SIAM J. SCI. COMPUT. © 2021 Society for Industrial and Applied Mathematics
Vol. 43, No. 5, pp. S1–S20

ANDERSON ACCELERATION FOR A CLASS OF NONSMOOTH FIXED-POINT PROBLEMS*

WEI BIAN†, XIAOJUN CHEN‡, AND C. T. KELLEY§

Abstract. We prove convergence of Anderson acceleration for a class of nonsmooth fixed-point problems for which the nonlinearities can be split into a smooth contractive part and a nonsmooth part which has a small Lipschitz constant. These problems arise from compositions of completely continuous integral operators and pointwise nonsmooth functions. We illustrate the results with two examples.

Key words. nonsmooth equatioins, Anderson acceleration, integral equations, nonlinear equations, fixed-point problems

AMS subject classifications. 65H10, 45G10

DOI. 10.1137/20M132938X

1. Introduction.

In this paper we prove convergence of Anderson acceleration [1] for a class of nonsmooth fixed-point problems.

Anderson acceleration was originally designed for integral equations and is now very common in electronic structure computations (see [6] and many references since then). Anderson acceleration is essentially the same as direct inversion on the iterative subspace (DIIS) [18, 19, 26, 27], nonlinear GMRES [2, 21, 23, 32], and interface quasi-Newton [7, 13, 20]. It is also closely related to Pulay mixing [25], also known as Commutator DIIS [10, 15, 16, 26].

Convergence analysis has been reported in the literature only recently, and most of that work assumes at least continuous differentiability of the fixed-point map. There are convergence results for the linear case [30, 31], the continuously differentiable case [3], the Lipschitz continuously differentiable case [29, 30], and even smoother cases [8, 24].

In this paper we assume that nonlinearities can be split into a smooth part and a nonsmooth part with a small Lipschitz constant. The splittings we use in this paper are similar to ones used in nonsmooth nonlinear equations [5, 14, 17]. In those papers the norm of the nonsmooth part was small enough so that using the derivative of the smooth part led to a rapidly convergent Newton-like iteration. In this paper the splitting is only used in the analysis, and the algorithm does not change. However, the classes of problems to which the methods apply are very similar.

*Received by the editors April 6, 2020; accepted for publication (in revised form) November 23, 2020; published electronically January 20, 2021.
https://doi.org/10.1137/20M132938X
Funding: This work of the first author was partially supported by the NSF foundation of China through grants 11871178 and 61773136. The work of the second author was partially supported by Hong Kong Research Grant Council grant 15300219. The work of the third author was partially supported by Army Research Office grant W911NF-16-1-0504 and National Science Foundation grants OAC-1740309, DMS-1745654, and DMS-1906446.
†School of Mathematics and Institute of Advanced Study in Mathematics, Harbin Institute of Technology, Harbin 150001, China (bianweilvse520@163.com).
‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China (maxjchen@polyu.edu.hk).
§Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 USA (tim_kelley@ncsu.edu).

1.1. Notation and problem setting. In this paper we use bold faced fonts for vectors and operators which are finite dimensional or generic vectors and operators which can be either finite or infinite dimensional. We will use standard fonts for operators and (in section 3) vectors which are only defined in infinite dimensional function spaces.

The objective is to solve fixed-point problems of the form

$$\mathbf{u} = \mathbf{G}(\mathbf{u}), \tag{1.1}$$

where $\mathbf{G}$ is a Lipschitz continuous function defined on a Banach space X. We will make the following assumptions on $\mathbf{G}$ throughout this paper.

Assumption 1.1. $\mathbf{G}$ is a contraction with contractivity constant $c \in (0,1)$ in a closed convex set B in a Banach space X. $\mathbf{u}^*$ is the fixed point of $\mathbf{G}$ in B.

The Anderson acceleration algorithm is as follows:

`Anderson(m)`$(\mathbf{u}_0, \mathbf{G}, m)$
  $\mathbf{u}_1 = \mathbf{G}(\mathbf{u}_0)$; $\mathbf{F}_0 = \mathbf{G}(\mathbf{u}_0) - \mathbf{u}_0$.
  for $k = 1, \dots$ **do**
    Choose $m_k \le \min(m, k)$.
    $\mathbf{F}_k = \mathbf{G}(\mathbf{u}_k) - \mathbf{u}_k$.
    Minimize $\| \sum_{j=0}^{m_k} \alpha_j^k \mathbf{F}_{k-m_k+j} \|$ subject to $\sum_{j=0}^{m_k} \alpha_j^k = 1$.
    $\mathbf{u}_{k+1} = \sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}(\mathbf{u}_{k-m_k+j})$.
  end for

The *depth* m is the amount of storage needed beyond that of Anderson(0), which is simple Picard iteration:

$$\mathbf{u}_{k+1} = \mathbf{G}(\mathbf{u}_k).$$

We call the α's the *coefficients*.

The algorithm does not specify any norm, and the theory, for the most part, is independent of the choice of norm. Some results for Anderson(1) (see section 1.2.2) require a Hilbert space norm. In the case of a Hilbert space norm, the optimization problem can be formulated as a linear least squares problem [1]. For L^1 and L^∞ norms in finite dimension, the optimization problem can be formulated as a linear programming problem [30]. The examples in section 3 use the L^2 and the L^∞ norms.

The first convergence results for Anderson acceleration were reported in [30]. We state Theorem 1.1, one of the results from that paper, as generalized in [3], in order to compare it to the main results in this paper.

We allow for several ways to solve the optimization problem and also for different formulations (see section 1.2.1). Hence, following [30], we make an assumption on the optimization problem for the coefficients and its solution.

Assumption 1.2. The solution $\{\alpha_j^k\}$ of the optimization problem satisfies

1. $\| \sum_{j=0}^{m_k} \alpha_j^k \mathbf{F}(\mathbf{u}_{k-m_k+j}) \| \le \|\mathbf{F}(\mathbf{u}_k)\|$,
2. $\sum_{j=0}^{m_k} \alpha_j^k = 1$, and
3. there is M_α such that for all $k \ge 0$, $\sum_{j=1}^{m_k} |\alpha_j^k| \le M_\alpha$.

The first two parts of Assumption 1.2 simply state that the optimization problem finds an objective function value no larger than that for Picard iteration ($m = 0$ or

$\alpha^k_{m_k} = 1$) and that the constraints hold. To see this write the optimization problem as

$$\min_{\overline{\alpha} \in Q} \phi(\overline{\alpha}),$$

where

$$Q = \left\{ \overline{\alpha} \in R^{m_k+1} \mid \sum_{j=0}^{m_k} \alpha_j^k = 1 \right\}.$$

Let

$$\overline{\alpha}^* = argmin_{\overline{\alpha} \in Q} \phi(\overline{\alpha}).$$

Since $\phi(\bar{\alpha}^*) \leq \phi(\overline{\alpha})$ for all $\overline{\alpha} \in Q$, we have $\phi(\bar{\alpha}^*) = \min_{\overline{\alpha} \in Q} \phi(\overline{\alpha}) \leq \phi((0, 0, \ldots, 1)) = \|\mathbf{F}(\mathbf{u}_k)\|$.

The third part is generally not a consequence of the optimization problem formulation (unless $m = 1$ and $\| \cdot \|$ is a Hilbert space norm, or we add a nonnegativity constraint) and is critical in the proof. We have never observed that the bound of the L^1 norm of the coefficients is problematic (see [30] where we looked at this numerically).

As is standard, we denote the error $\mathbf{u} - \mathbf{u}^*$ by $\mathbf{e}$.

Theorem 1.1 ([3, 30]). *Let Assumptions* 1.1 *and* 1.2 *hold. Let* $\mathbf{G}$ *be continuously differentiable in*

$$B(\overline{\rho}) = \{\mathbf{u} \,|\, \|\mathbf{u} - \mathbf{u}^*\| < \overline{\rho}\} \subset B$$

for some $\overline{\rho} > 0$*. Let* $c < 1$ *be the contractivity constant from Assumption* 1.1*. Then if* $\|\mathbf{e}_0\|$ *is sufficiently small, the Anderson(m) iteration remains in* $B(\overline{\rho})$ *and converges to* $\mathbf{u}^*$ *r-linearly with r-factor c,*

$$\limsup_{k \to \infty} \left(\frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_0)\|} \right)^{1/k} \leq c, \tag{1.2}$$

which implies

$$\limsup_{k \to \infty} \left(\frac{\|\mathbf{e}_k\|}{\|\mathbf{e}_0\|} \right)^{1/k} \leq c. \tag{1.3}$$

1.2. Previous results for nonsmooth nonlinearaties. While the formulation of Anderson acceleration does not involve derivatives, there has been very little analysis of the method for nonsmooth $\mathbf{G}$. In this section we will discuss the results for general Lipschitz contractions. Those results, which we review in sections 1.2.1 and 1.2.2, are unsatisfactory because the estimate of the convergence rate is larger than c. Theorem 1.2 is a global convergence result, and the poor convergence rate is only a problem when the iteration is far from the solution. This is the result we extend in subsection 2.2.

