

## NEWTON-ANDERSON AT SINGULAR POINTS

MATT DALLAS AND SARA POLLOCK

**Abstract.** In this paper we develop convergence and acceleration theory for Anderson acceleration applied to Newton’s method for nonlinear systems in which the Jacobian is singular at a solution. For these problems, the standard Newton algorithm converges linearly in a region about the solution; and, it has been previously observed that Anderson acceleration can substantially improve convergence without additional a priori knowledge, and with little additional computation cost. We present an analysis of the Newton-Anderson algorithm in this context, and introduce a novel and theoretically supported safeguarding strategy. The convergence results are demonstrated with the Chandrasekhar H-equation and a variety of benchmark examples.

**Key words.** Anderson acceleration, Newton’s method, safeguarding, singular problems.

### 1. Introduction

Given a nonlinear function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and root  $x^*$  for which  $f(x^*) = 0$ , it is well known that if the derivative of  $f$  at the root is nonsingular, then Newton’s method exhibits quadratic convergence in a sufficiently small ball centered at the root. On the other hand, if the derivative is singular at  $x^*$ , e.g., at a bifurcation point [35], Newton’s method converges linearly in a star-like region containing the root [18]. This singular setting has been studied in great detail [8, 9, 18, 32, 33], and a number of acceleration schemes have been proposed and analyzed [4, 5, 8, 10, 19, 20, 22, 34]. For example, the Levenberg-Marquardt method featured in [4, 5, 20] is known to be effective for solving nonlinear systems with singular Jacobians under the local-error bound condition. The focus of this paper is the analysis and demonstration of an extrapolation scheme called Anderson acceleration, sometimes called Anderson mixing, e.g., [36], applied to Newton’s method, for singular problems. We will show that with the proposed safeguarding strategy, the method is both theoretically sound and can be beneficial in practice for singular problems.

Anderson acceleration was first proposed in [2], in the context of integral equations, to improve the convergence of fixed-point iterations. Anderson acceleration is an attractive method to improve the convergence of linearly converging fixed-point iterations due to its low computational cost, ease of implementation, and track record of success over a wide range of problems. The method recombines the  $m$  most recent update steps and iterates to form an accelerated iterate at each stage of a given fixed-point method, where the particular combination is generally given as the solution to a least-squares problem. Here  $m$  may be referred to as the algorithmic depth, which is often chosen small, say less than five [37], but may sometimes benefit from being taken substantially larger [31, 38]. The method has been found beneficial in diverse applications, such as the computation of canonical tensor decompositions [39], the study of block copolymer systems [36], geometry optimization and simulation [27], flow problems [24, 29], and electronic-structure computations [1, 14], to name a few. Substantial advances in understanding the

method in relation to generalized Broyden methods and (nonlinear) GMRES are developed in [13, 14, 38]. Recently, significant effort has been devoted to analyzing Anderson acceleration applied to contractive and noncontractive operators with certain nondegeneracy assumptions [12, 29, 31, 37].

Here, we will focus on the analysis of Anderson acceleration applied to Newton's method for a problem of the form  $f(x) = 0$ , when the derivative  $f'$  is singular at a root  $x^*$ . Rapid convergence of the accelerated scheme in comparison with standard Newton has been demonstrated numerically in this singular case [30], where it is also observed that it is generally both sufficient and preferable to set the algorithmic depth to  $m = 1$ . It was also found in [12] in a nondegenerate setting that Anderson accelerated Newton iterations with algorithmic depth  $m = 1$  could converge where Newton iterations failed, but that increasing  $m$  only slowed convergence. In the remainder, we will consider Anderson acceleration with depth  $m = 1$  applied to Newton iterations, which we will refer to simply as Newton-Anderson. In comparison to the accelerated Newton methods of [8, 10, 22], Newton-Anderson may be seen as advantageous as it does not require explicit knowledge of the order of the root (defined in section 7), or construction of an approximate projection mapping onto the null space of  $f'(x)$ . In further contrast to these predictor-corrector methods, Newton-Anderson requires a single linear solve per iteration. An analysis of Newton-Anderson in the one-dimensional singular case is presented in [28]; however, to our knowledge no previous convergence theory has been developed for dimension  $n > 1$ . The goal of this paper is to provide such a theory.

The remainder of the paper is organized as follows. The underlying foundation of the analysis relies on a technique for approximating the inverse of the derivative near a given point as developed in [8]. We discuss this technique in section 2, and in section 3 apply it to a Newton-Anderson step to obtain an expansion of the error at step  $k$ . We then analyze this expansion in sections 4 and 5 with a one-step analysis of the error based on previous consecutive error-pairs, revealing the mechanism behind the changes in convergence rate demonstrated by the method. The main challenge of proving convergence for any Newton-like method in the singular case is that the geometry of the region of invertibility is more complex. To handle this, in section 6 we introduce a novel safeguarding scheme, which we call  $\gamma$ -safeguarding. This technique leads to the main results of this paper: when the null space of the derivative at the root is one-dimensional, then under the same conditions implying local convergence of the standard Newton method e.g., [9, Theorem 1.2], Newton-Anderson with  $\gamma$ -safeguarding exhibits local convergence, and in general the rate of convergence is improved. We extend these results to high order roots in section 7. These results are then demonstrated numerically in section 8 with several standard benchmark problems, both singular and nonsingular, including the Chandrasekhar H-equation [6, 21]. The introduced  $\gamma$ -Newton-Anderson is further shown to perform favorably in comparison to existing methods developed for the problem class both in terms of robustness and efficiency.

## 2. Preliminaries

Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a  $C^3$  function such that  $f(x^*) = 0$  with  $x^* \in \mathbb{R}^n$ . This regularity assumption is standard for the problem class; see, for example [8, 9, 10, 16, 32, 33]. Suppose  $N = \text{null}(f'(x^*))$  is nontrivial, let  $R = \text{range}(f'(x^*))$ , and let  $\mathbb{R}^n = N \oplus R$ . Throughout this paper,  $B_r(x)$  denotes a ball of radius  $r > 0$  centered at  $x$ ,  $P_N$  and  $P_R$  denote the orthogonal projections onto  $N$  and  $R$  respectively. Denote the error by  $e_k = x_k - x^*$ , and the Newton update step by

$w_{k+1} = -f'(x_k)^{-1}f(x_k)$ . We take the *singular set*  $S$  to be the set of all  $x \in \mathbb{R}^n$  such that  $\det(f'(x)) = 0$ . When  $\|\cdot\| = \|\cdot\|_2$ , the *Newton-Anderson algorithm* reads as follows.

---

**Algorithm 1** Newton-Anderson

---

Choose  $x_0 \in \mathbb{R}^n$ . Set  $w_1 = -f'(x_0)^{-1}f(x_0)$ , and  $x_1 = x_0 + w_1$ .  
**for**  $k=1,2,\dots$  **do**  
 $w_{k+1} \leftarrow -f'(x_k)^{-1}f(x_k)$   
 $\gamma_{k+1} \leftarrow (w_{k+1} - w_k)^T w_{k+1} / \|w_{k+1} - w_k\|_2^2$   
 $x_{k+1} \leftarrow x_k + w_{k+1} - \gamma_{k+1}(x_k - x_{k-1} + w_{k+1} - w_k)$   
**end for**

---

Let  $\hat{D}_N(x)(\cdot) := P_N f''(x^*)(P_N(x - x^*), P_N(\cdot))$ . Here we're writing  $f''(x^*)$  as the bilinear map  $f''(x^*)(\cdot, \cdot)$  on  $\mathbb{R}^n \times \mathbb{R}^n$ . Hence  $\hat{D}_N(x)(\cdot)$  is a linear map from  $N$  to  $N$ . If  $\hat{D}_N(x)$  is invertible as a map on  $N$  whenever  $P_N(x - x^*) \neq 0$ , one can show that there exists constants  $\hat{\rho} > 0$  and  $\hat{\sigma} > 0$  for which  $f'(x)$  is invertible in the region

$$(1) \quad \hat{W} := W(\hat{\rho}, \hat{\sigma}, x^*) = \{x \in \mathbb{R}^n : \|x - x^*\| < \hat{\rho}, \|P_N(x - x^*)\| < \hat{\sigma}\|P_N(x - x^*)\|\},$$

and that the standard Newton iterates remain in this region for sufficiently small  $\|e_0\|$ . Further,  $f'(x)^{-1} = \hat{D}_N(x)^{-1} + \mathcal{O}(1)$  and  $\|f'(x)^{-1}\| \leq c\|x - x^*\|^{-1}$ . See [8, 9], or [18] for details. When accelerating Newton's method, care must be taken to ensure the accelerated iterates remain in  $W(\rho, \sigma, x^*)$ . We won't insist that the iterates lie in  $W(\rho, \sigma, x^*)$  in section 3 through section 5, as the primary focus is a one-step analysis. Rather, we'll make the more relaxed assumption that  $\hat{D}(x_i)(\cdot) = P_N f''(x^*)(e_i, P_N(\cdot))$  is invertible as a map on  $N$  in  $B_r(x^*) \setminus S$ , and the error  $\|e_i\|$  is so small that  $f'(x_i)$  is invertible so that  $f'(x_i)^{-1} = \hat{D}(x_i)^{-1} + \mathcal{O}(1)$  holds with  $i = k, k - 1$ . Note that near  $N$ ,  $\hat{D}(x_i) = P_N f''(x^*)(e_i, P_N)$  is invertible if and only if  $\hat{D}_N(x_i)$  is invertible. In section 6, we'll show that with safeguarding and sufficiently small  $\|e_0\|$ , the Newton-Anderson iterates remain in  $\hat{W}$  if  $x_0 \in \hat{W}$ . We remark that  $\hat{D}_N(x)$  fails to be invertible for all  $P_N(x - x^*) \neq 0$  if  $\dim N > 1$  and odd (see [17, p.148]). In these cases, one may instead assume that there exists a  $\varphi \in N$  for which the linear map  $\hat{D}(\varphi) = P_N f''(x^*)(\varphi, P_N)$  is invertible, and then work with the set

$$(2) \quad W(\rho, \sigma, \xi, x^*) = W \cap \{x \in \mathbb{R}^n : \|(P_N - P_\varphi)(x - x^*)\| < \xi\|P_\varphi(x - x^*)\|\}.$$

Here  $P_\varphi$  denotes the projection onto the one-dimensional subspace of  $N$  spanned by  $\varphi$  and  $W = W(\rho, \sigma, x^*)$ . This approach may be found in [8, 22, 33]. A more general analysis of regions of invertibility may be found in [18]. For our purposes studying rates of convergence, the sets  $W(\rho, \sigma, x^*)$  and  $W(\rho, \sigma, \xi, x^*)$  suffice. In section 6 we focus on the case when  $\dim N = 1$ , where we are able to work with the larger set  $W(\rho, \sigma, x^*)$ . For reference, we explicitly state

**Assumption 2.1.** *The linear operator  $\hat{D}(x)$  is invertible as a map on  $N$  in the ball  $B_{\hat{r}}(x^*) \setminus S$ , and  $\|x - x^*\| < \hat{r}$  implies  $f'(x)^{-1} = \hat{D}(x)^{-1} + \mathcal{O}(1)$ .*

Next we consider a standard error expansion for the analysis of Newton's method for singular problems, and show how the behavior of Newton-Anderson iterates differs from the behavior of Newton iterates without acceleration.

### 3. Error Expansion

**Theorem 3.1.** *Let Assumption 2.1 hold, and let  $x_k \in B_{\hat{r}}(x^*) \setminus S$ . Then*

$$(3) \quad e_k + w_{k+1} = \frac{1}{2}P_N e_k + \frac{1}{2}\hat{D}^{-1}(x_k)f''(x_k)(e_k, P_R e_k) + \mathcal{O}(\|e_k\|^2)$$

$$(4) \quad w_{k+1} = -\frac{1}{2}P_N e_k + \left( \frac{1}{2}\hat{D}^{-1}(x_k)f''(x_k)(e_k, \cdot) - I \right) P_R e_k + \mathcal{O}(\|e_k\|^2).$$

Bounds similar to (3) may be found in [8, 9, 33]. Note that (4) follows from (3) by subtracting  $e_k$  to the right hand side, writing  $e_k = P_N e_k + P_R e_k$ , and grouping like terms.

Henceforth, we will let  $T_k(\cdot) = (1/2)\hat{D}^{-1}(x_k)f''(x_k)(e_k, \cdot)$ . Let  $\gamma_{k+1}$  be the coefficient computed in Algorithm 1. For  $x, y \in \mathbb{R}^n$ , define the function

$$(5) \quad L_{k+1}(x, y) := (1 - \gamma_{k+1})x + \gamma_{k+1}y.$$

We will call any term of the form  $L_{k+1}(x, y)$  the *Newton-Anderson sum* of  $x$  and  $y$ . When the inputs are indexed, such as  $x_k$  and  $x_{k-1}$ , we write  $L_{k+1}(x_k, x_{k-1}) = x_k^\alpha$ . As in [31], we define the *optimization gain*

$$(6) \quad \theta_{k+1} := \|w_{k+1}^\alpha\|/\|w_{k+1}\| = \|w_{k+1} - \gamma_{k+1}(w_{k+1} - w_k)\|/\|w_{k+1}\|.$$

In [31],  $\theta_{k+1}$  was shown to be the key quantity determining the acceleration from a Newton-Anderson step in the nonsingular case. Here, in the singular case, it will be shown to be the key quantity determining the acceleration of the error along the null component. It was also shown in [31] that  $\theta_{k+1} = |\sin(w_{k+1} - w_k, w_k)|$ , where  $\sin(w_{k+1} - w_k, w_k)^2 = 1 - ((w_{k+1} - w_k)^T w_k)^2 / (\|w_{k+1} - w_k\|^2 \|w_k\|^2)$  is the direction sine between  $w_{k+1} - w_k$  and  $w_k$ . Therefore,  $\theta_{k+1}$  is small when  $w_{k+1} - w_k$  is nearly parallel to  $w_{k+1}$ .

With the notation described in the preceding paragraph, (3) and (4) become

$$(7) \quad e_k + w_{k+1} = \frac{1}{2}P_N e_k + T_k P_R e_k + \mathcal{O}(\|e_k\|^2)$$

$$(8) \quad w_{k+1} = -\frac{1}{2}P_N e_k + (T_k - I)P_R e_k + \mathcal{O}(\|e_k\|^2).$$

The following proposition provides the error expansions for a Newton-Anderson step that will be fundamental to our analysis.

