - \Delta t L\left(\delta T\right) = T^{n+1,m} - T^n + \Delta t N\left(T^{n+1,m}\right),$$

$$T^{n+1,m+1} = T^{n+1,m} + \delta T.$$

The convergence of the linear problem is tied to 10^{-2} times that of the nonlinear problem in our work here. This limits the amount of work that used to produce solutions that poorly approximate the nonlinear solution.

Newton's method uses the multigrid preconditioned Krylov (GMRES) for the solution. The sequence is identical to that used in the Picard iteration, except that the operator on the LHS of the equation is an approximation to the true Jacobian (in a matrix-free sense).

$$\delta T - \Delta t N\left(\delta T\right) = T^{n+1,m} - T^{n} + \Delta t N\left(T^{n+1,m}\right),$$

$$T^{n+1,m+1} = T^{n+1,m} + \delta T.$$

This method is used as an inexact Newton's iteration with the same relationship between the nonlinear and linear convergence as the Picard iteration.

3 Newton-Krylov Methods

Our goal is to compute an inexact Newton iteration using a preconditioned Krylov method to calculate the updates to the dependent variables by approximately solving,

$$J\left(\mathbf{x}^{m}\right)\delta\mathbf{x} = -F\left(\mathbf{x}^{m}\right), \tag{5}$$

and

$$\mathbf{x}^{m+1} = \mathbf{x}^{m} + \alpha\delta\mathbf{x} \tag{6}$$

to solve $F\left(\mathbf{x}\right) = 0$. The under-relaxation factor α is defined by $\alpha = \min\left(1, 1/\left\|\delta T/T\right\|\right)$. We can do this in a matrix-free manner without forming the full Jacobian via a finite difference approximation,

$$J\mathbf{v} \approx \frac{F\left(\mathbf{x}+\epsilon\mathbf{v}\right) - F\left(\mathbf{x}\right)}{\epsilon} \tag{7}$$

where $\epsilon = \rho\left(1 + \left\|\mathbf{v}\right\|\right)$ and $\rho = 10^{-7}$ here.

The properties of GMRES make it advantageous for use as the Krylov method here (conversely the properties of other methods such as CGS, BiCGStab, and other similar methods are problematic). Additionally, GMRES has the property of finite termination and is more robust as a consequence. This is offset to some degree by the increased storage and work requirements imposed by GMRES. As noted before preconditioning the linear problem is essential for efficiency. Standard ILU(n) preconditioning lose efficiency as the problem size grows and the corresponding growth in the number of GMRES iterations creates storage (and work) needs that limit problem size. We will employ a multigrid algorithm developed below to overcome this difficulty.

4 A Multigrid Algorithm

Previously [7, 6] a multigrid algorithm was developed for solving the pressure equation in a variable density incompressible flow integrator. This multigrid method uses standard smoothers (weighted Jacobi, Gauss-Seidel) and piecewise constant restriction and prolongation (for dealing with discontinuous coefficients). It was found to work well for $D_{\text{high}}/D_{\text{low}} < 10^{4}$.

Occasionally it failed to effectively converge (less than 10 V-cycles for 2 pre- and post-smooths). Often these failures were characterized by two things: large discontinuities in

effective diffusion coefficient and fine scale structure in the effective diffusion coefficients form.

Typically, the answer to such problems is to resort to writing the multigrid algorithm in a more rigorous form by employing Galerkin coarse grid operators which are in turn operator dependent [3, 4]. A simpler course of action was taken here that has proven to be extremely effective. In the vast majority of cases using the above algorithm to precondition a conjugate gradient algorithm returns the method to high efficiency and linear scaling characteristic of multigrid [9, 1].

Based on the results of using this multigrid algorithm to precondition the conjugate gradient algorithm, a study of preconditioning the Krylov iteration in a Newton-Krylov iteration was undertaken. Previously, we have discussed the improvement this approach provided in solving boundary value problem such as Burgers' equation [?].

For Newton-Krylov the linearized operator (like that used in the semi-implicit method) is used to precondition the nonlinear problem. This is approach in effect uses a Picard iteration to precondition the inexact Newton's method. The Krylov method is implemented in a matrix-free fashion. While the linearized operator is SPD, the Jacobian is nonsymmetric and can be indefinite.

We anticipate scaling similar to that observed with the conjugate gradient method preconditioned by the multigrid. Next, we will show some results demonstrating the efficacy of our approach.

5 Results

Here we will show results for several of the issues presented above: the multigrid algorithm stand alone, the Newton-Krylov method and a comparison of the Newton Krylov and Picard iteration.