The second result in subsubsection 1.2.2 is only for Anderson(1) and imposes the strong restriction $c < 2-\sqrt{3}$. This result is interesting for two reasons. The first is that the original form of this result in [30] assumed differentiability, but that assumption is not necessary. Our proof in the nondifferentiable case is new but borrows heavily from the analysis in [30]. Secondly, the proof we give motivates the one for result in subsection 2.1, where we show q-linear convergence with q-factor c for Anderson(1) for a class of nonsmooth problems.

1.2.1. The energy DIIS (EDIIS) algorithm. The EDIIS [18] algorithm adds a nonnegativity constraint to the optimization problem. The new optimization problem is

$$\text{Minimize } \left\| \mathbf{F}_k - \sum_{j=0}^{m_k-1} \alpha_j^k (\mathbf{F}_{k-m_k+j} - \mathbf{F}_k) \right\|_2^2$$

subject to

$$\sum_{j=0}^{m_k-1} \alpha_j^k = 1, \alpha_j^k \geq 0.$$

This problem is harder to solve than the linear least squares problem one must solve for Anderson acceleration, but one can obtain convergence from initial iterates in a larger set. Note that the solution of the EDIIS optimization problem satisfies all three parts of Assumption 1.2 by construction with $M_\alpha = \sum_{j=0}^{m_k-1} \alpha_j^k = 1$.

The result from [3] is the following.

THEOREM 1.2. *If* $\mathbf{G}$ *is Lipschitz continuous with Lipschitz constant* $c \in (0,1)$ *in a convex set* B*, then the EDIIS iteration converges for any* $\mathbf{u}_0 \in B$*, and*

$$\|\mathbf{e}_k\| \leq c^{k/(m+1)} \|\mathbf{e}_0\|. \tag{1.4}$$

Moreover, if $\mathbf{G}$ *is continuously differentiable, then the local convergence rate is no worse than that of Picard iteration, i.e.,*

$$\limsup_{k\to\infty} \left(\frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_0)\|} \right)^{1/k} \leq c. \tag{1.5}$$

The estimate (1.4) is valid for any Lipschitz continuous contraction but has a very pessimistic convergence rate. Continuous differentiability was necessary for the proof of (1.5). One contribution of this paper is to show that (1.5) holds for a class of nonsmooth problems.

1.2.2. Local convergence for Anderson(1). The proof of Theorem 1.3, the result in this section, is a direct extension of a proof in [28, 30] (Theorem 2.4, page 812 in [30]) of a similar result for the differentiable case. As we said earlier, the proof in [30] used continuous differentiability but really did not need it. We give the proof here in detail both for completeness and to illustrate the primary components in the new results in the paper. The convergence rate in Theorem 1.3 is q-linear rather than r-linear. In [30, Corollary 2.5, page 814], smoothness is used in an important way to obtain q-linear convergence with q-factor c for all $c \in (0,1)$. Theorem 2.1 in subsection 2.1 in this paper extends that result to a class of nonsmooth problems.

THEOREM 1.3. *Let* X *be a Hilbert space with scalar product* $(\cdot,\cdot)$. *Assume that the optimization problem is solved in the norm of* X. *Let* $\mathbf{G}$ *be Lipschitz continuous with Lipschitz constant* $c < 2 - \sqrt{3}$ *in a ball of radius* $\overline{\rho}$ *about a fixed point* u^*. *Then for* $\mathbf{u}_0$ *sufficiently close to* $\mathbf{u}^*$, *the Anderson(1) residuals converge q-linearly to* $\mathbf{u}^*$ *with q-factor*

$$\hat{c} \equiv \frac{3c - c^2}{1 - c} < 1$$

in the sense that for all k *sufficiently large*

$$\|F(u_{k+1})\| \leq \hat{c} \|F(u_k)\| \tag{1.6}$$

and $u_k \to u^*$ *r-linearly in the sense that*

$$\limsup_{k\to\infty} \left(\frac{\|e_k\|}{\|e_0\|}\right)^{1/k} \le \hat{c}. \tag{1.7}$$

Proof. We proceed by induction and allow for a "warm start" which may have an inferior convergence rate as EDIIS could. For example, this could be the final k_0+1 iterations of a longer EDIIS initialization phase or several Picard iterations. Assume that, for $0 \le j \le k_0$,

$$\mathbf{u}_j \in B(\rho) \equiv \{\mathbf{u} \,|\, \|\mathbf{u}-\mathbf{u}^*\| \le \rho\},$$

and for $0 \le j < k$ and some $\hat{c} \le \tilde{c} < 1$,

$$\|\mathbf{F}(\mathbf{u}_{j+1})\| \le \tilde{c}\|\mathbf{F}(\mathbf{u}_j)\|. \tag{1.8}$$

This assumption is clearly satisfied if $\mathbf{u}_1 = \mathbf{G}(\mathbf{u}_0)$ and $k_0 = 0$.

Note that if $u \in B(\overline{\rho})$, then

$$(1-c)\|\mathbf{e}\| \le \|\mathbf{F}(\mathbf{u})\| = \|\mathbf{G}(\mathbf{u})-\mathbf{u}\| = \|\mathbf{G}(\mathbf{u})-\mathbf{G}(\mathbf{u}^*)-(\mathbf{u}-\mathbf{u}^*)\| \le (1+c)\|\mathbf{e}\|. \tag{1.9}$$

We now show that (1.6) holds for all $k \ge k_0$ if (1.8) (which is implied by (1.6)) holds for all smaller k. The optimization problem can be solved in closed form for $m=1$. We have

$$\mathbf{u}_{k+1} = (1-\alpha^k)\mathbf{G}(\mathbf{u}_k) + \alpha^k\mathbf{G}(\mathbf{u}_{k-1}), \tag{1.10}$$

where

$$\alpha^k = \frac{(\mathbf{F}(\mathbf{u}_k), \mathbf{F}(\mathbf{u}_k)-\mathbf{F}(\mathbf{u}_{k-1}))}{\|\mathbf{F}(\mathbf{u}_k)-\mathbf{F}(\mathbf{u}_{k-1})\|^2}.$$

We estimate α^k using the induction hypothesis:

$$\begin{aligned} |\alpha^k| &\le \tfrac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_k)-\mathbf{F}(\mathbf{u}_{k-1})\|} \\ &\le \tfrac{\tilde{c}\|\mathbf{F}(\mathbf{u}_{k-1})\|}{(1-\tilde{c})\|\mathbf{F}(\mathbf{u}_{k-1})\|} \le \bar{\alpha} \equiv \tfrac{\tilde{c}}{1-\tilde{c}}. \end{aligned} \tag{1.11}$$

Our first task is to show that if $\|e_0\| < \overline{\rho}$ is sufficiently small, then $u_{k+1} \in B(\overline{\rho})$. The formula (1.10) implies that

$$\mathbf{e}_{k+1} = (1-\alpha^k)(\mathbf{G}(\mathbf{u}_k)-\mathbf{G}(\mathbf{u}^*)) + \alpha^k(\mathbf{G}(\mathbf{u}_{k-1})-\mathbf{G}(\mathbf{u}^*))$$

and hence

$$\|\mathbf{e}_{k+1}\| \le c(1+\bar{\alpha})\|\mathbf{e}_k\| + c\bar{\alpha}\|\mathbf{e}_{k-1}\|.$$

The induction hypothesis and (1.9) imply that, for $0 \le j \le k$,

$$\|\mathbf{e}_j\| \le \frac{\|\mathbf{F}(\mathbf{u}_j)\|}{1-c} \le \frac{\tilde{c}^j}{1-c}\|\mathbf{F}(\mathbf{u}_0)\| \le \frac{\tilde{c}^j(1+c)}{1-c}\|\mathbf{e}_0\|.$$

Hence,

$$\begin{aligned} \|\mathbf{e}_{k+1}\| &\le c(1+\bar{\alpha})\|\mathbf{e}_k\| + c\bar{\alpha}\|\mathbf{e}_{k-1}\| \\ &\le c(1+\bar{\alpha})\tfrac{\tilde{c}^k(1+c)}{1-c}\|\mathbf{e}_0\| + c\bar{\alpha}\tfrac{\tilde{c}^{k-1}(1+c)}{1-c}\|\mathbf{e}_0\| \\ &= \tfrac{c\tilde{c}^{k-1}(1+c)}{1-c}(\bar{\alpha}+(1+\bar{\alpha})\tilde{c})\|\mathbf{e}_0\|. \end{aligned}$$

Since $\tilde{c}, c < 1$, we have $\bar{\alpha} + (1+\bar{\alpha})\tilde{c} \leq (1+2\bar{\alpha})$ and $c\tilde{c}^{k-1} < 1$. Hence

$$\|\mathbf{e}_{k+1}\| \leq \frac{(1+c)(1+2\bar{\alpha})}{1-c}\|\mathbf{e}_0\| < \rho,$$

if

$$\|\mathbf{e}_0\| < \frac{(1-c)\rho}{(1+c)(1+2\bar{\alpha})},$$

which we will assume throughout.

