Fourteenth International Conference on Domain Decomposition Methods
Editors: Ismael Herrera , David E. Keyes, Olof B. Widlund, Robert Yates

8. Nonlinearly Preconditioned Newton's Method

S. H. Lui[1]

1. Introduction. Many challenging problems in science and engineering are large and nonlinear. Typically they are solved by Newton's method or its many variations. If parallel computers are available, the solution process can be sped up by the use of domain decomposition techniques. The traditional domain decomposition approach for nonlinear PDEs is to use the classical Newton's method and apply classical domain decomposition techniques such as the additive Schwarz preconditioner ([6]) to the resulting linear systems. This is often referred to as the Newton-Krylov-Schwarz method ([1], [2]). For most nonlinear equations, this works very well. However, for more difficult problems, the lack of a good initial guess means that the Newton-Krylov-Schwarz iteration may not converge or may converge very slowly. Often, the failure may be traced to boundary layers, singularities (corners/cusps) in the domain, and/or multi–physics domains (fluid–structure interaction problems for instance). These problematic regions slow down global convergence or cause stagnation in the iteration. There are of course many papers on the application of domain decomposition methods to nonlinear problems, especially those in fluid mechanics. Many references can be found in the proceedings of the annual conference on domain decomposition methods, starting with [7].

Meanwhile, other workers have begun to look at applying Schwarz methods directly on the nonlinear subdomain problems: [5], [17], [18], [11], [12], and [13]. These nonlinear Schwarz methods have the nice property that difficult regions are isolated in a small number of subdomains where special techniques (finer grid, asymptotics, etc.) may be brought to bear without interfering with the convergence in other parts of the domain. However, they still require a good initial guess for convergence and their rate of convergence is usually slow (linear).

Recently, Cai and Keyes ([3]) have proposed a new method which is a marriage of the Newton-Krylov-Schwarz and nonlinear Schwarz methods. Their idea is to nonlinearly precondition the given nonlinear equations $F(u) = 0$ so that the resultant equations $\mathcal{F}(u) = 0$ are closer to linear equations and so amenable to solution by Newton's method without the necessity of a good initial guess. The nonlinear preconditioner is a nonlinear additive Schwarz preconditioner which requires the solution of a nonlinear subdomain PDE. The new system $\mathcal{F}(u) = 0$ is solved using a modified Newton's method where the Jacobian has the same form as in the Newton-Krylov-Schwarz algorithm. In particular, it reduces to the additive Schwarz algorithm when F is linear. In [3], they illustrate the impressive robustness of this new method with the driven cavity flow problem where Newton's method stagnates at a moderate Reynolds number while the nonlinearly preconditioned method is able to compute to a considerably larger Reynolds number and still maintain fast quadratic convergence.

In this paper, we carry out some preliminary convergence analysis of this nonlinearly preconditioned method as well as estimate crudely its radius of quadratic convergence, that is, the radius of the ball where the iterates converge quadratically. This is compared to the corresponding quantity for the classical Newton's method. The discussion is in the context of semilinear elliptic PDEs which are described in the next section. In section three, we shall examine two types of convergence theories: classical q-quadratic convergence and r-quadratic convergence assuming data only at the initial guess. In the last section, we carry out some numerical experiments on some quasi-linear two-point boundary value problems and conclude.

[1]Department of Mathematics, University of Manitoba, Winnipeg, Manitoba, Canada luish@cc.umanitoba.ca. This work was supported in part by a grant from NSERC of Canada.

2. Nonlinearly preconditioned PDEs. In this section, we apply the nonlinear preconditioner to a class of semilinear elliptic PDEs and see its relation with the Newton-Krylov-Schwarz and nonlinear Schwarz methods.

Let Ω be a bounded domain in $\mathbf{R}^N$ with a smooth boundary. We consider the PDE

$$-\triangle u = f(x,u) \text{ on } \Omega \tag{2.1}$$

for the solution $u \in H_0^1(\Omega)$. For simplicity, we write $f(u)$ for $f(x,u)$. Throughout this paper, we assume that this PDE has the unique solution u.

Suppose for some fixed integer $m > 1$, $\Omega = \Omega_1 \cup \cdots \cup \Omega_m$, where the subdomains Ω_i have smooth boundaries and are overlapping, meaning that $H_0^1(\Omega) = H_0^1(\Omega_1) + \cdots + H_0^1(\Omega_m)$. In this paper, a function in $H_0^1(\Omega_i)$ is considered as a function in $H_0^1(\Omega)$ by extension by zero. Let $\|\cdot\|$ be the norm on $H_0^1(\Omega)$, that is,

$$\|v\|^2 = \int_\Omega |\nabla v|^2$$

and $\|\cdot\|_{-1}$ be the norm on the dual space $H^{-1}(\Omega)$. Let P_i denote the projection $P_i H_0^1(\Omega) = H_0^1(\Omega_i)$ in the $H_0^1(\Omega)$–norm.