**Proposition 3.1.** *Let  $q_{k-1}^k$  denote a term for which  $\|q_{k-1}^k\| \leq c(|1 - \gamma_{k+1}| \|e_k\|^2 + |\gamma_{k+1}| \|e_{k-1}\|^2)$ . Under the assumptions of theorem (3.1), if  $x_{k+1}$  is the  $(k+1)$ -st Newton-Anderson iterate, then we can expand  $e_{k+1}$  and  $w_{k+1}^\alpha$  as*

$$(9) \quad e_{k+1} = \frac{1}{2}P_N e_k^\alpha + (T_k P_R e_k)^\alpha + q_{k-1}^k$$

$$(10) \quad w_{k+1}^\alpha = -\frac{1}{2}P_N e_k^\alpha + ((T_k - I)P_R e_k)^\alpha + q_{k-1}^k.$$

*Proof.* Given  $x_k$  and  $x_{k-1}$ , a Newton-Anderson step takes the form

$x_{k+1} = L_{k+1}(x_k + w_{k+1}, x_{k-1} + w_k) = (1 - \gamma_{k+1})(x_k + w_{k+1}) + \gamma_{k+1}(x_{k-1} + w_k)$ . It follows that

$$(11) \quad e_{k+1} = (1 - \gamma_{k+1})(e_k + w_{k+1}) + \gamma_{k+1}(e_{k-1} + w_k).$$

Applying (7) to  $e_k + w_{k+1}$  and  $e_{k-1} + w_k$ , and grouping up appropriate terms yields

$$(12) \quad e_{k+1} = \frac{(1 - \gamma_{k+1})}{2} P_N e_k + \frac{\gamma_{k+1}}{2} P_N e_{k-1} + (1 - \gamma_{k+1}) T_k P_R e_k + \gamma_{k+1} T_{k-1} P_R e_{k-1} + (1 - \gamma_{k+1}) \mathcal{O}(\|e_k\|^2) + \gamma_{k+1} \mathcal{O}(\|e_{k-1}\|^2).$$

Writing  $q_{k-1}^k = (1 - \gamma_{k+1}) \mathcal{O}(\|e_k\|^2) + \gamma_{k+1} \mathcal{O}(\|e_{k-1}\|^2)$  and using the  $\alpha$  notation for a Newton-Anderson sum, we have

$$(13) \quad e_{k+1} = \frac{1}{2} P_N e_k^\alpha + (T_k P_R e_k)^\alpha + q_{k-1}^k.$$

By analogous reasoning, applying (8) to  $(1 - \gamma_{k+1})w_{k+1} + \gamma_{k+1}w_k$  yields (10).  $\square$

The structure of  $f$  leads to a simple upper bound on  $P_R e_{k+1}$ . Apply  $P_R$  to eq. (9). Since the range of  $T_k$  lies in  $N$ , the only term remaining on the right hand side is  $q_{k-1}^k$ . Thus

$$(14) \quad \|P_R e_{k+1}\| \leq |1 - \gamma_{k+1}| C_1 \|e_k\|^2 + |\gamma_{k+1}| C_2 \|e_{k-1}\|^2.$$

The constants  $C_1$  and  $C_2$  are independent of  $k$  and depend on  $f$ . The bound in (14) resembles the result of Lemma 1 in [30] for depth  $m = 1$ . There, the Jacobian is assumed to be nonsingular at the solution  $x^*$ , and the  $\gamma$  coefficients are assumed bounded. We do not assume  $\gamma_{k+1}$  is bounded for our one-step analysis in sections 4 and 5. When we consider convergence in section 6,  $\gamma$ -safeguarding will ensure that the  $\gamma_{k+1}$ 's remain bounded in the region of convergence. The point here is that like a standard Newton step, the range component of the error from a Newton-Anderson step behaves as if the Jacobian were nonsingular at the solution. Thus, when the Jacobian is singular at  $x^*$ , the source of slow convergence must come from the null space component. Conversely, if Newton-Anderson is seen to accelerate a given Newton sequence, and  $f'(x^*)$  is singular, then it must accelerate the null component. If  $P_R e_{k+1}$  is the only accelerated component, then the source of the linear convergence remains unaltered, and linear convergence would still be observed. Therefore much of the analysis focuses on the null component error, the foundation of which is the notion of *pair-types*. This method introduces a new technique to the one-step analysis of Anderson acceleration, and explains how the convergence rate changes at different steps for Newton-Anderson.

**3.1. Pair Types.** We may consider the Newton-Anderson algorithm as acting on ordered pairs  $(x_k, x_{k-1})$ , where  $x_k$  and  $x_{k-1}$  are the previous two Newton-Anderson iterates. We then analyze the output,  $x_{k+1}$ , based on where the vectors  $x_k$  and  $x_{k-1}$  lie in  $\mathbb{R}^n$ , e.g.,  $x_k$  and  $x_{k-1}$  both lie near  $N$ , or  $x_k$  lies near  $N$  and  $x_{k-1}$  lies near  $R$ . This leads to the notion of pair types, which we now define. Here, “dominant” means greatest in norm relative to the other terms on the right hand side of (7). We’ll make this more explicit in the next section.

**Definition 3.1.** Let  $\{x_k\}$  be a sequence of Newton-Anderson iterates.

- (1)  $(x_k, x_{k-1})$  is an N-pair if  $(1/2)P_N e_i$  is the dominant term in (7) for  $i = k, k - 1$ ,
- (2)  $(x_k, x_{k-1})$  is an R-pair if  $T_i P_R e_i$  is the dominant term in (7) for  $i = k, k - 1$ ,
- (3)  $(x_k, x_{k-1})$  is an NR-pair if  $(1/2)P_N e_i$  is the dominant term for  $i = k$ , and  $T_i P_R e_i$  is the dominant term for  $i = k - 1$ , and
- (4)  $(x_k, x_{k-1})$  is an RN-pair if  $T_i P_R e_i$  is the dominant term for  $i = k$ , and  $(1/2)P_N e_i$  is the dominant term for  $i = k - 1$ .

Each respective pair is a strong pair if the corresponding Newton-Anderson sum is dominant in equation (9).

Strong pairs are those for which the dominant terms in  $x_k$  and  $x_{k-1}$  remain dominant in the Newton-Anderson step. Our analysis focuses on strong pairs, and of particular interest is how Newton-Anderson acts on strong N-pairs, since near  $N$  the standard Newton method exhibits linear convergence. One of the most interesting, but perhaps not surprising, results is that strong N-pairs are the most responsive to a successful optimization step in the Newton-Anderson algorithm. In other words, the pairs that are “close” to the null space, the “slow” region for standard Newton, stand to gain the most from Anderson acceleration.

Figure 1 below summarizes the results of the technical analysis in sections 4 and 5. The left most path is optimal in the sense that we have effectively recovered the bound from the nonsingular case. The right most path is the most Newton-like case, where there is little or no acceleration of the null components. The middle path, in which  $(x_k, x_{k-1})$  is compatible (see definition 4.1), lies between these two cases. As  $\theta_{k+1} \rightarrow 1$ ,  $x_{k+1}$  tends to look more like a standard Newton step, and as  $\theta_{k+1} \rightarrow 0$ , we see acceleration of  $x_{k+1}$ . In this sense,  $\theta_{k+1}$  interpolates between the nonsingular case where we have superlinear order, and the singular case with linear order. Note that theorem 5.2 can be interpreted as saying that no more than two consecutive pairs can follow the right most path in fig. 1.

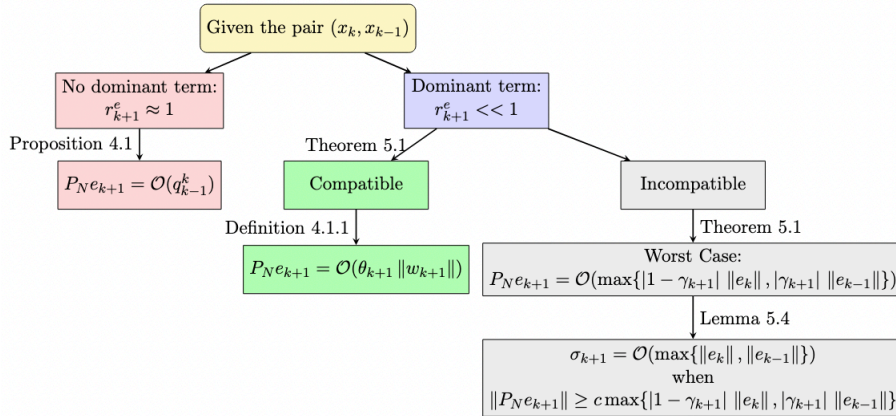


FIGURE 1. Summary of results from sections section 4 and section 5, with labels noting where each relation is proven or defined.

#### 4. The Null Space Component and Compatibility

Our goal in this section is to derive a general relation between the null component error and the optimization gain  $\theta_{k+1}$ . Such a relation will allow us to rigorously describe how the optimization step in algorithm 1 accelerates the null component. This relation is given in proposition (4.1). The strategy is to take the largest term (with respect to the norm) in the expansion of  $P_N e_{k+1}$  from (9) and find conditions under which this term is bounded by  $\theta_{k+1} ||w_{k+1}||$ . We’ll then see (in the next section) when each strong pair-type satisfies these conditions. First, consider each term in expansion (9). Namely,

$$(15) \quad \frac{(1-\gamma_{k+1})P_N e_k}{2}, \frac{\gamma_{k+1}P_N e_{k-1}}{2}, (1-\gamma_{k+1})T_k P_R e_k, \gamma_{k+1}T_{k-1} P_R e_{k-1}, \text{ and } P_N q_{k-1}^k.$$

Any sum of a given strict subset of these terms could be the dominant term in equation (9). We'll let  $S_{k+1}^e$  denote the set of all such sums. For example, if  $(x_k, x_{k-1})$  is a strong N-pair as defined in Definition (3.1), then  $(1 - \gamma_{k+1})P_N e_k + \gamma_{k+1}P_N e_{k-1}$  is the dominant term. We'll write

$$(16) \quad R_{k+1}^e := \{ \|P_N e_{k+1} - \tilde{E}_{k+1}\| / \|\tilde{E}_{k+1}\| : \tilde{E}_{k+1} \in S_{k+1}^e \},$$

and  $E_{k+1} := \operatorname{argmin} R_{k+1}^e$ . We then define

$$(17) \quad r_{k+1}^e := \min R_{k+1}^e = \|P_N e_{k+1} - E_{k+1}\| / \|E_{k+1}\|.$$

Observe that if  $q \in R_{k+1}^e$ , then  $1/q \in R_{k+1}^e$ . This gives the following lemma.

**Lemma 4.1.** *The minimum ratio  $r_{k+1}^e$  satisfies  $r_{k+1}^e \leq 1$  for all  $k \geq 0$ .*

Analogous notation will be used for the expansion  $P_N w_{k+1}^\alpha$  in (10). Namely,  $S_{k+1}^w$  will denote the set of sums of strict subsets of the terms

$$(18) \quad -\frac{(1-\gamma_{k+1})}{2}P_N e_k, -\frac{\gamma_{k+1}P_N e_{k-1}}{2}, (1 - \gamma_{k+1})T_k P_R e_k, \gamma_{k+1}T_{k-1}P_R e_{k-1}, \text{ and } P_N q_{k-1}^k,$$

We'll write  $R_{k+1}^w = \{ \|P_N w_{k+1}^\alpha - \tilde{F}_{k+1}\| / \|\tilde{F}_{k+1}\| : \tilde{F}_{k+1} \in S_{k+1}^w \}$ ,  $F_{k+1} := \operatorname{argmin} R_{k+1}^w$ , and  $r_{k+1}^w = \min R_{k+1}^w$ .

We now state and prove the proposition referenced at the beginning of this section that provides a bound on the null space component of the error in terms of the optimization gain  $\theta_{k+1}$  (defined in (6)).

**Proposition 4.1.** *Let the assumptions of theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ . Let  $q_{k-1}^k$  denote a term for which  $\|q_{k-1}^k\| \leq c(|1 - \gamma_{k+1}| \|e_k\|^2 + |\gamma_{k+1}| \|e_{k-1}\|^2)$ , and suppose*

$$(19) \quad \|E_{k+1}\| \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - r_{k+1}^e}.$$

Then

$$(20) \quad \|P_N e_{k+1}\| \leq \left( \frac{1 + r_{k+1}^e}{1 - r_{k+1}^e} \right) \theta_{k+1} \|w_{k+1}\|.$$

*Proof.* Adding and subtracting  $E_{k+1}$  to  $P_N e_{k+1}$  gives  $\|P_N e_{k+1}\| \leq (1 + r_{k+1}^e) \|E_{k+1}\| \leq (1 + r_{k+1}^e)(1 - r_{k+1}^e)^{-1} \|P_N w_{k+1}^\alpha\|$ . Then observing that

$$\left( \frac{1 + r_{k+1}^e}{1 - r_{k+1}^e} \right) \|P_N w_{k+1}^\alpha\| \leq \left( \frac{1 + r_{k+1}^e}{1 - r_{k+1}^e} \right) \|w_{k+1}^\alpha\| = \left( \frac{1 + r_{k+1}^e}{1 - r_{k+1}^e} \right) \theta_{k+1} \|w_{k+1}\|$$

completes the proof. □

We'll soon state definition 4.1, which is motivated by proposition 4.1, but first we'll prove two related propositions. The following gives sufficient conditions for (19) to hold.

**Proposition 4.2.** *The relation (19) holds if  $r_{k+1}^e < 1$ , and there is an element  $\tilde{F}_{k+1} \in S_{k+1}^w$  such that  $\|\tilde{F}_{k+1}\| = \|E_{k+1}\|$ , and  $\|P_N e_{k+1} - E_{k+1}\| = \|P_N w_{k+1}^\alpha - \tilde{F}_{k+1}\|$ .*

*Proof.* Suppose  $\|\tilde{F}_{k+1}\| = \|E_{k+1}\|$  and  $\|P_N e_{k+1} - E_{k+1}\| = \|P_N w_{k+1}^\alpha - \tilde{F}_{k+1}\|$ . Then  $r_{k+1}^e = \|P_N w_{k+1}^\alpha - \tilde{F}_{k+1}\| / \|\tilde{F}_{k+1}\|$ , and  $r_{k+1}^e \leq 1$  by lemma 4.1. If  $r_{k+1}^e < 1$ , then

$$(21) \quad \|P_N w_{k+1}^\alpha\| \geq \|\tilde{F}_{k+1}\| - \|P_N w_{k+1}^\alpha - \tilde{F}_{k+1}\| = (1 - r_{k+1}^e) \|E_{k+1}\|.$$

□

As  $r_{k+1}^e$  approaches 1, the denominator in (20) approaches zero, resulting in a poor bound. However, the following proposition shows that  $r_{k+1}^e \approx 1$  implies  $P_N e_{k+1}$  was accelerated.