Now we obtain the asymptotic result (1.6). Write

$$\mathbf{F}(\mathbf{u}_{k+1}) = \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{u}_{k+1} = A_k + B_k,$$

where

$$A_k = \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{G}((1-\alpha^k)\mathbf{u}_k + \alpha^k \mathbf{u}_{k-1})$$

and

$$B_k = \mathbf{G}((1-\alpha^k)\mathbf{u}_k + \alpha^k \mathbf{u}_{k-1}) - \mathbf{u}_{k+1}. \tag{1.12}$$

We next estimate $\|A_k\|$ and $\|B_k\|$ separately.

The estimation for $\|A_k\|$ is straightforward, as it will be throughout the paper.

$$\begin{aligned} \|A_k\| &= \|\mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{G}((1-\alpha^k)\mathbf{u}_k + \alpha^k \mathbf{u}_{k-1})\| \\ &\leq c\|\mathbf{u}_{k+1} - (1-\alpha^k)\mathbf{u}_k - \alpha^k \mathbf{u}_{k-1}\| \\ &= c\|(1-\alpha^k)(\mathbf{G}(\mathbf{u}_k) - \mathbf{u}_k) + \alpha^k(\mathbf{G}(\mathbf{u}_{k-1}) - \mathbf{u}_{k-1})\| \\ &= c\|(1-\alpha^k)\mathbf{F}(\mathbf{u}_k) + \alpha^k \mathbf{F}(\mathbf{u}_{k-1})\| \leq c\|\mathbf{F}(\mathbf{u}_k)\|, \end{aligned} \tag{1.13}$$

where the last inequality follows from optimality of the coefficients.

The estimate for $\|B_k\|$ is where differentiability was used, but not really needed, in [3, 30]. The analysis in those papers used the fundamental theorem of calculus to estimate the left side of (1.14) in terms of the errors and, in the case of [30], the Lipschitz constant of the Jacobian. The more recent paper [3] used the modulus of continuity of the Jacobian, and we employ similar logic in the proof of Theorem 2.1 (see (2.5)).

We begin by using (1.12) and (1.10) to obtain

$$\begin{aligned} B_k &= \mathbf{G}((1-\alpha^k)\mathbf{u}_k + \alpha^k \mathbf{u}_{k-1}) - (1-\alpha^k)\mathbf{G}(\mathbf{u}_k) - \alpha^k \mathbf{G}(\mathbf{u}_{k-1}) \\ &= \mathbf{G}(\mathbf{u}_k + \alpha^k \delta_k) - \mathbf{G}(\mathbf{u}_k) + \alpha^k(\mathbf{G}(\mathbf{u}_k) - \mathbf{G}(\mathbf{u}_{k-1})). \end{aligned} \tag{1.14}$$

Using contractivity, we obtain

$$\|B_k\| \leq 2c|\alpha^k|\,\|\delta_k\|,$$

where $\delta_k = \mathbf{u}_{k-1} - \mathbf{u}_k$. The next step is to estimate the product $|\alpha^k|\|\delta_k\|$.

The difference in residuals is

$$\mathbf{F}(\mathbf{u}_k) - \mathbf{F}(\mathbf{u}_{k-1}) = \mathbf{G}(\mathbf{u}_k) - \mathbf{G}(\mathbf{u}_{k-1}) + \delta_k.$$

Using contractivity $\|\mathbf{G}(\mathbf{u}_k) - \mathbf{G}(\mathbf{u}_{k-1})\| \le c\|\delta_k\|$ we obtain

$$\|\mathbf{F}(\mathbf{u}_k) - \mathbf{F}(\mathbf{u}_{k-1})\| \ge (1-c)\|\delta_k\|.$$

Hence

$$\|\delta_k\| \le \|\mathbf{F}(\mathbf{u}_k) - \mathbf{F}(\mathbf{u}_{k-1})\|/(1-c). \tag{1.15}$$

Finally, we use the formula for α^k to obtain

$$|\alpha^k|\|\delta_k\| \le \frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_k) - \mathbf{F}(\mathbf{u}_{k-1})\|}\|\delta_k\| \le \frac{\|\mathbf{F}(\mathbf{u}_k)\|}{1-c}. \tag{1.16}$$

So

$$\begin{aligned} \|\mathbf{F}(\mathbf{u}_{k+1})\| &\le c\|\mathbf{F}(\mathbf{u}_k)\| + \tfrac{2c\|\mathbf{F}(\mathbf{u}_k)\|}{1-c} \\ &= \tfrac{3c-c^2}{1-c}\|\mathbf{F}(\mathbf{u}_k)\| = \hat{c}\|\mathbf{F}(\mathbf{u}_k)\|. \end{aligned}$$ ∎

This completes the proof.

The important point for this paper in the proof of Theorem 1.3 is the decomposition of $\mathbf{F}(\mathbf{u}_{k+1})$ into A_k and B_k. In the results in section 2 we use the same decomposition, and, as in the proof of Theorem 1.3, the estimate of $\|A_k\|$ only uses the contractivity of $\mathbf{G}$. The estimate for $\|B_k\|$, however, is new and uses the structure of the nonsmoothness, which we describe in the next section.

2. Splitting-based results for nonsmooth problems.

The results in this section depend on Assumption 2.1, which states that $\mathbf{G}$ can be locally split into smooth ($\mathbf{G}_S$) and nonsmooth ($\mathbf{G}_N$) parts, with the nonsmooth part having a small Lipschitz constant. The motivation for this is a class of nonsmooth compact fixed-point problems, which we fully describe in section 3. We will also assume that Assumption 1.1 and (except for the Hilbert space case with $m = 1$) Assumption 1.2 hold.

Assumption 2.1. There is $\overline{\rho}$ such that $B(\overline{\rho}) \subset B$. There are nonincreasing nonnegative functions σ and ω defined on $(0,1)$ such that for any $0 < \rho < \overline{\rho}$,

1. $\lim_{t\to 0} \omega(t) = 0$,
2. $\lim_{t\to 0} \sigma(t) = 0$,
3. $\mathbf{G} = \mathbf{G}_S^\rho + \mathbf{G}_N^\rho$,
4. $\mathbf{G}_S^\rho$ is uniformly (in ρ) continuously differentiable in the sense that

$$\|(\mathbf{G}_S^\rho)'(\mathbf{u}) - (\mathbf{G}_S^\rho)'(\mathbf{v})\| \le \omega(\|\mathbf{u} - \mathbf{v}\|)$$

   for all $\mathbf{u}, \mathbf{v} \in B(\overline{\rho})$, and
5. $\mathbf{G}_N^\rho$ is Lipschitz continuous in $B(\rho)$ with Lipschitz constant $\sigma(\rho)$.

As we said in the introduction, the splitting is only exploited in the analysis. The algorithm is unchanged. The construction in this paper is different from the one used in nonlinear equations [5, 14, 17] in that we need the nonsmooth part to have a small Lipschitz constant, not a small norm. The examples in section 3 are compositions of nonsmooth substitution operators and integral operators and fit nicely with Assumption 2.1.

As was the case in [30], we are able to prove q-linear convergence of the residual norms only for $m = 1$. We obtain r-linear convergence for $m > 1$.

2.1. Anderson(1). In this section we extend Corollary 2.5 from [30, page 814]. That result was from the proof of Theorem 2.4 [30, page 812] in that paper. We extended that result to the nonsmooth case in Theorem 1.3 in section 1.2.2 in the present paper.

THEOREM 2.1. *Let X be a Hilbert space with scalar product $(\cdot,\cdot)$. Assume that the optimization problem is solved in the norm of X. Let Assumptions* 1.1 *and* 2.1 *hold. Then for $\mathbf{u}_0$ sufficiently close to $\mathbf{u}^*$, the Anderson(*1*) residuals converge q-linearly to* 0 *with q-factor c in the sense that*

$$\limsup_{k\to\infty} \frac{\|\mathbf{F}(\mathbf{u}_{k+1})\|}{\|\mathbf{F}(\mathbf{u}_k)\|} \le c. \tag{2.1}$$

Proof. As in the proof of Theorem 1.3 we allow for a warm start and assume that (1.8) holds for some $\rho < \bar{\rho}$, $\tilde{c} < 1$, and all $0 \le j \le k_0$. Most of the analysis we need in this proof can be taken directly from the proof of Theorem 1.3 or Corollary 2.5 from [30].