It is more convenient to express the PDE as the nonlinear operator equation

$$F(u) \equiv u + \triangle^{-1} f(u) = 0,$$

where $F : H_0^1(\Omega) \to H_0^1(\Omega)$, and $\triangle^{-1} : H^{-1}(\Omega) \to H_0^1(\Omega)$ denotes the inverse of the Laplacian operator on Ω with homogeneous Dirichlet boundary conditions. Define the new nonlinear equations ([3])

$$\mathcal{F}(u) \equiv \sum_{i=1}^m T_i(u) = 0$$

where $T_i : H_0^1(\Omega) \to H_0^1(\Omega_i)$ satisfies

$$P_i F(v + T_i(v)) = 0, \qquad v \in H_0^1(\Omega).$$

It is assumed that this solution exists and is unique given v. One can think of $T_i(v)$ as a correction to the current guess v obtained by solving a nonlinear subdomain PDE. Let $y_i = T_i(v)$ and using the definition of F, we obtain

$$y_i + P_i \triangle^{-1} f(v + y_i) = -P_i v \tag{2.2}$$

or

$$-\triangle_i y_i - f(v + y_i) = \triangle v \text{ on } \Omega_i. \tag{2.3}$$

This nonlinear subdomain PDE is very much like that in nonlinear Schwarz algorithms mentioned above.

The nonlinearly preconditioned method solves the new nonlinear equations using Newton's method. That is given $u^{(0)}$, it produces the sequence

$$u^{(n+1)} = u^{(n)} - \mathcal{F}'(u^{(n)})^{-1} \mathcal{F}(u^{(n)}).$$

In practice, a Krylov subspace method such as GMRES ([15]) is used to solve the above linear equations. These methods only require that we supply a procedure to compute $\mathcal{F}'(u^{(n)})w$ for an arbitrary $w \in H_0^1(\Omega)$. Let us look at this in a little more detail. Let

$$\mathcal{F}'(u^{(n)})w = \sum_{i=1}^m z_i, \qquad z_i = \frac{\partial T_i(u^{(n)})}{\partial v} w \in H_0^1(\Omega_i).$$

From (2.2) and abbreviating $T_i(u^{(n)})$ by y_i,

$$z_i + \triangle_i^{-1} f'(u^{(n)} + y_i)(w + z_i) = -P_i w$$

or equivalently

$$[-\triangle_i - f'(u^{(n)} + y_i)]\, z_i = [\triangle + f'(u^{(n)} + y_i)]\, w. \tag{2.4}$$

This scheme will be referred to as the nonlinearly preconditioned Newton's method (NP1). Other variations are possible. The original scheme of Cai and Keyes (henceforth called NP0) replaces $\mathcal{F}'(u^{(n)})$ by $\mathcal{F}'(\tilde{u}^{(n)})$ whose action on w yields

$$[-\triangle_i - f'(u^{(n)})]\, z_i = [\triangle + f'(u^{(n)})]\, w.$$

Note that this has the same form as applying the additive Schwarz preconditioner to solve a linear system for operator $F'(u^{(n)})$. While this gives a nice connection to the well-understood Newton-Krylov-Schwarz algorithm, it does not use the most up-to-date information ($T_i(u^{(n)})$) and this sometimes compromises the robustness of the algorithm. For some examples, see the section on numerical experiments.

A third variation (NP2) replaces (2.4) by

$$[-\triangle_i - f'(u^{(n)} + y)]\, z_i = [\triangle + f'(u^{(n)} + y)]\, w$$

where $y = \sum_{i=1}^{m} y_i = \mathcal{F}(u^{(n)})$. The reasoning here is that y incorporates information from neighboring subdomains and may lead to a better estimate. We assume that z_i exists and is unique in all three cases.

The following is a version of the partition lemma ([14], [10]) for bounded linear operators.

