

Some Results on the Convergence of Anderson Acceleration

by

Adam Smith

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2024

© Adam Smith 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Anderson acceleration (AA), also known as Anderson mixing, is an extrapolation technique used to accelerate the convergence of fixed-point iterations $\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k)$, $k = 0, 1, \dots$, with $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{x}_k \in \mathbb{R}^n$. AA was first introduced by D.G. Anderson in the context of solving integral equations but has since been adapted to fixed-point iteration problems in general. Despite relatively little being known about its convergence properties, AA has seen considerable usage in several areas such as electronic structure computations and machine learning. This thesis presents a broad overview of the current convergence literature for AA and introduces a variety of new results concerning properties of AA, such as its asymptotic convergence rate, the possibility of stagnation, and an analysis of its coefficients. Additionally, some variations of the AA iteration are proposed with an accompanying analysis and comparison to the classical AA algorithm.

Acknowledgements

I am sincerely grateful to my supervisor, Hans De Sterck, as well as Oliver Krzysik and Yunhui He for their consistent support and exceptional guidance throughout my Master's Research. Their knowledge and advice were integral to the completion of this thesis.

Additionally, I am deeply appreciative of the love and support I received throughout this journey from my partner Laura Roach, my parents Dan and Annette Smith, as well as the rest of my family.

Table of Contents

Author's Declaration	ii
Abstract	iii
Acknowledgements	iv
List of Figures	vii
1 Introduction	1
2 Background	3
2.1 Generalized Minimal Residual Method	3
2.2 General Fixed-point Iterations	6
2.2.1 Newton's Method	6
2.2.2 Power Iteration	7
2.2.3 Contraction Mappings	8
2.3 Anderson Acceleration	9
2.3.1 Relationship with Other Iterations	10
2.3.2 Notions of Convergence	12
2.3.3 Existing Convergence Literature	13

3	Asymptotic Convergence Rates	17
3.1	Investigation of a Simple Linear Setting	19
3.2	Asymptotic Convergence Factors for Certain Linear Maps	22
3.2.1	Convergence for W_λ	25
3.2.2	Convergence of M_λ	27
4	Stagnation	30
4.1	Stagnation for General Problems	33
4.2	Stagnation for Linear Problems	36
4.2.1	Skew-symmetric Matrices	37
4.2.2	Symmetric Matrices	39
4.3	Nonlinear Example	40
5	Acceleration Coefficients for Linear Problems	43
5.0.1	Coefficients for AA(1)	43
5.0.2	Coefficients for AA(2)	48
6	Restarted Anderson Acceleration	53
6.1	Convergence of rAA(2)	54
6.2	Convergence of arAA	59
7	Conclusion	65
	References	67

List of Figures

3.1	Acceleration coefficients β_k as a function of iteration number (left) and the approximate convergence factor $\ \mathbf{r}_k\ ^{1/k}$ as a function of iteration number (right) for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$, and $\mathbf{r}_0 = (-0.1285, 0.6687)^T$	20
3.2	Asymptotic convergence factor as a function of initial residual $\mathbf{r}_0 \in [-1, 1]^2$ for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$. A 128×128 grid was used for the initial residuals and convergence factors were numerically found after generating the residuals according to (3.3) for each grid point.	21
3.3	Worst-case asymptotic convergence factor for AA(1) (blue) applied to (3.5) as a function of λ_2 with λ_1 fixed compared with the associated convergence factor for the underlying fixed-point iteration (red) and the bound for the convergence factor (yellow) given by Theorem 2.3.1.	21
3.4	Acceleration coefficients β_k as a function of iteration number for AA(1) applied to affine function with $M = M_{0.7}$ (left) and $M = W_{0.7}$ (right). . . .	23
4.1	Normed iterates $\ \mathbf{x}_k\ $ as a function of iteration number k for AA(m) and FP applied to a random skew-symmetric A with random initial iterate, comparing $m = 1, 2, 3$. Stagnation periods of length $m+1$, separated by fixed-point iterations, can be observed for each choice of m	38
4.2	Normed iterates $\ \mathbf{x}_k\ $ as a function of iteration number k ($0 \leq k \leq 13$ in left plot and $0 \leq k \leq 100$ in right plot) for AA applied to $\mathbf{q}(\mathbf{x})$ of (4.18). Yellow and purple plots take initial iterate according to [21] ($x_0 = (3, -1, 0, 1)^T$).	41
5.1	Acceleration coefficients β_k as a function of $y_k \in [0, \ M\]$ and $\phi_k \in [0, \pi]$ according to (5.4)	45