We show that if (1.8) holds for all $0 \le j \le k$ with $k \ge k_0$, then

$$\|\mathbf{F}(\mathbf{u}_{k+1})\| \le \|\mathbf{F}(\mathbf{u}_k)\|(c + \epsilon_k),$$

where $\epsilon_k \to 0$ as $k \to \infty$. This will imply that (2.1) holds. Our proof will give an explict formula for ϵ_k.

We begin by finding ρ_k so that

$$\mathbf{u}_k + t\alpha^k \delta_k \in B(\rho_k/2) \text{ and } \mathbf{u}_k + t\delta_k \in B(\rho_k/2) \tag{2.2}$$

for all $t \in [0,1]$. This will allow us to use the splitting in our estimate of $\|\mathbf{F}(\mathbf{u}_{k+1})\|$.

Using (1.9) and (1.8) we see that for $j = k-1, k$,

$$\|\mathbf{e}_j\| \le \|\mathbf{F}(\mathbf{u}_j)\|/(1-c) \le \tilde{c}^j \|\mathbf{F}(\mathbf{u}_0)\|/(1-c) \le \tilde{c}^{k-1} \|\mathbf{F}(\mathbf{u}_0)\|/(1-c). \tag{2.3}$$

Therefore, for all $t \in [0,1]$,

$$\begin{aligned} \|\mathbf{e}_k + t\alpha^k \delta_k\| &\le \|\mathbf{e}_k\| + \bar{\alpha}(\|\mathbf{e}_k\| + \|\mathbf{e}_{k-1}\|) \\ &\le \tilde{c}^{k-1}(1 + 2\bar{\alpha})\|\mathbf{F}(\mathbf{u}_0)\|/(1-c). \end{aligned} \tag{2.4}$$

We simplify the notation for the splitting by writing $\mathbf{G}_S = \mathbf{G}_S^{\rho_k}$ and $\mathbf{G}_N = \mathbf{G}_N^{\rho_k}$, where

$$\rho_k = 2\tilde{c}^{k-1}(1 + 2\bar{\alpha})\|\mathbf{F}(\mathbf{u}_0)\|/(1-c).$$

With this choice, (2.4) implies (2.2).

We split $\mathbf{F}(\mathbf{u}_{k+1})$ into three parts:

$$\mathbf{F}(\mathbf{u}_{k+1}) = \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{u}_{k+1} = A_k + C_k + D_k.$$

Here

$$A_k = \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{G}((1-\alpha^k)\mathbf{u}_k + \alpha^k \mathbf{u}_{k-1}).$$

We use (1.14) to split $B_k = C_k + D_k$, where

$$C_k = \mathbf{G}_S(\mathbf{u}_k + \alpha^k \delta_k) - \mathbf{G}_S(\mathbf{u}_k) + \alpha^k(\mathbf{G}_S(\mathbf{u}_k) - \mathbf{G}_S(\mathbf{u}_{k-1}))$$

and

$$D_k = \mathbf{G}_N(\mathbf{u}_k + \alpha^k \delta_k) - \mathbf{G}_N(\mathbf{u}_k) + \alpha^k(\mathbf{G}_N(\mathbf{u}_k) - \mathbf{G}_N(\mathbf{u}_{k-1})).$$

The estimate for $\|A_k\|$ is unchanged:

$$\|A_k\| \le c\|\mathbf{F}(\mathbf{u}_k)\|.$$

The estimate for $\|C_k\|$ is done exactly the same way as in [30] or [3]. We use differentiability of $\mathbf{G}_S$ to get the estimate (see equation (2.27), page 813, in [30])

$$\|C_k\| \le |\alpha^k|\|\delta_k\| \int_0^1 \|\mathbf{G}'_S(\mathbf{u}_k + t\alpha_k\delta_k) - \mathbf{G}'_S(\mathbf{u}_k + t\delta_k)\| \, dt. \tag{2.5}$$

We invoke Assumption 2.1 and the estimates (2.2), (2.3), and (1.16) to obtain

$$\begin{aligned}\|C_k\| &\le |\alpha^k|\|\delta_k\|\omega(|1-\alpha_k|\delta_k)\\ &\le \|\mathbf{F}(\mathbf{u}_k)\|\tfrac{\omega(\xi_k)}{1-c},\end{aligned}$$

where

$$\xi_k = 2(1+\bar{\alpha})\tilde{c}^{k-1}\|\mathbf{F}(\mathbf{u}_0)\|/(1-c).$$

Finally we estimate $\|D_k\|$, which is the new part of the analysis. We have, using (1.16),

$$\begin{aligned}\|D_k\| &\le \|\mathbf{G}_N(\mathbf{u}_k + \alpha^k\delta_k) - \mathbf{G}_N(\mathbf{u}_k)\| + |\alpha^k|\|\mathbf{G}_N(\mathbf{u}_k) - \mathbf{G}_N(\mathbf{u}_{k-1}))\|\\ &\le 2\sigma(\rho_k)|\alpha^k|\|\delta_k\| \le 2\sigma(\rho_k)\|\mathbf{F}(\mathbf{u}_k)\|/(1-c).\end{aligned}$$

Hence,

$$\|\mathbf{F}(\mathbf{u}_{k+1})\| \le \|\mathbf{F}(\mathbf{u}_k)\|(c + (\omega(\xi_k) + 2\sigma(\rho_k))/(1-c)).$$

This will complete the proof with

$$\epsilon_k = (\omega(\xi_k) + 2\sigma(\rho_k))/(1-c).$$ ∎

2.2. The case $m \ge 1$. In this section we prove a nonsmooth analogue of Theorem 1.2. As was the case in subsection 2.1, we split $\mathbf{G}(\mathbf{u}_{k+1})$ and analyze the parts separately. Many parts of the proof are taken from the proof of Theorem 1.2 in [3], and we will simply refer to the relevant pages in [3] for that material rather than copy the details.

The main result is Theorem 2.2.

THEOREM 2.2. *Let Assumptions* 1.1, 2.1, *and* 1.2 *hold. Then if* $\|\mathbf{e}_0\|$ *is sufficiently small the Anderson(m) iterations converge and* (1.5) *holds.*

Proof. We will allow for a warm start and assume that (1.8) holds for $0 \le j \le k$ with $k \ge k_0$. As before, this assumption is clearly satisfied if $k_0 = 0$ and $\mathbf{u}_1 = \mathbf{G}(\mathbf{u}_0)$, a cold start. We assume that $\mathbf{u}_j \in B(\overline{\rho})$ for $0 \le j \le k$.

Let $\hat{c} \in (c, 1)$ be given. We will show that

$$\limsup_{k\to\infty} \left(\frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_0)\|}\right)^{1/k} \le \hat{c} \tag{2.6}$$

by showing that there is L such that

$$\|\mathbf{F}(\mathbf{u}_k)\| \le L\hat{c}^k\|\mathbf{F}(\mathbf{u}_0)\|, \tag{2.7}$$

which implies (2.6) since $\lim_{k\to\infty} L^{1/k} = 1$. This will complete the proof of (1.5) as $\hat{c} \in (c, 1)$ is arbitrary.

We may, without loss of generality, assume that $\tilde{c} \in (\hat{c}, 1)$, where $\tilde{c}$ is the convergence rate from (1.8). The estimate (2.7) holds for $k \le k_0$ if we use $L = (\tilde{c}/\hat{c})^m$, which will begin an induction on k.

We assume that (2.7) holds for k and all $j < k$. We also assume that

$$\|\mathbf{e}_0\| < \frac{\overline{\rho} c^m (1-c)}{L M_\alpha (1+c)}, \tag{2.8}$$

where M_α is the bound from Assumption 1.2.

First note that (2.7) will imply that $\mathbf{u}_k \in B(\overline{\rho})$ because $\mathbf{u}_0 \in B(\overline{\rho})$ and (2.8) implies that

$$\|\mathbf{e}_0\| \le \frac{\overline{\rho}(1-c)}{L(1+c)}.$$

We use the formula for the Anderson iteration

$$\mathbf{u}_{k+1} = \sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}(\mathbf{u}_{k-m_k+j})$$

to split $\mathbf{F}(\mathbf{u}_{k+1})$. We have, following [3],

$$\begin{aligned}\mathbf{F}(\mathbf{u}_{k+1}) &= \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{u}_{k+1} \\ &= \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{G}(\textstyle\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}) + \mathbf{G}(\textstyle\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}) - \mathbf{u}_{k+1}.\end{aligned}$$

We begin with the usual splitting $\mathbf{F}(\mathbf{u}_{k+1}) = A_k + B_k$, where

$$A_k = \mathbf{G}(\mathbf{u}_{k+1}) - \mathbf{G}\left(\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}\right)$$

and

$$\begin{aligned}B_k &= \mathbf{G}(\textstyle\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}) - \mathbf{u}_{k+1} \\ &= \mathbf{G}(\textstyle\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}) - \textstyle\sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}(\mathbf{u}_{k-m_k+j}).\end{aligned}$$

The proof that

$$\|A_k\| \le c\|\mathbf{F}(\mathbf{u}_k)\| \le Lc\hat{c}^k \|\mathbf{F}(\mathbf{u}_0)\| \tag{2.9}$$

carries over unchanged from (1.13) in this paper or from equation (2.15) on page A372 of [3].