Lemma 2.1 *Let A be a bounded linear operator on a Hilbert space H. Suppose $H = H_1 + \cdots + H_m$ and $A = A_1 + \cdots + A_m$ where H_i are Hilbert spaces and A_i are bounded linear operators on H_i. Then there is some constant C_m such that*

$$\|A\| \geq C_m \sum_{i=1}^{m} \|A_i\|.$$

Finally, we collect together the definitions of all constants which will appear later. Let

1. r denote the radius of $B_r(u)$, the open ball with center at u;
2. $\alpha_0(u^{(0)})$ denote the eigenvalue of $F'(u^{(0)}) = I + \triangle^{-1} f'(u^{(0)})$ of smallest magnitude and $\alpha_0 = \alpha_0(u)$;
3. γ denote the Lipschitz constant for f':
$$\|f'(w) - f'(v)\|_{-1} \leq \gamma \|w - v\|, \qquad w, v \in B_r(u);$$
4. α_i denote the eigenvalue of $I + \triangle_i^{-1} f'(u)$ on Ω_i of smallest magnitude with corresponding eigenfunction $\phi_i \in H_0^1(\Omega_i)$ and $|\alpha_{max}| = \max\limits_{1 \leq i \leq m} |\alpha_i|$;
5. $\alpha_i(u^{(0)})$ denote the eigenvalue of $I + \triangle_i^{-1} f'(u^{(0)} + T_i(u^{(0)}))$ on Ω_i of smallest magnitude and $|\alpha_{max}(u^{(0)})| = \max\limits_{1 \leq i \leq m} |\alpha_i(u^{(0)})|$; note $\alpha_i = \alpha_i(u)$;
6. $\rho_i = \sup\limits_{w \in B_r(u)} \| [I + \triangle_i^{-1} f'(w)]^{-1} P_i [I + \triangle^{-1} f'(w)] \|$ and $\rho_{max} = \max\limits_{1 \leq i \leq m} \rho_i$;
7. $\beta_i = \sup\limits_{w \in B_r(u)} \| [I + \triangle_i^{-1} f'(w)]^{-1} \|$ and $\beta_{max} = \max\limits_{1 \leq i \leq m} \beta_i$.

Note that r must be sufficiently small so that the Newton iteration is well defined and all iterates remain in $B_r(u)$.

3. Convergence Theory. Newton's method is one of the oldest, simplest and most efficient methods for solving nonlinear equations. Most of the best algorithms today are modifications of Newton's method. It is not surprising that many types of convergence theories exist, depending on the hypotheses and the convergence result. We shall examine two such theories. Recall that iterates $\{e^{(n)}\}$ converge q-quadratically to 0 if $\|e^{(n+1)}\| \leq c \|e^{(n)}\|^2$ for some constant c while it converges r-quadratically to 0 if $\|e^{(n)}\| \leq c_n$ where $\{c_n\}$ converges q-quadratically to 0.

The first theory is well known and is concerned with the q-quadratic convergence of Newton's method. The second theory is rather special in that all assumptions are at one point, the initial iterate – there is no Lipschitz condition in a region which is required in the other theory. It is unfortunate that we are unable to do much analysis for the nonlinearly preconditioned method in regard to this theory and must resort to some numerical experiments. For the first theory, we attempt to contrast the rate of convergence of the nonlinearly preconditioned method versus that of the classical Newton's method, and the radii of quadratic convergence of the two methods.

3.1. q-quadratic convergence. The first convergence theory assumes that $F'(u)$ has a bounded inverse with $\|F'(u)^{-1}\| \leq \eta$ and $\|F'(u) - F'(v)\| \leq \Gamma \|u - v\|$ for all $v \in B_r(u)$. Then the (classical) Newton iterates $\{u^{(n)}\}$ for $F(u) = 0$ are well defined and the error $e_N^{(n)} = u^{(n)} - u$ satisfies

$$\|e_N^{(n+1)}\| \leq \eta\Gamma \|e_N^{(n)}\|^2$$

provided the initial iterate $u^{(0)} \in B_{\epsilon_N}(u)$, where

$$\epsilon_N = \min\left(r, \frac{1}{2\eta\Gamma}\right).$$

See, for instance, [4] or [9]. We use the subscript N to describe the relevant quantity for the classical Newton's method. Hence ϵ_N is a lower bound of the radius of the ball where quadratic convergence takes place.

Applying the classical Newton's method to our semilinear elliptic PDE $F(u) = 0$, we find that $\eta = |\alpha_0|^{-1}$ and for any $v \in B_r(u)$,

$$\begin{aligned} \|F'(u) - F'(v)\| &= \|\triangle^{-1}(f'(u) - f'(v))\| \\ &= \|f'(u) - f'(v)\|_{-1} \\ &\leq \gamma \|u - v\|. \end{aligned}$$

Thus provided $u^{(0)} \in B_{\epsilon_N}(u)$,

$$\|e_N^{(n+1)}\| \leq \frac{\gamma}{|\alpha_0|} \|e_N^{(n)}\|^2, \qquad \epsilon_N = \min\left(r, \frac{|\alpha_0|}{2\gamma}\right). \tag{3.1}$$