5.2	Acceleration coefficients β_k as a function of iteration number for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$, and $\mathbf{r}_0 = (-0.1285, 0.6687)^T$. The dotted blue lines correspond to the upper and lower bounds for β_k as determined by Theorem 5.0.1 with upper bound given by $7/3$ and lower bound given by $-7/17$	46
5.3	Illustration of AA(2) residual bounds from Theorem 5.0.3 applied to affine $\mathbf{q}$ with $M = \text{diag}(0.7, -0.5, 0.3)$ and $\mathbf{r}_0$ chosen randomly. The red points correspond to the values of $y_{k+1} = \ \mathbf{r}_{k+1}\ / \ \mathbf{r}_k\ $ generated by AA(2) and the black bars correspond to the bound given for each y_{k+1} in Theorem 5.0.3. The iterations considered in the above plot range from $k = 2$ to $k = 21$. . .	51
6.1	Convergence of rAA(2) for symmetric $M \in \mathbb{R}^{2 \times 2}$ with eigenvalues m_1 and m_2 . (Left) Worst-case asymptotic convergence factor of rAA(2) according to Corollary 6.1.1. (Right) Ratio of worst-case asymptotic convergence factor of rAA(2) and convergence factor of fixed-point iteration.	59
6.2	Worst-case asymptotic convergence factor for AA(1) (blue) applied to (3.5) as a function of λ_2 with λ_1 fixed compared with the associated convergence factor for the underlying fixed-point iteration (red), the bound for the convergence factor (yellow) given by Theorem 2.3.1, and the worst-case asymptotic convergence factor for rAA(2) (purple) according to Corollary 6.1.1. .	60
6.3	Approximate r-linear convergence factor $\ \mathbf{r}_k\ ^{1/k}$ as a function of iteration number k for arAA applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$ for a random initial residual $\mathbf{r}_0$	64

Chapter 1

Introduction

Anderson acceleration (AA) is an iterative method first proposed in [1] to solve nonlinear integral equations. Although first discussed in the context of integral equations, it was soon adapted and applied to a variety of applications. Such applications include accelerating the self-consistent field method for electronic structure computations [23], geometry optimization [28], radiation diffusion and nuclear physics [40], and machine learning [38].

The term *acceleration* is derived from the observation that AA often improves the convergence speed of fixed-point iterations. That is, given a function $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with fixed point x^* , the fixed-point iteration given in (2.7) with initial guess $\mathbf{x}_0$ can potentially see significantly faster convergence with the application of AA. Additionally, AA can expand the domain of convergence to include initial iterates that would otherwise not see convergence in the underlying fixed-point iteration. To contextualize these points, the AA iteration is now introduced. Consider solving for a fixed point of the function $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, then the Anderson iteration formula is given by

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \sum_{j=1}^k \beta_j^{(k)} (\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-j})) \quad k = 0, 1, 2, \dots, \quad (1.1)$$

where the coefficients $\{\beta_j^{(k)}\}_{j=1}^k$ are determined from the optimization problem

$$\min_{\{\beta_j^{(k)}\}} \left\| \mathbf{r}_k + \sum_{j=1}^k \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right\|^2. \quad (1.2)$$

Here, $\mathbf{r}_k = \mathbf{x}_k - \mathbf{q}(\mathbf{x}_k)$ denotes the residual at the k -th iteration of AA. Currently, the convergence theory for AA has been relatively unsuccessful in quantifying the asymptotic

convergence rate. The goal of this thesis is to explore various convergence properties of AA.