Note that (2.7) and (2.8) imply that

$$\mathbf{u}_j \in B(\rho_k) \text{ for } j = k - m_k, \ldots, k+1,$$

and

$$\mathbf{w}_k = \sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j} \in B(\rho_k).$$

Here,

$$\rho_k = L M_\alpha \hat{c}^{k-m_k} \|\mathbf{F}(\mathbf{u}_0)\|/(1-c) \le M_\alpha \hat{c}^{k-m} L(1+c)\|\mathbf{e}_0\|/(1-c). \tag{2.10}$$

Equation (2.8) implies that $\rho_k < \overline{\rho}$.

This allows us to split B_k as we did in the Anderson(1) case:

$$B_k = C_k + D_k,$$

where

$$C_k = \mathbf{G}_S\left(\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}\right) - \sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}_S(\mathbf{u}_{k-m_k+j})$$

and

$$D_k = \mathbf{G}_N\left(\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}\right) - \sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}_N(\mathbf{u}_{k-m_k+j}).$$

The estimate for $\|C_k\|$ uses exactly the same analysis as in [3, pages A372–A374]. We obtain

$$\|C_k\| \le 2M_\alpha\omega(\rho_k)\rho_k \le (2M_\alpha^2\omega(\rho_k)L\hat{c}^{k-m})\|F(u_0)\|/(1-c) \le \frac{2M_\alpha^2\omega(\rho_k)}{\hat{c}^m(1-c)}L\hat{c}^k\|\mathbf{F}(\mathbf{u}_0)\|.$$

Reduce $\|\mathbf{e}_0\|$ if necessary so that

$$\frac{2M_\alpha^2\omega(\rho_k)}{\hat{c}^m(1-c)} < (\hat{c}-c)/2. \tag{2.11}$$

Finally, write

$$D_k = \left(\mathbf{G}_N\left(\sum_{j=0}^{m_k} \alpha_j^k \mathbf{u}_{k-m_k+j}\right) - \mathbf{G}_N(\mathbf{u}^*)\right) - \left(\sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}_N(\mathbf{u}_{k-m_k+j}) - \mathbf{G}_N(\mathbf{u}^*)\right)$$

to obtain

$$\begin{aligned}\|D_k\| &\le 2\sigma(\rho_k)M_\alpha \max\nolimits_{0\le j\le m_k}\|\mathbf{e}_{k-m_k+j}\| \\ &\le \tfrac{2\sigma(\rho_k)M_\alpha}{1-c}\max\nolimits_{0\le j\le m_k}\|\mathbf{F}(\mathbf{u}_{k-m_k+j})\| \qquad (2.12)\\ &\le \tfrac{2\sigma(\rho_k)M_\alpha}{(1-c)\hat{c}^m}L\hat{c}^k\|\mathbf{F}(\mathbf{u}_0)\|.\end{aligned}$$

Reduce $\|\mathbf{e}_0\|$ if necessary to make

$$\frac{2\sigma(\rho_k)M_\alpha}{(1-c)\hat{c}^m} < (\hat{c}-c)/2. \tag{2.13}$$

This completes the proof since (2.11) and (2.13) imply that

$$\|\mathbf{F}(\mathbf{u}_{k+1})\| \le \|A_k\| + \|C_k\| + \|D_k\| < L\hat{c}^{k+1}\|\mathbf{F}(\mathbf{u}_0)\|. \qquad \square$$

2.3. Approximations. If X is finite dimensional, as it will be for discretizations of problems in function space, then part 2 of Assumption 2.1 may not hold. However, as we illustrate in the examples in section 3, we will still have a small (but generally nonzero) $\limsup \sigma(t)$. We replace part 2 of Assumption 2.1 with

$$\limsup_{t\to 0} \sigma(t) = \overline{\sigma}. \tag{2.14}$$

For any $q \in (0,1)$ and $\overline{\sigma}$ sufficiently small, we will still obtain r-linear convergence with r-factor $c + \overline{\sigma}^q$. We summarize the results for Anderson(m) in the following theorem.

Theorem 2.3. *Let Assumptions* 1.1, 1.2, *and* 2.1 *hold with part* 2 *replaced by* (2.14) *and*

$$\overline{\sigma} < \min\left((1-c)^{1/q}, \left(\frac{(1-c)c^m}{8M_\alpha} \right)^{1/(1-q)} \right) \tag{2.15}$$

for some $q \in (0,1)$. Then if $\|\mathbf{e}_0\|$ is sufficiently small, then the Anderson(m) iterations converge and

$$\limsup_{k\to\infty} \left(\frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_0)\|} \right)^{1/k} \le c + \overline{\sigma}^q < 1. \tag{2.16}$$

Proof. We will reduce $\overline{\sigma}$ in the course of the proof. Set $\hat{c} = c + \overline{\sigma}^q < 1$. We can then use the proof of Theorem 2.2 with very little change. We let $\tilde{L} = (\tilde{c}/c)^m$, which will play the role of L from the proof of Theorem 2.2.

We decompose the residual

$$\mathbf{F}(\mathbf{u}_{k+1}) = A_k + C_k + D_k$$

and use the estimates (2.9) and (2.11) without change (reducing $\|\mathbf{e}_0\|$ as needed).

The only difference is the estimate for D_k. Let $\|\mathbf{e}_0\|$ be small enough so that $\sigma(t) \le 2\overline{\sigma}$ for all $t \le \|\mathbf{e}_0\|$. We have, as before,

$$\begin{aligned} \|D_k\| &\le \tfrac{4\overline{\sigma}M_\alpha}{(1-c)\hat{c}^m} \tilde{L}\hat{c}^k \|\mathbf{F}(\mathbf{u}_0)\| \\ &\le \tfrac{4\overline{\sigma}M_\alpha}{(1-c)c^m} \tilde{L}\hat{c}^k \|\mathbf{F}(\mathbf{u}_0)\|. \end{aligned}$$

Then (2.15) implies that

$$\frac{4\overline{\sigma}M_\alpha}{(1-c)c^m} \le \overline{\sigma}^q/2 = (\hat{c}-c)/2.$$

∎

This estimate completes the proof exactly as it did in the proof of Theorem 2.2.

The result for Anderson(1) is similar, and we omit the proof, which is essentially the same as that for Theorem 2.3.

Theorem 2.4. *Let X be a Hilbert space with scalar product $(\cdot,\cdot)$. Assume that the optimization problem is solved in the norm of X. Let Assumptions* 1.1 *and* 2.1 *hold with part* 2 *replaced by* (2.14). *Let $q \in (0,1)$ be given. Then if $\overline{\sigma} \in (0,(1-c)^{1/q})$ is sufficiently small and $\mathbf{u}_0$ is sufficiently close to $\mathbf{u}^*$, the Anderson(1) residuals converge q-linearly to $\mathbf{u}^*$ with q-factor $c + \overline{\sigma}^q$ in the sense that*

$$\limsup_{k\to\infty} \frac{\|\mathbf{F}(\mathbf{u}_{k+1})\|}{\|\mathbf{F}(\mathbf{u}_k)\|} \le c + \overline{\sigma}^q. \tag{2.17}$$

3. Examples.

Our examples are compositions of nonsmooth substitution operators and nonlinear Hammerstein integral operators.

We let $C = C([0,1])$ be the space of continuous functions on $[0,1]$ with the usual L^∞ norm and $L^2 = L^2([0,1])$. We have two examples. The one in subsection 3.1 is in L^2 and the other, in subsection 3.2, is in C.

We let $g \in C([0,1]\times[0,1])$ and let $\mathcal{G}$ be the integral operator given by

$$\mathcal{G}(u)(x) = \int_0^1 g(x,y)u(y)\,dy.$$

In all the examples in this paper g is the Green's function for the negative Laplacian in one space dimension with zero boundary conditions. We discretize with the standard second-order central difference scheme and realize the product of $\mathcal{G}$ with a vector via a tridiagonal solver. We used a grid of $N = 100$ interior grid points and the composite trapezoid rule for integration.

The important properties of $\mathcal{G}$ are that

- $\mathcal{G}$ is a bounded operator on L^2 and
- $\mathcal{G}$ is a bounded operator from L^2 to C:

$$\|\mathcal{G}(u)\|_\infty \leq \|g\|_\infty \|u\|_2.$$

The maps in this section are compositions of nonsmooth substitution operators and nonlinear integral operators of the form

$$G_I(u)(x) = \mathcal{G}(f(u))(x) = \int_0^1 g(x,y) f(u(y))\, dy. \tag{3.1}$$

G_I maps L^2 to C if $f(\xi) = O(|\xi|)$ for large $|\xi|$ and is Fréchet differentiable if f' is bounded. In that case $G_I'(u)$ is the linear integral operator defined by

$$(G_I'(u)w)(x) = \int_0^1 g(x,y) f'(u(y)) w(y)\, dy.$$

G_I' is a compact linear operator from L^2 to C.

