Numerical Algorithms 0 (2000) 1–15

Newton Iteration for Partial Differential Equations and the Approximation of the Identity

Gregory E. Fasshauer [a], Eugene C. Gartland, Jr. [b,*] and Joseph W. Jerome [c,**]

[a] *Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616*
E-mail: fass@amadeus.math.iit.edu
[b] *Department of Mathematics and Computer Science, Kent State University, Kent, OH 44242*
E-mail: gartland@mcs.kent.edu
[c] *Department of Mathematics, Northwestern University, Evanston, IL 60208*
E-mail: jwj@math.nwu.edu

It is known that the critical condition which guarantees quadratic convergence of approximate Newton methods is an approximation of the identity condition. This requires that the composition, of the numerical inversion of the Fréchet derivative with the derivative itself, approximate the identity to an accuracy calibrated by the residual. For example, the celebrated quadratic convergence theorem of Kantorovich can be proven when this holds, subject to regularity and stability of the derivative map. In this paper, we study what happens when this condition is not evident 'a priori' but is observed 'a posteriori'. Through an in-depth example involving a semilinear elliptic boundary value problem, and some general theory, we study the condition in the context of dual norms, and the effect upon convergence. We also discuss the connection to Nash iteration.

Keywords: Newton methods, partial differential equations, approximation of the identity, Nash iteration

AMS Subject classification: Primary 47J25, 65J15; Secondary 65N30

1. Introduction

In previous works [11,7,8], we have studied Nash iteration, which consists of an approximate Newton iteration, followed by smoothing, as part of an iterative procedure to solve the operator equation,

$$\mathbf{F}(u) = \mathbf{0}.$$

The smoothing was intended to deal with the case when a finite dimensional numerical linear operator $\mathbf{G}_h$ is employed to approximate $\mathbf{G}$, where the latter represents the actual inverse of $\mathbf{F}'$.

* Research supported by the National Science Foundation under grant DMS-9870420.
** Research supported by the National Science Foundation under grant DMS-9704458.

The reason that smoothing is seen as theoretically necessary is due to standard arguments in approximation theory. For example, if operator derivatives are computed at a point v in a Banach space, if w is arbitrary, then each differentiation of

$$[\mathbf{G}(v) - \mathbf{G}_h(v)]w$$

leads to a loss of order one in the convergence order. In particular, the approximation of the identity,

$$[\mathbf{F}'(v)\mathbf{G}_h(v) - I]w = [\mathbf{F}'\mathbf{G}_h(v) - \mathbf{F}'\mathbf{G}(v)]w$$

is of order $O(1)$ in standard norms, and thus experiences a loss of derivatives. Classical theories, such as the Kantorovich theory [10], are based on use of an approximation of the identity of the order of the residual, which translates in numerical methods to a polynomial function of the grid size, chosen adaptively according to the current residual. The present study deals with the effective implementation of approximate Newton methods for a differential map, defined via a numerical discretization. We study, in the process, the approximation of the identity condition with respect to dual norms, and discuss the impact upon convergence.

There is an alternative route, which we do not explicitly explore in this paper. This is to make use of the compact fixed point map, via the Krasnosel'skii calculus developed in [14]. This calculus, which is a strict logical extension of the Babuška-Aziz inf-sup theory (see [1]), has been shown to admit Newton linearization of the fixed point map which does *not* encounter a loss of derivatives (see [12] for details). Naturally, the major obstacle to this approach is at the level of numerical linear algebra. Sparse matrix computations are familiar when linearization is based upon the differential map. However, issues related to computational complexity can still be circumvented with this approach (see [13]) by use of Arnoldi iteration.

One of the motivating ideas in this work is the integration of the multigrid method [9] of solving the finite element equations with near optimal computational complexity. In some sense, we may view this as the inner iteration of a strategy which attempts outer superlinear or quadratic convergence via approximate Newton iteration.

2. A Partial Differential Equation Example

We illustrate some of these concepts in the context of a model semilinear elliptic partial differential equation in two dimensions. Our numerical solution is via a conventional finite-element method.

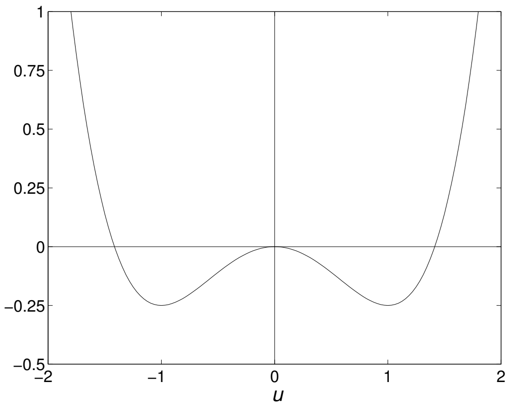

Figure 1. Profile of the double-well potential $g(u) = -u^2/2 + u^4/4$.