**Proposition 4.3.** *Let  $r_{k+1}^e$  be defined as in (17) and let the assumptions of theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ . If  $0 \leq \varepsilon < 1$  and  $r_{k+1}^e = 1 - \varepsilon$ , then  $P_N e_{k+1} = ((2 - \varepsilon)/(1 - \varepsilon))\mathcal{O}(\max\{\|e_k\|^2, \|e_{k-1}\|^2\})$ .*

*Proof.* Since  $r_{k+1}^e \leq \|P_N e_{k+1} - \tilde{E}_{k+1}\|/\|\tilde{E}_{k+1}\|$  for any  $\tilde{E}_{k+1} \in S_{k+1}^e$ , we have

$$(22) \quad 1 - \varepsilon = r_{k+1}^e \leq \frac{\|P_N e_{k+1} - \tilde{E}_{k+1}\|}{\|\tilde{E}_{k+1}\|}$$

for all  $\tilde{E}_{k+1} \in S_{k+1}^e$ . Taking  $\tilde{E}_{k+1} = (1/2)P_N e_k^\alpha + T_k P_R e_k^\alpha$ , it follows that

$$(23) \quad 1 - \varepsilon \leq \frac{\|q_{k-1}^k\|}{\|(1/2)P_N e_k^\alpha + T_k P_R e_k^\alpha\|}.$$

Thus  $(1 - \varepsilon)\|(1/2)P_N e_k^\alpha + T_k P_R e_k^\alpha\| \leq \|q_{k-1}^k\|$ , and it follows that  $\|P_N e_{k+1}\| \leq ((2 - \varepsilon)/(1 - \varepsilon))\|q_{k-1}^k\|$ . □

Now we state definition 4.1.

**Definition 4.1.** *Let  $\{x_k\}$  be a sequence of Newton-Anderson iterates. We say that  $x_{k+1}$  is compatible or a compatible step if there exists a moderate constant  $C > 0$  independent of  $k$  such that  $\|P_N e_{k+1}\| \leq C\theta_{k+1}\|w_{k+1}\|$ , in which case we'll write  $P_N e_{k+1} = \mathcal{O}(\theta_{k+1}\|w_{k+1}\|)$ . Otherwise,  $(x_k, x_{k-1})$  is an incompatible pair, and  $x_{k+1}$  is incompatible or an incompatible step*

We note that this particular use of  $\mathcal{O}$  is common, e.g., [40].

*Compatible* here is suggestive of the result of proposition 4.1. If  $(x_k, x_{k-1})$  is a pair that satisfies proposition 4.1, then  $P_N e_{k+1} = \mathcal{O}(\theta_{k+1}\|w_{k+1}\|)$ . Hence a successful optimization step in Newton-Anderson implies acceleration of  $P_N e_{k+1}$ , and therefore of  $e_{k+1}$ . This can be seen by applying (8) to bound  $w_{k+1}$  in terms of  $e_k$ . In particular, when  $(x_k, x_{k-1})$  is a strong N-pair, (8) implies that  $\|w_{k+1}\| \leq (1/2)(1 + c_1\sigma_k + c_2(1 + \sigma_k)\|e_k\|)\|P_N e_k\|$ , where  $\sigma_k = \|P_R e_k\|/\|P_N e_k\|$ , and  $c_1$  and  $c_2$  are constants determined by  $f$ . Combining this with compatibility gives

$$(24) \quad \|P_N e_{k+1}\| \leq C\theta_{k+1}(1/2)(1 + c_1\sigma_k + c_2(1 + \sigma_k)\|e_k\|)\|P_N e_k\|.$$

The error for the null component of an analogous standard Newton step, i.e., one where  $(1/2)\|P_N e_k\|$  is norm-dominant on the right-hand-side of (7), is given by

$$(25) \quad \|P_N(e_k + w_{k+1})\| \leq (1/2)(1 + c_1\sigma_k + c_2(1 + \sigma_k)\|e_k\|)\|P_N e_k\|.$$

In  $\hat{W}$ , the region where the Jacobian is invertible defined in section 1, we can bound  $\sigma_k \leq \hat{\sigma}$  and  $\|e_k\| \leq \hat{\rho}$ , and we can expect  $C$  in (24) to be moderate in size. In appendix B, with  $\gamma$ -safeguarding, we obtain a bound of the form (24) with  $C < 1$  for sufficiently small  $\hat{\sigma}$  and  $\hat{\rho}$ . Hence the bound in (24) for a Newton-Anderson step in  $\hat{W}$  is essentially the bound seen in (25) for a standard Newton step in  $\hat{W}$  scaled by  $\theta_{k+1}$ . Since  $\theta_{k+1} \leq 1$ , this implies that in  $\hat{W}$ , a Newton-Anderson step will be no worse than a Newton step, and in the case of a successful optimization step,



i.e., when  $\theta_{k+1}$  is small, we have acceleration. Compatible steps also nicely mirror the Anderson theory developed in [31] under certain nondegeneracy assumptions, where it's shown that small  $\theta_{k+1}$  results in acceleration, and  $\theta_{k+1} \approx 1$  results in a standard, non-accelerated step. In this sense, compatible pairs are those that behave like nonsingular pairs, where the mechanism behind the acceleration is  $\theta_{k+1}$ . When  $(x_k, x_{k-1})$  is *incompatible*, i.e.,  $P_N e_{k+1} \neq \mathcal{O}(\theta_{k+1} \|w_{k+1}\|)$ , acceleration may still occur. However, this is not guaranteed, and a successful optimization step does not imply acceleration of  $P_N e_{k+1}$ .

So far, it has been shown that the expansion of  $P_N e_{k+1}$  in (9) either has a dominant term, and this term is  $\mathcal{O}(\theta_{k+1} \|w_{k+1}\|)$  when  $(x_k, x_{k-1})$  is a compatible pair, or there is no dominant term and  $P_N e_{k+1} = \mathcal{O}(\max\{\|e_k\|^2, \|e_{k-1}\|^2\})$ .

**5. Analysis of Pair Types**

In this section, the results of section 4 are applied to pair types. As stated in the paragraph following definition 3.1, we focus on strong pair types. If a given pair  $(x_k, x_{k-1})$  is not strong, and there is a dominant term in equation (9), then this term can be analyzed analogously to the strong terms analyzed here. Moreover, strong pairs, in particular strong N-pairs, are most relevant for the convergence theory developed in section 6, and theorem 5.2 proven later in this section essentially says that after two consecutive steps with little improvement in the null component of the error, the next step will either be a strong N-pair or be bounded only by higher order terms.

**5.1. Compatibility Conditions.** Our aim here is to establish conditions for each strong pair type under which compatibility is assured. That is, conditions under which  $P_N e_{k+1} = \mathcal{O}(\theta_{k+1} \|w_{k+1}\|)$ . Evidently, each pair type can be compatible if certain alignment conditions are met. These vary by pair type, with strong mixed pairs having the most stringent alignment conditions. On the other hand, strong N-pairs are automatically compatible.

**Lemma 5.1.** *Let the assumptions in theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ , and suppose  $r_{k+1}^e < 1$ . If  $(x_k, x_{k-1})$  is a strong N-pair, then  $(x_k, x_{k-1})$  is a compatible pair.*

*Proof.* Let  $E_{k+1}$  and  $F_{k+1}$  be defined as in the discussion preceding proposition 4.1. To show that a strong N-pair is compatible, it suffices to prove (19). Here, the condition from proposition 4.2 is used. Suppose  $(x_k, x_{k-1})$  is a strong N-pair, so that  $E_{k+1} = (1/2)P_N e_k^\alpha$ . Let  $F_{k+1} = -(1/2)P_N e_k^\alpha$ . Then  $\|E_{k+1}\| = \|F_{k+1}\|$ , and

$$(26) \quad \|P_N w_{k+1}^\alpha - F_{k+1}\| = \|T_k P_R e_k^\alpha + P_N q_{k-1}^k\| = \|P_N e_{k+1} - E_{k+1}\|.$$

Relation (19) then follows by proposition 4.2, by which  $\|P_N e_{k+1}\| = \mathcal{O}(\theta_{k+1} \|w_{k+1}\|)$ . Thus  $(x_k, x_{k-1})$  is a compatible pair.  $\square$

**Remark 5.1.** *Lemma 5.1 is significant because it says that if  $x_k$  and  $x_{k-1}$  are close to  $N$ , and the resulting Newton-Anderson step remains near  $N$ , the region in which standard Newton is slowest, then  $(x_k, x_{k-1})$  is guaranteed to be compatible, and therefore  $P_N e_{k+1}$  is controlled by  $\theta_{k+1} \|w_{k+1}\|$  which implies acceleration when  $\theta_{k+1}$  is small as described in the paragraph following definition 4.1. This justifies the statement at the end of section 3 that pairs near  $N$  are the most responsive to a successful optimization step in the Newton-Anderson algorithm. In section 6 we introduce  $\gamma$ -safeguarding, which will ensure the iterates remain near  $N$  if  $x_0$  is chosen near  $N$ , thereby guaranteeing the iterates remain well-defined and compatible.*

Lemma 5.2 gives conditions in which a strong R-pair is compatible, and a bound for the incompatible case. Note that in the case of an incompatible strong R-pair, we obtain a quadratic bound. However, we have no guarantee that the new iterate generated from this strong R-pair will be well-defined, as the Jacobian may not be invertible at this iterate.

**Lemma 5.2.** *Let the assumptions in theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ , let  $(x_k, x_{k-1})$  be a strong R-pair, and suppose  $r_{k+1}^e < 1$ . Then  $(x_k, x_{k-1})$  is compatible if  $(P_N e_k^\alpha)^T q_{k-1}^k \geq 0$ . Otherwise, if  $(x_k, x_{k-1})$  is incompatible, then*

$$(27) \quad \|P_N e_{k+1}\| \leq (1 + r_{k+1}^e) \left( |1 - \gamma_{k+1}| C_1 q_{k-2}^{k-1} + |\gamma_{k+1}| C_2 q_{k-3}^{k-2} \right),$$

where  $q_{i-1}^i$  is a term such that  $\|q_{i-1}^i\| \leq |1 - \gamma_{i+1}| C_3 \|e_i\|^2 + C_4 \mathcal{O}(\|e_{i-1}\|^2)$  for  $i \in \{k-2, k-1\}$ , where  $C_3$  and  $C_4$  are constants determined by  $f$ .

*Proof.* First, it's shown that  $(P_N e_k^\alpha)^T q_{k-1}^k \geq 0$  implies  $(x_k, x_{k-1})$  is compatible. Since  $(x_k, x_{k-1})$  is strong R-pair, the dominant term  $E_{k+1}$  is  $T_k P_R e_k^\alpha$ . By proposition 4.1, it suffices to show

$$(28) \quad \|T_k P_R e_k^\alpha\| \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - r_{k+1}^e},$$

where  $r_{k+1}^e$  is defined as in (17). Here  $r_{k+1}^e = \|(1/2)P_N e_k^\alpha + q_{k-1}^k\| / \|T_k P_R e_k^\alpha\|$ .

For any vectors  $u$  and  $v$  in  $\mathbb{R}^n$ , the polarization identity  $4u^T v = \|u+v\|^2 - \|u-v\|^2$  implies that if  $u^T v \geq 0$ , then  $\|u-v\| \leq \|u+v\|$ . Letting  $u = q_{k-1}^k$  and  $v = (1/2)P_N e_k^\alpha$ , it follows that if  $(P_N e_k^\alpha)^T q_{k-1}^k \geq 0$ , then  $\|q_{k-1}^k - (1/2)P_N e_k^\alpha\| \leq \|q_{k-1}^k + (1/2)P_N e_k^\alpha\|$ . By (10),

$$(29) \quad \|T_k P_R e_k^\alpha\| \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - \frac{\|q_{k-1}^k - (1/2)P_N e_k^\alpha\|}{\|T_k P_R e_k^\alpha\|}} \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - \frac{\|q_{k-1}^k + (1/2)P_N e_k^\alpha\|}{\|T_k P_R e_k^\alpha\|}}.$$

The last inequality follows because  $\|q_{k-1}^k - (1/2)P_N e_k^\alpha\| \leq \|q_{k-1}^k + (1/2)P_N e_k^\alpha\|$ . Since  $r_{k+1}^e = \|(1/2)P_N e_k^\alpha + q_{k-1}^k\| / \|T_k P_R e_k^\alpha\|$ ,  $(x_k, x_{k-1})$  is compatible by proposition 4.1. When  $(x_k, x_{k-1})$  is incompatible, we still have  $\|P_N e_{k+1}\| \leq (1 + r_{k+1}^e) \|T_k P_R e_k^\alpha\|$ . After applying the triangle inequality and (14) to  $\|T_k P_R e_k^\alpha\|$  we arrive at

$$(30) \quad \|T_k P_R e_k^\alpha\| \leq |1 - \gamma_{k+1}| C_1 q_{k-2}^{k-1} + |\gamma_{k+1}| C_2 q_{k-3}^{k-2}.$$

This gives the bound in (27).  $\square$

Now we come to strong mixed pairs, the case with the most stringent alignment conditions. As in proposition 4.1,  $E_{k+1} = \operatorname{argmin} R_{k+1}^e$ , where  $R_{k+1}^e$  is defined in (16). We will also use the *Newton-Anderson sum* notation  $L_k(x, y)$  defined in (5).

**Lemma 5.3.** *Let the assumptions of theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ , and suppose  $r_{k+1}^e < 1$ . Suppose  $(x_k, x_{k-1})$  is a strong mixed pair. There are two cases.*

(1) (Strong NR-pair) *If  $E_{k+1} = L_{k+1}(P_N e_k, T_{k-1} P_R e_{k-1})$ , then  $(x_k, x_{k-1})$  is compatible if*

- (a)  $((1 - \gamma_{k+1})P_N e_k)^T (\gamma_{k+1} T_{k-1} P_R e_{k-1}) \leq 0$ , and
- (b)  $(\gamma_{k+1} P_N e_{k-1})^T ((1 - \gamma_{k+1})T_k P_R e_k + q_{k-1}^k) \geq 0$ .

*When  $(x_k, x_{k-1})$  is not compatible,*

$$(31) \quad \|P_N e_{k+1}\| \leq C(1 + r_{k+1}^e) \max\{|1 - \gamma_{k+1}| \|P_N e_k\|, |\gamma_{k+1}| \|T_{k-1} P_R e_{k-1}\|\}.$$

- (2) (Strong RN-pair) If  $E_{k+1} = L_{k+1}(T_k P_R e_k, P_N e_{k-1})$ , then  $(x_k, x_{k-1})$  is compatible if
- (a)  $((1 - \gamma_{k+1})T_k P_R e_k)^T (\gamma_{k+1} P_N e_{k-1}) \leq 0$ , and
  - (b)  $((1 - \gamma_{k+1} P_N e_k)^T (\gamma_{k+1} T_{k-1} P_R e_{k-1} + q_{k-1}^k) \geq 0$ .
- When  $(x_k, x_{k-1})$  is not compatible,

$$(32) \quad \|P_N e_{k+1}\| \leq C(1 + r_{k+1}^e) \max\{|1 - \gamma_{k+1}| \|T_k P_R e_k\|, (|\gamma_{k+1}|/2) \|P_N e_{k-1}\|\}.$$

In both cases,  $C$  denotes a constant determined by  $f$ .