Now we compute these same quantities for the nonlinear preconditioned Newton's method which employs the classical Newton's method to solve $\mathcal{F}(u) = 0$. By some straightforward calculations,

$$\|\mathcal{F}'(u)\| \geq C_m |\alpha_0| \sum_{i=1}^{m} \frac{1}{|\alpha_i|}$$

and for $v \in B_r(u)$,

$$\|\mathcal{F}'(u) - \mathcal{F}'(v)\| \leq \gamma \sum_{i=1}^{m} \frac{(1 + \rho_i)^2}{|\alpha_i|} \|u - v\|.$$

Putting everything together, we obtain the error relation for the nonlinearly preconditioned Newton's method

$$\|e_{NP1}^{(n+1)}\| \leq \frac{\gamma \sum_{i=1}^{m} \frac{(1+\rho_i)^2}{|\alpha_i|}}{C_m \, |\alpha_0| \sum_{i=1}^{m} \frac{1}{|\alpha_i|}} \|e_{NP1}^{(n)}\|^2 \leq \frac{\gamma}{|\alpha_0|} \frac{(1+\rho_{max})^2}{C_m} \|e_{NP1}^{(n)}\|^2,$$

provided that $u^{(0)} \in B_{\epsilon_{NP1}}(u)$, where

$$\epsilon_{NP1} = \min\left(r, \frac{|\alpha_0|}{2\gamma} \frac{C_m}{(1+\rho_{max})^2}\right).$$

These can be compared directly with (3.1), unfortunately to the detriment of NP1. To obtain a sharper estimate, we believe that it is necessary to restrict the class of PDEs. Note if $f' \equiv 0$, then $\rho_{max} = 1$ while if $f' \leq 0$, then $\rho_{max} \leq C$ for some constant C. With a suitable finite element discretization, C is independent of the mesh size but can increase with the number of subdomains.

It is not difficult to deduce similar estimates for NP0, the original scheme of Cai and Keyes:

$$\|e_{NP0}^{(n+1)}\| \leq K\left(1 + \gamma \sum_{i=1}^{m} \beta_i (1+\rho_i)\rho_i\right) \|e_{NP0}^{(n)}\|^2$$

for some constant K.

We have examined two other r-quadratic convergence theories that are similar to the first theory above. They differ in the Lipschitz condition ([19], [21]) or the assumption that $F'(u^{(0)})$ is invertible (rather than $F'(u)$) ([8], [20]). The results of the analysis are similar to those of the first theory and will be reported elsewhere.

3.2. r-quadratic convergence. In this theory due to Smale [16], we no longer assume a Lipschitz condition in a ball. Instead, all assumptions are at the initial point $u^{(0)}$ of the iteration. However, we need to assume that F is an analytic operator. Define

$$\omega(u^{(0)}) = \|F'(u^{(0)})^{-1} F(u^{(0)})\| \sup_{j>1} \left\| \frac{F'(u^{(0)})^{-1} F^{(j)}(u^{(0)})}{j!} \right\|^{\frac{1}{j-1}},$$

where $F^{(j)}$ denotes the jth derivative of F. If $\omega(u^{(0)}) < \omega_0 = .13\cdots$ which is a universal constant, then Newton's method for $F(u) = 0$ with initial guess $u^{(0)}$ converges quadratically in the manner

$$\|u^{(n)} - u\| \leq \left(\frac{1}{2}\right)^{2^{n-1}} \frac{7\,\|u^{(1)} - u^{(0)}\|}{4}.$$

This theory is extremely interesting. It is more practical in the sense that no Lipschitz condition in a region is necessary. However, the computation of ω can be a daunting task. For some problems, the nonlinearity is quadratic (Navier-Stokes equations, for instance) and the supremum in the definition of ω is taken over $j = 2$ only.

For Newton's method applied to our semilinear elliptic PDE,

$$\omega(u^{(0)}) \leq \| \, [I + \triangle^{-1} f'(u^{(0)})]^{-1} \, [u^{(0)} + \triangle^{-1} f(u^{(0)})] \, \| \sup_{j>1} \left\| \frac{f^{(j)}(u^{(0)})}{|\alpha_0(u^{(0)})| \, j!} \right\|_{-1}^{\frac{1}{j-1}}$$

which can usually be worked out in practice. However, for the nonlinear preconditioned Newton's method,

$$\omega(u^{(0)}) = \|\mathcal{F}'(u^{(0)})^{-1} \mathcal{F}(u^{(0)})\| \sup_{j>1} \left\| \frac{\mathcal{F}'(u^{(0)})^{-1} \mathcal{F}^{(j)}(u^{(0)})}{j!} \right\|^{\frac{1}{j-1}}$$