2.1. Model Problem

Consider the "energy functional"

$$J(u) := \int_\Omega \left\{ \frac{\varepsilon^2}{2}|\nabla u|^2 - \frac{1}{2}u^2 + \frac{1}{4}u^4 \right\}, \quad \Omega := (0,1)^2.$$

We seek minimizers of J over a class of admissible functions that are sufficiently smooth and which vanish on the boundary $\partial\Omega$:

$$\min_{u \in V} J(u), \quad V := H_0^1(\Omega).$$

This problem is a simple prototype of models arising out of the Landau theory of phase transitions (e.g., the Ginzburg-Landau model for superconductivity [5] and the Landau-de Gennes model for liquid crystals [4]).

The symmetric double-well potential $g(u) := -u^2/2 + u^4/4$ has minima at $u = \pm 1$ and a relative maximum at $u = 0$. It is depicted in Figure 1. The energy $J(u)$ expresses a competition between the "elastic energy" density $|\nabla u|^2$ (which penalizes spatial variations in u) and the potential $g(u)$ (which tries to coerce u to take on the values 1 or -1); the boundary condition imposes $u = 0$ on the boundary of the square. The factor ε^2 is a "coupling coefficient"; it controls the relative strength of the influences of the elastic versus the potential terms. The parameter ε gives a *length scale* for the pointwise changes in u.

Stationary points solve the variational problem (*weak formulation*)

$$J'(u)v = \int_\Omega \left\{ \varepsilon^2 \nabla u \cdot \nabla v + \left(u^3 - u\right) v \right\} = 0, \quad \forall v \in V.$$

Smooth critical points satisfy the *strong formulation* (Euler-Lagrange equation)

$$-\varepsilon^2 \nabla^2 u - u + u^3 = 0, \quad \text{in } \Omega$$

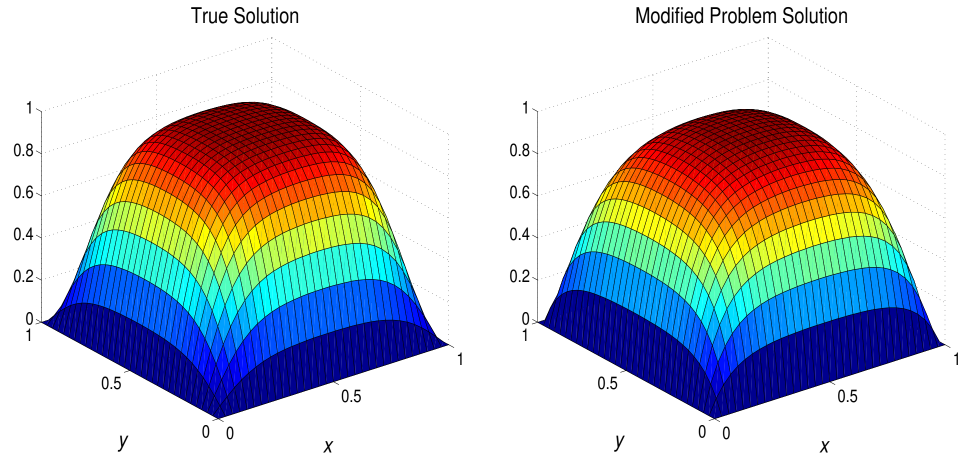

Figure 2. Surface plots of the true solution of the original problem ((1), accurate numerical approximation, left) versus the "qualitatively correct" solution of the modified problem ((2), known analytical solution, right), both for $\varepsilon = .1$.

$$u = 0, \quad \text{on } \partial\Omega. \tag{1}$$

The constant function $u = 0$ is always a stationary point; it is locally stable for

$$\varepsilon > \varepsilon_c := \frac{1}{\sqrt{2}\,\pi} \doteq .225$$

and unstable for $\varepsilon < \varepsilon_c$.

The problem enjoys a reflection symmetry: if u is a solution, then so is $-u$. Thus we expect for small ε to have two stable solutions, one satisfying $u \approx 1$ in the interior and another its mirror image ($u \approx -1$ in the interior). For ε small, the problem is also *singularly perturbed*, with $O(\varepsilon)$ boundary layers near the edges.

An analytical solution to this problem is not known, and so for some of our numerical experiments we utilize the *modified equation*

$$-\varepsilon^2\nabla^2 u - u + u^3 = f \tag{2}$$

with the source term f constructed so that the solution is given by

$$u(x, y) = \psi(x)\psi(y),$$

where ψ is defined

$$\psi(t) := 1 + \exp(-1/\varepsilon) - \exp(-t/\varepsilon) - \exp((t-1)/\varepsilon).$$

While this solution is artificial, it does capture well the qualitative features of the true solution of the original problem, as is illustrated in Figure 2. The two surface plots in this figure are barely distinguishable visually. They differ (in terms of absolute errors) by roughly 3% in the interior, 5% near the edges, and 8% near the corners.

The solution of the modified problem will not capture subtler aspects of the true solution, such as singularities of higher derivatives in the corners. However, it enables us to calculate errors in our numerical approximations and to study the convergence of the global inexact Newton iteration.

2.2. Numerical Approximation

We wish to explore the convergence properties of a *global* Newton iteration (taking place in the infinite-dimensional function space V)—as opposed to Newton linearization of a *fixed* discretization of our problem. Thus we require a discretization procedure that produces approximate problems that can be interpreted in the setting of the infinite-dimensional function space V and approximate solutions that reside in V. We then can (and shall) use various norms of V to analyze the convergence properties of the resulting approximate Newton iteration. For this purpose, we are naturally led to consider *projection methods* (Galerkin, spectral, collocation, ...). We have chosen to implement a conforming finite-element method, with bilinear rectangular elements ("type (1) rectangles" in the terminology of Ciarlet [3, §2.2]). Thus at each stage, our approximate solution and Newton correction (on any mesh) will both always reside in our space of admissible functions.

