AN ADAPTIVE NEWTON ALGORITHM BASED ON NUMERICAL INVERSION: REGULARIZATION AS POSTCONDITIONER

BY

JOSEPH W. JEROME

IMA Preprint Series # 76

June 1984

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

514 Vincent Hall

206 Church Street S.E.

Minneapolis, Minnesota 55455

Recent IMA Preprints

#	Author(s)	Title
1		Workshop Summaries from the September 1982 workshop on Statistical Mechanics, Dynamical Systems and Turbulence
2	Raphael De laLlave	A Simple Proof of C. Siegel's Center Theorem
3	H. Simpson, S. Spector	On Copositive Matrices and Strong Ellipticity for Isotropic Elastic Materials
4	George R. Sell	Vector Fields in the Vicinity of a Compact Invariant Manifold
5	Milan Miklavcic	Non-linear Stability of Asymptotic Suction
6	Hans Weinberger	A Simple System with a Continuum of Stable Inhomogeneous Steady States
7	Bau-Sen Du	Period 3 Bifurcation for the Logistic Mapping
8	Hans Weinberger	Optimal Numerical Approximation of a Linear Operator
9	L.R. Angel, D.F. Evans, B. Ninham	Three Component Ionic Microemulsions
10	D.F. Evans, D. Mitchell, S. Mukherjee, B. Ninham	Surfactant Diffusion; New Results and Interpretations
11	Leif Arkeryd	A Remark about the Final Aperiodic Regime for Maps on the Interval
12	Luis Magalhaes	Manifolds of Global Solutions of Functional Differential Equations
13	Kenneth Meyer	Tori in Resonance
14	C. Eugene Wayne	Surface Models with Nonlocal Potentials: Upper Bounds
15	Kathy Pericak-Spector	On Stability and Uniqueness of Fluid Flow Through a Rigid Porous Medium
16	George R. Sell	Smooth Linearization Near a Fixed Point
17	David Wollkind	A Nonlinear Stability Analysis of a Model Equation for Alloy Solidification
18	Pierre Collet	Local C^{∞} Conjugacy on the Julia Set for some Holomorphic Perturbations of $z \to z^2$
19	Henry C. Simpson Scott J. Spector	On the Modified Bessel Functions of the First Kind On Barrelling for a Material in Finite Elasticity
20	George R. Sell	Linearization and Global Dynamics
21	P. Constantin C. Foias	Global Lyapunov Exponents, Kaplan-Yorke Formulas and the Dimension of the Attractors for 2D Navier-Stokes Equations
22	Milan Miklavcic	Stability for Semilinear Parabolic Equations with Noninvertible Linear Operator
23	P. Collet, H. Epstein	Perturbations of Geodesic Flows on Surfaces of Constant Negative Curvature and their Mixing Properties
24	J.E. Dunn, J. Serrin	On the Thermomechanics of Interstitial working
25	Scott J. Spector	On the Absence of Bifurcation for Elastic Bars in Uniaxial Tension
26	W.A. Coppel	Maps on an Interval
27	James Kirkwood	Phase Transitions in the Ising Model with Traverse Field
28	Luis Magalhaes	The Asymptotics of Solutions of Singularly Perturbed Functional Differential Equations: and Concentrated Delays are Different
29	Charles Tresser	Homoclinic Orbits for Flow in R^3
30	Charles Tresser	About some Theorems by L.P. Sil'nikov
31	Michael Aizenmann	On the renormalized Coupling Constant and the Susceptibility in ϕ_4^4 Field theory and the Ising Model in Four Dimensions
32	Gene Wayne	The KAM Theory of Systems with Short Range Interations I
33	M. Slemrod J.E. Marsden	Temporal and Spatial Chaos in a Van der Waals Fluid due to Periodic Thermal Fluctuations
34	J. Kirkwood, C.E. Wayne	Percolation in Continuous Systems
35	Luis Magalhaes	Invariant Manifolds for Functional Differential Equations Close to Ordinary Differential Equations
36	C. Eugene Wayne	The KAM Theory of Systems with Short Range Interactions II
37	Jean De Canniere	Passive Quasi-Free States of the Noninteracting Fermi Gas
38	Elias C. Aifantis	Maxwell and van der Waals Revisited
39	Elias C. Aifantis	On the Mechanics of Modulated Structures

AN ADAPTIVE NEWTON ALGORITHM BASED ON NUMERICAL INVERSION: REGULARIZATION AS POSTCONDITIONER

by

Joseph W. Jerome
Department of Mathematics
Northwestern University
Evanston, Illinois 60201

Research supported by National Science Foundation grant MCS-8218041. This is the second part of work in progress during the author's visit to the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis.

AN ADAPTIVE NEWTON ALGORITHM BASED ON NUMERICAL INVERSION: REGULARIZATION AS POSTCONDITIONER

Abstract

We examine a class of approximate inversion processes, satisfying estimates similar to those defined by finite element or truncated spectral approximations; these are to be used as approximate right inverses for Newton iteration methods. When viewed at the operator level, these approximations introduce a defect, or "loss of derivatives," of order one or more. Regularization is introduced as a form of defect correction. A superlinearly convergent, approximate Newton iteration is thereby obtained by using the numerical inversion adaptively, i.e., with spectral or grid parameters correlated to the magnitude of the current residual in an intermediate norm defined by the defect. This adaptive choice makes possible ascribing an order to the convergent process, and this is identified as essentially optimal, relative to complexity. The design of the algorithm involves multi-parameter selection, thereby opening up interesting avenues for numerical experimentation. This applies also to the regularization, which may be carried out in the Fourier transform space, and is band-limited in the language of Whittaker-Shannon sampling theory. The norms employed in the analysis are of Hölder space type; the iteration is an adaptation of Nash-Moser iteration; and, the complexity studies use Vituškin's theory of information processing.

1. Introduction

This paper introduces a class of adaptive, approximate Newton methods, based upon Nash-Moser iteration, and studies their computational complexity, via the theory of information processing due to Vituškin. Approximate Newton methods have long held a secure position in numerical analysis, where they have typically been realized as outer iterations in sequential linearization schemes, for the solution of finite dimensional nonlinear systems. Recently, it has been realized (cf. Bank and Rose [1]) that application at the infinite dimensional, frame-space level can prove to be an effective way of studying nonlinear elliptic systems and other operator equations. In fact, approximate Newton methods, based upon approximate right inverses of Fréchet derivatives, can yield traditional quadratic convergence estimates, via the Kantorovich recurrence relations (cf. Jerome [7]). The associated requirement, that the inverse approximation, composed with the linearized map, approximates the identity in a manner directly proportional to the currently computed residual, fails, however, for typical right inverses defined by numerical computing procedures such as the finite element method or truncated spectral expansions. The net result is a situation analogous to the "loss of derivatives" phenomenon, overcome by regularization in the superlinearly convergent, Nash-Moser iteration scheme. The regularization speeds are tied to the superlinear convergence parameter, the acceleration parameter, and the initial residual.