	N	NP0	NP1	NP2
f_1	19, 19	5, 7	4, 6	4, 9
f_2	F, F	8, F	6, 6	6, 6
f_3	12, 12	6, 6	4, 5	4, 5
f_4	8, 8	F, F	4, 4	4, 4
f_5	40, 40	F, F	F, F	F, F
f_6	15, 15	9, 7	8, 6	8, 5

Table 4.1: Comparison of the number of Newton iterations to convergence. F denotes not converged after 100 iterations. The first entry of each pair refers to the number of iterations for an overlap of one point while the second refers to that for an overlap of 10 points.

and we are unable to give a more explicit expression.

Another related result in [16] states that if

$$\|e_N^{(0)}\| < \chi, \qquad \chi \equiv \frac{3-\sqrt{7}}{2}\left(\sup_{j>1}\left\|\frac{F'(u)^{-1}F^{(j)}(u)}{j!}\right\|^{\frac{1}{j-1}}\right)^{-1}, \tag{3.2}$$

then the Newton iteration converges r-quadratically:

$$\|e_N^{(n)}\| \le \left(\frac{1}{2}\right)^{2^{n-1}} \|e_N^{(0)}\|.$$

We shall evaluate χ numerically in the next section.

4. Numerical Experiments and Discussions. We have performed some numerical experiments in MATLAB to solve two-point boundary value problems of the form

$$-u'' = f(x, u, u') \text{ on } (0,1) \tag{4.1}$$

with homogeneous Dirichlet boundary conditions. The ODEs are discretized using the usual second-order finite difference scheme with step size $h = 1/160$ and the resultant nonlinear equations are solved using four methods: classical Newton's method (N), and the three variations of the nonlinearly preconditioned Newton's methods NP0, NP1, and NP2. For the latter three, the domain is split into two overlapping subdomains. Two domain decompositions were tested: one with an overlap of one grid point and the other with an overlap of 10 grid points. Throughout, we employ Newton's method (rather than an inexact Newton's method in [3]) and a simple backtracking algorithm where the length of the Newton step is halved until a sufficient decrease in the residual (Algorithm 6.3.5 in [4]). ([3] uses cubic backtracking.) For a fair comparison, all methods use the same stopping criteria: the nonlinear residual $\|v'' + f(x, v, v')\|_{L^2} < 10^{-8}$ and the L^2-norm of the Newton step is smaller than $h^2 \approx 4 \times 10^{-5}$. (It would be more natural for the nonlinearly preconditioned methods to base the stopping criteria on $\mathcal{F}$ rather than on F.) The initial iterate is always the zero function.

We display the results for six functions

1. $f_1 = (10\sin(10x) - u^3u')/.02$;

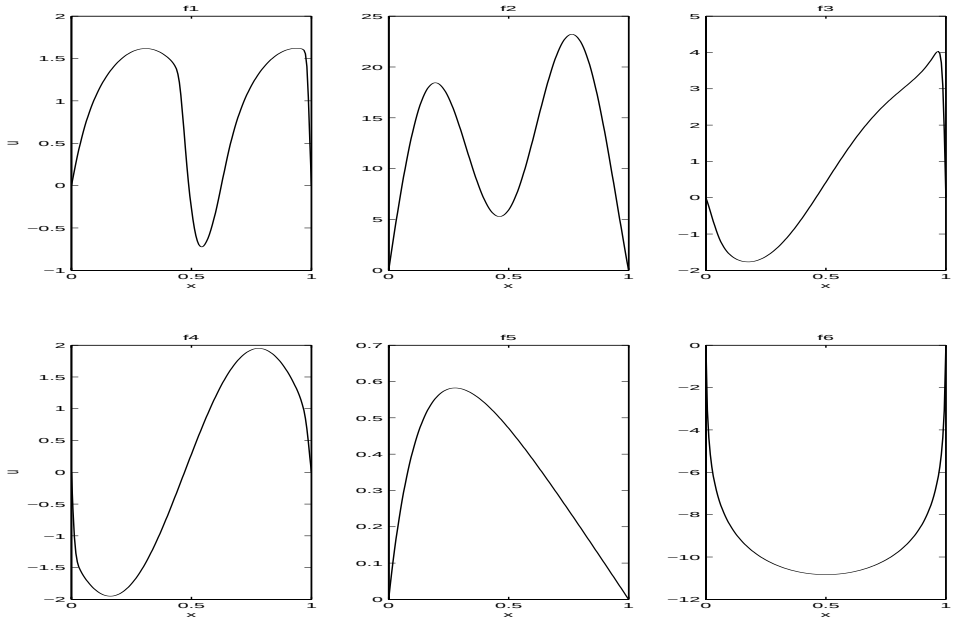


Figure 4.1: Solutions of boundary value problems

2. $f_2 = 100e^{\cdot 1u/(1+u'^2)} + 1000\sin(10x)$;