We interpret our operator equation in the *weak sense*, in the dual space $V' = H^{-1}(\Omega)$ of bounded linear functionals on V ($= H_0^1(\Omega)$), that is, the nonlinear residual functional $\mathbf{F} := J'$ satisfies $\mathbf{F}(u) = 0$ in V', or

$$\mathbf{F}(u)\varphi = \int_\Omega \left\{ \varepsilon^2 \nabla u \cdot \nabla \varphi + \left(u^3 - u\right) \varphi \right\} = 0, \quad \forall \varphi \in V.$$

Given an approximate solution u, the Newton correction v solves the linearized problem $\mathbf{F}'(u)v = -\mathbf{F}(u)$, or

$$\int_\Omega \left\{ \varepsilon^2 \nabla v \cdot \nabla \varphi + \left(3u^2 - 1\right) v\varphi \right\} = \int_\Omega \left\{ -\varepsilon^2 \nabla u \cdot \nabla \varphi + \left(u - u^3\right) \varphi \right\}, \quad \forall \varphi \in V.$$

The Galerkin approximation v_h to this correction is from the finite-element subspace $V_h \subset V$ of piecewise bilinear polynomials (span$\{1, x, y, xy\}$, represented in terms of nodal values) on a mesh of $n \times n$ uniform squares of edge lengths $h = 1/n$ (which vanish on the boundary $\partial\Omega$). It satisfies $\mathbf{F}'(u)v_h = -\mathbf{F}(u)$ in V_h, or

$$\int_\Omega \left\{ \varepsilon^2 \nabla v_h \cdot \nabla \varphi_h + \left(3u^2 - 1\right) v_h \varphi_h \right\} = \int_\Omega \left\{ -\varepsilon^2 \nabla u \cdot \nabla \varphi_h + \left(u - u^3\right) \varphi_h \right\}, \quad \forall \varphi_h \in V_h. \tag{3}$$

A further approximation employs a nodal quadrature rule:

$$\int_0^h \int_0^h g(x,y)\,dx\,dy \approx \frac{h^2}{4}\left[g(0,0)+g(h,0)+g(0,h)+g(h,h)\right].$$

This quadrature rule is natural and convenient for our application. It satisfies the conditions of the "Strang Lemma" [3, §4.1], and so it is sufficient for the standard finite-element convergence results for this element. This leads to the linear system of equations for the interior nodal values of v_h

$$-\varepsilon^2 \nabla_h^2 v_{ij}^h + \left(3u_{ij}^2 - 1\right) v_{ij}^h = \varepsilon^2 \nabla_h^2 u_{ij} + u_{ij} - u_{ij}^3, \quad i,j = 1,\ldots,n-1. \tag{4}$$

Here ∇_h^2 denotes the standard *discrete Laplacian.* Notice that this is precisely the same system of equations one obtains from standard finite differencing of the strong form of the equation for the Newton correction

$$-\varepsilon^2 \nabla^2 v + \left(3u^2 - 1\right) v = \varepsilon^2 \nabla^2 u + u - u^3.$$

2.3. Mesh Refinement and Attainable Global Convergence Rates

Let u_h denote the Galerkin approximate solution from V_h ($\mathbf{F}(u_h)\varphi_h = 0$, $\forall \varphi_h \in V_h$), and denote the error $e_h := u - u_h$. Finite-element theory predicts best possible approximation errors of $\|e_h\|_{H^1} = O(h)$ and $\|e_h\|_{L^2} = O(h^2)$ for general $u \in H^2(\Omega) \cap H_0^1(\Omega)$ for the linear positive-definite symmetric problem. We expect similar rates for our problem. Thus to obtain a globally convergent Newton iteration requires successive mesh refinement, and the maximum attainable convergence rates of the Newton iteration will depend on the rate of mesh refinement.

For our iteration, we use a nested sequence of meshes and proceed as follows. We start from an initial mesh h_0 and approximate solution $u_0 \in V_{h_0}$. At each stage, we select a refined mesh h_k, interpolate the nodal values of our current approximate solution u_k onto the finer grid, solve the linearized equations (4) for the corrections v_k, and update ($u_{k+1} = u_k + v_k$). The nesting of the meshes leads to an embedding of the finite-element subspaces: $V_{h_0} \subset V_{h_1} \subset \cdots$. So the interpolation merely "fills in" values for the linear-system solver; no additional approximation errors are incurred.