Since f' is bounded, f is Lipschitz continuous with Lipschitz constant L_f. This implies that G_I is a Lipschitz continuous map from L^2 to C. In fact, for $u, v \in L^2$ and $x \in [0,1]$, we may apply the Cauchy–Schwarz inequality to obtain

$$\begin{aligned} |G_I(u)(x) - G_I(v)(x)| &\leq \|g\|_\infty L_f \textstyle\int_0^1 |u(y) - v(y)|\, dy \\ &\leq \|g\|_\infty L_f \|u - v\|_2. \end{aligned} \tag{3.2}$$

After integration of (3.2) we obtain

$$\|G_I(u) - G_I(v)\|_\infty \leq \|g\|_\infty L_f \|u - v\|_2.$$

We consider nonsmooth substitution maps Φ that are based on point evaluation. Examples include

$$\Phi(u)(x) = \max(u(x) + b(x), 0),$$

where $b \in C$ is given. In general we assume the following.

Assumption 3.1. There is a real-valued function β and $b \in C$ such that

$$\Phi(u)(x) = \beta(u(x) + b(x)) \tag{3.3}$$

and β is Lipschitz continuous and differentiable except for finitely many points.

In our examples the function β will be differentiable except at one point.

If β is differentiable, then Φ is defined and Fréchet differentiable on both $C[0,1]$ and $L^2[0,1]$ if

- $|\beta(\xi)| = O(|\xi|)$ for $|\xi|$ large and
- β' is bounded.

In that case the Fréchet derivative $\Phi'(u)$ of Φ at u is the operator of multiplication by $\beta'(u+b)$ i.e.,

$$\Phi'(u)w(x) = \beta'(u(x) + b(x))w(x).$$

In the examples β is nondifferentiable only at $w = 0$ and is uniformly Lipschitz continuously differentiable away from $w = 0$. We formalize this as the following.

Assumption 3.2. β is Lipschitz continuous with Lipschitz constant L_β. There is $\gamma_\beta > 0$ such that if u and v have the same sign, then

$$|\beta'(u) - \beta'(v)| \leq \gamma_\beta |u - v|.$$

For example, if $\beta(u) = |u|$, then $\gamma_\beta = 0$.

3.1. A class of integral operators. We consider fixed-point maps of the form

$$u = G(u) = \Phi(G_I(u)). \tag{3.4}$$

We will work in L^2 in this example. We use the fact that G_I maps L^2 to C in the analysis in a significant way.

We will assume that f is a real-valued Lipschitz continuously differentiable function and that f' has Lipschitz constant γ_f.

We assume that Assumption 1.1 holds and that

$$B(\overline{\rho}) = \{u \,|\, \|u - u^*\|_2 \leq \overline{\rho}\} \subset B.$$

If $\rho \leq \overline{\rho}$ and $u \in B(\rho)$, then (3.2) implies that

$$\|G_I(u) - G_I(u^*)\|_\infty \leq \|g\|_\infty L_f \|u - u^*\|_2 \leq \|g\|_\infty L_f \rho \equiv \epsilon(\rho).$$

We can now construct the splitting. This will motivate the assumptions of our convergence result. Let

$$\Omega_\rho = \{x \,|\, |G_I(u^*)(x) + b(x)| < 2\epsilon(\rho)\},$$

and let χ_ρ be the characteristic function of Ω_ρ.

We define

$$G_N^\rho(u)(x) = \chi_\rho(x) G(u)(x)$$

and

$$G_S^\rho(u)(x) = G(u)(x) - G_N^\rho(u)(x) = (1 - \chi_\rho(x))G(u)(x).$$

Suppose $u \in B(\rho)$. Then $G_I(u)(x) + b(x)$ has the same sign as $G_I(u^*)(x) + b(x)$ for all $x \in \Omega_\rho^c$, the complement of Ω_ρ. This implies that G_S^ρ is differentiable at u, and for all $w \in L^2$ and $x \notin \Omega_\rho$,

$$\begin{aligned} (G_S^\rho)'(u)w(x) &= \beta'(G_I(u)(x) + b(x))(G_S^\rho)'(u)w)(x) \\ &= \beta'(G_I(u)(x) + b(x)) \textstyle\int_0^1 g(x,y) f'(u(y))w(y)\,dy. \end{aligned} \tag{3.5}$$

For $x \in \Omega_\rho$, $(G_S^\rho)'(u)w(x) = 0$. Moreover, if $v \in B^\infty(\rho)$, then

$$\|(G_S^\rho)'(u) - (G_S^\rho)'(v)\|_2 \leq \gamma_\beta \|g\|_\infty \gamma_f \|u - v\|_2. \tag{3.6}$$

As for the nonsmooth part, note that for $x \in [0,1]$ we may use (3.2) to obain

$$\begin{aligned} |G_N^\rho(u)(x) - G_N^\rho(v)(x)| &\leq \chi_\rho(x)|\Phi(G_I(u)(x)) - \Phi(G_I(v)(x))| \\ &\leq \chi_\rho(x) L_\beta \|g\|_\infty L_f \|u-v\|_2. \end{aligned}$$

Hence, using the Cauchy–Schwarz inequality again,

$$\|G_N^\rho(u) - G_N^\rho(v)\|_2 \leq \|g\|_\infty L_f L_\beta \sqrt{\mu(\Omega_\rho)} \|u-v\|_2,$$

because the L^2 norm of the characteristic function of Ω_ρ is $\sqrt{\mu(\Omega_\rho)}$ where μ is a Lebesgue measure.

The critical assumption is the splitting method in [14, 17] is that the support of nonsmoothness for u^* is small. In the setting of this paper, we assume that

$$\lim_{\rho \to 0} \mu(\Omega_\rho) = 0.$$

So we have the splitting with

$$\sigma(\rho) = \|g\|_\infty L_f L_\beta \sqrt{\mu(\Omega_\rho)} \text{ and } \omega(\rho) = \gamma_\beta \|g\|_\infty \gamma_f \rho.$$

3.1.1. Norms in finite dimension. In the computations we must, of course, approximate the integrals by quadratures. We use the composite trapezoid rule. A more subtle point is that we must scale the norm so that discretizations of constant functions have the same norm independently of N. Hence we use the discrete ℓ^2 norm

$$\|\mathbf{u}\|_2 = \frac{1}{\sqrt{N}} \sqrt{\sum_{j=1}^N u_j^2}$$

and ℓ^1 norm

$$\|\mathbf{u}\|_1 = \frac{1}{N} \sum_{j=1}^N |u_j|.$$

Using the scaled norm does not matter in Anderson acceleration because the scaling is irrelevant in the optimization problem and cancels in the relative residuals. However, it does matter when computing the Lipschitz constant. In the example in subsection 3.1.2, $G_I(u^*)(x) + b(x) = 0$ at only two points. For the approximate finite dimensional problem, this means that the set Ω_ρ, for ρ sufficently small, contains at most two grid points. The correct computation of $\mu(\Omega_\rho)$ is to use the discrete L^1 norm, and, therefore, to apply Theorem 2.3 to this example we would use

$$\overline{\sigma} \leq L_f L_\beta \sqrt{2/N}.$$

3.1.2. Obstacle Bratu problem. The equation in this section is an integral equation formulation of the obstacle Bratu problem [22]:

$$u = \min(\lambda \mathcal{G}(e^u), \alpha). \tag{3.7}$$

Here α is a given function of x. In the example here $\lambda = 5$ and

$$\alpha(x) = 1 + \sin(2\pi x)/2.$$

The right side of Figure 3.1 is a plot of the solution and the upper bound. One can see that the $\lambda\mathcal{G}(e^u)$ is equal to α at only two points. The left of the plot is the iteration history. We have tuned λ to make Picard iteration perform poorly. The Anderson(m) iterations for $m = 1, 2, 3$ perform equally well and significantly better than Picard iteration.