3. $f_3 = (10\cos(10x) - uu' + e^u)/.03$;

4. $f_4 = (10\cos(10x) - uu')/.01$;

5. $f_5 = \left[u'^2 - u + 1 - \left(\frac{e^{\frac{-x}{\sqrt{\epsilon}}}}{\sqrt{\epsilon}} + e^{\frac{-1}{\sqrt{\epsilon}}} - 1\right)^2 + x(e^{\frac{-x}{\sqrt{\epsilon}}} - 1)\right]/\epsilon, \qquad \epsilon = .02$;

6. $f_6 = -10^6 e^u$;

whose solutions are illustrated in Figure 4.1. Actually we tried other functions too. Most of them were too easy and all four methods converged rapidly. Table 4.1 shows the number of Newton iterations for the different methods and functions. Tables 4.2 tabulates the average number of GMRES iterations to solve each global linear system (ignoring the number of GMRES iterations in solving nonlinear subdomain problems). The number of Newton iterations to solve each nonlinear subdomain ODE is typically four or five. Figure 4.2 shows the convergence history of the methods for f_1.

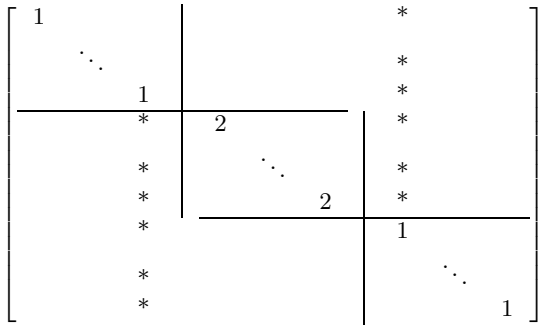
For f_2, the Newton iteration failed to converge after 100 iterations. The residuals decreased at an extremely slow rate. NP0 also failed to converge here (for an overlap of ten) as well as failing for f_4 due to non-convergence of the nonlinear subdomain ODE solver. Here, the algorithm neglects the most up-to-date data ($T_i(u^{(0)})$) causing one iterate to stray too far away. For f_5, Newton's method had some difficulty but eventually converged while all three nonlinearly preconditioned methods failed. The cause of the failures was that the Newton iteration for the subdomain nonlinear equation did not converge, mainly because the initial iterate is too far from the exact solution. Note that Newton's method fails to converge if the

	N	NP0	NP1	NP2
f_1	3.0, 4.0	3.0, 4.0	3.0, 4.0	3.0, 4.0
f_2	F, F	3.0, F	3.0, 4.0	3.0, 4.0
f_3	3.0, 4.0	3.0, 4.0	3.0, 4.0	3.0, 4.0
f_4	3.0, 4.0	F, F	3.0, 4.0	3.0, 4.0
f_5	3.0, 4.0	F, F	F, F	F, F
f_6	2.5, 2.9	2.9, 2.9	3.0, 3.0	3.0, 3.0

Table 4.2: Comparison of the average number of GMRES iterations per Newton step.

constant .02 in f_5 is replaced by .01. Except for f_5, NP1 and NP2 converge with between 1/2 and 1/4 of the iterations required by Newton's method.

In general, the number of GMRES iterations increases from three to four as the overlap increases from one to ten. This can be explained as follows. The matrix approximation of $\mathcal{F}'$ has a rather simple structure:



with non-zero diagonal entries plus two non-zero columns indicated by $*$. Note that the middle block corresponds to the unknowns in the overlapping region. This matrix has at most four distinct eigenvalues, including 1 and 2. Thus GMRES converges in at most four iterations. In the special case that the overlap is one, the the middle block does not appear and so the matrix has a 2×2 block structure and has at most three distinct eigenvalues, including 1. (Note that some authors ([3]) call this case the non-overlapping case.) We stress that this is independent of the step size h.

Next, we numerically evaluate the radii of quadratic convergence for the first convergence theory (q-quadratic convergence). We choose the ODE

$$-u'' = f(u), \qquad f(u) \equiv -\lambda(u+1)(u+2) \tag{4.2}$$

to facilitate this calculation. Initially, we take $\lambda = 100$. For this nonlinearity, γ can be evaluated analytically, equal to $200/\pi^2$. Thus from (3.1),

$$\epsilon_N = \frac{|\alpha_0|\pi^2}{400}.$$

Note that r cannot be much larger than 1.7 because the Jacobian for N can become singular beyond this point. As for NP1, the matrix approximation of $\mathcal{F}'(u)$ is computed explicitly while the Lipschitz constant is estimated numerically. The interval $[0, 1]$ is divided into 160 subintervals in this calculation and overlaps of two and twenty points are considered.