In order to focus the ideas more clearly, we recall the typical estimate, for an approximation U_n, in solving the second order, uniformly elliptic, operator boundary value problem, $Lu = f$ on Ω in $\mathbb{R}^N$, where L has positive eigenvalues $0 < \lambda_1 \leqslant \lambda_2 \leqslant \dots$, by truncated spectral expansions of order n:

$$\|D^{\alpha}(u - U_n)\|_{L^2} \leqslant C\lambda_{n+1}^{(|\alpha|-2)/2} \|f\|_{L^2} \quad , \qquad 0 \leqslant |\alpha| \leqslant 2 \; .$$

In particular, if $U_n = T_n f$ and $u = Tf$, we have

$$\|LT_n f - f\|_{L^2} \leqslant C\|f\|_{L^2} \quad , \tag{1.1}$$

which expresses a zero order approximation to the identity, by LT_n, in terms of reciprocal eigenvalues. For operator equations $F(u) = 0$, the superlinear convergence of the approximate Newton iteration scheme $\{u_k\}$, based upon operators such as $T_n(u_k) \approx [F'(u_k)]^{-1}$, requires, as in the case of quadratic convergence, the estimation of the operator, $F'(u_k)T_n(u_k) - I$, as a function of the residual. One way of achieving such an estimate is to modify the r.h.s. norm in (1.1) via an estimate of the form, say,

$$\|(LT_n - I)f\|_{L^2} \leqslant C\lambda_{n+1}^{-1/2} \|f\|_{H^1} \tag{1.2}$$

if $f \in H^1$. Since the H^1 norm of f, rather than the L^2 norm of f, appears on the right hand side of (1.2), we speak of a "loss of derivatives" of order one in this case. However, the parameter $\lambda_{n+1}^{-1/2}$, or $\lambda_{n+1}^{-\gamma/2}$ for a loss of derivatives of order γ, permits the deployment of an adaptive computing strategy, in which the magnitude of the residual determines n. If we set $h = \lambda_{n+1}^{-1/2}$, then the constant in (1.2) is of the form $\tau(h) = Ch$, or, $\tau(h) = Ch^{\gamma}$ if $\gamma > 1$.

The iteration of the Nash-Moser implicit function theorem employs regularization as a way of defining and estimating $\|u_k - u_{k-1}\|$; the usual residual domination of this quantity in estimation does not directly apply, because the norm of the residual is "smoother" than the norm measuring the Newton increment, as a consequence of the loss of derivatives. This fact necessitates monitoring the residual in norms corresponding to the endpoints of a smoothness gauge, with interpolation necessary for intermediate norms corresponding to the order of derivative loss. The regularization may be thus viewed as a "catching up" procedure: it is applied as a postconditioner to the numerical inversion. The interpolation has a fundamental consequence: smooth norms of $F \circ v$ must depend linearly upon the corresponding smooth norm of v, under certain conditions on lower norms of v. The corollary of this, aside from the restriction to smooth nonlinear problems, is that the numerical analyst is encouraged to work in the rings of Hölder type spaces. The usual Sobolev norms of $F \circ v$ have polynomial growth in $\|v\|$, corresponding to the order of smoothness inherent in the norm, and are therefore unacceptable for estimation without some corresponding compensation increasing the order still more.

We are led, then, to pose the following question. Should Hilbert space projections, such as least squares, minimum energy and spectral projections, provide the basis for numerical computations analyzed in terms of pointwise norms? The answer seems to be a guarded "yes." The studies of de Boor [2], Douglas, Dupont, and Wahlbin [4], and Schatz and Wahlbin [12] reveal the stability, or near stability, of L^2 and H^1 spline projections as measured in L^∞ norms. Even spectral projection norms, which might grow logarithmically in L^∞, can be expected to be compensated by a loss of derivatives inequality analogous to (1.2). In order to maintain a discreet objectivity on these questions, however, we shall simply assume the availability of two computing processes: (i) an approximate right inverse, satisfying a "loss of derivatives" property expressed through an inequality involving a monotone, continuous function τ, with $\tau(0) = 0$ (the parameter is adaptively proportional to the evaluation of τ^{-1} at the current residual); and, (ii) a regularization procedure satisfying, in the Hölder spaces, inequalities similar to the Jackson and Bernstein inequalities of approximation theory. The latter are known to be stafisfied by standard mollification. In practice, since mollification is defined here by band-limited convolution, one can make use of the band-limited discrete Fourier transform, relating convolutions and products. Also, the approximate right inverse may be more amenable to spectral methods, since low degree finite elements are excluded.

The plan of this paper is as follows. In Section 2 we describe the framework and result, including a detailed analysis of the regularization, and we provide the proof, in Section 3, for the convergence of the superlinear iterations. Superlinear convergence is achieved by modification of hypotheses and proof given by Nirenberg [10] for the result of Nash-Moser. The purpose of this paper is really to raise questions: the author is not aware of any other studies in which numerical processes have been characterized via such defects, and associated corrections.

Adaptive, regularizing computation, as a correlative, is the fundamental insight we offer. Still, as presented, the algorithm offers much leeway for creative choice, since multi-parameter selection is required.

In practice, this method can be used in conjunction with continuation (cf. [7]) to guarantee the initial residual condition. This issue will be taken up in the sequel to this paper. We make two comments on the computing scheme related to the right inverse approximation: (i) at some crossover point, τ^{-1}(residual) $\approx$ SMALL(h) may be too expensive a criterion to implement, and one may wish to keep h "constant," while ensuring linear convergence in this case; and, (ii) the boundedness and approximation of the inversions would be expected to hold for operator equations which are compact perturbations of the identity such that no bifurcation can occur at or near the root. Self-adjointness and positive definiteness are not necessary, however.

The final issue we address here in the Introduction is the computational efficiency of the sequential linearization scheme we are presenting. Is the scheme profligate, in the sense that computing accuracy, ϵ, is achieved via wasteful procedures, as measured by the yardstick of linear tabulation? In the final section, we answer this question, negatively, by sketching arguments based upon ϵ-entropy, information processing, and the Kolmogorov n-dimensional diameters. In fact, the scheme is essentially optimal, with any deviations accounted for by use of Hilbert space algorithms (i.e., Galerkin methods) in Hölder space. In this framework, convergence and complexity are analyzed in tandem, and superlinear convergence is not viewed as inferior to quadratic convergence.

2. Framework, Postconditioning Iteration, and Convergence

We denote by $C^a(K)$ the real Hölder spaces. Here a is a real, nonnegative number, and K is a compact, convex subset with interior points in $\mathbb{R}^N$. The norm $\|\cdot\|_0$ in $C^0(K)$ is the usual supremum norm in the space of continuous functions, and, for $a > 0$, the norm $\|\cdot\|_a$ is the sum of the norm in $C^{[a]}(K)$ with the Hölder norm of order $a - [a]$; the quantity $[a]$ denotes the largest integer strictly less than a. The inequality, $\|\cdot\|_a \geqslant \|\cdot\|_k$, for k integral and $k < a$, will be used. The deeper, fundamental properties of these spaces which we make use of are the following (cf. Hormander [6, Appendix A]).