5.1 Stand Alone Multigrid Performance

The performance of the multigrid algorithm is critical to the overall solution strategy pursued here. Earlier uses of this multigrid method have shown that the multigrid's performance could degrade significantly with increases in the ratio of diffusion coefficients and material topology of a problem. Here we show that a similar phenomena occurs with the Marshak wave problem solved here.

The solution of a single linearized system is shown in Figure 3. While the first V-cycle reduces the residual by over an order of magnitude, the next order of magnitude reduction in residual norm requires more than 15 V-cycles. Table 1 shows these results for a wider variety of conditions and grid sizes. It should be noted that each of these problems uses a fairly small time step and when the time step size is increased the multigrid algorithm will fail to converge. This problem bounds the simpler end of the spectrum of problem difficulty thus indicating that something else must be considered for this problem. At lower resolution, the 32^2 grid, the effects of under-resolving the material boundary are indicated by the poor multigrid performance (which will not effect the combined multigrid conjugate gradient method).

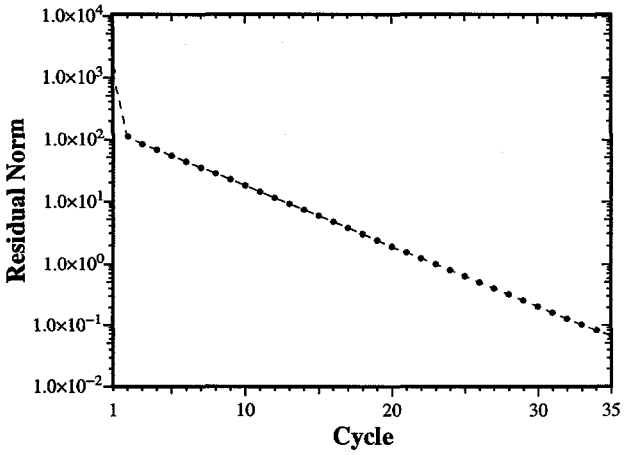


Figure 3: Standalone multigrid performance for a Marshak wave problem ($D \propto T^3$) and a 32^2 grid.

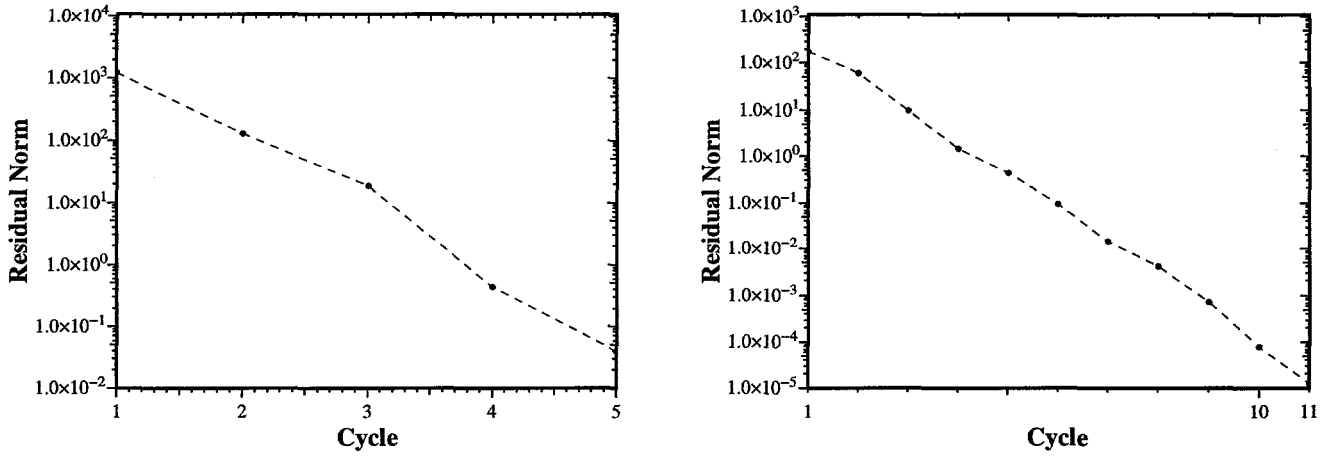
Table 1: Average number of multigrid cycles as a function of grid size and diffusion coefficient nonlinearity.

Grid	$D \propto T^0$	$D \propto T^1$	$D \propto T^2$	$D \propto T^3$
32^2	5.2	7.7	12.6	38.9
64^2	4.9	7.7	12.7	22.6
128^2	5.0	7.4	12.3	16.9
256^2	5.2	7.6	11.9	16.4

Next, we consider the performance of using this multigrid to precondition a conjugate gradient algorithm.