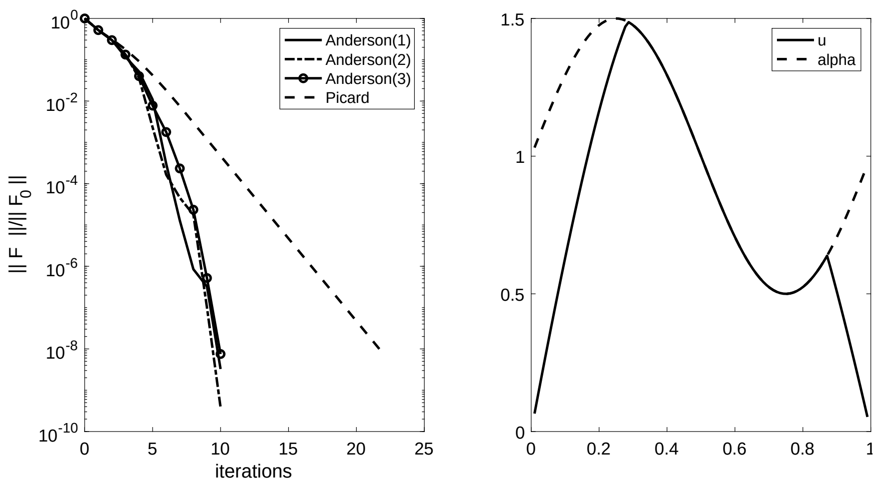

FIG. 3.1. *Example 1: Obstacle Bratu Problem.*

TABLE 3.1
Convergence rates for the Bratu problem.

Picard	Anderson 1	Anderson 2	Anderson 3
4.27e-01	1.42e-01	1.14e-01	1.54e-01

We can quantify the observations in Figure 3.1 by estimating the r-factors for the four methods. As we did in [3] we estimate the r-factor by

$$\left(\frac{\|\mathbf{F}(\mathbf{u}_{\bar{k}})\|}{\|\mathbf{F}(\mathbf{u}_0)\|} \right)^{1/\bar{k}}, \tag{3.8}$$

where $\bar{k}$ is the final iteration index. $\bar{k}$ varies over the method-problem combinations. In Table 3.1 we see that the estimate rates are consistent with Figure 3.1.

3.2. Compositions of the form $\mathbf{G} = \boldsymbol{\mathcal{G}}(\boldsymbol{\Phi})$. In this section we consider problems of the form

$$u = G(u) = \mathcal{G}(\Phi(u)). \tag{3.9}$$

We can now construct the splitting. We do this via an example which readily extends to the general case. We will solve the optimization problem in the L^∞ norm for this example.

For this case we let

$$\Omega_\rho = \{x \,|\, |u^*(x) + b(x)| < 2\rho\}.$$

We define

$$G_N^\rho(u)(x) = \int_{\Omega_\rho} g(x,y)\Phi(u)(y)\, dy = \int_{\Omega_\rho} g(x,y)\beta(u(y) + b(y))\, dy$$

and

$$G_S^\rho(u) = G(u) - G_N^\rho(u) = \int_{\Omega_\rho^c} g(x,y)\beta(u(y)+b(y))\,dy,$$

where Ω_ρ^c is the complement of Ω_ρ in $[0,1]$. Suppose $u \in B^\infty(\rho)$; then $u(x)+b(x)$ has the same sign as $u^*(x)+b(x)$ for all $x \in \Omega_\rho^c$. This implies that G_S^ρ is differentiable at u and that for all $w \in C$,

$$(G_S^\rho)'(u)w = \int_{\Omega_\rho^c} g(x,y)\beta'(u(y)+b(y))w(y)\,dy. \tag{3.10}$$

Moreover, if $v \in B^\infty(\rho/2)$, then

$$\|(G_S^\rho)'(u) - (G_S^\rho)'(u)\|_\infty \le \|g\|_\infty \gamma_\beta \|u-v\|_\infty. \tag{3.11}$$

As for the nonsmooth part, note that

$$G_N^\rho(u) - G_N^\rho(v) = \int_{\Omega_\rho} g(x,y)(\beta(u(y)+b(y)) - \beta(v(y)+b(y)))\,dy.$$

So, by the Hölder inequality,

$$\begin{aligned} \|G_N^\rho(u) - G_N^\rho(v)\| &\le \|g\|_\infty L_\beta \textstyle\int_{\Omega_\rho} |u(y)-v(y)|\,dy \\ &\le \|g\|_\infty L_\beta \mu(\Omega_\rho)\|u-v\|_\infty. \end{aligned}$$

The critical assumption for the splitting method in [14, 17] is that the support of nonsmoothness for u^* is small. In the setting for this paper, we assume that

$$\lim_{\rho\to 0}\mu(\Omega_\rho) = 0.$$

We have constructed the splitting with

$$\sigma(\rho) = \|g\|_\infty L_\beta \mu(\Omega_\rho) \text{ and } \omega(t) = \|g\|_\infty \gamma_\beta t.$$

The comments in subsection 3.1.1 are relevant here as well. In this case we need the discrete measure of Ω_ρ which converges to 0 as $N \to \infty$. In the example in subsection 3.2, this set contains only one point, so

$$\overline{\sigma} \le L_f L_\beta \frac{1}{N}.$$

3.2.1. Nonsmooth Dirichelet problem. The example, taken from [4], is

$$-v'' = \lambda \max(v-\alpha, 0),\ v(0) = v_0, v(1) = v_1. \tag{3.12}$$

In this problem the nonsmoothness is in the forcing term.

We convert (3.12) to a compact fixed-point problem by setting $v = u + \phi$, where $\phi(x) = v_1 x + (1-x)v_0$, letting $\mathcal{G}$ be the integral operator which inverts $-d^2/dx^2$ with zero boundary conditions and then multiplying the equation by G.

We obtain a nonlinear compact fixed-point problem:

$$u = G(u) \equiv \lambda\mathcal{G}(\max(u+\phi-\alpha, 0)).$$

In the numerical experiment we use central differences with 100 interior grid points and solve the problem with Anderson(m) for $m = 0, 1, 2, 3$.

In the computation we used $v_0 = 1$, $v_1 = .5$, $\lambda = 11.65$, and $\alpha = .8$. The value of λ was tuned to make the contractivity constant large so that Picard iteration performed very poorly.

We report two sets of results, one for L^2 optimization (Figure 3.2) and the other (Figure 3.3) for L^∞ optimization. We plot iteration histories and graphs of the solution v and $-v'' = \lambda \max(v-\alpha, 0)$. The plot of $-v''$ clearly shows that v'' is nonsmooth at the solution at only one point.

The L^∞ optimization problem can be expressed as a linear programming problem [9]. We solved that with the CVX MATLAB software [11, 12]. We used the SeDuMi solver and set the `precision` in cvx to `high`. Solving the optimization problem in L^2 is much easier, requiring only the solution of a linear least squares problem. It is temping to do the optimization problem in L^2 even though the theory requires an L^∞ optimization. In Figure 3.2 we do exactly that. On the right side of Figure 3.2 we plot graphs of v and $-v'' = \lambda \max(v - \alpha, 0)$. The plot of $-v''$ clearly shows that v'' is nonsmooth at the solution at only one point. On the left we plot the results using an L^2 optimization rather than the L^∞ optimization that the theory requires.

In Figure 3.3 we use the L^∞ norm for the optimization problem for the coefficients and show on the left the residual norms in the L^2 norm to best compare the two approaches. On the right we show the residual L^∞ norms. The figures show that the results are very similar and that the norm used for the optimization makes little difference.

We use (3.8) to estimate the r-factors for both L^2 and L^∞ optimization. The estimates in Table 3.2 are consistent with the results in Figures 3.2 and 3.3. In particular, we see that Picard is slowly convergent in this example and that there is little difference between the two norms used for optimization.

TABLE 3.2
Convergence rates for the Dirichlet problem.

Picard	Anderson 1	Anderson 2	Anderson 3
	L^2 optimization		
8.91e-01	2.34e-01	1.70e-01	1.56e-01
	L^∞ optimization		
8.91e-01	2.01e-01	1.77e-01	1.52e-01