*Proof.* The proof is very similar to that of lemma 5.2. We only discuss the part that uses the additional alignment conditions, namely conditions (a) in the lemma statement. Further, the proofs are identical for strong NR-pairs and strong RN-pairs. We'll focus on strong NR-pairs. By proposition 4.1, it suffices to show  $\|E_{k+1}\| \leq \|P_N w_{k+1}^\alpha\|/(1 - r_{k+1}^e)$ . Suppose  $(x_k, x_{k-1})$  is a strong NR-pair. Then  $E_{k+1} = (1 - \gamma_{k+1})P_N e_k/2 + \gamma_{k+1}T_{k-1}P_R e_{k-1}$ . Write (10) as

$$(33) \quad \begin{aligned} P_N w_{k+1}^\alpha &= -\frac{1}{2}(1 - \gamma_{k+1})P_N e_k + \gamma_{k+1}T_{k-1}P_R e_{k-1} \\ &\quad - \frac{1}{2}\gamma_{k+1}P_N e_{k-1} + (1 - \gamma_{k+1})T_k P_R e_k + P_N q_{k-1}^k. \end{aligned}$$

Applying the reverse triangle inequality gives

$$(34) \quad \begin{aligned} \|P_N w_{k+1}^\alpha\| &\geq \left\| -\frac{1}{2}(1 - \gamma_{k+1})P_N e_k + \gamma_{k+1}T_{k-1}P_R e_{k-1} \right\| \\ &\quad - \left\| -\frac{1}{2}\gamma_{k+1}P_N e_{k-1} + (1 - \gamma_{k+1})T_k P_R e_k + P_N q_{k-1}^k \right\|. \end{aligned}$$

Factoring  $\|-\frac{1}{2}(1 - \gamma_{k+1})P_N e_k + \gamma_{k+1}T_{k-1}P_R e_{k-1}\|$  and dividing gives

$$(35) \quad \|\gamma_{k+1}T_{k-1}P_R e_{k-1} - \frac{(1-\gamma_{k+1})}{2}P_N e_k\| \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - \frac{\|(1-\gamma_{k+1})T_k P_R e_k - (\gamma_{k+1}/2)P_N e_{k-1} + q_{k-1}^k\|}{\|\gamma_{k+1}T_{k-1}P_R e_{k-1} - ((1-\gamma_{k+1})/2)P_N e_k\|}}.$$

Since  $((1 - \gamma_{k+1})P_N e_k)^T (\gamma_{k+1}T_{k-1}P_R e_{k-1}) \leq 0$ , the polarization identity implies that  $\|\gamma_{k+1}T_{k-1}P_R e_{k-1} + ((1 - \gamma_{k+1})/2)P_N e_k\| \leq \|\gamma_{k+1}T_{k-1}P_R e_{k-1} - ((1 - \gamma_{k+1})/2)P_N e_k\|$ . Hence

$$(36) \quad \|((1 - \gamma_{k+1})/2)P_N e_k + \gamma_{k+1}T_{k-1}P_R e_{k-1}\| \leq \frac{\|P_N w_{k+1}^\alpha\|}{1 - \frac{\|(1-\gamma_{k+1})T_k P_R e_k - (\gamma_{k+1}/2)P_N e_{k-1} + q_{k-1}^k\|}{\|\gamma_{k+1}T_{k-1}P_R e_{k-1} + ((1-\gamma_{k+1})/2)P_N e_k\|}}.$$

Now apply  $(\gamma_{k+1}P_N e_{k-1})^T ((1 - \gamma_{k+1})T_k P_R e_k + q_{k-1}^k) \geq 0$  to the right hand side of (36), and proceed as in the proof of lemma 5.2. The proof of the strong RN-pair case is identical.  $\square$

**Remark 5.2.** *In principal, as long as  $f'(x_0)$  and  $f'(x_1)$  are invertible,  $x_0$  and  $x_1$  can be chosen arbitrarily, but this strategy may lead to a mixed pair  $(x_0, x_1)$ . The stringent alignment conditions on strong mixed pairs suggest the better strategy is to chose  $x_0$  near  $N$  and take  $x_1 = x_0 + w_1$  as written in algorithm 2. This way both  $x_1$  and  $x_0$  will be near  $N$ , and in section 6 this will be shown to imply compatibility when  $\gamma$ -safeguarding is used.*

The results of lemmas 5.1-5.3 together are summarized in the following theorem.

**Theorem 5.1.** *Let the assumptions of theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ . Suppose  $(x_k, x_{k-1})$  is a strong pair. If  $(x_k, x_{k-1})$  is a strong N-pair, then  $(x_k, x_{k-1})$  is compatible. Otherwise, there are alignment conditions under which  $(x_k, x_{k-1})$  is*

compatible. If  $(x_k, x_{k-1})$  is an incompatible strong  $R$ -pair, then  $P_N e_{k+1}$  is bounded by higher order terms. If  $(x_k, x_{k-1})$  is an incompatible strong mixed pair, then  $P_N e_{k+1}$  is no worse than  $\mathcal{O}(\max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\})$ .

**5.2. Incompatible Pairs.** In this section, we investigate the incompatible case further. We will use the quantity  $\sigma_{k+1} := \|P_R e_{k+1}\| / \|P_N e_{k+1}\|$  introduced in section 2. One can think of  $\sigma_{k+1}$  as measuring the angle between  $e_{k+1}$  and the null space  $N$ , but it can also be interpreted as measuring the acceleration of  $\|P_N e_{k+1}\|$  relative to the standard Newton algorithm. In a standard Newton step,  $P_N e_{k+1} = \mathcal{O}(\|e_k\|)$ , and  $P_R e_{k+1} = \mathcal{O}(\|e_k\|^2)$ . Hence  $\sigma_{k+1} = \mathcal{O}(\|e_k\|)$ . So we can think of any Newton-Anderson step as “Newton-like” when  $\sigma_{k+1} = \mathcal{O}(\max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\})$ , and as accelerated, if  $\sigma_{k+1} = \mathcal{O}(\max\{|1 - \gamma_{k+1}| \|e_k\|^\ell, |\gamma_{k+1}| \|e_{k-1}\|^\ell\})$  with  $0 \leq \ell < 1$ . From this perspective, the worst-case for a Newton-Anderson step is when  $P_N e_{k+1}$  is incompatible and bounded below by first order in the sense that  $\|E_{k+1}\| > C \max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\}$ . This can occur if there is a single term in (15) that is much larger than the others in norm, or if  $(x_k, x_{k-1})$  is an incompatible strong mixed pair since  $P_N e_{k+1} = \mathcal{O}(|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|)$  in the worst case by lemma 5.3.

**Lemma 5.4.** *Let the assumptions of theorem 3.1 hold for  $x_k$  and  $x_{k-1}$ . As in proposition 4.1, let  $E_{k+1} = \operatorname{argmin} R_{k+1}^e$ . If*

$$(37) \quad \|E_{k+1}\| > C \max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\},$$

where  $C > 0$  is some constant dependent on  $f$ , and  $1 - r_{k+1}^e \neq 0$ , then  $\sigma_{k+1} = \mathcal{O}(\max\{\|e_k\|, \|e_{k-1}\|\})$ .

*Proof.* By (14),  $\|P_R e_{k+1}\| \leq C \max\{|1 - \gamma_{k+1}| \|e_k\|^2, |\gamma_{k+1}| \|e_{k-1}\|^2\}$ . Let  $M_{k,k-1} = \max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\}$ . The assumption on  $E_{k+1}$  allows us to bound  $\|P_N e_{k+1}\|$  below:

$$(38) \quad \begin{aligned} \|P_N e_{k+1}\| &= \|P_N e_{k+1} - E_{k+1} + E_{k+1}\| \\ &\geq (1 - r_{k+1}^e) \|E_{k+1}\| \\ &\geq C(1 - r_{k+1}^e) M_{k,k-1}. \end{aligned}$$

Combining the constants from the upper bound on  $\|P_R e_{k+1}\|$  and the lower bound on  $\|P_N e_{k+1}\|$ , we have

$$(39) \quad \begin{aligned} \sigma_{k+1} = \frac{\|P_R e_{k+1}\|}{\|P_N e_{k+1}\|} &\leq \frac{C}{1 - r_{k+1}^e} \frac{\max\{|1 - \gamma_{k+1}| \|e_k\|^2, |\gamma_{k+1}| \|e_{k-1}\|^2\}}{\max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\}} \\ &\leq \frac{C}{1 - r_{k+1}^e} \max\{\|e_k\|, \|e_{k-1}\|\}. \end{aligned}$$

The last inequality holds since  $\max\{|1 - \gamma_{k+1}| \|e_k\|, |\gamma_{k+1}| \|e_{k-1}\|\} \geq |1 - \gamma_{k+1}| \|e_k\|$  and  $|\gamma_{k+1}| \|e_{k-1}\|$ .  $\square$

**Remark 5.3.** *Under the same assumptions, we can make the slightly stronger statement: Suppose  $\|E_{k+1}\| > C \max\{|1 - \gamma_{k+1}| \|e_k\|^\ell, |\gamma_{k+1}| \|e_{k-1}\|^\ell\}$ , where  $\ell \in \mathbb{R}$  and  $0 < \ell < 2$ , then  $\sigma_{k+1} \leq (C/(1 - r_{k+1}^e)) \max\{\|e_k\|^{2-\ell}, \|e_{k-1}\|^{2-\ell}\}$ . In this paper, however, we'll only use the case  $\ell = 1$  discussed in lemma 5.4.*

By definition,  $\|P_R e_{k+1}\| = \sigma_{k+1} \|P_N e_{k+1}\|$ . So if  $\sigma_{k+1} = \mathcal{O}(\max\{\|e_k\|, \|e_{k-1}\|\})$ , then

$$(40) \quad \|P_R e_{k+1}\| \leq C \max\{\|e_k\|, \|e_{k-1}\|\} \|P_N e_{k+1}\|.$$

This says that if  $x_{k+1}$  is a Newton-like step and  $\max\{\|e_k\|, \|e_{k-1}\|\} \ll 1$ , then  $\|P_R e_{k+1}\| \ll \|P_N e_{k+1}\|$ . Therefore  $x_{k+1}$  is close to  $N$ . That is, Newton-like steps send iterates towards  $N$ , the region in which compatibility is guaranteed by lemma 5.1. Moreover, the smaller  $\max\{\|e_k\|, \|e_{k-1}\|\}$  is, the closer Newton-like steps are to  $N$ . One may think that after a finite number of these Newton-like steps, the iterates will be sufficiently clustered around  $N$  to yield a compatible pair. This is indeed the case, as is shown in the following theorem.

**Theorem 5.2.** *Let Assumption 2.1 hold, and let  $x_i \in B_{\bar{r}}(x^*) \setminus S$  for  $i = k, k - 1, k - 2$ , and  $k - 3$ . Let  $E_k = \operatorname{argmin} R_k^e$  and  $E_{k-1} = \operatorname{argmin} R_{k-1}^e$ . Suppose  $x_j$  and  $x_{j-1}$  are incompatible with  $\|E_j\| > C \max\{|1 - \gamma_j| \|e_{j-1}\|, |\gamma_j| \|e_{j-2}\|\}$ , and  $r_j^e$  is bounded away from one for  $j = k, k - 1$ . Then  $P_N e_{k+1}$  is either  $\mathcal{O}(\theta_{k+1} \|w_{k+1}\|)$ , i.e.,  $(x_k, x_{k-1})$  is compatible, or consists only of higher order terms.*

*Proof.* Apply the triangle inequality to (9) to get

$$(41) \quad \begin{aligned} \|P_N e_{k+1}\| &\leq \frac{1}{2} \|P_N e_k^\alpha\| + |1 - \gamma_{k+1}| C \|P_R e_k\| \\ &\quad + |\gamma_{k+1}| C \|P_R e_{k-1}\| + \|q_{k-1}^k\|. \end{aligned}$$

By lemma 5.4 and (40),  $\|P_R e_i\| \leq C \|e_i\| \max\{\|e_{i-1}\|, \|e_{i-2}\|\}$  for  $i = k, k - 1$ . Grouping the higher order terms into a single term denoted  $Q_{k+1}$ ,<sup>1</sup> (41) becomes

$$(42) \quad \|P_N e_{k+1}\| \leq \frac{1}{2} \|P_N e_k^\alpha\| + Q_{k+1}.$$

There are now two cases. If  $(1/2) \|P_N e_k^\alpha\| > \|Q_{k+1}\|$ , then

$$(43) \quad \begin{aligned} \|P_N e_{k+1}\| &\leq \left(1 + \frac{\|Q_{k+1}\|}{(1/2) \|P_N e_k^\alpha\|}\right) (1/2) \|P_N e_k^\alpha\| \\ &\leq \left(\frac{1 + \frac{\|Q_{k+1}\|}{(1/2) \|P_N e_k^\alpha\|}}{1 - \frac{\|Q_{k+1}\|}{(1/2) \|P_N e_k^\alpha\|}}\right) \theta_{k+1} \|w_{k+1}\|. \end{aligned}$$

The last inequality follows from applying the reverse triangle inequality to the expansion of  $P_N w_{k+1}^\alpha$  from (10) as in the proof of proposition 4.2. Otherwise, if  $(1/2) \|P_N e_k^\alpha\| \leq \|Q_{k+1}\|$ . Then  $\|P_N e_{k+1}\| \leq 2 \|Q_{k+1}\|$ , so that  $P_N e_{k+1}$  is bounded only by higher order terms.  $\square$

From lemma 5.3, we know that incompatible pairs do not necessarily lead to improved error. However, we are assured by theorem 5.2 that no more than two steps with little or no decrease in error may occur, at which point we can expect improved error (locally) either by compatibility, for sufficiently small  $\theta_{k+1}$ , or from higher order terms dominating. Indeed, if we have two consecutive steps where there is little decrease (if any at all) in the null component of the error, then we can bound  $\|E_j\|$  as seen in the statement of theorem 5.2. Then as in the proof we'd find that  $\|P_N e_{k+1}\|$  is compatible, implying acceleration for small  $\theta_{k+1}$ , or that it's bounded only by higher order terms, which also means acceleration near the solution  $x^*$ .

In the next section we introduce the safeguarding strategy called  $\gamma$ -safeguarding, which leads to our main result and proof of convergence, theorem 6.1.