(I) Functional substitution is essentially of linear growth:

$$\|\Phi(D^\alpha v)\|_a \leqslant C(a, \gamma, m, \Phi, v_0)(\|v - v_0\|_{m+a} + \|v - v_0\|^a_{m+\gamma} + 1) \tag{2.1}$$

for any fixed v_0, $a \geqslant 1$, $1 \leqslant \gamma \leqslant a$ and $\{\alpha: |\alpha| \leqslant m\}$. Here, the subscript indicates the appropriate space, and Φ is in C^a on a fixed compact, convex set containing the range of the vector $D^\alpha v$.

(II) There exists a smoothing $S_t: C^{a_0}(K) \to \bigcap_{a>0} C^a(K)$ for $a_0 \geqslant 0$ fixed and $t \geqslant 1$:

$$\|S_t v - v\|_{a_0} \to 0 \ , \quad \text{as } t \to \infty \ ; \tag{2.2a}$$

$$\|S_t v\|_p \leqslant C_{rp} t^{p-r} \|v\|_r \ , \quad a_0 \leqslant r \leqslant p \ ; \tag{2.2b}$$

$$\|S_t v - v\|_r \leqslant C_{rp} t^{-(p-r)} \|v\|_p \ , \quad a_0 \leqslant r \leqslant p \ . \tag{2.2c}$$

(III) Norm interpolation is possible:

$$\|v\|_r \leqslant C(q, r_1, r_2) \|v\|^{1-q}_{r_1} \|v\|^q_{r_2} \ , \tag{2.3}$$

$0 \leqslant q \leqslant 1$, $0 \leqslant r_1 \leqslant r_2$, $r = (1-q)r_1 + qr_2$.

Note: The constants in (2.1), (2.2), and (2.3) may be taken independent of a, r, p, r_1, and r_2 if these numbers are bounded from above.

Remark 2.1. The operators S_t may be defined as follows, independently of a_0. Choose $\psi \in C_0^\infty(\mathbb{R}^N)$ such that $\psi = 1$ in a neighborhood of the origin. Set $\phi = \check{\psi}$, where $\vee$ denotes the inverse Fourier transform. Set $\phi_t(x) = t^N \phi(tx)$, $t \geqslant 1$. For a_0 fixed, and $v \in C^{a_0}(K)$, with support in K, define, for $x \in K$,

$$(2.4) \qquad S_t v(x) = v * \phi_t(x) = \int_{\mathbb{R}^N} v(x-y)\phi_t(y)dy \quad .$$

For $v \in C^{a_0}(K)$, the reduction to the previous case is achieved by a partition of unity of K^0. A major problem, of course, in the implementation of the method to follow, is the efficient construction of the smoothing operators S_t. The construction of (2.4) requires that v be smoothly extendible outside of K, either as zero, or by some reflection principle, so that the integral is convergent. Otherwise, the partition of unity is theoretically required. In practice, one would multiply v by a function ω in $C_0^{\infty}(K^0)$, similar to ψ as defined below. By using the Fourier transform, $\wedge$, one has from (2.4):

$$(2.5a) \qquad \widehat{S_t v}(x) = \hat{v}(x)\hat{\phi}_t(x) = \hat{v}(x)\psi(x/t) \quad .$$

Clearly, the smoothing can proceed in the transform space. Here the maps $\{\hat{}\,\check{}\}$ are defined for rapidly decreasing functions w by

$$(2.5b) \qquad \hat{w}(x) = \int_{\mathbb{R}^N} w(y)e^{-ix\cdot y}dy \ , \qquad \check{w}(x) = (2\pi)^{-N}\int_{\mathbb{R}^N} w(y)e^{ix\cdot y}dy \ .$$

Moreover, an acceptable choice of ψ is:

$$(2.6) \qquad \psi(x) = \begin{cases} 1 \ , & |x| \leqslant \pi/4 \ , \\ 1-\exp\left\{-\dfrac{1}{(\tan|x|)^2-1}\right\}, & \pi/4 < |x| < \pi/2 \ , \\ 0 \ , & |x| \geqslant \pi/2 \ . \end{cases}$$

In the language of sampling theory, $S_t v$ is band limited; the support of $\widehat{S_t v}$ is restricted to $t\prod_1^N [-\pi/2, \pi/2]$. Such functions can be reconstructed from their values at fractional multi-integer mesh points. Indeed, if g is band limited, with support in $W\prod_1^N [-\pi, \pi]$, then the Whittaker-Shannon sampling identity, or, cardinal series, holds (cf. Butzer [3] and Triebel [13]):

$$(2.7)\qquad g(x) = \sum_j g(j/W) \prod_{k=1}^{N} \frac{\sin \pi(x_k/W - j_k)}{\pi(x_k/W - j_k)}$$

for $x = (x_1, \ldots, x_N)$, and $j = (j_1, \ldots, j_N)$ an arbitrary multi-integer. The reader will also note that $\hat{v}$ is band-limited if v has support in K. Thus, by appropriate choices of W, both $S_t v$ and $\hat{v}$ can be computed from (2.7), thereby facilitating the calculation of (2.4).

In the case of periodic functions, Fourier series transforms can be used. In this case, incidentally, results of Grosse [5] become applicable; reconstruction via B-splines is possible. In the Fourier series case, one alters the Newton construction to follow by using a smoothing S_ℓ depending on a discrete parameter $\ell \geqslant 1$, instead of the continuous parameter t. S_ℓ may be defined, for $K = \prod^N [-d, d]$, by

$$(2.8a)\qquad S_\ell v(x) = (1/2d)^N \left(\prod^N \int_{-d/\ell}^{d/\ell}\right) \phi_\ell(x-y) v(y) dy\ , \qquad x \in K\ ,$$

where, for $j = (j_1, \ldots, j_N)$,

$$(2.8b)\qquad \phi_\ell(y) = \ell^N \sum_j a_j e^{i\pi\ell j \cdot y/d}\ , \quad a_j = \left(1 + \sum_{k=1}^{N} j_k^2\right)^{-1} .$$

Simple computations show that the discrete analogs of (2.2) hold and that

$$(2.9)\qquad \widehat{S_\ell v}(\ell j) = \hat{v}(\ell j) \left(1 + \sum_{k=1}^{N} j_k^2\right)^{-1}$$

with $\widehat{S_\ell v}(j') = 0$ if $j' \neq \ell j$. Here $\wedge$ is defined by

$$(2.10)\qquad \hat{w}(j) = (1/2d)^N \left(\prod^N \int_{-d}^{d}\right) w(y)\, e^{-i\pi j \cdot y/d} dy\ .$$

We emphasize that this smoothing presupposes 2d-periodicity in each variable. Such periodicity can be artificially effected by a partition of unity, or multiplication by $\omega \in C_0^{\infty}(K)$.

Definition 2.1. For a_0 and m fixed nonnegative real numbers, m an integer, set $X_\sigma = C^{a_0+\sigma}(K)$, $Y_\sigma = C^{a_0+m+\sigma}(K)$ for each $\sigma \geqslant 0$. Let $\Phi(D^\alpha) = F : B_{\eta,0} \subset Y_0 \rightarrow X_0$ be a differential mapping of order m, from $B_{\eta,0} = \{v \in Y_0 : \|v - v_0\|_{Y_0} \leqslant \eta\}$ to X_0, which satisfies the following conditions.