Any mesh sequence $\{h_k\}$ that converges to zero can in principle lead to a convergent global Newton iteration; however the rate of convergence can be arbitrarily slow. Successive mesh halving should be capable of producing q-linear convergence (cf., [15, Ch. 9]):

$$h_k = \frac{h_{k-1}}{2} \;\Rightarrow\; \frac{e_k}{e_{k-1}} \approx \frac{h_0/2^k}{h_0/2^{k-1}} = \frac{1}{2} < 1.$$

Table 1
Growth of n for different mesh refinements. Problem size is $O(n^2)$. The final values in each column represent sizes beyond the capabilities of our code/workstation.

k	halving $n_k = 2n_{k-1}$	factorial $n_k = (k+1)n_{k-1}$	$p = 3/2$ $n_k = \lceil n_{k-1}^{3/2} \rceil$	$p = 2$ $n_k = n_{k-1}^2$
0	2	2	2	2
1	4	4	3	4
2	8	12	6	16
3	16	48	15	256
4	32	240	59	*65,536
5	64	1,440	454	
6	128	*10,080	*9,674	
7	256			
8	512			
9	1,024			
10	2,048			
11	*4,096			

An accelerated mesh refinement is needed to enable q-superlinear convergence, for example, "factorial refinement":

$$h_k = \frac{h_{k-1}}{k} \quad \Rightarrow \quad \frac{e_k}{e_{k-1}} \approx \frac{h_0/k!}{h_0/(k-1)!} = \frac{1}{k} \to 0, \quad \text{as } k \to \infty.$$

Such convergence will be q-superlinear but not q-order p for any $p > 1$.

To obtain q-order p ($p > 1$) requires $h_k = O(h_{k-1}^p)$. Thus to obtain q-quadratic convergence (of which Newton is capable) requires successive mesh *squaring*. Such refinements quickly lead to intractable problem sizes. We illustrate this in Table 1. We have performed extensive numerical experiments utilizing the mesh-halving and factorial refinements, as well as some with mesh squaring. Refinements involving non-integral powers p (e.g., $p = 3/2$ in the table) are not convenient for *nested* meshes. A difficulty with the mesh-squaring refinement ($p = 2$ in the table) is that it does not provide a sufficient number of iterative steps to make inferences about the asymptotic convergence behavior.

2.4. Numerical Experiments

To solve the linear system for the Newton correction, we have implemented a Conjugate Gradient algorithm preconditioned by a fast Poisson solver (which utilizes fast discrete Fourier sine transforms, calculated using NAG Library routine `C06HAF`). This was relatively easy to implement and adequate to solve our largest problems ($n = 2,048$) in an acceptable amount of time (on the order of a few minutes). The core image of the largest calculation was around 300 megabytes in execution. All experiments were done with the parameter value $\varepsilon = .1$.

Table 2
Errors and q-convergence factors ($q = \|e_k\|/\|e_{k-1}\|$) for Galerkin approximations on levels associated with *mesh-halving* refinement.

k	n	$\|e_k\|_{L^2}$	q	$\|e_k\|_{H^1}$	q
1	2	.357(−1)		.584(1)	
2	4	.413(−1)	1.16	.409(1)	.70
3	8	.255(−1)	.62	.229(1)	.56
4	16	.726(−2)	.28	.117(1)	.51
5	32	.182(−2)	.25	.584(0)	.50
6	64	.456(−3)	.25	.292(0)	.50
7	128	.114(−3)	.25	.146(0)	.50
8	256	.285(−4)	.25	.731(−1)	.50
9	512	.712(−5)	.25	.366(−1)	.50
10	1,024	.178(−5)	.25	.183(−1)	.50
11	2,048	.445(−6)	.25	.914(−2)	.50

Table 3
Errors and q-convergence factors ($q = \|e_k\|/\|e_{k-1}\|$) for Galerkin approximations on levels associated with *factorial* refinement.

k	n	$\|e_k\|_{L^2}$	q	$\|e_k\|_{H^1}$	q
1	12	.127(−1)		.155(1)	
2	24	.324(−2)	.26(0)	.779(0)	.50
3	72	.360(−3)	.11(0)	.260(0)	.33
4	288	.225(−4)	.63(−1)	.650(−1)	.25
5	1,440	.900(−6)	.40(−1)	.130(−1)	.20

k	n	$\|e_k\|_{L^2}$	q	$\|e_k\|_{H^1}$	q
1	16	.726(−2)		.117(1)	
2	32	.182(−2)	.25(0)	.584(0)	.50
3	96	.202(−3)	.11(0)	.195(0)	.33
4	384	.127(−4)	.63(−1)	.487(−1)	.25
5	1,920	.506(−6)	.40(−1)	.975(−2)	.20

2.4.1. Errors in the Modified Problem

For a benchmark, we have calculated the errors in the nonlinear Galerkin approximations on each fixed level (iterated to convergence) to the solutions of the modified problem (2). Tables 2, 3, and 4 contain this data. The predicted degrees of approximation are evident in these tables, with L^2-errors $O(h^2)$ and H^1-errors $O(h)$. The superlinear convergence rates are also consistent with expectations, with L^2-errors decaying like $1/k^2$ (q-factors 1/4, 1/9, 1/16, and 1/25) and H^1-errors like $1/k$ in Table 3. The mesh-squaring sequence (Table 4) is very short, but the data at least appear to be consistent with q-quadratic convergence.