<sup>1</sup>That is,  $Q_{k+1}$  denotes a term such that  $\|Q_{k+1}\| \leq |1 - \gamma_{k+1}| C \|e_k\| \max\{\|e_{k-1}\|, \|e_{k-2}\|\} + |\gamma_{k+1}| C \|e_{k-1}\| \max\{\|e_{k-2}\|, \|e_{k-3}\|\} + q_{k-1}^k$ .

## 6. Convergence

Here we restrict our attention to the case when  $\dim N = 1$ . Recall from section 2 that if  $\hat{D}_N(x)$  is invertible as a map on  $N$  whenever  $P_N(x - x^*) \neq 0$ , then  $f'(x)$  is invertible for all  $x \in \hat{W}$ . We'll show that with an appropriate safeguarding scheme, which we call  $\gamma$ -safeguarding, Newton-Anderson iterates remain in  $\hat{W}$  if  $x_0 \in \hat{W}$ , and converge locally under the same conditions that imply local convergence of the standard Newton method. This safeguarding scheme provides an automated way to decide how to scale  $\gamma_{k+1}$  so that  $\nu_{k+1}$  in equation 44 below remains bounded away from one. This will prove useful in the proof of theorem 6.1. Another interpretation is that by scaling  $\gamma_{k+1}$  towards zero when appropriate, and therefore taking a more "Newton-like" step, we prevent the null-space component from accelerating too much and possibly leaving the region of invertibility. We now prove lemma 6.1, which provides theoretical justification for  $\gamma$ -safeguarding.

**Lemma 6.1.** *For  $\lambda \in (0, 1]$ , let*

$$(44) \quad \nu_{k+1} = \frac{\min\{|1 - \lambda\gamma_{k+1}| \|P_N(e_k + w_{k+1})\|, |\lambda\gamma_{k+1}| \|P_N(e_{k-1} + w_k)\|\}}{\max\{|1 - \lambda\gamma_{k+1}| \|P_N(e_k + w_{k+1})\|, |\lambda\gamma_{k+1}| \|P_N(e_{k-1} + w_k)\|\}}.$$

*Fix  $0 < r < 1$ . Assume  $|\gamma_{k+1}| < 1$  and nonzero. Then given  $\sigma_i < \hat{\sigma}$  and  $\rho_i < \hat{\rho}$ ,  $i = k, k-1$ , there exists a number  $\lambda \in (0, 1]$  such that  $\nu_{k+1} \leq r < 1$  and  $\min\{|1 - \lambda\gamma_{k+1}| \|P_N(e_k + w_{k+1})\|, |\lambda\gamma_{k+1}| \|P_N(e_{k-1} + w_k)\|\} = |\lambda\gamma_{k+1}| \|P_N(e_{k-1} + w_k)\|$ .*

*Proof.* The key relation is

$$(45) \quad \frac{|\gamma_{k+1}| \|P_N(e_{k-1} + w_k)\|}{|1 - \gamma_{k+1}| \|P_N(e_k + w_{k+1})\|} \leq \frac{\left( \frac{1/2 + c_3\sigma_k + c_2(1 + \sigma_k)\|e_k\|}{1/2 - c_1\sigma_k - c_2(1 + \sigma_k)\|e_k\|} \right)}{\left( \frac{1/2 - c_3\sigma_{k-1} - c_2(1 + \sigma_{k-1})\|e_{k-1}\|}{1/2 + c_1\sigma_{k-1} + c_2(1 + \sigma_{k-1})\|e_{k-1}\|} \right)} \frac{|\gamma_{k+1}| \|w_k\|}{|1 - \gamma_{k+1}| \|w_{k+1}\|}.$$

To prove (45), it suffices to show

$$(46) \quad \frac{(1/2 - c_1\sigma_k - c_2(1 + \sigma_k)\|e_k\|)}{(1/2 + c_3\sigma_k + c_2(1 + \sigma_k)\|e_k\|)} \|w_{k+1}\| \leq \|P_N(e_k + w_{k+1})\| \leq \frac{(1/2 + c_1\sigma_k + c_2(1 + \sigma_k)\|e_k\|)}{(1/2 - c_3\sigma_k - c_2(1 + \sigma_k)\|e_k\|)} \|w_{k+1}\|$$

where the constants  $c_1, c_2$ , and  $c_3$  are determined by  $f$ . From (7), it follows that

$$(47) \quad \|P_N(e_k + w_{k+1})\| \leq \left( 1/2 + \|T_k\|\sigma_k + c \frac{\|e_k\|^2}{\|P_N e_k\|} \right) \|P_N e_k\|.$$

We also have  $\|e_k\| \leq (1 + \sigma_k)\|P_N e_k\|$ . Therefore

$$(48) \quad \|P_N(e_k + w_{k+1})\| \leq (1/2 + \|T_k\|\sigma_k + c(1 + \sigma_k)\|e_k\|) \|P_N e_k\|.$$

We'll now use (8) to bound  $\|P_N e_k\|$  in terms of  $\|w_{k+1}\|$ . Indeed, applying the reverse triangle inequality to the right hand side of (8) gives

$$(49) \quad \|P_N e_k\| \leq (1/2 - \|T_k - I\|\sigma_k - c(1 + \sigma_k)\|e_k\|)^{-1} \|w_{k+1}\|.$$

Hence

$$(50) \quad \begin{aligned} \|P_N(e_k + w_{k+1})\| &\leq (1/2 + \|T_k\|\sigma_k + c(1 + \sigma_k)\|e_k\|) \|P_N e_k\| \\ &\leq \frac{(1/2 + \|T_k\|\sigma_k + c(1 + \sigma_k)\|e_k\|)}{(1/2 - \|T_k - I\|\sigma_k - c(1 + \sigma_k)\|e_k\|)} \|w_{k+1}\|. \end{aligned}$$

Since  $f \in C^3$ ,  $\|T_k\| \leq c_1$  independent of  $k$  in  $W(\hat{\rho}, \hat{\sigma}, x^*)$ , and  $\|T_k - I\| \leq c_3$ . This gives the second inequality in (46). To obtain the first inequality, use (7) to obtain a lower bound on  $\|P_N(e_k + w_{k+1})\|$  in terms of  $\|P_N e_k\|$ , then use (8) to bound  $\|P_N e_k\|$  below in terms of  $\|w_{k+1}\|$ . Now, since  $\sigma_i < \hat{\sigma}$  and  $\|e_i\| < \hat{\rho}$ , the first

ratio on the right-hand-side of (45) is bounded by a constant  $c_4 < 1 + \varepsilon$ , with  $\varepsilon > 0$ , for sufficiently small  $\hat{\sigma}$  and  $\hat{\rho}$ . Therefore, to prove the lemma it suffices to show that there exists a number  $\lambda \in (0, 1]$  such that  $|\lambda\gamma_{k+1}|/|1 - \lambda\gamma_{k+1}| \leq r\|w_{k+1}\|/((1 + \varepsilon)\|w_k\|)$ . Let  $\beta_{k+1} = r\|w_{k+1}\|/((1 + \varepsilon)\|w_k\|)$ . Since  $|\gamma_{k+1}| < 1$ ,  $|1 - \lambda\gamma_{k+1}| = 1 - \lambda\gamma_{k+1}$ . There are two cases to consider. If  $\gamma_{k+1} > 0$ , then it suffices to take  $\lambda \leq \beta_{k+1}/(\gamma_{k+1}(1 + \beta_{k+1}))$ . Now suppose  $\gamma_{k+1} < 0$ . This gives  $-\lambda\gamma_{k+1}/(1 - \lambda\gamma_{k+1}) \leq \beta_{k+1}$ , and rearranging yields  $\lambda\gamma_{k+1}(\beta_{k+1} - 1) \leq \beta_{k+1}$ . Note that if  $\beta_{k+1} \geq 1$ , then  $\lambda\gamma_{k+1}(\beta_{k+1} - 1) \leq \beta_{k+1}$  holds for any positive  $\lambda$  since  $\beta_{k+1} \geq 0$ . In particular, we may take  $\lambda = 1$ . Hence no safeguarding is necessary. If  $\beta_{k+1} < 1$ , then it suffices to take  $\lambda \leq \beta_{k+1}/(\gamma_{k+1}(\beta_{k+1} - 1))$ . In any of these cases, we may take  $\lambda \leq 1$ .  $\square$

Based on Lemma 6.1, the safeguarded version of Algorithm 1 is summarized as follows.

---

**Algorithm 2**  $\gamma$ -safeguarding

---

- 1: Given  $x_k, x_{k-1}, w_{k+1}, w_k$ , and  $\gamma_{k+1}$ . Set  $r \in (0, 1)$  and  $\lambda = 1$ .
  - 2:  $\beta \leftarrow r\|w_{k+1}\|/\|w_k\|$
  - 3: **if**  $\gamma_{k+1} = 0$  **or**  $\gamma_{k+1} \geq 1$  **then**
  - 4:      $x_{k+1} \leftarrow x_k + w_{k+1}$
  - 5: **else**
  - 6:     **if**  $|\gamma_{k+1}|/|1 - \gamma_{k+1}| > \beta$  **then**
  - 7:         **if**  $\gamma_{k+1} > 0$  **and**  $\beta/(\gamma_{k+1}(1 + \beta)) < 1$  **then**
  - 8:              $\lambda \leftarrow \beta/(\gamma_{k+1}(1 + \beta))$
  - 9:         **end if**
  - 10:         **if**  $\gamma_{k+1} < 0$  **and**  $0 \leq \beta/(\gamma_{k+1}(\beta - 1)) < 1$  **then**
  - 11:              $\lambda \leftarrow \beta/(\gamma_{k+1}(\beta - 1))$
  - 12:         **end if**
  - 13:     **end if**
  - 14: **end if**
  - 15:  $\gamma_{k+1} \leftarrow \lambda\gamma_{k+1}$
  - 16:  $x_{k+1} \leftarrow x_k + w_{k+1} - \gamma_{k+1}(x_k - x_{k-1} + w_{k+1} - w_k)$
- 

If  $\gamma_{k+1} = 0$  or  $\gamma_{k+1} \geq 1$ , we just take a standard Newton step, i.e.,  $x_{k+1} = x_k + w_{k+1}$ . This condition is similar to that of the safeguarding strategy seen in [30]. There, the authors take a Newton step if the direction cosine between  $w_{k+1}$  and  $w_k$  exceeds 0.942. The condition used in this paper, taking a Newton step when  $\gamma_{k+1} \geq 1$ , is a stronger condition than  $\cos(w_{k+1}, w_k) > 0.942$  in the sense that if  $\cos(w_{k+1}, w_k) \geq 0.942$ , then  $|\gamma_{k+1}| < 2$ . On the other hand,  $\gamma$ -safeguarding is weaker in the sense that, depending on  $\|w_{k+1}\|/\|w_k\|$ , it's possible that  $|\gamma_{k+1}| < 1$  even if  $\cos(w_{k+1}, w_k) > 0.942$ . It is also important to note that bounding  $(|\lambda\gamma_{k+1}|\|w_k\|)/(|1 - \lambda\gamma_{k+1}|\|w_{k+1}\|)$  away from one also bounds  $(|\lambda\gamma_{k+1}|\|P_N e_{k-1}\|)/(|1 - \lambda\gamma_{k+1}|\|P_N e_k\|)$  away from one by (8). Indeed, (8) implies

$$(51) \quad \begin{aligned} \|w_{k+1}\| &\leq (1/2 + c_1\sigma_k + c_2(1 + \sigma_k)\|e_k\|)\|P_N e_k\|, \\ \|w_{k+1}\| &\geq (1/2 - c_1\sigma_k - c_2(1 + \sigma_k)\|e_k\|)\|P_N e_k\|. \end{aligned}$$

This fact will be used in the proof of theorem 6.1. Analogous to (44), we'll let

$$(52) \quad \tilde{\nu}_{k+1} = \frac{\min\{|1 - \lambda\gamma_{k+1}| \|P_N e_k\|, |\lambda\gamma_{k+1}| \|P_N e_{k-1}\|\}}{\max\{|1 - \lambda\gamma_{k+1}| \|P_N e_k\|, |\lambda\gamma_{k+1}| \|P_N e_{k-1}\|\}}.$$

Some notation: let  $\lambda_{k+1}$  be the value computed by algorithm 2 for the chosen parameter  $r$  at step  $k + 1$ ,  $\theta_{k+1}^\lambda = \|w_{k+1} - \lambda_{k+1}\gamma_{k+1}(w_{k+1} - w_k)\|/\|w_{k+1}\|$ , and  $L_{k+1}^\lambda(x, y) = (1 - \lambda_{k+1}\gamma_{k+1})x + \lambda_{k+1}\gamma_{k+1}y$ . We now state the main convergence result. The proof may be found in appendix B.

**Theorem 6.1.** *Let  $\dim N = 1$ , and let  $\hat{D}_N(x)$  be invertible as a map on  $N$  for all  $P_N(x - x^*) \neq 0$ . Let  $W_k = W(\|e_k\|, \sigma_k, x^*)$ . If  $x_0$  is chosen so that  $\sigma_0 < \hat{\sigma}$  and  $\|e_0\| < \hat{\rho}$ , for sufficiently small  $\hat{\sigma}$  and  $\hat{\rho}$ ,  $x_1 = x_0 + w_1$ , and  $x_{k+1} = L_{k+1}^\lambda(x_k + w_{k+1}, x_{k-1} + w_k)$  for  $k \geq 1$ , then  $W_{k+1} \subset W_0$  for all  $k \geq 0$  and  $x_k \rightarrow x^*$ . That is,  $\{x_k\}$  remains well-defined and converges to  $x^*$ . Furthermore, there exist constants  $c_4 > 0$  and  $\kappa \in (1/2, 1)$  such that*

$$(53) \quad \|P_R e_{k+1}\| \leq c_4 \max\{|1 - \lambda_{k+1}\gamma_{k+1}| \|e_k\|^2, |\lambda_{k+1}\gamma_{k+1}| \|e_{k-1}\|^2\}$$

$$(54) \quad \|P_N e_{k+1}\| < \kappa \theta_{k+1}^\lambda \|P_N e_k\|$$

for all  $k \geq 1$ .

### 7. High Order Roots

In this section, we generalize the results from the previous sections to higher order roots, which we define now. Let  $p$  denote the smallest integer such that  $f^{(p+1)}(x^*)(x - x^*)^p \neq 0$ . Here  $(x - x^*)^p$  means that the first  $p$  arguments of the  $(p + 1)$ -linear map  $f^{(p+1)}(x^*)$  are  $x - x^*$ . The *order* of the root  $x^*$  is the smallest integer  $d$  such that  $P_N f^{(d+1)}(x^*)(x - x^*)^d \neq 0$ . We define the order of  $x^*$  this way since  $d$  determines the rate of convergence of  $P_N x_k$  when  $\{x_k\}$  is generated by Newton's method. With this definition, we can say that the analysis thus far has focused on first order roots. Now we're interested in roots of order greater than one, that is, roots  $x^*$  such that  $d > 1$ . The discussion in section 2 generalizes in a reasonably straightforward way to higher order roots. See [8] for details on invertibility of  $f'(x)$  in this context. We'll continue to assume  $\dim N = 1$ , and we also assume the following throughout this section. Recall that  $S$  is the set in which  $f'(x)$  is singular.