(i) F is continuously Lipschitz (Fréchet) differentiable on an open set U containing $B_{\eta,0}$:

$$\|F'(v) - F'(w)\|_{Y_0, X_0} \leqslant 2M\|v - w\|_{Y_0}, \quad v, w \in B_{\eta,0}, \tag{2.11a}$$

$$\|F'(v)\|_{Y_0, X_0} \leqslant M, \qquad v \in B_{\eta,0}. \tag{2.11b}$$

(ii) There exist numerical inversion operators with loss of derivatives. Specifically, there is a constant $\gamma \geqslant 1$ and a family $T_h(v) : X_\gamma \rightarrow Y_0$, of linear mappings dpending on parameters h and v, and a continuous monotone increasing function $\tau: [0,b] \rightarrow [0,\infty)$, $\tau(0) = 0$, such that (cf. the analog (1.2)), for all $w \in X_\gamma$,

$$\|[F'(v)T_h(v) - I]w\|_{X_0} \leqslant \tau(h)\,\|w\|_{X_\gamma}, \quad \text{for all } v \in B_{1,\gamma} \subset Y_\gamma. \tag{2.12}$$

(iii) $F : B_{\eta,0} \cap Y_\sigma \rightarrow X_\sigma$ is a well-defined map for $0 \leqslant \sigma \leqslant s$ and s sufficiently large (cf. (2.19c)). This includes the assumption that $F = \Phi(D^\alpha)$, with Φ as in (2.1) for $a = a_0 + s$.

(iv) The maps $\{T_h(v)\}$ are uniformly bounded in h and v from X_σ to $Y_{\sigma-\gamma}$ for $\gamma \leqslant \sigma \leqslant s$:

$$\|T_h(v)w\|_{Y_{\sigma-\gamma}} \leqslant M\|w\|_{X_\sigma}. \tag{2.13}$$

Note that h and v are restricted by $h \in [0,b]$ and $v \in B_{1,\gamma}$.

Remark 2.2. In this remark we discuss the approximate Newton method with postconditioning. We also discuss the fundamental parameters. We begin with u_0, the initial iterate. For ease of notation, since computations proceed with reference to u_0, we identify u_0 with the center v_0 of $B_{\eta,0}$. More generally, u_0 will be the center of the closed balls $B_{c,\sigma} \subset Y_\sigma$ of radius c. We assume

$$u_0 \in Y_s \ , \qquad f_0 = F(u_0) \in X_s \ , \quad \|f_0\|_{X_0} \leqslant \rho^{-\lambda} \ , \tag{2.14}$$

where $\rho > 0$ and $1 < \lambda < s$. We select a superlinear convergence parameter, β, satisfying $1 < \beta < 2$. In terms of these quantities, we require

$$\|f_0\|_{X_s} \leqslant M\rho^{(s-\lambda)/\beta} \ . \tag{2.15}$$

We select an acceleration parameter, $\theta > 1$, and define the smoothing speeds by

$$t_k = \rho^{\theta\beta^k} \ , \qquad k \geqslant 1 \ . \tag{2.16}$$

We assume inductively that $u_{k-1} \in Y_s \cap B_{1,\gamma}$ has been defined for $k \geqslant 1$. Then u_k is defined by Newton/postconditioning iteration:

$$u_k - u_{k-1} = -S_{t_k} T_{h_k}(u_{k-1}) F(u_{k-1}) \ . \tag{2.17}$$

Here, the parameter h_k is subject to specification (cf. (2.21) below), so that $\tau(h_k)$ is of the order of $F(u_{k-1})$. The iteration is required to converge in Y_μ for specified $\mu \geqslant 1$. The relations among γ, s, λ, β, θ, and μ are given in Definition 2.2 to follow. For ease of estimation, we assume that M is chosen so that $M \geqslant 1$ and

$$M \geqslant \sup\Big(3C(a_0 + s, \gamma, m, F, u_0),\ C_{r,p},\ C(q, r_1, r_2)\Big) \ , \tag{2.18}$$

where the constants appear in (2.1), (2.2), and (2.3). The supremem is understood to be taken over $r, r_1, r_2, p \leqslant s + a_0 + m$ and $0 \leqslant q \leqslant 1$.

Definition 2.2. The relations satisfied by the exponents and parameters are presented here. We require:

(2.19a) $$1 < \beta < 2\,, \quad 1 < \theta\,, \quad 1 \leqslant \gamma \leqslant \mu < \lambda < s\,,$$

(2.19b) $$\lambda \geqslant \max\left\{2\beta\gamma(2-\beta)^{-1}, \quad \beta(\gamma+\mu\beta)\right\} + \lambda_0(2-\beta)^{-1}\,,$$

(2.19c) $$s \geqslant \max\left\{\theta\gamma(\theta-1)^{-1}, \lambda+\theta\gamma(\beta-1)^{-1}\right\} + s_0 \max\left\{(\theta-1)^{-1}, (\beta-1)^{-1}\right\}\,,$$

for arbitrary $\lambda_0 > 0$, $s_0 > 0$. The number ρ is required to satisfy

(2.20a) $$\rho \geqslant e$$

(2.20b) $$M^9 \rho^{-\lambda_0} \leqslant 1/4\,,$$

(2.20c) $$M^4 \rho^{-s_0} \leqslant 1/2\,,$$

(2.20d) $$M^4 \rho^{-(\lambda-\beta\gamma)} \leqslant \eta/2\,,$$

(2.20e) $$\rho^{\lambda_0 \ln \beta} \geqslant e^{1/3}\,.$$

The numbers h_k are selected so that

(2.21) $$\tau(h_k) \leqslant M^5 \|F(u_{k-1})\|_{X_\gamma}\,, \qquad k = 1, \ldots\,.$$

Theorem 2.1. Suppose the hypotheses of Definitions 2.1 and 2.2 are satisfied, and that the initial iterate $u_0 \in B_{\eta,0} \cap Y_s$ satisfies (2.14) and (2.15). Then the Newton iterates $\{u_k\}$, defined by the adaptive, postconditioning procedure (2.17), converge to a root u of F in $B_{\eta,0} \cap B_{1,\mu}$. The superlinear convergence in Y_0 is described by the estimate

(2.22) $$\|u - u_k\|_{Y_0} \leqslant M^4 \rho^{-(\lambda-\beta\gamma)\beta^{k-1}} \Big/ \beta^{k-1}\,,$$

for $k = 1, 2, \ldots$. The superlinear convergence in Y_μ is described by

(2.23) $$\|u - u_k\|_{Y_\mu} \leqslant M^5 \rho^{-\lambda_0 \beta^{k-1}} \Big/ \beta^{k-1}\,,$$

for $k = 1, 2, \ldots$.

Remark 2.3. The choices

$$(2.24) \qquad \theta = 1.16\ , \quad \beta = 1.4\ , \quad \mu = \gamma\ , \quad \lambda = \frac{29}{6}\gamma\ , \quad s = \frac{47}{6}\gamma\ ,$$

satisfy (2.19a, b, c), for $\lambda_0 = .1$ and $s_0 = .016\gamma$.

These are not the best choices, but do reflect realistic magnitudes for the parameters. Some independent experimental calculations suggest that $\theta = 1.1435$, $\lambda = 3.2952$, $s = 7.9702$ and $\beta = 1.2446$, for $\mu = \gamma = 1$, might serve as first approximations for these parameters if the decrease of λ and s is preferred to the increase of β.