Table 4
Errors and q-order-2 convergence factors ($q^2 = \|e_k\|/\|e_{k-1}\|^2$) for Galerkin approximations on levels associated with *mesh-squaring* refinement.

k	n	$\|e_k\|_{L^2}$	q^2	$\|e_k\|_{H^1}$	q^2
1	6	.355(−1)		.297(1)	
2	36	.144(−2)	1.14	.520(0)	.059
3	1,296	.111(−5)	.54	.144(−1)	.053

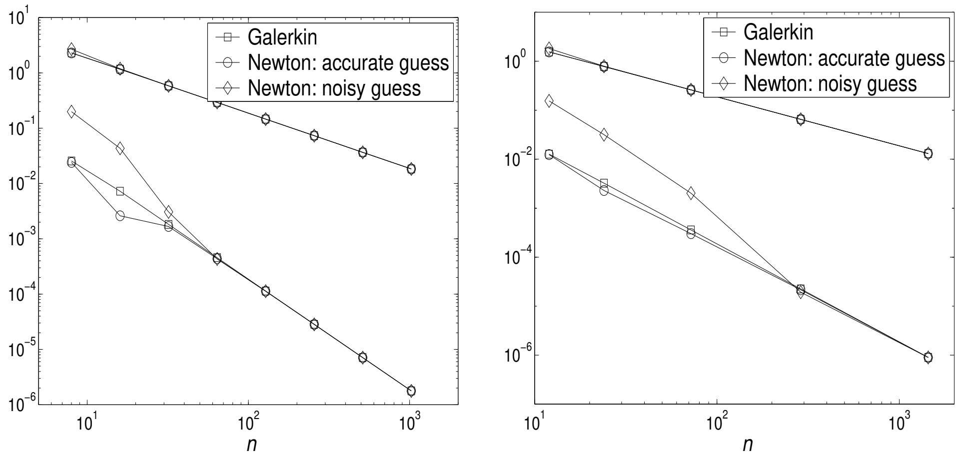

Figure 3. Errors in the global Newton iteration with *accurate initial guess* ("circles") versus *inaccurate initial guess* ("diamonds," noise level = .5) versus benchmark errors in the *Galerkin approximation* ("squares") on each level. Upper curves give $|e|_{H^1}$; lower curves $\|e\|_{L^2}$. *Mesh-halving* refinement (n = 8, 16, ..., 1024) left; *factorial* (n = 12, 24, 72, 288, 1440) right.

Next we performed global Newton iterations with these three types of refinements, to see if loss of derivatives (discussed in the Introduction above) provided any impediment to attaining optimal convergence rates. We started on relatively coarse meshes with initial guesses provided by the nodal values of the true analytical solution (of the continuous, modified problem (2)) with added random noise (of a controlled amplitude). Thus with the noise level set to zero, we have a very accurate initial guess; whereas with large noise levels, we are able to generate very inaccurate initial guesses with very rough initial errors.

The results of some typical experiments are presented in Figures 3 and 4. For the mesh-halving and factorial refinements (Figure 3), we see that the errors in the Newton iterates from the accurate initial guess remain very close to those of the Galerkin approximation on each level—they are even a little more accurate at step $k = 2$. The iterates from the noisy initial guess struggle for a little bit, but once they "home in," they closely track the Galerkin approximation as well. This is the case for both of these types of refinement.

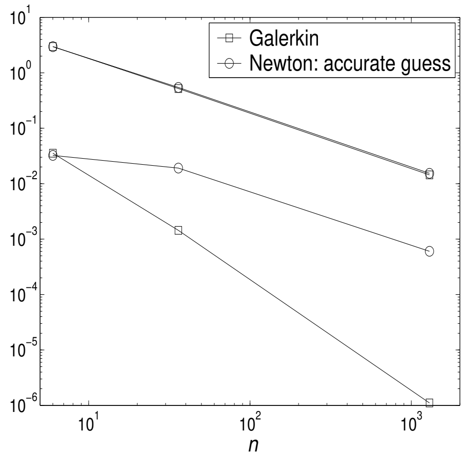

Figure 4. Errors in the global Newton iteration with *accurate initial guess* ("circles") versus the benchmark errors in the *Galerkin approximation* ("squares") on each level. Upper curves give $|e|_{H^1}$; lower curves $\|e\|_{L^2}$. *Mesh-squaring* refinement ($n = 6$, 36, 1296).

It is difficult to infer much from the data with mesh-squaring refinement (Figure 4). The L^2 errors in these three iterates are not that close yet to the optimal values; although the trends look good. An initial mesh of $n = 6$ is *very* coarse. Other refinement strategies behave similarly when started on such levels, struggling in the early going, yet homing in and tracking the optimal approximation eventually.

2.4.2. Corrections and Residuals in Original Problem

For our original problem (1), no analytical solution is known, and so we are not able to compute errors in our Newton approximants. Some indication of the convergence behavior in this case can be obtained by examining the *corrections* and the *residuals*. For the mesh-halving refinement, the corrections are quite "jittery" and rich in high-frequency content, as is illustrated in Figure 5. These oscillations do not grow and do not indicate an instability. They are reflective of the fact that the nodal errors in the approximate solution interpolated onto the finer mesh are highly oscillatory, with high-frequency grid modes excited in the bilinear interpolation—this phenomenon is relevant to multigrid methods and is discussed in that literature (cf., [2, Ch. 5] or [9]). Thus a highly oscillatory correction is needed to update the Newton approximation and produce an improved approximate solution, with a smooth error (which we observe). The other, accelerated refinements do not manifest this phenomenon of "jittery corrections" to such an extent, because for those cases, the wavelengths of the excited modes are longer (in terms of numbers of grid points on the refined meshes).