**Assumption 7.1.** *The operator  $\hat{D}^d(x) = (d!)^{-1} P_N f^{(d+1)}(x^*)(x - x^*)^d P_N$  is invertible as a map on  $N$  in the ball  $B_R(x^*) \setminus S$ , and  $R$  is sufficiently small to ensure  $f'(x)^{-1}$  exists and  $f'(x)^{-1} = \mathcal{O}(\|x - x^*\|^{-d})$ .*

With this assumption, we have lemma 7.1. The expansions in (55) and (56) are similar to those found in [8]. Expansion (57) follows from adding (55) to (56) and then isolating  $w_{k+1}$ . We've let  $\hat{T}_k(\cdot)$  denote a linear map on  $\mathbb{R}^n$  whose bounded independent of  $x_k$ .

**Lemma 7.1.** *Let assumption (7.1) hold and let  $x_k \in B_{\hat{r}}(x^*) \setminus S$ . Then*

$$(55) \quad P_N(e_k + w_{k+1}) = \frac{d}{d+1} P_N e_k + \hat{T}_k P_R e_k + \mathcal{O}(\|e_k\|^2)$$

$$(56) \quad P_R(e_k + w_{k+1}) = \mathcal{O}(\|e_k\|^{d+1})$$

$$(57) \quad w_{k+1} = -\frac{1}{d+1} P_N e_k + (\hat{T}_k - I) P_R e_k + \mathcal{O}(\|e_k\|^2).$$



Applying lemma (7.1) to a Newton-Anderson step yields

$$(58) \quad P_N e_{k+1} = \left(\frac{d}{d+1}\right) P_N e_k^\alpha + (\hat{T}_k P_R e_k)^\alpha + q_{k-1}^k.$$

**Lemma 7.2.** *Under the assumptions of lemma (7.1), strong N-pairs are automatically compatible. Further, writing  $x_{k+1} = (1 - \gamma_{k+1})(x_k + w_{k+1}) + \gamma_{k+1}(x_{k-1} + w_k)$  and  $e_k = x_k - x^*$ , the following bound holds*

$$(59) \quad \|P_N e_{k+1}\| \leq \theta_{k+1} \left(\frac{d}{d+1}\right) \left(\frac{1+r_{k+1}^e}{1-r_{k+1}^e}\right) \|P_N e_k\|.$$

*Proof.* Assume  $r_{k+1}^e \ll 1$ . Then if  $(x_k, x_{k-1})$  is a strong N-pair, we have

$$(60) \quad \|P_N e_{k+1}\| \leq (1+r_{k+1}^e) \left(\frac{d}{d+1}\right) \|P_N e_k^\alpha\|.$$

By lemma (7.1), it follows that

$$(61) \quad \|P_N e_{k+1}\| \leq d \left(\frac{1+r_{k+1}^e}{1-r_{k+1}^e}\right) \|w_{k+1}^\alpha\| = d\theta_{k+1} \left(\frac{1+r_{k+1}^e}{1-r_{k+1}^e}\right) \|w_{k+1}\|.$$

Applying (57) completes the proof. □

While  $d$  may grow very large for very high-order roots, we still feel it’s reasonable to say that strong N-pairs are automatically compatible because the final bound on  $P_N e_{k+1}$  seen in (59) contains  $d(d+1)^{-1}$  rather than  $d$  alone. This ensures that  $\|P_N e_{k+1}\| = \mathcal{O}(\theta_{k+1} \|P_N e_k\|)$  where the constant is of moderate size. Thus  $\theta_{k+1}$  can lead to significant acceleration just as in the order one case. It’s worth noting that this acceleration can occur without any a priori knowledge of the order of the root. Theorem 7.1 is the generalization of 6.1 for higher order roots. The proofs are nearly identical due to the similar structure of 55 and 7, and therefore we simply state 7.1. Analogously to the first order root case, we let  $\hat{D}_N^d(x) = (d!)^{-1} P_N f^{(d+1)}(x^*) P_N (x - x^*)^d P_N$ .

**Theorem 7.1.** *Let  $\dim N = 1$ , and let  $\hat{D}_N^{d+1}(x)$  be invertible as a map on  $N$  for all  $P_N(x - x^*) \neq 0$ . Let  $W_k = W(\|e_k\|, \sigma_k, x^*)$ . If  $x_0$  is chosen so that  $\sigma_0 < \hat{\sigma}$  and  $\|e_0\| < \hat{\rho}$ , for sufficiently small  $\hat{\sigma}$  and  $\hat{\rho}$ ,  $x_1 = x_0 + w_1$ , and  $x_{k+1} = L_{k+1}^\lambda(x_k + w_{k+1}, x_{k-1} + w_k)$  for  $k \geq 1$ , then  $W_{k+1} \subset W_0$  for all  $k \geq 0$  and  $x_k \rightarrow x^*$ . That is,  $\{x_k\}$  remains well-defined and converges to  $x^*$ . Furthermore, there exist constants  $c > 0$  and  $\kappa \in (d/(d+1), 1)$  such that*

$$(62) \quad \|P_R e_{k+1}\| \leq c \max\{1 - \lambda_{k+1} \gamma_{k+1} \|e_k\|^{d+1}, |\lambda_{k+1} \gamma_{k+1}| \|e_{k-1}\|^{d+1}\}$$

$$(63) \quad \|P_N e_{k+1}\| \leq \kappa \theta_{k+1}^\lambda \|P_N e_k\|.$$

**8. Numerical Results**

In this section we compare the performance of a few variations of Newton-Anderson (N.Anderson), including standard N.Anderson, and the projected Levenberg-Marquardt method from [20]. The variations on N.Anderson are  $\gamma$ -N.Anderson, that is, N.Anderson with  $\gamma$ -safeguarding from algorithm 2, N.Anderson with an Armijo linesearch (see, for instance, [3]), and  $\gamma$ -Armijo-N.Anderson, that is, we use algorithm 2 in tandem with a linesearch. The linesearch is only carried out if a given step fails to reduce the residual by a factor of 0.99. It should be noted that in practice,  $\gamma$ -safeguarding does not appear necessary for convergence, as is demonstrated below. Though it can recover and even improve convergence for

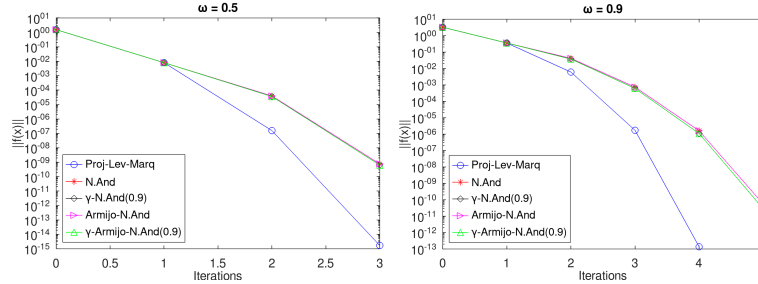


FIGURE 2. Left: Residual history for the Chandrasekhar H-equation with parameter  $\omega = 0.5$ . Right: Residual history for Chandrasekhar H-equation with  $\omega = 0.9$ .

some problems. There is only one problem for which N.Anderson did not converge. Both Armijo-N.Anderson and  $\gamma$ -N.Anderson converged well for that problem.

Our first two test problem feature nonlinear systems of order  $n = 10^4$ : the Chandrasekhar H-equation from [6], and a modified version of the multivariate polynomial seen in example three on page 45 of [23]. The Chandrasekhar H-equation is a familiar benchmark problem from the literature concerning Newton's method, in particular Newton's method in the presence of singularities. The multivariate polynomial is a scaleable singular problem where we can easily adjust the order of the root in order to demonstrate the theory for higher order roots. Here we mean the order of the root as defined in 7. We then apply the methods to various smaller-scale benchmark problems, both singular and nonsingular, from the literature. All test problems are square systems.

All computations were performed in Octave on an M1 Macbook Pro. The iterations were terminated when  $\|f(x_k)\| < 10^{-8}$  or the number of iterations exceeded fifty, in which case we say the algorithm failed to converge. The parameters for the projected Levenberg-Marquardt method are mostly those of [20]. Our implementation differs from theirs in that we terminate when  $\|f(x_k)\| < 10^{-8}$ , and our Armijo linesearch parameters may differ. In our search, the step size is  $1/2$  and the damping parameter is  $10^{-4}$ . Thus we seek the smallest  $j$  such that  $g(x + (1/2)^j d) \leq g(x) + 10^{-4}(1/2)^j g'(x)d$ , where  $g(x) = \|f(x)\|^2$  and  $d$  is the search direction. For the N.Anderson method and variants,  $d_k = w_{k+1} - \gamma_{k+1}(x_k - x_{k-1} + w_{k+1} - w_k)$ , and we scale the step size  $1/2$  by  $3/10$  at each iteration for each problem except for Dayton10. That is, we seek the smallest  $j$  such that  $g(x + (1/2)(3/10)^j d) \leq g(x) + 10^{-4}(1/2)(3/10)^j g'(x)d$ . For Dayton10, we took the step size to be  $4/5$  and scaled by  $3/10$  at each iteration. When we want to emphasize the  $\gamma$ -safeguarding parameter  $r$ , we write  $\gamma$ -N.Anderson( $r$ ).

## 8.1. Singular Problems.

**8.1.1. The Chandrasekhar H-Equation.** The Chandrasekhar H-equation is an important benchmark problem for Newton and Newton-like methods (see [10, 21, 22, 30]) defined by the integral equation

$$(64) \quad F(H)(\mu) := H(\mu) - \left(1 - \frac{\omega}{2} \int_0^1 \frac{\mu H(\nu) d\nu}{\mu + \nu}\right)^{-1} = 0.$$

As discussed in [21], eq. (64) admits real solutions for parameter  $\omega \in [0, 1]$ . The derivative at each solution is invertible for  $\omega \in [0, 1)$ , and features a one dimensional

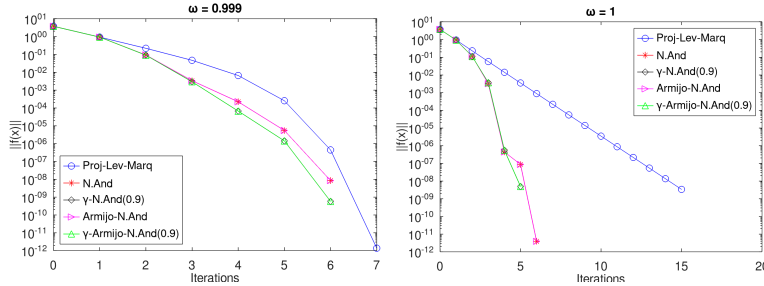


FIGURE 3. Left: Residual history for the Chandrasekhar H-equation with parameter  $\omega = 0.999$ . Right: Residual history for Chandrasekhar H-equation with  $\omega = 1$ .

nullspace at the bifurcation point,  $\omega = 1$ . As shown in [10], (64) with  $\omega = 1$  satisfies the assumptions of theorem 6.1. Following [21], we discretize (64) by the composite midpoint rule with  $n = 10^4$  nodes, yielding a discrete system in  $\mathbb{R}^n$ . We take  $x^0$  to be the vector of ones. The results are shown in Figures 2 and 3.

No line search steps were taken for the N.Anderson variations. Consequently, the residual history for N.Anderson and Armijo-N.Anderson are identical, as are  $\gamma$ -N.Anderson and  $\gamma$ -Armijo-N.Anderson. We observe that Projected-Levenberg-Marquardt converges moderately faster than the N.Anderson variations up to  $\omega = 0.999$ , however the iteration count to converge to the given tolerance varies by only one. For  $\omega = 1$  the problem is singular, and the N.Anderson variations display substantially better performance. For each case, the variations of N.Anderson all perform similarly to each other with  $\gamma$ -N.Anderson slightly outperforming N.Anderson for  $\omega = 0.999$ . For comparison, we note that from the same initial iterate, the standard Newton method converged in 4 iterations for  $\omega = 0.5$ , 5 for  $\omega = 0.9$ , 8 for  $\omega = 0.999$ , and 17 for  $\omega = 1$ .

**8.1.2. A Multivariate Polynomial.** Here we apply the methods mentioned at the beginning of this section to

$$(65) \quad f(x_1, \dots, x_n) = (x_1^2 + x_1 - x_2^k, x_2^2 + x_2 - x_3^k, \dots, x_{n-1}^2 + x_{n-1} - x_n^k, x_n^k), \quad n = 10^4.$$

The zero vector is a root of order  $k-1$ . That is,  $P_N f^{(j)}(x^*) = 0$  for  $j = 0, 1, \dots, k-1$ , and is nonzero for  $j = k$ . Here  $x^*$  is the zero vector. We choose  $x^0$  such that  $x_n^0 = 0.9$ , and  $x_j^0 = 0.3$  for all  $j = 1, 2, \dots, n-1$ .

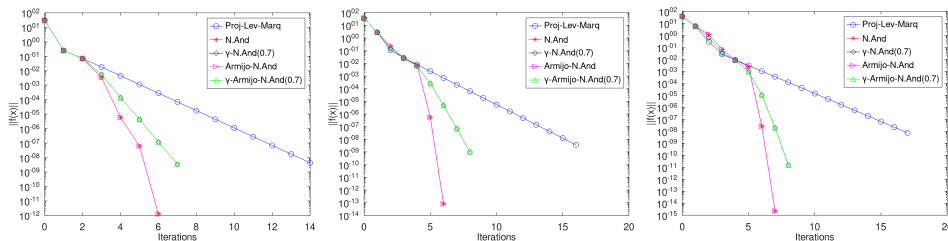


FIGURE 4. Left: Residual history for  $k = 2$ , in which case the zero vector is a root of order one. Middle: Residual history for  $k = 3$ . Right: Residual history for  $k = 7$ .