3. Proof of the Superlinear Convergence

This section is devoted entirely to a proof of Theorem 2.1. The verification proceeds by induction on k.

Proof of Theorem 2.1: We establish the following statements recursively for $k \geqslant 0$:

$$(3.1a) \qquad \|F(u_k)\|_{X_0} \leqslant \rho^{-\lambda\beta^k}\ ,$$

$$(3.1b) \qquad \|F(u_k)\|_{X_s} \leqslant M\rho^{(s-\lambda)\beta^{k+1}}\ ,$$

$$(3.1c) \qquad \|F(u_k)\|_{X_\gamma} \leqslant M^2\rho^{-(\lambda-\beta\gamma)\beta^k}$$

$$(3.1d) \qquad \|u_{k+1} - u_k\|_{Y_s} \leqslant \frac{1}{2}\rho^{(s-\lambda)\beta^{k+2}}\ ,$$

$$(3.1e) \qquad \|u_{k+1} - u_0\|_{Y_s} \leqslant \rho^{(s-\lambda)\beta^{k+2}}\ ,$$

$$(3.1f) \qquad \|u_{k+1} - u_k\|_{Y_0} \leqslant M^4\rho^{-(\lambda-\beta\gamma)\beta^k}\ ,$$

$$(3.1g) \qquad \|u_{k+1} - u_0\|_{Y_0} \leqslant \eta\ ,$$

$$(3.1h) \qquad \|u_{k+1} - u_k\|_{Y_\delta} \leqslant M^5\rho^{-\lambda_0\beta^k}\ , \qquad \gamma \leqslant \delta \leqslant \mu\ ,$$

(3.1i) $$\|u_{k+1} - u_0\|_{Y_\delta} \leqslant 1 \quad , \qquad \gamma \leqslant \delta \leqslant \mu \ .$$

Suppose, then, that (3.1a-i) hold for $k < j$, where $j \geqslant 0$ is an arbitrary fixed integer. To verify (3.1a), for $k = j$, we write, for $j > 0$,

(3.2a) $$F(u_j) = -[F'(u_{j-1})T_{h_j}(u_{j-1}) - I]F(u_{j-1}) + R(u_{j-1}, u_j) + F'(u_{j-1})[I - S_{t_j}]T_{h_j}(u_{j-1})F(u_{j-1})$$

where

(3.2b) $$R(u_{j-1}, u_j) = F(u_j) - F(u_{j-1}) - F'(u_{j-1})(u_j - u_{j-1}) \ .$$

Now an application of (2.12), with $\tau(h)$ selected according to (2.21), yields

$$\|[F'(u_{j-1})T_{h_j}(u_{j-1}) - I]F(u_{j-1})\|_{X_0} \leqslant M^5\|F(u_{j-1})\|^2_{X_\gamma} \quad ,$$

and (2.11a) gives

$$\|R(u_{j-1}, u_j)\|_{X_0} = \left\|\int_0^1 [F'(u_{j-1} + r(u_j - u_{j-1})) - F'(u_{j-1})](u_j - u_{j-1})dr\right\|_{X_0}$$
$$\leqslant M\|u_j - u_{j-1}\|^2_{Y_0} \quad .$$

Finally, applications of (2.11b), of (2.2c), with $r = a_0 + m$ and $p = s - \gamma + a_0 + m$, and of (2.13), with $\sigma = s$, yield

$$\|F'(u_{j-1})[I - S_{t_j}]T_{h_j}(u_{j-1})F(u_{j-1})\|_{X_0} \leqslant M\|[I - S_{t_j}]T_{h_j}(u_{j-1})F(u_{j-1})\|_{Y_0}$$
$$\leqslant M^2 t_j^{-(s-\gamma)}\|T_{h_j}(u_{j-1})F(u_{j-1})\|_{Y_{s-\gamma}}$$
$$\leqslant M^3 t_j^{-(s-\gamma)}\|F(u_{j-1})\|_{X_s} \quad .$$

We now estimate, separately, by use of the induction hypotheses. Applications of (3.1c) to the first of the three derived inequalities, of (3.1f) to the second, and of (3.1b) to the third, all with $k = j-1$ according to the induction hypotheses, give

$$(3.3) \qquad \left\|F(u_j)\right\|_{X_0} \leqslant 2M^9 \rho^{-2(\lambda-\beta\gamma)\beta^{j-1}} + M^4 \rho^{-\theta(s-\gamma)\beta^j} \rho^{(s-\lambda)\beta^j} .$$

When the inequalities

$$2(\lambda - \beta\gamma) \geqslant \lambda\beta + \lambda_0 , \qquad \theta(s-\gamma) - (s-\lambda) \geqslant \lambda + s_0 ,$$

which follow from (2.19b) and (2.19c) respectively, are applied to (3.3), we obtain

$$(3.4) \qquad \left\|F(u_j)\right\|_{X_0} \leqslant 2M^9 \rho^{-\lambda_0} \rho^{-\lambda\beta^j} + M^4 \rho^{-s_0} \rho^{-\lambda\beta^j} .$$

This yields (3.1a), with $k = j$, upon use of (2.20b) and (2.20c). Notice that $\rho \geqslant 1$ was used in deriving (3.4). Also, the case $j = 0$ in (3.1a) is immediate from the hypothesis (2.14).

Inequality (3.1b), for $k = j > 0$, follows from the induction hypothesis: $\left\|u_j - u_0\right\|_{Y_s} \leqslant \rho^{(s-\lambda)\beta^{j+1}}$, and from (2.1), with $a = a_0 + s$, $v_0 = u_0$, if account is taken of the induction hypothesis: $\left\|u_j - u_0\right\|_{Y_\gamma} \leqslant 1$, and of (2.18). Again, note that $\rho \geqslant 1$. Inequality (3.1b), for $j = 0$, is immediate from the hypothesis (2.15).

To establish (3.1c), for $k = j \geqslant 0$, we begin with the interpolation inequality (2.3), taking $q = \gamma/s$, $r_1 = a_0$, and $r_2 = a_0 + s$, and obtain

$$(3.5) \qquad \left\|F(u_j)\right\|_{X_\gamma} \leqslant M \left\|F(u_j)\right\|_{X_0}^{1-\frac{\gamma}{s}} \left\|F(u_j)\right\|_{X_s}^{\gamma/s} .$$

Applying the estimates for the residual which we have just derived, i.e., (3.1a,b) for $k = j$, we obtain from (3.5) the estimate

$$\|F(u_j)\|_{X_\gamma} \leqslant M^2 \rho^{-\lambda\beta^j(1-\frac{\gamma}{s})} \rho^{(s-\lambda)\beta^{j+1}(\gamma/s)}$$

$$= M^2 \rho^{-\beta^j(\lambda-\beta\gamma+(\frac{\lambda\gamma}{s})(\beta-1))} \leqslant M^2 \rho^{-(\lambda-\beta\gamma)\beta^j} ,$$

since $\rho, M \geqslant 1$. This is just (3.1c) for $k = j$.