At any given stage, our approximate solution u_k is in H^1 but *not* in H^2; so our "residual" does not exist in a strong form, as an L^2 function, say, but only as

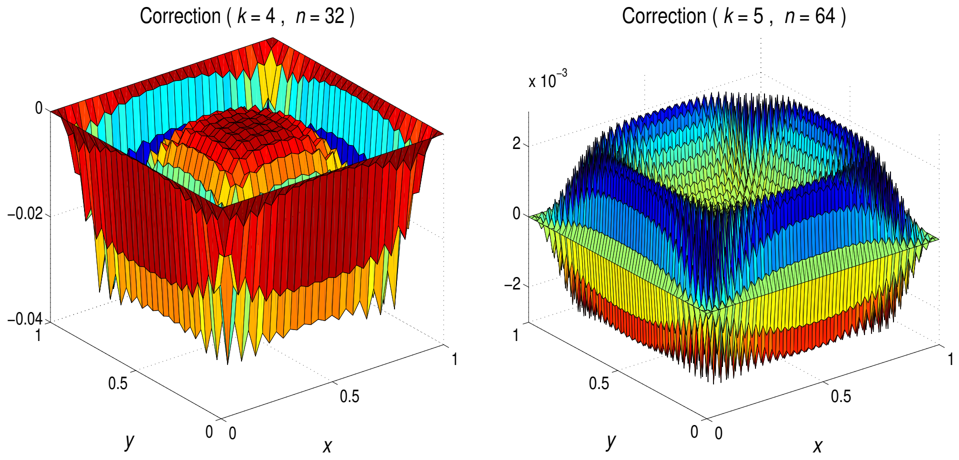

Figure 5. Newton corrections at iterations 4 and 5 ($n = 32$ and 64) for approximation of original problem (1) from an *accurate initial guess*. *Mesh-halving* refinement: $n = 4, 8, 16, 32, 64, \dots$.

an element of H^{-1}. However, a norm of this residual functional, $\mathbf{F}(u_k)$, is readily available, as we now show.

Define the u-dependent bilinear form

$$a_u(v, w) := \int_\Omega \left\{ \varepsilon^2 \nabla v \cdot \nabla w + \left(3u^2 - 1\right) vw \right\} . \tag{5}$$

Then (3) (the equation for the Galerkin approximation, $v_h \in V_h$, to the Newton correction) can be written

$$a_u(v_h, \varphi_h) = -\mathbf{F}(u)\varphi_h, \quad \forall \varphi_h \in V_h,$$

and we see that the (approximate) Newton correction at the k-th iteration, v_k, can be interpreted as the representer (in V_{h_k}) of the residual functional $-\mathbf{F}(u_k)$ with respect to the inner product $a_{u_k}(\cdot, \cdot)$. Using this norm, we then have

$$\|\mathbf{F}(u)\|^2 = \sup_{\varphi \in V} \frac{|\mathbf{F}(u)\varphi|^2}{a_u(\varphi, \varphi)} = \sup_{\varphi \in V} \frac{a_u(v, \varphi)^2}{a_u(\varphi, \varphi)} = a_u(v, v),$$

where v here is the true Newton correction, the true representer of $-\mathbf{F}(u)$ in V. So we have the convenient approximation $\|\mathbf{F}(u_k)\| \approx a_{u_k}(v_k, v_k)^{1/2}$.

Now the bilinear form $a_u(\cdot, \cdot)$ will be positive definite for *any* u if $\varepsilon > \varepsilon_c$. It should be positive definite for smaller ε's if u is sufficiently close to either of the $u \approx \pm 1$ stationary points, assuming these are indeed isolated local minimizers of J, as we expect (and observe). In these circumstances, $a_u(\cdot, \cdot)$ will be equivalent to the H^1 inner product, and the norm $\|\mathbf{F}(u)\|$ calculated above will be equivalent to $\|\mathbf{F}(u)\|_{H^{-1}}$.

Our numerical experiments (Tables 5 and 6) indicate an $O(h)$ behavior in this residual norm. We also observe a consistent superlinear convergence pattern

Table 5
Residual norms and q-convergence factors ($q = \|\mathbf{F}(u_k)\|/\|\mathbf{F}(u_{k-1})\|$) for global Newton approximant to solution of original problem ((1), no known analytical solution). *Accurate initial guess. Mesh-halving* refinement, left; *factorial* refinement, right.

	halving			factorial		
k	n	$\|\mathbf{F}(u_k)\|$	q	n	$\|\mathbf{F}(u_k)\|$	q
1	8	.261(−1)		16	.586(−1)	
2	16	.102(0)	3.920	32	.336(−1)	.574
3	32	.560(−1)	.548	96	.179(−1)	.534
4	64	.290(−1)	.517	384	.612(−2)	.341
5	128	.146(−1)	.504	1,920	.155(−2)	.253
6	256	.731(−2)	.501			
7	512	.366(−2)	.500			
8	1,024	.183(−2)	.500			