The results in figure 4 demonstrate the efficacy of N.Anderson and its variations in the presence of higher order roots. As indicated in the legends, the result of running  $\gamma$ -N.Anderson with  $\gamma$ -safeguarding parameter  $r = 0.7$  are shown. We note that upon setting  $r = 0.9$ , the convergence was very similar to that of standard N.Anderson. No line search steps were taken for the N.Anderson variations. Consequently, the residual history for N.Anderson and Armijo-N.Anderson are identical, as are  $\gamma$ -N.Anderson and  $\gamma$ -Armijo-N.Anderson. The Projected-Levenberg-Marquardt method is seen to converge linearly, and converge slightly slower as the order of the root increases. The convergence of the variations of N.Anderson, however, is less affected by the order of the root. With  $\gamma$ -safeguarding parameter  $r = 0.7$ ,  $\gamma$ -N.Anderson converges slower than standard N.Anderson, but it still exhibits fast convergence compared to the Projected-Levenberg-Marquardt method as the order of the root increases. From the same initial iterate, the standard Newton method converged in 15 iterations when  $k = 2$ , 17 when  $k = 3$ , and 18 when  $k = 7$ .

**8.2. Small-Scale Benchmark Problems.** Numerical results for the small-scale problems can be found in appendix A. The test problems in table 1 correspond to test problems 14.1.1 – 14.2.6 in [15]. Following [20], we take  $x_0$  to be the vector of lower bounds found in [15] for these problems. The test problems in table 2 are taken from various sources in the literature, which we include by the name of the problem in the first column. The largest of these problems has dimension  $n = 8$ . The initial iterate is taken from the source for each problem except Dayton10, for which we take  $x_0 = (0.9, 0.9, 0.9, 0.9)^T$ . The fifth column  $\|f(x)\|$  is the value of the residual upon termination. The final column, following [20], records the number of standard projected Levenberg-Marquardt steps taken, the number of Armijo linesearch steps, and the number of projected gradient step (see Algorithm 3.12 and the first paragraph of section 4 in [20]). We also use this column to count the number of Armijo linesearch steps taken in Armijo-N.Anderson and  $\gamma$ -Armijo-N.Anderson. The number of function evaluations, denoted by  $f$ -evals, is not recorded for N.Anderson or  $\gamma$ -N.Anderson since for these algorithms the number of function evaluations is always the number of iterations plus one.

For the nonsingular problems, as shown in table 1,  $\gamma$ -N.Anderson performs comparably to the Levenberg-Marquardt method in all but two problems: Eq-Combustion and Robot Kin. Sys. All methods perform similarly with the Ferraris-Tronconi function. For the Bullard-Biegler function, for our chosen initial iterate, N.Anderson fails to converge. We recover convergence by applying  $\gamma$ -safeguarding, a linesearch, or both, and  $\gamma$ -N.Anderson outperforms Armijo-N.Anderson. In fact,  $\gamma$ -N.Anderson is seen to outperform or match Armijo-N.Anderson for each of the nonsingular test problems.

For the singular problems, as shown in table 2, N.Anderson and its variations all outperform Projected-Levenberg-Marquardt, except for Dayton10 where  $\gamma$ -Armijo-N.Anderson failed to converge. For most of the problems, applying  $\gamma$ -safeguarding, a linesearch, or both yields similar results. Although the linesearch can yield better results than  $\gamma$ -safeguarding, the latter requires fewer function evaluations than a linesearch; and in some cases, far fewer.

## 9. Conclusion

We have presented an analysis of and convergence theory for the N.Anderson method in a general Euclidean space when the derivative has a nontrivial null space at the root. Efficient and robust methods for this problem class are of continuing

importance as they naturally arise at bifurcation points in parameter-dependent mathematical models. As motivated by previous numerical benchmarking results where greater algorithmic depths were explored but generally not found beneficial, we restricted our attention to an algorithmic depth of one for Anderson accelerated Newton iterations. Greater algorithmic depths and methods for problems with nullspaces of higher dimension may be explored in future work if they are found to be advantageous in relevant applications.

We showed that near the null space, the component of the error along the null space is accelerated by a factor determined by the success of the minimization step in the N.Anderson algorithm, which is measured by the optimization gain. In other words, it is shown that the region in which standard Newton exhibits linear convergence is precisely the region most susceptible to acceleration, demonstrated by a substantial decrease in the convergence rate when the optimization gain is small. We developed a novel and theoretically supported safeguarding scheme that when applied to N.Anderson ensures local convergence under the same conditions as standard Newton; moreover, the rate of convergence is improved in general, depending on the optimization gain. The theory was demonstrated with benchmark problems from the literature.

### Acknowledgements

MD and SP are supported in part by the National Science Foundation NSF-DMS 2011519.

### References

- [1] Hengbin An, Xiaowei Jia, and Homer F. Walker. Anderson acceleration and application to the three-temperature energy equations. *Journal of Computational Physics*, 347:1–19, 2017.
- [2] D. G. Anderson. Iterative procedures for nonlinear integral equations. *J. Assoc. Comput. Mach.*, 12(4):547–560, 1965.
- [3] Larry Armijo. Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, 16(1), 1966.
- [4] R. Behling and A. Fischer. A unified local convergence analysis of inexact constrained levenberg-marquardt methods. *Optimization Letters*, 6:927–940, 2012.
- [5] S. Bellavia and B. Morini. Strong local convergence properties of adaptive regularized methods for nonlinear least squares. *IMA Journal of Numerical Analysis*, 35, 2015.
- [6] S. Chandrasekhar. *Radiative transfer*. Engineering special collection. Dover Publications, 1960.
- [7] Barry H. Dayton and Zhonggang Zeng. Computing the multiplicity structure in solving polynomial systems. In *Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation*, ISSAC '05, page 116123, New York, NY, USA, 2005. Association for Computing Machinery.
- [8] D. W. Decker, H. B. Keller, and C. T. Kelley. Convergence rates for Newtons method at singular points. *SIAM J. Numer. Anal.*, 20(2):296–314, 1983.
- [9] D. W. Decker and C. T. Kelley. Newtons method at singular points I. *SIAM J. Numer. Anal.*, 17(1):66–70, 1980.
- [10] D. W. Decker and C. T. Kelley. Convergence acceleration for Newtons method at singular points. *SIAM J. Numer. Anal.*, 19(1):219–229, 1982.
- [11] D.W. Decker and C.T. Kelley. Newton’s method at singular points II. *SIAM J. Numer. Anal.*, 17(3):465–471, 1980.
- [12] C. Evans, S. Pollock, L. 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(1):788–810, 2020.
- [13] V. Eyert. A comparative study on methods for convergence acceleration of iterative vector sequences. *Journal of Computational Physics*, 124(2):271–285, 1996.
- [14] H. Fang and Y. Saad. Two classes of multiseccant methods for nonlinear acceleration. *Numer. Linear Algebra Appl.*, 16(3):197–221, 2009.

- [15] C.A. Floudas, P.M. Pardalos, C.S. Adjiman, W.R. Esposito, Z.H. Gümmüs, S.T. Harding, J.L. Klepeis, C.A. Meyer, and C.A. Schweiger. *Handbook of Test Problems in Local and Global Optimization*. Kluwer Publishers, 1999.
- [16] A. Griewank and M. R. Osborne. Analysis of newton’s method at irregular singularities. *SIAM J. Numer. Anal.*, 20(4):747–773, 1983.
- [17] Andreas Griewank and M. R. Osborne. Newtons method for singular problems when the dimension of the null space is  $> 1$ . *SIAM J. Numer. Anal.*, 18(1):145–149, 1981.
- [18] A.O. Griewank. Starlike domains of convergence for Newton’s method at singularities. *Numer. Math.*, 35:95–111, 1980.
- [19] Jos L. Hueso, Eulalia Martinez, and Juan R. Torregrosa. Modified Newtons method for systems of nonlinear equations with singular Jacobian. *J. Comput. Appl. Math.*, 224(1):77–83, 2008.
- [20] Christian Kanzow, Nobuo Yamashita, and Masao Fukushima. Levenbergmarquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *J. Comput. Appl. Math.*, 172:375–397, 2004.
- [21] C.T. Kelley. Numerical methods for nonlinear equations. *Acta Numer.*, 27(1):207–287, 2018.
- [22] C.T. Kelley and R Suresh. A new acceleration method for Newton’s method at singular points. *SIAM J. Numer. Anal.*, 20(5):1001–1009, 1983.
- [23] Nan Li and Lihong Zhi. Improved two-step newton’s method for computing simple multiple zeros of polynomial systems. *Numerical Algorithms*, 91:19–50, 2022.
- [24] P. A. Lott, H. F. Walker, C. S. Woodward, and U. M. Yang. An accelerated Picard method for nonlinear systems related to variably saturated flow. *Adv. Water Resour.*, 38:92–101, 2012.
- [25] Takeo Ojika. A numerical method for branch points of a system of nonlinear algebraic equations. *Applied Numerical Mathematics*, 4:419–430, 1988.
- [26] Takeo Ojika, Satoshi Watanabe, and Taketomo Mitsui. Deflation algorithm for the multiple roots of a system of nonlinear equations. *Journal of Mathematical Analysis and Applications*, 96(2):463–479, 1983.
- [27] Y. Peng, B. Deng, J. Zhang, F. Geng, W. Qin, and L. Liu. Anderson acceleration for geometry optimization and physics simulation. *ACM Trans. Graph.*, 37(4), 2018.
- [28] S. Pollock. Anderson acceleration for degenerate and nondegenerate problems. In Baasansuren Jadamba, Akhtar A. Khan, Stanisaw Migrski, and Miguel Sama, editors, *Deterministic and Stochastic Optimal Control and Inverse Problems*. CRC Press, 2021.
- [29] S. Pollock, L. Rebholz, and M. Xiao. Anderson-accelerated convergence of Picard iterations for incompressible Navier-Stokes equations. *SIAM J. Numer. Anal.*, 57(2):615–637, 2019.
- [30] S. Pollock and H. Schwartz. Benchmarking results for the NewtonAnderson method. *Results Appl. Math.*, 8:100095, 2020.
- [31] Sara Pollock and Leo G Rebholz. Anderson acceleration for contractive and noncontractive operators. *IMA J. Numer. Anal.*, 41(4):2841–2872, 2021.
- [32] G.W. Reddien. On Newton’s method for singular problems. *SIAM J. Numer. Anal.*, 15(5):993–996, 1978.
- [33] G.W. Reddien. Newton’s method and high order singularities. *Comput. Math. Appl.*, 5(2):79–86, 1979.
- [34] Robert B. Schnabel and Paul D. Frank. Tensor methods for nonlinear equations. *SIAM J. Numer. Anal.*, 21(5):815–843, 1984.
- [35] R. Seydel. Numerical computation of branch points in ordinary differential equations. *Numer. Math.*, 32:51–68, 1979.
- [36] R. B. Thompson, K. O/. Rasmussen, and T. Lookman. Improved convergence in block copolymer self-consistent field theory by Anderson mixing. *The Journal of Chemical Physics*, 120(1):31–34, 2004.
- [37] A. Toth and C. T. Kelley. Convergence analysis for Anderson acceleration. *SIAM J. Numer. Anal.*, 53(2):805–819, 2015.
- [38] H. F. Walker and P. Ni. Anderson acceleration for fixed-point iterations. *SIAM J. Numer. Anal.*, 49(4):1715–1735, 2011.
- [39] Dawei Wang, Yunhui He, and Hans De Sterck. On the asymptotic linear convergence speed of Anderson acceleration applied to ADMM. *J. Sci. Comput.*, 88(2):38, 2021.
- [40] S. Wright. Stability of linear equations solvers in interior-point methods. *SIAM J. Matrix Anal. Appl.*, 16:1287–1307, 1995.

Appendix A. Numerical Results for Small-Scale Problems

TABLE 1. Nonsingular Benchmark Problems.

Problem	Algorithm	Iterations	F-vals	$\ f(x)\ $	LM/LS/PG
Himmelbau	Proj-Lev-Marq	6	7	2.842e-14	6/0/0
	N.Anderson	8	-	7.105e-15	-
	$\gamma$ -N.Anderson(0.5)	6	-	5.309e-12	-
	Armijo-N.Anderson	8	9	7.105e-15	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.5)	7	8	5.309e-12	-/0/-
Eq-Combustion	Proj-Lev-Marq	11	28	8.200e-14	8/3/0
	N.Anderson	35	-	8.510e-12	-
	$\gamma$ -N.Anderson(0.5)	17	-	3.092e-09	-
	Armijo-N.Anderson	18	56	2.136e-10	-/2/-
	$\gamma$ -Armijo-N.Anderson(0.5)	17	18	3.092e-09	-/0/-
Bullard-Biegler	Proj-Lev-Marq	13	26	2.270e-10	10/3/0
	N.Anderson	F	-	-	-
	$\gamma$ -N.Anderson(0.5)	11	-	1.799e-12	-
	Armijo-N.Anderson	20	202	1.212e-10	-/12/-
	$\gamma$ -Armijo-N.Anderson(0.5)	13	34	1.629e-11	-/6/-
Ferraris-Tronconi	Proj-Lev-Marq	4	5	5.339e-14	4/0/0
	N.Anderson	4	-	6.937e-11	-
	$\gamma$ -N.Anderson(0.5)	4	-	6.008e-11	-
	Armijo-N.Anderson	4	5	6.937e-11	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.5)	4	5	6.008e-11	-/0/-
Brown's Al. Lin.	Proj-Lev-Marq	9	10	1.638e-14	9/0/0
	N.Anderson	19	-	5.137e-10	-
	$\gamma$ -N.Anderson(0.5)	11	-	1.441e-11	-
	Armijo-N.Anderson	11	19	1.286e-10	-/2/-
	$\gamma$ -Armijo-N.Anderson(0.5)	11	12	1.441e-11	-/0/-
Robot Kin. Sys.	Proj-Lev-Marq	5	6	6.404e-10	5/0/0
	N.Anderson	9	-	7.390e-14	-
	$\gamma$ -N.Anderson(0.5)	8	-	4.290e-14	-
	Armijo-N.Anderson	9	10	7.390e-14	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.5)	8	9	4.290e-14	-/0/-

TABLE 2. Singular Benchmark Problems.