5.2 Multigrid Preconditioning Conjugate Gradient

If this multigrid is instead used to precondition a conjugate gradient algorithm the story changes. In Figure 4a the same problem that was shown in Figure 3 is given. The combination of multigrid with a conjugate gradient method converges in 5 iterations (5 V-cycles). The rate of convergence is relatively uniform iteration-to-iteration with each being roughly as effective as the first stand alone multigrid V-cycle. If the difficulty of the problem is increased by increasing the time step size to the extent that the problem is effectively a steady-state problem, the method still converges efficiently as shown in Figure 4b. Table 2 shows that for smaller time step sizes the order of the nonlinearity play nearly no role. If



(a) Multigrid Preconditioned Conjugate Gradient Convergence for $D \propto T^3$ and small time step.

(b) Multigrid Preconditioned Conjugate Gradient Convergence for $D \propto T^3$ and large time step.

Figure 4: The convergence of the multigrid preconditioned conjugate gradient algorithm on two Marshak wave problems.

Table 2: Average number of iterations of the multigrid preconditioned conjugate gradient algorithm for Marshak wave problems where the time step is small.

Grid	$D \propto T^0$	$D \propto T^1$	$D \propto T^2$	$D \propto T^3$
32^2	4.2	4.3	4.3	4.4
64^2	4.4	4.5	4.5	4.5
128^2	4.2	4.4	4.5	4.5
256^2	4.4	4.6	4.6	4.7

the time step size is increased to the extent of making an effective steady-state problem, the method no longer scales independent of problem size, but retains its relatively small dependence on nonlinearity as shown in Table 3 scaling at about $N^{1.06}$.

Now we consider the efficiency of the full nonlinear time dependent solution algorithm.

5.3 Performance of the Nonlinear Implicit Solution

The full nonlinear solution will employ two methods: a Picard iteration and a inexact Newton's method. The linear algebra for each inexact Newton iteration is a Krylov method (GMRES) preconditioned with a multigrid method. This preconditioner is exactly the linear solution from the Picard iteration.

Table 3: Average number of iterations of the multigrid preconditioned conjugate gradient algorithm for Marshak wave problems where the time step is very large.

Grid	$D \propto T^0$	$D \propto T^1$	$D \propto T^2$	$D \propto T^3$
32^2	8.4	8.3	7.8	7.6
64^2	9.5	8.8	8.5	8.3
128^2	9.8	9.4	8.9	8.5
256^2	11.0	10.2	9.9	9.7

Table 4: The Picard Iteration, $D \propto T^0$, for the full nonlinear problem with counts of the nonlinear/linear iterations used to converge the solution to a specified residual tolerance.

Nonlinear ϵ	$\Delta t = 1$	$\Delta t = 0.1$	$\Delta t = 0.01$
10^{-4}	22.2/100.3	18.2/63.2	16.7/37.8
10^{-8}	43.8/208.3	32.2/116.5	29.5/67.8

Table 5: The inexact Newton-Krylov Iteration, $D \propto T^0$, for the full nonlinear problem with counts of the nonlinear/linear iterations used to converge the solution to a specified residual tolerance.

Nonlinear ϵ	$\Delta t = 1$	$\Delta t = 0.1$	$\Delta t = 0.01$
10^{-4}	6.0/21.2	7.3/27.0	6.6/23.5
10^{-8}	7.9/29.3	8.9/33.6	8.3/30.7

First, we compare the two methods on the simple problem where $D \propto T^0$ at two convergence tolerances. While the differences are most acute at large time steps ($\Delta t = 1$), the differences persist for all time steps and grow substantially at tighter tolerances. This growth is indicative that the Newton iteration enters into the quadratic region of convergence for the problem. A summary of the results are given in Tables 4 and 5.