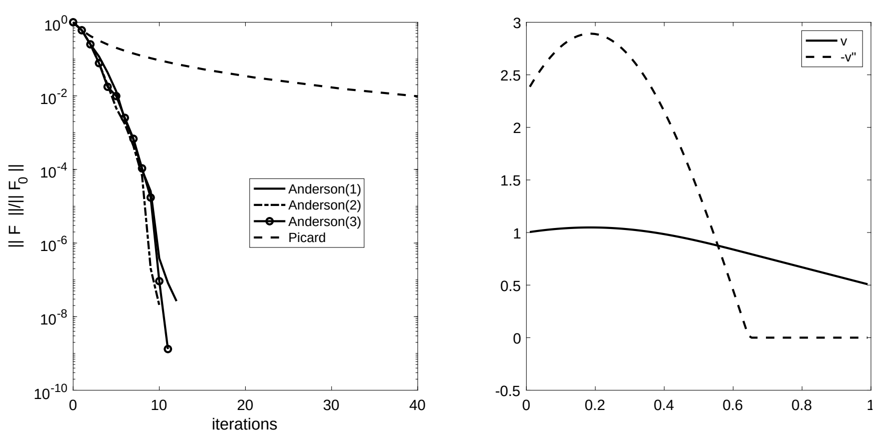

FIG. 3.2. *Example 2: Nonsmooth forcing term, L^2 optimization.*

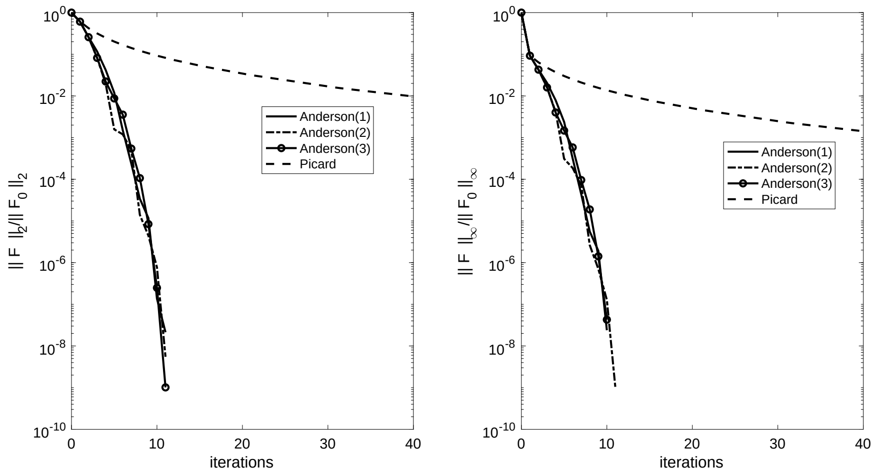

FIG. 3.3. *Example* 2*: Nonsmooth forcing term,* L^∞ *optimization.*

4. Conclusions. In this paper we prove convergence of Anderson acceleration for a class of nonsmooth fixed-point problems. Compositions of nonsmooth substitution operators and integral operators are examples of such problems. We illustrate the theoretical results with examples.

Acknowledgments. The authors are grateful to two referees and the associate editor for their very helpful comments and suggestions. Wei Bian and C. T. Kelley also appreciate the hospitality of the Department of Applied Mathematics of Hong Kong Polytechnic University while working on this paper.

REFERENCES

[1] D. G. Anderson, *Iterative procedures for nonlinear integral equations*, J. ACM, 12 (1965), pp. 547–560, https://doi.org/10.1145/321296.321305.

[2] N. N. Carlson and K. Miller, *Design and application of a gradient weighted moving finite element code* I*: In one dimension*, SIAM J. Sci. Comp., 19 (1998), pp. 766–798.

[3] X. Chen and C. T. Kelley, *Analysis of the EDIIS algorithm*, SIAM J. Sci. Comp., 41 (2019), pp. A365–A379, https://doi.org/10.1137/18M1171084.

[4] X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal., 38 (2001), pp. 1200–1216.

[5] X. Chen and T. Yamamoto, *Convergence domains of certain iterative methods for solving nonlinear equations*, Numer. Funct. Anal. Optim., 10 (1989), pp. 37–48.

[6] P. H. Dedrichs and R. Zeller, *Self-consistency iterations in electronic-structure calculations*, Phys. Rev. B, 28 (1983), pp. 5462–5472.

[7] J. Degroote, K.-J. Bathe, and J. Vierendeels, *Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction*, Comput. Structures, 97 (2009), pp. 793–801.

[8] C. Evans, S. Pollock, L. G. Rebholz, and M. Xiao, *A proof that Anderson acceleration improves the convergence rate in linearly converging fixed point methods (but not in those converging quadratically)*, SIAM J. Numer. Anal., 58 (2020), pp. 788–810.

[9] M. Ferris, O. Mangasarian, and S. Wright, *Linear Programming with MATLAB*, SIAM, Philadelphia, 2007.

[10] A. J. Garza and G. E. Scuseria, *Comparison of self-consistent field convergence acceleration techniques*, J. Chem. Phys., 137 (2012), 054110, https://doi.org/10.1063/1.4740249.

[11] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lect. Notes Control Inf. Sci., Springer-Verlag, Cham, 2008, pp. 95–110.

[12] M. Grant and S. Boyd, *CVX: MATLAB Software for Disciplined Convex Programming*, Version 2.0, http://cvxr.com/cvx, Aug. 2012.

[13] R. Haelterman, J. Degroote, D. van Heule, and J. Vierendeels, *The quasi-Newton least squares method: A new and fast secant method analyzed for linear systems*, SIAM J. Numer. Anal., 47 (2009), pp. 2347–2368.

[14] M. Heinkenschloss, C. T. Kelley, and H. T. Tran, *Fast algorithms for nonsmooth compact fixed point problems*, SIAM J. Numer. Anal., 29 (1992), pp. 1769–1792.

[15] S. Høst, J. Olsen, B. Jansík, L. Thøgersen, P. Jørgensen, and T. Helgaker, *The augmented Roothaan-Hall method for optimizing Hartree-Fock and Kohn-Sham density matrices*, J. Chem. Phys., 129 (2008), 124106, https://doi.org/10.1063/1.2974099.

[16] W. Hu, L. Lin, and C. Yang, *Projected commutator DIIS method for accelerating hybrid functional electronic structure calculations*, J. Chem. Theory Comput., 13 (2017), pp. 5458–5467, https://doi.org/10.1021/acs.jctc.7b00892.

[17] C. T. Kelley, *Identification of the support of nonsmoothness*, in Large Scale Optimization: State of the Art, W. W. Hager, D. W. Hearn, and P. Pardalos, eds., Kluwer Academic Publishers, Norwell, MA, 1994, pp. 192–205.

[18] K. N. Kudin, G. E. Scuseria, and E. Cancès, *A black-box self-consistent field convergence algorithm: One step closer*, J. Chem. Phys., 116 (2002), pp. 8255–8261, https://doi.org/10.1063/1.1470195.

[19] L. Lin and C. Yang, *Elliptic preconditioner for accelerating the self-consistent field iteration in Kohn–Sham density functional theory*, SIAM J. Sci. Comp., 35 (2013), pp. S277–S298.

[20] F. Lindner, M. Mehl, K. Scheufele, and B. Uekermann, *A comparison of various quasi-Newton schemes for partitioned fluid-structure interaction*, in ECCOMAS Coupled Problems in Science and Engineering, Venice, B. A. Schrefler, E. Oñate, and M. Papadrakakis, eds., CIMNE, Barcelona, 2015, pp. 477–488.

[21] K. Miller, *Nonlinear Krylov and moving nodes in the method of lines*, J. Comp. Appl. Math., 183 (2005), pp. 275–287.

[22] H. D. Mittelmann, *The obstacle Bratu problem*, in Computational Solution of Nonlinear Systems of Equations, E. L. Allgower and K. Georg, eds., AMS Lect. Appl. Math. 26, American Mathematical Society, Providence, RI, 1990, pp. 747–748.

[23] C. W. Oosterlee and T. Washio, *Krylov subspace accleration for nonlinear multigrid schemes*, SIAM J. Sci. Comput., 21 (2000), pp. 1670–1690.

[24] S. Pollock, L. G. Rebholz, and M. Xiao, *Anderson-accelerated convergence of Picard iterations for incompressible Navier–Stokes equations*, SIAM J. Numer. Anal., 57 (2019), pp. 615–637, https://doi.org/10.1137/18M1206151.

[25] P. Pulay, *Improved SCF convergence acceleration*, J. Comp. Chem., 3 (1982), pp. 556–560.

[26] T. Rohwedder and R. Schneider, *An analysis for the DIIS acceleration method used in quantum chemistry calculations*, J. Math. Chem., 49 (2011), pp. 1889–1914.

[27] R. Schneider, T. Rohwedder, A. Neelov, and J. Blauert, *Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure*, J. Comp. Math., 27 (2008), pp. 360–387.

[28] A. Toth, *A Theoretical Analysis of Anderson Acceleration and Its Application in Multiphysics Simulation for Light-Water Reactors*, Ph.D. thesis, North Carolina State University, Raleigh, NC, 2016.

[29] A. Toth, J. A. Ellis, T. Evans, S. Hamilton, C. T. Kelley, R. Pawlowski, and S. Slattery, *Local improvement results for Anderson acceleration with inaccurate function evaluations*, SIAM J. Sci. Comp., 39 (2017), pp. S47–S65, https://doi.org/10.1137/16M1080677.

[30] A. Toth and C. T. Kelley, *Convergence analysis for Anderson acceleration*, SIAM J. Numer. Anal., 53 (2015), pp. 805–819, https://doi.org/10.1137/130919398.

[31] H. W. Walker and P. Ni, *Anderson acceleration for fixed-point iterations*, SIAM J. Numer. Anal., 49 (2011), pp. 1715–1735.

[32] T. Washio and C. Oosterlee, *Krylov subspace acceleration for nonlinear multigrid schemes*, Electron. Trans. Numer. Anal., 6 (1997), pp. 271–290.