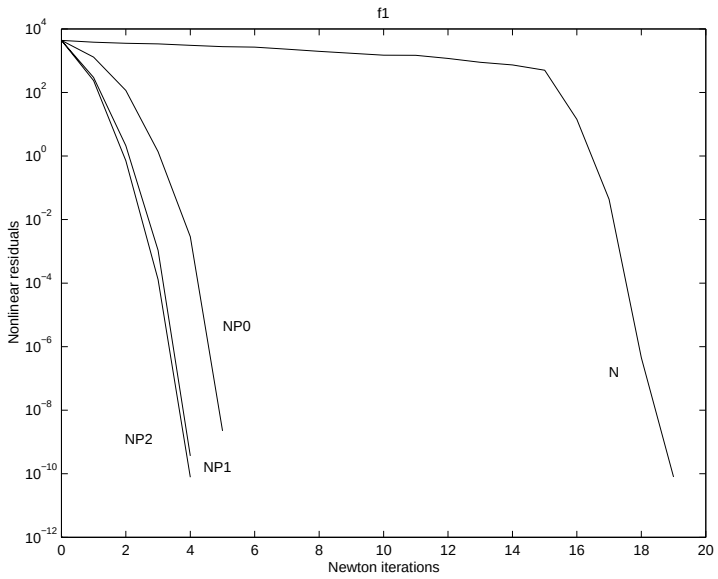


Figure 4.2: Convergence history.

The results are shown in the left diagram of Figure 4.3, which indicates that the radius of quadratic convergence of NP1 is larger than that of Newton's method. This observation should be viewed with some caution because these radii are only lower bounds for the true radii of quadratic convergence. It would be desirable to come up with a sharp upper bound of these radii for comparison.

We also repeated the calculation for $\lambda = 1$ (right diagram in Figure 4.3). Note that for a small overlap, the radius of quadratic convergence of NP1 is actually smaller than that of Newton's method. Any theory must take this into account.

Finally, we report on numerical evaluations of some quantities in Smale's theory for (4.2) with $\lambda = 100$. The main difficulty is in the computation of the supremum term in χ. Currently, we compute all terms up to $j = 20$ in (3.2) and then extrapolate the result (a least squares fit of a rational function) to infinity, a highly speculative process! We obtain $\chi \approx .3$ for NP1 in contrast with the corresponding value of χ for N which is .07. Thus, the estimated radius of quadratic convergence of NP1 is four to five times larger than that of the classical Newton's method. For $\lambda = 1$, the results are qualitative similar, in contrast with the first convergence theory. This may indicate the result of the first theory is not as sharp as Smale's.

Based on these limited experiments, the classical Newton's method does well. Note that each iteration of a nonlinearly preconditioned method costs about twice as much as one iteration of a classical Newton's method in terms of execution time because of the extra nonlinear subdomain solves. NP1 and NP2 are better than NP0 in terms of both speed and robustness. Assuming a parallel computing environment where each processor is assigned to a subdomain, then the addition of y_i in (2.4) involves no communication while replacing y_i by y (NP2) entails communications with all adjacent neighbors. This should not be of much concern since y has to be formed anyway because it is the nonlinear residual $\mathcal{F}(u^{(n)})$. Clearly, many more numerical experiments on nonlinear PDEs are necessary before any definitive conclusion can be reached.