For (3.1d), we use the definition (2.17) in conjunction with (2.2b), for $p = a_0 + m + s$, and $r = a_0 + m + s - \gamma$. When this is coupled with (2.13), for $\sigma = s$, we obtain

$$\|u_{j+1} - u_j\|_{Y_s} \leqslant M^2 (t_{j+1})^\gamma \|F(u_j)\|_{X_s} ,$$

and this, in turn, is estimated by (2.16) and (3.1b), for $k = j$. Altogether, we have

$$\|u_{j+1} - u_j\|_{Y_s} \leqslant M^3 \rho^{\theta\gamma\beta^{j+1}} \rho^{(s-\lambda)\beta^{j+1}} ,$$

and the right hand side of this expression is bounded by $M^3 \rho^{-s_0} \rho^{(s-\lambda)\beta^{j+2}}$, upon use of the inequality

$$s - \lambda + \theta\gamma \leqslant (s-\lambda)\beta - s_0 .$$

The latter follows from (2.19c). If we now apply (2.20c), we obtain (3.1d), with $k = j \geqslant 0$.

We now verify (3.1e) for $k = j > 0$. Indeed, we have

$$\|u_{j+1} - u_0\|_{Y_s} \leqslant \|u_{j+1} - u_j\|_{Y_s} + \sum_{k=1}^{j} \|u_k - u_{k-1}\|_{Y_s}$$

(3.6)

$$\leqslant \frac{1}{2} \rho^{(s-\lambda)\beta^{j+2}} + \frac{1}{2} \sum_{k=1}^{j} \rho^{(s-\lambda)\beta^{k+1}} ,$$

upon use of (3.1d) and the induction hypothesis. The second sum is estimated, via the lower sums, by the elementary technique of dominating it by the integral $I(j+1)$, where

$$I(b) = \frac{1}{2}\int_1^b \rho^{(s-\lambda)\beta^{x+1}}\,dx \quad .$$

An evident comparison of areas yields

$$(3.7) \qquad I(j+1) \geqslant \frac{1}{2}\sum_{k=1}^{j} \rho^{(s-\lambda)\beta^{k+1}} \quad ,$$

since the function

$$f(x) = \rho^{(s-\lambda)\beta^{x+1}} \quad ,$$

is (strictly) increasing on $[1, j+1]$. In order to estimate $I(j+1)$, we compute

$$f'(x) = \rho^{(s-\lambda)\beta^{x+1}}\beta^{x+1}\,\ln(\rho^{(s-\lambda)})\,\ln\beta \quad ,$$

so that, for $1 \leqslant x \leqslant j+1$,

$$(3.8) \qquad f(x) \leqslant f'(x)\Big/\left(\beta^2(s-\lambda)\ln\rho\,\ln\beta\right) \quad .$$

If (3.8) is integrated over the interval $[1, j+1]$, we have

$$(3.9) \quad I(j+1) = \frac{1}{2}\int_1^{j+1} f(x)dx \leqslant \frac{1}{2}\int_1^{j+1} f'(x)dx \leqslant \frac{1}{2}f(j+1) = \frac{1}{2}\rho^{(s-\lambda)\beta^{j+2}} \quad .$$

Here we have used the inequality (cf. (2.20a))

$$(3.10) \qquad \beta^2(s-\lambda)\ln\rho\,\ln\beta \geqslant 1 \quad ,$$

which follows the observation that $\beta \mapsto \beta^2\ln\beta/(\beta-1)$ increases from 1 on $[1,2]$. Combining (3.6), (3.7), and (3.9) gives (3.1e) for $k = j$.

The verification of (3.1f) reduces to the inequality

$$(3.11) \qquad \|u_{j+1} - u_j\|_{Y_0} \leqslant M^2\|F(u_j)\|_{X_\gamma} \quad ,$$

as is seen by comparison of (3.1c) and (3.1f). However, if we begin with (2.17), and apply (2.2b), with $r = p = a_0 + m$, and (2.13), with $\sigma = \gamma$, we obtain (3.11). Inequality (3.1g) is verified similarly to (3.1e), by making use of the decreasing function

$$g(x) = \rho^{-(\lambda-\beta\gamma)\beta^x}, \qquad 0 \leqslant x \leqslant j .$$

With this substitution, the estimate

$$\text{(3.12)} \qquad M^{-4}\|u_{j+1} - u_0\|_{Y_0} \leqslant \sum_{k=0}^{j} \rho^{-(\lambda-\beta\gamma)\beta^k} \leqslant \rho^{-(\lambda-\beta\gamma)} + \int_0^j g(x)\,dx$$

$$\leqslant \rho^{-(\lambda-\beta\gamma)} + [(\lambda-\beta\gamma)\ell n\,\rho\,\ell n\,\beta]^{-1}\int_0^j \big(-g'(x)\big)\,dx$$

$$\leqslant \left\{1 + [(\lambda-\beta\gamma)\,\ell n\,\rho\,\ell n\,\beta]^{-1}\right\}\rho^{-(\lambda-\beta\gamma)}$$

$$\leqslant 2\rho^{-(\lambda-\beta\gamma)} \leqslant \eta \quad ,$$

upon use of (2.20a, d) and (2.19b).

We next prove (3.1h). We begin with an interpolation inequality similar to (3.5), for δ satisfying $\gamma \leqslant \delta \leqslant \mu$:

$$\|u_{j+1} - u_j\|_{Y_\delta} \leqslant M\|u_{j+1} - u_j\|_{Y_0}^{1-\frac{\delta}{s}}\|u_{j+1} - u_j\|_{Y_s}^{\delta/s} \quad ,$$

and estimate the right hand side terms by (3.1f) and (3.1d). This gives

$$\|u_{j+1} - u_j\|_{Y_\delta} \leqslant M^5\rho^{-[(\lambda-\beta\gamma)(1-\frac{\delta}{s}) - \beta^2(s-\lambda)(\frac{\delta}{s})]\beta^j} .$$

If the inequality,

$$(\delta/s)(\beta^2\lambda - \lambda + \beta\gamma) \geqslant 0 \quad ,$$

is applied to the preceding inequality, we obtain

$$\|u_{j+1} - u_j\|_{Y_\delta} \leqslant M^5 \rho^{-(\lambda-\beta\gamma-\beta^2\delta)\beta^j} \leqslant M^5 \rho^{-\lambda_0\beta^j} ,$$

where we have used (2.19b). This establishes (3.1h), with $k = j \geqslant 0$.

It remains to establish (3.1i). By an estimate similar to (3.12) we obtain

$$M^{-5}\|u_{j+1} - u_0\|_{Y_\gamma} \leqslant \sum_{k=0}^{j}{}' \rho^{-\lambda_0\beta^k} \leqslant \rho^{-\lambda_0} + [\lambda_0 \ln\rho \ln\beta]^{-1} \rho^{-\lambda_0} ,$$

and inequality (3.1i) follows from (2.20b, e).