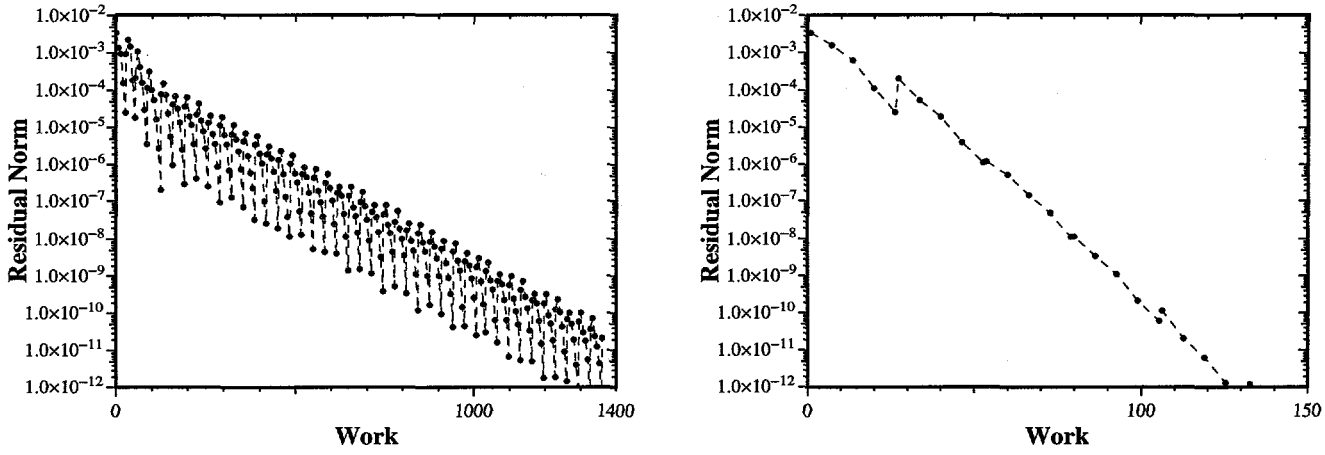
The situation becomes more favorable for the inexact-Newton-Krylov algorithm as the nonlinearity of the problem increases. This is seen in the compilation of results shown in Tables 6 and 7. The convergence of two cases given here are shown in Figure 5. In Figure 5a the Picard iteration is shown and it is clear that the linear iteration is quite wasteful in terms of its work because of the slow convergence of the nonlinear iteration. The opposite is

Table 6: The Picard iteration for $D \propto T^3$.

Nonlinear ϵ	$\Delta t = 0.001$	$\Delta t = 0.0001$	$\Delta t = 0.00001$
10^{-4}	54.3/174.9	29.5/63.9	14.5/25.9
10^{-8}	97.4/326.2	52.9/114.6	26.2/49.0

Table 7: The Newton-Krylov iteration for $D \propto T^3$

Nonlinear ϵ	$\Delta t = 0.001$	$\Delta t = 0.0001$	$\Delta t = 0.00001$
10^{-4}	15.3/85.6	8.4/35.5	5.7/15.1
10^{-8}	16.8/94.2	10.2/44.1	7.1/20.2



(a) Mulitgrid Picard Iteration

(b) Multigrid Newton-Krylov

Figure 5: The convergence of the multigrid preconditioned Picard and Newton-Krylov algorithms for $\epsilon = 10^{-8}$, $\Delta t = 10^{-5}$ and $D \propto T^3$.

true with Newton's method (Figure 5b) where the convergence of the linear and nonlinear problem are aligned. In that case the convergence is nearly monotone with respect to the work done. The conclusion that can be drawn from this study is that the Newton-Krylov combination should be more robust than the Picard iteration.

5.4 Closing Remarks

In summary, multigrid Newton-Krylov methods appear to be attractive for nonlinear initial value problems. The multigrid algorithm is critical to the efficient solution and using some sort of Krylov acceleration improves the robustness of the multigrid so that it can be used for this type of problem. Newton's method is significantly more efficient than a Picard iteration in providing accurate nonlinear solutions for this problem. The Picard iteration is effective in providing a preconditioning strategy for nonlinear problems.

References

[1] S.F. Ashby and R.D. Falgout. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow systems. *Nuc. Sci. and Eng.*, 124:145, 1996.

[2] P.N. Brown and A.C. Hindmarsh. Matrix-free methods for stiff systems of ode's. *SIAM J. Num. Anal.*, 23:610–638, 1986.

[3] J.E. Dendy. Black box multigrid. *J. Comput. Phys.*, 48:366, 1982.

[4] J.E. Dendy. Black box multigrid for nonsymmetric problems. *Appl. Math. Comput.*, 13:261, 1983.

[5] P.R. McHugh and D.A. Knoll. Inexact Newton's method solution to the incompressible navier-stokes and energy equations using standard and matrix-free implementations. *AIAA J.*, 32(12), 1994.

[6] E. G. Puckett, A. S. Almgren, J. B. Bell, D. L. Marcus, and W. J. Rider. A second-order projection method for tracking fluid interfaces in variable density incompressible flows. 130:269–282, 1997.

[7] W. J. Rider, D. B. Kothe, S. J. Mosso, J. H. Cerruti, and J. I. Hochstein. Accurate solution algorithms for incompressible multiphase fluid flows. Technical Report AIAA 95–0699, AIAA, 1995. Presented at the 33rd Aerospace Sciences Meeting and Exhibit.

[8] Y. Saad and M.H. Schultz. GMRES: A generalized minimal residual algorithm for solving non-symetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856, 1986.

[9] O. Tatebe. The multigrid preconditioned conjugate gradient method. In N. D. Melson, T. A. Manteuffel, and S. F. McCormick, editors, *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods*, pages 621–634, 1993.