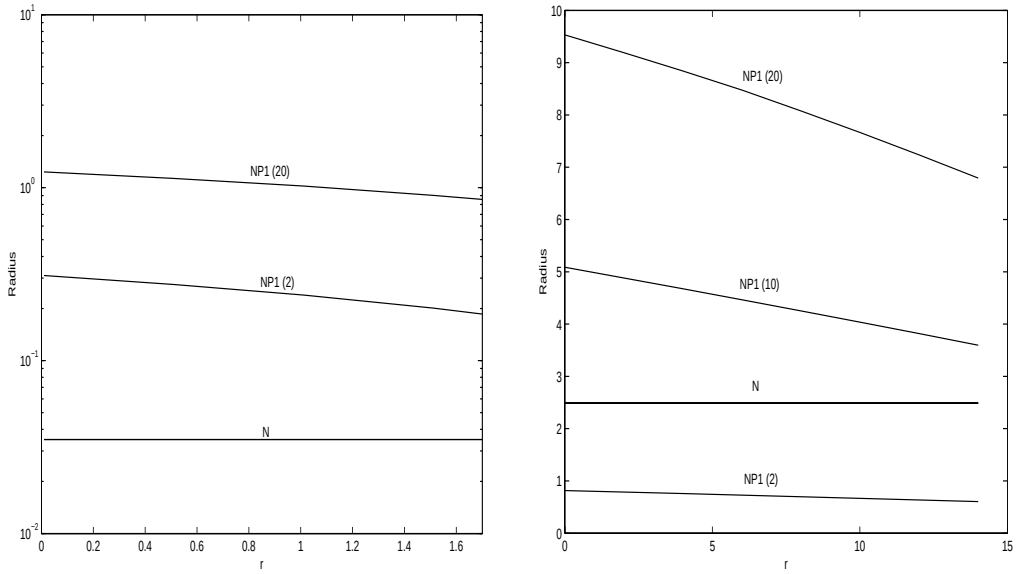


Figure 4.3: Numerical evaluation of ϵ_N and ϵ_{NP1}, the radii of q-quadratic convergence for N and NP1 (with overlap of two, ten, and twenty points) for $u^{(0)} \in B_r(u)$ and $\lambda = 100$ (left), $\lambda = 1$ (right).

While nonlinearly preconditioned Newton's methods are undoubtedly more robust for some problems, they can breakdown when the classical Newton's method works. The main reason is they require the solution of nonlinear subdomain problems which typically involves another Newton's iteration where there is a chance of non-convergence. This can be due to the lack of a good initial guess or may be the subdomain nonlinear problem has no solution or multiple solutions! It is not difficult to write down specific examples where NP1 will fail in the first iteration. This will be discussed in a future report.

REFERENCES

[1] X. C. Cai and M. Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In D. Keyes and J. Xu, editors, *Domain decomposition methods in scientific and engineering computing*, pages 335–360. AMS, 1994.

[2] X. C. Cai, W. D. Gropp, D. E. Keyes, R. G. Melvin, and D. P. Young. Parallel Newton-Krylov-Schwarz algorithms for the transonic full potential equation. *SISC*, 19:245–265, 1998.

[3] X. C. Cai and D. Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SISC*, 24:183–200, 2002.

[4] J. E. Dennis, Jr and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[5] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, pages 296–311, 1997.

[6] M. Dryja and O. B. Widlund. An additive variant of the Schwarz alternating method for the case of many subregions. Technical Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, 1987.

[7] R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors. *Domain Decomposition Methods for Partial Differential Equations*, Philadelphia, PA, 1988. SIAM. Proceedings of

the First International Symposium on Domain Decomposition Methods for Partial Differential Equations, Paris, France, January 1987.

[8] L. V. Kantorovich and G. P. Akilov. *Functional Analysis*. Pergamon Press, New York, 1982.

[9] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. Frontiers in applied mathematics. SIAM, 1995.

[10] P.-L. Lions. On the Schwarz alternating method. I. In R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42, Philadelphia, PA, 1988. SIAM.

[11] S. H. Lui. On Schwarz alternating methods for nonlinear elliptic PDEs. *SISC*, 21:1506–1523, 2000.

[12] S. H. Lui. On Schwarz alternating methods for the incompressible Navier-Stokes equations. *SISC*, 22:1974–1986, 2001.

[13] S. H. Lui. On linear monotone and Schwarz methods for nonlinear elliptic PDEs. *Numer. Math.*, 2002.

[14] A. M. Matsokin and S. V. Nepomnyaschikh. A Schwarz alternating method in a subspace. *Soviet Mathematics*, 29(10):78–84, 1985.

[15] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Company, 1996.

[16] S. Smale. Newton's method estimates from data at one point. In R. Ewing, K. Gross, and C. Martin, editors, *The Merging of Disciplines: New Directions in Pure, Applied and Computational Math*, pages 185–196, NY, 1986. Springer.

[17] X.-C. Tai and M. S. Espedal. Rate of convergence of some space decomposition method for linear and non-linear elliptic problems. *SIAM J. Numer. Anal.*, 35:1558–1570, 1998.

[18] X. C. Tai and J. Xu. Global convergence of subspace correction methods for convex optimization problems. *Math. Comput.*, 2001.

[19] J. Traub and H. Wozniakowski. Convergence and complexity of Newton iteration. *J. Assoc. Comput. Math.*, 29:250–258, 1979.

[20] X. Wang. Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comp.*, 68:169–186, 1999.

[21] X. Wang. Convergence of Newton's method and uniqueness of the solution of equations in Banach space. *IMA J. Numer. Anal.*, 20:123–134, 2000.