Problem	Algorithm	Iterations	F-vals	$\ f(x)\ $	LM/LS/PG
Decker1[8]	Proj-Lev-Marq	15	16	3.559e-09	15/0/0
	N.Anderson	9	-	2.698e-12	-
	$\gamma$ -N.Anderson(0.9)	8	-	4.187e-09	-
	Armijo-N.Anderson	9	10	2.698e-12	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.9)	8	9	4.187e-09	-/0/-
Decker2[11]	Proj-Lev-Marq	16	17	3.356e-09	16/0/0
	N.Anderson	7	-	3.118e-09	-
	$\gamma$ -N.Anderson(0.9)	7	-	8.659e-09	-
	Armijo-N.Anderson	7	8	3.118e-09	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.9)	7	8	8.659e-09	-/0/-
Ojika1[25]	Proj-Lev-Marq	28	59	2.419e-09	18/10/0
	N.Anderson	19	-	1.620e-09	-
	$\gamma$ -N.Anderson(0.9)	17	-	5.990e-09	-
	Armijo-N.Anderson	17	69	4.214e-09	-/5/-
	$\gamma$ -Armijo-N.Anderson(0.9)	17	64	6.162e-09	-/4/-
Ojika2[26]	Proj-Lev-Marq	13	14	2.909e-09	13/0/0
	N.Anderson	7	-	3.096e-09	-
	$\gamma$ -N.Anderson(0.9)	7	-	7.165e-09	-
	Armijo-N.Anderson	7	8	3.096e-09	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.9)	7	8	7.165e-09	-/0/-
Pollock1[30]	Proj-Lev-Marq	14	15	3.991e-09	14/0/0
	N.Anderson	5	-	1.656e-10	-
	$\gamma$ -N.Anderson(0.9)	5	-	8.268e-10	-
	Armijo-N.Anderson	5	6	1.656e-10	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.9)	5	6	8.268e-10	-/0/-
Dayton10[7]	Proj-Lev-Marq	F	-	-	-
	N.Anderson	11	-	3.294e-10	-
	$\gamma$ -N.Anderson(0.5)	14	-	5.157e-09	-
	Armijo-N.Anderson	15	140	2.153e-11	-/4/-
	$\gamma$ -Armijo-N.Anderson(0.5)	F	-	-	-/-
Hueso1[19]	Proj-Lev-Marq	13	14	3.701e-09	13/0/0
	N.Anderson	12	-	5.417e-09	-
	$\gamma$ -N.Anderson(0.9)	12	-	4.918e-09	-
	Armijo-N.Anderson	9	39	7.184e-09	-/1/-
	$\gamma$ -Armijo-N.Anderson(0.9)	10	40	6.264e-09	-/1/-
Hueso6[19]	Proj-Lev-Marq	16	17	3.531e-09	16/0/0
	N.Anderson	6	-	3.459e-10	-
	$\gamma$ -N.Anderson(0.9)	6	-	6.590e-09	-
	Armijo-N.Anderson	6	7	3.459e-10	-/0/-
	$\gamma$ -Armijo-N.Anderson(0.9)	6	7	6.590e-09	-/0/-

## Appendix B. Proof of Convergence Theorem

We'll now prove theorem 6.1. We restate it here for convenience.

**Theorem B.1.** *Let  $\dim N = 1$ , and let  $\hat{D}_N(x)$  be invertible as a map on  $N$  for all  $P_N(x - x^*) \neq 0$ . Let  $W_k = W(\|e_k\|, \sigma_k, x^*)$ . If  $x_0$  is chosen so that  $\sigma_0 < \hat{\sigma}$  and  $\|e_0\| < \hat{\rho}$ , for sufficiently small  $\hat{\sigma}$  and  $\hat{\rho}$ ,  $x_1 = x_0 + w_1$ , and  $x_{k+1} = L_{k+1}^\lambda(x_k + w_{k+1}, x_{k-1} + w_k)$  for  $k \geq 1$ , then  $W_{k+1} \subset W_0$  for all  $k \geq 0$  and  $x_k \rightarrow x^*$ . That is,  $\{x_k\}$  remains well-defined and converges to  $x^*$ . Furthermore, there exist constants  $c_4 > 0$  and  $\kappa \in (1/2, 1)$  such that*

$$(B.1) \quad \|P_R e_{k+1}\| \leq c_4 \max\{1 - \lambda_{k+1} \gamma_{k+1} \|e_k\|^2, |\lambda_{k+1} \gamma_{k+1}| \|e_{k-1}\|^2\}$$

$$(B.2) \quad \|P_N e_{k+1}\| < \kappa \theta_{k+1}^\lambda \|P_N e_k\|$$

for all  $k \geq 1$ .

*Proof.* Throughout this proof, we assume  $\gamma$ -safeguarding is in use, and will simply write “ $\gamma_{k+1}$ ” in place of “ $\lambda_{k+1} \gamma_{k+1}$ ”. Further, it's assumed that  $\gamma_k < 1$  for all  $k$ . If  $\gamma_k = 1$ , we take a standard Newton step. We'll let  $c$  denote intermediate constants that will eventually be absorbed into another. The bound in (B.1) is just a restatement of (14) with  $\gamma$ -safeguarding. It will follow for all  $k$  once we prove that  $\{x_k\}$  is well-defined. The remainder of the proof proceeds as follows. We'll first show inductively that  $W_k \subset W_0$  for all  $k$ , thus ensuring  $\{x_k\}$  is well-defined. Along the way we'll establish (B.2), and define a sequence  $\{\eta_k\}$  such that  $\|e_{k+1}\| < \eta_k \|e_0\|$ . Then we'll show  $\eta_k \rightarrow 0$ , proving convergence. We begin with  $x_2$ .

It's known from standard Newton theory (cf [8]) that if  $P_N e_0 \neq 0$  and  $\sigma_0$  and  $\|e_0\|$  are sufficiently small there exists constants  $c > 0$ ,  $\psi_0 \in (0, 1)$ ,  $\eta_0 \in (0, 1)$ , and  $s_0 \in (0, 1)$  such that  $\|P_R e_1\| < c \|e_0\|^2$ ,  $\|P_N e_1\| < \psi_0 \|P_N e_0\|$ ,  $\|e_1\| < \eta_0 \|e_0\|$ , and  $\sigma_1 < s_0 \sigma_0$ . Moreover,  $\hat{D}_N(x_0)$  and  $\hat{D}_N(x_1)$  are invertible as a maps on  $N$ . Hence (9) and (10) hold. Apply  $P_N$  to (9) with  $k = 1$  and pull out  $\|(1/2)P_N e_1^\alpha\|$  to get

$$(B.3) \quad \|P_N e_2\| \leq \left(1 + \frac{\|T_1 P_R e_1^\alpha + q_0^1\|}{\|P_N e_1^\alpha/2\|}\right) \|P_N e_1^\alpha/2\| \\ \leq \left(1 + \frac{\|T_1 P_R e_1^\alpha + q_0^1\|}{(1 - \tilde{\nu}_2) \max\{|1 - \gamma_2| \|P_N e_1\|, |\gamma_2| \|P_N e_0\|\}/2}\right) \|P_N e_1^\alpha/2\|.$$

Lemma 6.1 guarantees that with  $\gamma$ -safeguarding,  $1/(1 - \tilde{\nu}_{k+1})$  remains bounded independent of  $k$ . Moreover,

$$(B.4) \quad \|T_1 P_R e_1^\alpha + q_0^1\| \leq c_5 \max\{|1 - \gamma_2| \|P_R e_1\|, |\gamma_2| \|P_R e_0\|, |1 - \gamma_2| \|e_1\|^2, |\gamma_2| \|e_0\|^2\}.$$

The constant  $c_5$  here depends on  $f$ . Since for any real numbers  $a$  and  $b$ ,  $1/\max\{a, b\} \leq \min\{1/a, 1/b\}$ , it follows that

$$(B.5) \quad \|T_1 P_R e_1^\alpha + q_0^1\|/\|P_N e_1^\alpha\| \leq c \max\{\|P_R e_i\|/\|P_N e_i\|, \|e_i\|^2/\|P_N e_i\|\},$$

with  $i = 0$  or  $i = 1$  depending on the value of the right hand side of (B.4). The constant  $c$  here depends on  $f$  and the value of  $r$  in algorithm 2. Noting that  $\|e_i\|^2/\|P_N e_i\| < (1 + \sigma_i)\|e_i\|$ , it follows that

$$(B.6) \quad \|P_N e_2\| \leq (1 + C_1(\sigma_{0,1}, \|e_{0,1}\|)) \|(1/2)P_N e_1^\alpha\|,$$

where  $C_1(\sigma_{0,1}, \|e_{0,1}\|) = c \max\{\sigma_0, \sigma_1, \|e_0\|, \|e_1\|\}$ . Through an analogous argument applied to (10) and the definition of  $\theta_{k+1}$ , we can bound  $\|(1/2)P_N e_1^\alpha\|$  in terms of



$\|w_2^\alpha\|$  to obtain  $\|P_N e_2\| \leq ((1 + C_1(\sigma_{0,1}, \|e_{0,1}\|))(1 - C_2(\sigma_{0,1}, \|e_{0,1}\|))^{-1})\theta_2\|w_2\|$ . Using (8) to bound  $\|w_2\|$  in terms of  $\|P_N e_1\|$  yields

$$(B.7) \quad \|P_N e_2\| \leq \left( \frac{(1/2)(1 + C_1(\sigma_{0,1}, \|e_{0,1}\|))(1 + C_3(\sigma_{0,1}, \|e_{0,1}\|))}{1 - C_2(\sigma_{0,1}, \|e_{0,1}\|)} \right) \theta_2 \|P_N e_1\|.$$

We'll write  $(1 + C_1(\sigma_{0,1}, \|e_{0,1}\|))(1 + C_3(\sigma_{0,1}, \|e_{0,1}\|)) = (1 + C_1(\sigma_{0,1}, \|e_{0,1}\|))$ . Reducing  $\|e_0\|$  and  $\sigma_0$  further if necessary, we have

$$(B.8) \quad \|P_N e_2\| < \kappa \theta_2 \|P_N e_1\|$$

where  $1/2 < \kappa < 1$ . Note that  $\|P_N e_1\| \leq \psi_0 \|P_N e_0\| \leq \psi_0(1 - \sigma_0)^{-1} \|e_0\|$ . Writing  $\psi_1 := \kappa \theta_2 \psi_0(1 - \sigma_0)^{-1}$  gives  $\|P_N e_2\| \leq \psi_1 \|e_0\|$ . Since we enforce  $|\gamma_k| \leq 1$  for all  $k$  with safeguarding, it follows that

$$(B.9) \quad \|P_R e_2\| \leq c \max\{\|e_1\|^2, \|e_0\|^2\}.$$

By standard Newton theory we have  $\|e_1\| < \|e_0\|$ , so in this case  $\|P_R e_2\| < c\|e_0\|^2$ . Equations (B.8) and (B.9) give  $\|e_2\| \leq \|P_R e_2\| + \|P_N e_2\| \leq (c\|e_0\| + \psi_1)\|e_0\|$ . Reducing  $\|e_0\|$  and  $\sigma_0$  if necessary, we can ensure  $\eta_1 := (c_1\|e_0\| + \psi_1) < 1$ . Hence  $\|e_2\| < \eta_1 \|e_0\|$ . Next we consider  $\sigma_2$ . Applying the reverse triangle inequality directly to  $L_2^\lambda(e_1 + w_2, e_0 + w_1)$  gives  $\|P_N e_2\| \geq c(1 - \nu_2) \max\{|1 - \gamma_2| \|P_N e_1\|, |\gamma_2| \|P_N e_0\|\}$ . It follows that

$$(B.10) \quad \sigma_2 \leq \frac{c \max\{|1 - \gamma_2| \|e_1\|^2, |\gamma_2| \|e_0\|^2\}}{(1 - \nu_2) \max\{|1 - \gamma_2| \|P_N e_1\|, |\gamma_2| \|P_N e_0\|\}} \leq \frac{c(1 + \sigma_0)}{1 - \nu_2} \|e_0\|,$$

The last inequality follows since  $\|e_1\| < \eta_0 \|e_0\| < \|e_0\|$ . By lemma 6.1,  $c(1 + \sigma_0)/(1 - \nu_2) < c_6$ . Once more reducing  $\|e_0\|$  if necessary, we can ensure

$$(B.11) \quad \sigma_2 < c_6 \|e_0\| < s_0 \sigma_0.$$

Hence  $W_2 \subset W_0$ , and  $x_3$  is well-defined. Thus (B.8), (B.9), and (B.11) all hold with each subscript increased by one since  $C_1(\sigma_{1,2}, \|e_{1,2}\|) = c \max\{\sigma_1, \sigma_2, \|e_1\|, \|e_2\|\} < c \max\{\sigma_0, \sigma_1, \|e_0\|, \|e_1\|\} = C_1(\sigma_{0,1}, \|e_{0,1}\|)$ . If we let  $\psi_2 := \kappa \theta_3 \psi_1$ , then  $\|e_3\| < (c \max\{\eta_0, \eta_1\}^2 \|e_0\| + \psi_2)\|e_0\|$ . Let  $\eta_2 := (c \max\{\eta_0, \eta_1\}^2 \|e_0\| + \psi_2)$  so that  $\|e_3\| < \eta_2 \|e_0\|$ . As  $\max\{\eta_0, \eta_1\} \leq \hat{\eta} < 1$ ,  $\eta_2 \leq \max\{\hat{\eta}, \kappa \theta_3\}(c\|e_0\| + \psi_1) = \max\{\hat{\eta}, \kappa \theta_3\} \eta_1 < \eta_1$  since  $\kappa \theta_3 < 1$ . Let  $\phi = \max\{\hat{\eta}, \kappa\} < 1$  so that  $\eta_2 < \phi \eta_1 < \hat{\eta}$ . Therefore  $\|e_3\| < \|e_0\|$  and  $W_3 \subset W_0$ . Proceeding inductively, it can be shown that for all that for all  $k \geq 1$ ,  $W_{k+1} \subset W_0$ ,  $\|P_N e_{k+1}\| < \kappa \theta_{k+1} \|P e_k\| < \psi_k \|e_0\|$ , and  $\|e_{k+1}\| < \eta_k \|e_0\|$ . This proves that  $x_k$  remains well defined, and establishes (B.1) and (B.2).

It remains to show that  $e_k$  goes to zero. It suffices to prove that  $\{\eta_k\}$  defined by  $\eta_k = c \max\{\eta_{k-1}, \eta_{k-2}\}^2 \|e_0\| + \psi_k$  converges to zero. Again by induction, one can show that  $\eta_k$  decreases, and for  $k \geq 7$  we have  $\eta_{k+1} < \phi^k \eta_1$ . Thus  $\eta_k \rightarrow 0$ , and therefore  $\|e_{k+1}\| \rightarrow 0$ . This completes the proof.  $\square$

Note that the essential parts of the proof rely on the structure of  $\|P_N e_k\|$  and  $\|P_R e_k\|$ . For any order root, the error expansions are of the form

$$(B.12) \quad P_N(e_k + w_{k+1}) = \delta P_N e_k + \hat{T}_k P_R e_k + \mathcal{O}(\|e_k\|^2)$$

$$(B.13) \quad P_R(e_k + w_{k+1}) = \mathcal{O}(\|e_k\|^\mu)$$

where  $\delta \in (0, 1)$  and  $\mu \geq 2$ . Thus an analogous argument can be used to prove 7.1. The only noticeable difference occurs at equation (B.8), where in the general case

$\kappa \in (d/(d+1), 1)$ . This, however, does not effect convergence since  $\theta_{k+1} \leq 1$  for all  $k$ .

Department of Mathematics, University of Florida, Gainesville, FL, 32611  
*E-mail:* mdallas1@ufl.edu and s.pollock@ufl.edu