Table 6
Residual norms and q-convergence factors ($q = \|\mathbf{F}(u_k)\|/\|\mathbf{F}(u_{k-1})\|$) for global Newton approximant to solution of original problem ((1), no known analytical solution). *Accurate initial guess. Mesh-squaring* refinement.

k	n	$\|\mathbf{F}(u_k)\|$	q
1	6	.362(−1)	
2	36	.160(0)	4.410
3	1,296	.366(−1)	.229

for the factorial refinement here, with the q factors for $k = 3$, 4, and 5 approximately 1/2, 1/3, and 1/4. It is difficult to infer much from the mesh-squaring refinement (Table 6). The Newton approximations lose accuracy from $k = 1$ to $k = 2$ (from $n = 6$ to $n = 36$), but then so does the mesh-halving refinement started from a comparably coarse initial mesh ($n = 8$ to $n = 16$, left side of Table 5).

2.5. Approximation of the Identity

As is indicated in the Introduction, in the analysis of inexact Newton methods, an important role is played by the degree of "approximation to the identity"

$$\| \left(\mathbf{F}'(u)\mathbf{G}_h(u) - I\right) w\|.$$

In particular, in order to obtain a maximally convergent global inexact Newton iteration, it is sufficient that this be of the same relative order as the nonlinear residual $\|\mathbf{F}(u)\|$. We can assert this in the context of our present scheme if an appropriate mesh refinement is utilized, as we now sketch.

Given an approximate solution $u \in V = H_0^1(\Omega)$ and a linear functional $\Lambda \in V' = H^{-1}(\Omega)$, we wish to estimate

$$(\mathbf{F}'(u)\mathbf{G}_h(u) - I)\,\Lambda,$$

where $\mathbf{G}_h(u)$ denotes our approximate numerical inversion of the linearized operator $\mathbf{F}'(u)$. For this we will again find it expedient to use the u-dependent inner product $a_u(\cdot,\cdot)$, (5), of §2.4.2.

Let v denote the representer of Λ with respect to this inner product, defined by

$$a_u(v,\varphi) = \Lambda(\varphi), \quad \forall \varphi \in V,$$

and let v_h denote the Galerkin approximation to v from V_h, which is characterized by

$$a_u(v_h,\varphi_h) = \Lambda(\varphi_h), \quad \forall \varphi_h \in V_h.$$

Then $v_h = \mathbf{G}_h(u)\Lambda$ and gives the a_u-orthogonal projection of v onto V_h. We thus have

$$\Lambda_h := \mathbf{F}'(u)\mathbf{G}_h(u)\Lambda = \mathbf{F}'(u)v_h = a_u(v_h,\cdot),$$

and

$$(\Lambda - \Lambda_h)(\varphi) = a_u(v - v_h,\varphi), \quad \forall \varphi \in V.$$

So (as in §2.4.2) with respect to the dual norm in V' induced by $a_u(\cdot,\cdot)$, we have

$$\|(\mathbf{F}'(u)\mathbf{G}_h(u) - I)\,\Lambda\| = \|\Lambda_h - \Lambda\| = a_u(v - v_h, v - v_h)^{1/2} = O\left(\|v - v_h\|_{H^1}\right)$$

(assuming again that u is such that $a_u(\cdot,\cdot)$ is equivalent to the H^1 inner product), and we see that the approximation to the identity can be assessed via the Galerkin approximation to v.

In the context of our inexact Newton iteration, at the k-th stage, we have $\Lambda = -\mathbf{F}(u_k)$, where $u_k \in V_{h_{k-1}}$ is the approximation to the true solution u at the start of iteration k. Thus the representer v above is the *true* Newton correction (in V) corresponding to an *exact* Newton step from u_k; while v_h above corresponds to v_k in our algorithm, namely the calculated inexact/approximate Newton step, in the refined subspace $V_{h_k} \supset V_{h_{k-1}}$—recall that in our notation $u_{k+1} = u_k + v_k$. We thus have

$$\|(\mathbf{F}'(u_k)\mathbf{G}_h(u_k) - I)\,\mathbf{F}(u_k)\| = a_{u_k}(v - v_k, v - v_k)^{1/2} = O(\|v - v_k\|_{H^1}),$$

and we see that the approximation to the identity here corresponds to the approximation of the exact Newton step by the inexact/approximate Newton step, as measured in the $\|\cdot\|_{a_{u_k}}$ (or the equivalent $\|\cdot\|_{H^1}$) norm.

The error in this projection ($\|v - v_k\|_{H^1}$) can be made as small as we wish by refining our new mesh size, h_k, sufficiently. In fact, this error should be $O(h_k)$. To see this, note that although v will not be in $H^2(\Omega)$—the functional $\mathbf{F}(u_k)$

cannot be identified with an L^2 function—it will be in H^2 *locally*. It just suffers jump discontinuities in its normal derivatives across the element boundaries on $\Omega_{h_{k-1}}$, similar to those of u_k. We can thus obtain

$$\|v - v_k\|_{H^1(\Omega)} \leq C h_k \left\{ \sum_{i,j} |v|^2_{H^2(\Omega_{ij})} \right\}^{1/2},$$

where Ω_{ij} denotes the ij-th mesh cell on $\Omega_{h_{k-1}}$ (cf., [3, §3.2 and Remark 3.2.2]).