We now prove that $\{u_k\}$ is a Cauchy sequence in Y_0 and in Y_μ, satisfying the estimates

(3.13a) $$\|u_{k+m} - u_k\|_{Y_0} \leqslant M^4 \rho^{-(\lambda-\beta\gamma)\beta^{k-1}} / \beta^{k-1} ,$$

(3.13b) $$\|u_{k+m} - u_k\|_{Y_\mu} \leqslant M^5 \rho^{-\lambda_0\beta^{k-1}} / \beta^{k-1} ,$$

for $k, m \geqslant 1$. An application of (3.1f) and (3.1h) gives

$$\|u_{k+m} - u_k\|_{Y_0} \leqslant M^4 \sum_{j=k}^{k+m-1}{}' \rho^{-(\lambda-\beta\gamma)\beta^j} ,$$

and

$$\|u_{k+m} - u_k\|_{Y_\mu} \leqslant M^5 \sum_{j=k}^{k+m-1}{}' \rho^{-\lambda_0\beta^j} .$$

Estimation of these sums by improper integrals gives

(3.14a) $$\|u_{k+m} - u_k\|_{Y_0} \leqslant M^4 \int_{k-1}^{\infty} \rho^{-(\lambda-\beta\gamma)\beta^x} dx ,$$

and

$$(3.14b) \qquad \left\| u_{k+m} - u_k \right\|_{Y_\mu} \leqslant M^5 \int_{k-1}^{\infty} \rho^{-\lambda_0 \beta^x} dx \quad .$$

Both integrals are of the form $\int_{k-1}^{\infty} \rho^{-\delta\beta^x} dx$, $\delta \geqslant \lambda_0$, which can be estimated as

$$(3.15) \qquad \int_{k-1}^{\infty} \rho^{-\delta\beta^x} dx \leqslant \left[\beta^{k-1} \delta \ln \rho \ln \beta\right]^{-1} \int_{k-1}^{\infty} \left[\frac{d}{dx} (-\rho^{-\delta\beta^x})\right] dx$$

$$\leqslant \left[\beta^{k-1} \delta \ln \rho \ln \beta\right]^{-1} \rho^{-\delta\beta^{k-1}} \leqslant \rho^{-\delta\beta^{k-1}} \Big/ \beta^{k-1} \quad .$$

Here we have used (2.20e).

The estimates (3.13a, b) follow from (3.14a, b) and (3.15). Since $\{u_k\}$ is a Cauchy sequence in Y_0 and Y_μ, and since the injection $Y_\mu \rightarrow Y_0$ is continuous, it follows that there is an element $u \in B_{\eta,0} \cap B_{1,\mu}$ such that $u_k \rightarrow u$ in both topologies. Inequalities (2.22) and (2.23) follow from (3.13a, b) by letting $m \rightarrow \infty$. The fact that u is a root, i.e., $F(u) = 0$, follows routinely from (3.1a) and the continuity of F on $B_{\eta,0}$. ∎

4. Computational Complexity and Summary/Conclusions

The new insights, furnished by this paper, are related to both: (i) formation of algorithms; and, (ii) analysis of algorithms. Concerning (i), we offer a new adaptive, Newton iteration of the form

$$(4.1) \qquad u_k - u_{k-1} = -S_{t_k} T_{h_k}(u_{k-1}) F(u_{k-1}) \quad ,$$

in order to solve the equation $F(u) = 0$ approximately. The operators $T_{h_k}(u_{k-1})$ are approximate inverses of $F'(u_{k-1})$, and are typical numerical inversions defined by finite element or spectral approximations. The parameter h_k is chosen so that $\tau(h_k)$, which measures the extent to which $F'(u_{k-1})T_{h_k}(u_{k-1})$ approximates the identity (cf. (2.12)), is of the order of the X_{γ_1}-norm of the residual $F(u_{k-1})$. The choice $\gamma = 1$, $a_0 = 0$, in which case $X_\gamma = C^1(K)$, requires evaluation of difference quotients of the residual. The function τ is typically of the form

$$\tau(h) = Ch^\gamma , \qquad \tau(h) = C(\ell n\, h)^r h^\gamma , \tag{4.2}$$

for some integer r. Here h may be interpreted as a finite element grid parameter, or, $n^{-1/N}$, in a spectral method in $\mathbb{R}^N$ with n terms. Fortunately, C in (4.2) can be large with respect to the fundamental constant M (cf. (2.21)), or, alternatively, the constraint on the size of h can be relaxed. As the computations proceed, one might wish to avoid the computation of $\|F(u_{k-1})\|_{X_\gamma}$ for the adaptive criterion. Alternatively, one can require, in place of (2.21),

$$\tau(h_k) \leqslant M^7 \rho^{-(\lambda-\beta\gamma)\beta^{k-1}} , \tag{4.3}$$

as suggested by (3.1c). This will also lead to the convergence estimates of Theorem 2.1, as is transparent from the proof. Although use of (4.3) is not as natural as that of (2.21), it has the merit of disengaging the computations from the norm structure. The operator S_{t_k} is a mollification operator, and can be computed in the Fourier transform space, by use of band-limited, continuous Fourier transforms (cf. (2.5) and (2.6) and the Whittaker-Shannon sampling identity (2.7)) or by use of Fourier series (cf. (2.9)). In both cases, as a prior adjunct to computation, smooth extension outside the domain K is necessary to ensure the proper application of the theory, perhaps achieved by a compact support function such as (2.6). The smoothing speeds t_k are defined in terms of the fundamental parameters in (2.16).

Concerning (ii), we have proved that the algorithm, constructed in (i), is superlinearly convergent in Y_0 and in Y_μ, where distinct numbers, $\rho_1 = \rho^{-(\lambda-\beta\gamma)}$ and $\rho_2 = \rho^{-\lambda_0}$, characterize the respective order $\rho_1^{\beta^{k-1}}$, $\rho_2^{\beta^{k-1}}$. Since, typically, (4.4) $$\|u-u_k\|_{Y_0} = O(\rho_1^{\beta^{k-1}}) = \|F(u_{k-1})\|_{X_\gamma} \approx \tau(h_k) ,$$

we may also speak of τ as describing the "big O" order of convergence of this recursive process. Here $f \approx g$ means f and g are of the same order. Somewhat curiously, a larger value of γ enhances the order; this seems puzzling until one realizes that λ increases with γ, and the initial residual must satisfy (2.14). In fact, the convergence process is best interpreted in terms of computational complexity as follows. Set

(4.5) $$\epsilon = \epsilon_k = \rho^{-(\lambda-\beta\gamma)\beta^{k-1}} .$$

If the ϵ-entropy, $H_\epsilon(B_{\eta,0} \cap B_{1,\gamma})$, of the intersection of the ball of radius η in Y_0 with the ball of radius 1 in Y_γ (both centered at u_0) is computed in Y_0, one finds (cf. Lorentz [9] for the basic result)

(4.6) $$H_\epsilon \approx \epsilon^{-N/\gamma} .$$

The solution u lies in this intersection, and the ϵ-entropy estimates the logarithm of the minimal number, N_ϵ, of sets, of diameter not exceeding 2ϵ, which cover the intersection. Heuristically, the error estimate (2.22) locates u and u_k in one of these scaled covering sets. For compact subsets of Hölder space, such as the above intersection, the ϵ-entropy serves as a lower bound for n, the minimal number of parameters (dimension) required for a linear computing process, to represent an arbitrary element of the compact set to within the error ϵ (cf. Vituškin [14] and Lorentz [9]). The integer n is juxtaposed in the literature with the Kolmogorov widths, often written d_n (for elaboration, cf. Jerome [8]). A slightly extended interpretation of Vituškin's result, applied to the present context, is:

(4.7) $$n \geqslant C\epsilon^{-N/\gamma} \Longleftarrow d_n \leqslant \epsilon .$$

However,

(4.8) $$h \approx \tau^{-1}(\epsilon)$$

by (4.4), and, typically,

$$h \approx n^{-1/N} \quad , \tag{4.9}$$

where n is the dimension of the range of a numerical, linear inversion map. If, for simplicity, we treat the case $r = 0$ of (4.2), so that

$$\tau(h) \approx h^{\gamma} , \tag{4.10}$$

we see that

$$n \approx h^{-N} \approx (\epsilon^{1/\gamma})^{-N} = \epsilon^{-N/\gamma} \quad .$$

When this result is compared to (4.7), we grasp a remarkable conclusion: the adaptive sequential linearization procedure is optimal from the standpoint of computational complexity. Of course, we have neglected the complexity calculations introduced by the regularization, but this does not directly affect the judgment concerning the linear inversion, per se.

We remind the reader of the rather disappointing fact that, if finite elements are employed, they must be quite smooth: $C^{s+a_0+m-\gamma}$, in fact (cf. (2.13)). The choice of a_0 in defining X_σ and Y_σ can be integral. Since loss of derivatives of order $\gamma \geqslant 1$ is required, we do not use, in any direct way, the Schauder theory hypothesis, requiring the Hölder indices to be fractional for isomorphism. The use of smooth finite element spaces would typically accompany estimates of the form (4.2), with $r = 0$; a fairly extensive literature has developed on the topic (e.g., Nitsche [11]).

References

[1] R. Bank and D. Rose, Global approximate Newton methods, Numer. Math. 37 (1981), 279-295.

[2] C. de Boor, On local linear functionals which vanish at all B-splines but one, Theory of Approximation, with Applications (Law and Sahny, eds.), Academic Press New York (1976), 120-145.

[3] P. Butzer, A survey of the Whittaker-Shannon sampling theorem, and some of its extensions, J. Math. Research and Exposition 3 (1983), 195-212.

[4] J. Douglas, T. Dupont, and L. Wahlbin, Optimal L^{∞} error estimates for Galerkin approximations to solutions of two-point boundary value problems, Math. Comp. 29 (1975), 475-483.

[5] E. Grosse, Ph.D. Dissertation, Department of Computer Science, Stanford University, Stanford, California (1980).

[6] L. Hormander, The boundary problems of physical geodesy, Arch. Rat. Mech. Anal. 62 (1976), 1-52.

[7] J. Jerome, Approximate Newton methods and homotopy for stationary operator equations, IMA preprint No. 61, University of Minnesota, Minneapolis, February, 1984.

[8] J. Jerome, Approximation of Nonlinear Evolution Systems, Academic Press, New York (1983).

[9] G. Lorentz, Metric entropy and approximation, Bull. Amer. Math. Soc. 72 (1966), 903-937.

[10] L. Nirenberg, Topics in Nonlinear Functional Analysis, Courant Institute of Mathematical Sciences, New York University, New York (1973-74).

[11] J. Nitsche, L_{∞} convergence of finite element approximation, Rennes Conf. Finite Elements, 2nd, Rennes, France (1975).

[12] A. Schatz and L. Wahlbin, On the quasi-optimality in L_{∞} of the H^1 projection into finite element spaces, Math. Comp. 38 (1982), 1-22.

[13] H. Triebel, Fourier Analysis and Function Spaces, Teubner, Leipzig (1977).

[14] A. Vituškin, Theory of Transmission and Processing of Information, Pergamon Press, New York (1961).

Recent IMA Preprints (continued)

#	Author(s)	Title
40	William Ruckle	The Strong ϕ Topology on Symmetric Sequence Spaces
41	Charles R. Johnson	A Characterization of Borda's Rule Via Optimization
42	Hans Weinberger, Kazuo Kishimoto	The Spatial Homogeneity of Stable Equilibria of Some Systems on Convex Domains
43	K.A. Pericak-Spector, W.O. Williams	On Work and Constraints in Mixtures
44	H. Rosenberg, E. Toubiania	Some Remarks on Deformations of Minimal Surfaces
45	Stephan Pelikan	The Duration of Transients
46	V. Capasso, K.L. Cooke, M. Witten	Random Fluctuations of the Duration of Harvest
47	E. Fabes, D. Stroock	The L^p-intergrability of Green's functions and fundamental solutions for elliptic and parabolic equations
48	H. Brezis	Semilinear Equations in R^N without conditions at infinity
49	M. Slemrod	Lax-Friedrichs and the Viscosity-Capillarity Criterion
50	C. Johnson, W. Barrett	Spanning Tree Extensions of the Hadamard-Fischer Inequalities
51	Andrew Postlewaite, David Schmeidler	Revelation and Implementation under Differential Information
52	Paul Blanchard	Complex Analytic Dynamics on the Riemann Sphere
53	G. Levitt, H. Rosenberg	Topology and Differentiability of Labyrinths in the Disc and Annulus
54	G. Levitt, H. Rosenberg	Symmetry of Constant Mean Curvature Hypersurfaces in Hyperbolic Space
55	Ennio Stacchetti	Analysis of a Dynamic, Decentralized Exchange Economy
56	Henry Simpson, Scott Spector	On Failure of the Complementing Condition and Nonuniqueness in Linear Elastostatics
57	Craig Tracy	Complete Integrability in Statistical Mechanics and the Yang-Baxter Equations
58	Tongren Ding	Boundedness of Solutions of Duffing's Equation
59		Abstracts for the Workshop on Price Adjustment, Quantity Adjustment, and Business Cycles
60	Rafael Rob	The Coase Theorem an Informational Perspective
61	Joseph Jerome	Approximate Newton Methods and Homotopy for Stationary Operator Equations
62	Rafael Rob	A Note on Competitive Bidding with Asymmetric Information
63	Rafael Rob	Equilibrium Price Distributions
64	William Ruckle	The Linearization Projection, Global Theories
65	Russell Johnson, Kenneth Palmer, George R. Sell	Ergodic Properties of Linear Dynamical Systems
66	Stanley Reiter	How a Network of Processors can Schedule its Work
67	R.N. Goldman, D.C. Heath	Linear Subdivision Is Strictly a Polynomial Phenomenon
68	R. Giachett, R. Johnson	The Floquet Exponent for Two-dimensional Linear Systems with Bounded Coefficients
69	S. Williams	Realization and Nash implementation: Two Aspects of Mechanism Design
70	S. Williams	Sufficient Conditions for Nash Implementation
71	William Zame	Equilibria in Banach Lattices Without Ordered Preferences
72	W. Harris, Y. Sibuya	The Reciprocals of Solutions of Linear Ordinary Differential Equations
73	S. Pelikan	A Dynamical Meaning of Fractal Dimension
74	D. Heath, W. Sudderth	Continuous-Time Portfolio Management: Minimizing the Expected time to Reach A G
75	J.S. Jordan	Information Flows Intrinsic to the Stability of Economic Equilibrium