This gives us the control we need. In order to satisfy the quadratic-convergence hypotheses of Kantorovich-type theorems, we need

$$\frac{\|(\mathbf{F}'(u_k)\mathbf{G}_h(u_k) - I)\mathbf{F}(u_k)\|}{\|\mathbf{F}(u_k)\|} = O(\|\mathbf{F}(u_k)\|)$$

or

$$\|v - v_k\|_{H^1} = O(\|\mathbf{F}(u_k)\|^2),$$

which implies $h_k = O(\|\mathbf{F}(u_k)\|^2)$. This can be used adaptively, or if our numerical observation that $\|\mathbf{F}(u_k)\| = O(h_{k-1})$ can be taken as a gauge, then it is suggested that a mesh-refinement strategy taking $h_k = O(h_{k-1}^2)$ should produce global quadratic convergence. On the basis of this, we believe that this is what we would see with our mesh-squaring refinement were it not for the limitations of intractable problem sizes.

2.6. Conclusions

Because of the relatively slow convergence of low-order finite-element and finite-difference approximations such as this (which are polynomial order, $O(h^p)$, some p), it is extremely difficult to explore numerically any globally quadratic convergence behavior. With *spectrally* convergent schemes, however, the situation should be better.

3. Some General Comments

Our purpose has been to study the convergence properties of a global approximate Newton iteration in the context of a particular semilinear elliptic PDE boundary value problem. In the process, we have seen the role played by the *mesh-refinement sequence* in determining convergence rates of the Newton iteration. One can consider using this information and such an approach in the context of an inexact Newton method [16, §6.4] to solve similar problems on *fixed* (but "fine") meshes, appropriately tuning the mesh-refinement to the efficiency of the linear solver being used to calculate the Newton correction on each level.

In §2.5, a dual space approximation of the identity condition was derived under certain assumptions. A Kantorovich type theory is possible under the further conditions that the derivative is Lipschitz continuous and the derivative maps are (locally) uniformly bounded (see [10] for development). However, the Lipschitz continuity does not immediately follow without 'a priori' uniform pointwise estimates. This technicality can often be circumvented by taking the domain to be $H^1 \cap L_\infty$, but we shall not pursue that here. It has been our intent to justify the observed quadratic convergence of piecewise linear finite element methods by using 'a posteriori' analysis. Although we have not carried out smoothing experiments for this paper, we would expect some correlation with the experience of [8], at least for certain approximation schemes, such as collocation/radial basis function schemes. Although the issue of the necessity of smoothing is not yet completely resolved, we believe we have clarified the issue by the introduction of dual norms.

References

[1] I. Babuška and A.K. Aziz, Survey lectures on the mathematical foundations of the finite element method, in: *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, Academic Press, 1972, pp. 5–359.

[2] W.L. Briggs, *A Multigrid Tutorial*, SIAM, 1987.

[3] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.

[4] T.A. Davis and E.C. Gartland, Jr., Finite element analysis of the Landau-de Gennes minimization problem for liquid crystals, SIAM J. Numer. Anal. 35 (1998) 336–362.

[5] Q. Du, M.D. Gunzburger and J.S. Peterson, Analysis and approximation of the Ginzburg-Landau model of superconductivity, SIAM Rev. 34 (1992) 54–81.

[6] G.E. Fasshauer, Solving partial differential equations by collocation with radial basis functions, in: *Surface Fitting and Multiresolution Methods*, eds. A. Le Méhauté, C. Rabut and L.L. Schumaker, Vanderbilt University Press, 1997, pp. 131–138.

[7] G.E. Fasshauer and J.W. Jerome, Multistep approximation algorithms: Improved convergence rates through postconditioning with smoothing kernels, Advances in Comp. Math. 10 (1999) 1–27.

[8] G.E. Fasshauer, E.C. Gartland, Jr. and J.W. Jerome, Nash iteration as a computational tool for differential equations, J. Comp. Appl. Math., 119 (2000) 161–183.

[9] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer, 1985.

[10] J.W. Jerome, Approximate Newton methods and homotopy for stationary operator equations, Constr. Approx. 1 (1985) 271–285.

[11] J.W. Jerome, An adaptive Newton algorithm based on numerical inversion: regularization as postconditioner, Numer. Math. 47 (1985) 123–138.

[12] J.W. Jerome, An asymptotically linear fixed point extension of the inf-sup theory of Galerkin approximation, Numer. Funct. Anal. Optim. 16 (1995) 345–361.

[13] T. Kerkhoven and Y. Saad, On acceleration methods for coupled nonlinear elliptic systems, Numer. Math. 60 (1992) 525–548.

[14] M.A. Krasnosel'skii, G.M. Vainikko, P.P. Zabreiko, Ya.B. Rititskii and V.Ya. Stetsenko, *Approximate Solution of Operator Equations*, Wolters-Noordhoff, 1972.

[15] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, 1970.

[16] W.C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, 2nd ed., CBMS-NSF Regional Conference Series in Applied Mathematics 70, SIAM, 1998.