This thesis is divided into six chapters. Background material on preliminary topics such as GMRES, fixed-point iterations, contraction mappings, and existing AA convergence literature is provided in [Chapter 2](#). This is followed by a discussion on the asymptotic convergence rates for AA applied to linear problems in [Chapter 3](#), with a special interest placed on two classes of linear problems. [Chapter 4](#) investigates the presence of stagnation in AA applied to both linear and nonlinear settings. In [Chapter 5](#), expressions are derived for the acceleration coefficients used in AA in the linear setting, which are applied to produce new residual bounds that make explicit the acceleration achieved at each iteration. An analysis of two new AA variations is given in [Chapter 6](#) and compared with the traditional AA method.

Chapter 2

Background

Before we dive into an analysis of the Anderson acceleration method, it is important for us to first establish some important background material. In particular, this chapter will concentrate on presenting the current convergence literature regarding Anderson acceleration as well as the relationship Anderson acceleration has with other iterative methods.

2.1 Generalized Minimal Residual Method

Consider solving the linear system

$$A\mathbf{x} = \mathbf{b}, \tag{2.1}$$

where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is nonsingular. Since A is nonsingular, there exists a unique solution to (2.1) which we will denote by $\mathbf{x}^*$, i.e. $A\mathbf{x}^* = \mathbf{b}$. Given an initial approximation $\mathbf{x}_0 \in \mathbb{R}^n$, a *projection method* constructs a sequence of approximate solutions $\mathbf{x}_k$, $k = 1, 2, \dots$, such that

$$\mathbf{x}_k \in \mathbf{x}_0 + \mathcal{S}_k, \quad \mathbf{r}_k \perp \mathcal{L}_k \tag{2.2}$$

where $\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b}$ are the residual vectors and $\mathcal{S}_k$ and $\mathcal{L}_k$ are k -dimensional subspaces of $\mathbb{R}^n$ termed the candidate space and constraint space respectively. The iterate update and the residual orthogonality constraint present in (2.2) is sometimes referred to as the *Petrov-Galerkin* conditions [36]. Some examples of projection methods include Steepest Descent [5], Minimal Residual (MINRES) [27], and the Full Orthogonalization Method (FOM) [35]. Perhaps the most notable example of such a projection method is known as the *Generalized Minimal RESidual* (GMRES) method [34], which seeks to minimize the

residual norm over all vectors in the candidate space, i.e. it finds $\mathbf{x}_k$ such that

$$\|\mathbf{r}_k\| = \|A\mathbf{x}_k - \mathbf{b}\| = \min_{\mathbf{z} \in \mathbf{x}_0 + \mathcal{S}_k} \|A\mathbf{z} - \mathbf{b}\|, \quad (2.3)$$

where $\mathcal{S}_k$ is taken to be the *Krylov Subspace* $\mathcal{K}_k$, which is defined as follows:

Definition 2.1.1. Let $A \in \mathbb{R}^{n \times n}$ and $\mathbf{r}_0 \in \mathbb{R}^n$, then the *Krylov subspace* generated by A and $\mathbf{r}_0$ is the subspace of $\mathbb{R}^n$ given by

$$\mathcal{K}_m(A, \mathbf{r}_0) = \text{Span}\{\mathbf{r}_0, A\mathbf{r}_0, A^2\mathbf{r}_0, \dots, A^{m-1}\mathbf{r}_0\}.$$

Throughout this thesis, the Krylov subspace $\mathcal{K}_m(A, \mathbf{r}_0)$ will be often be written as $\mathcal{K}_m$ when there is no ambiguity. These subspaces consist of vectors $\mathbf{v} \in \mathbb{R}^n$ such that $\mathbf{v} = p(A)\mathbf{r}_0$ where p is a polynomial of degree at most $m - 1$. With system (2.1) and Definition 2.1.1 in mind, and setting $\mathbf{r}_0 = A\mathbf{x}_0 - \mathbf{b}$ for some initial guess $\mathbf{x}_0$, the GMRES approximations are of the form

$$A^{-1}\mathbf{b} \approx \mathbf{x}_k = \mathbf{x}_0 + p_{k-1}(A)\mathbf{r}_0,$$

where p_{k-1} is a certain polynomial of degree $k - 1$. In particular, if $\mathbf{x}_0 = 0$, our approximation takes the form

$$\mathbf{x}_k = p_{k-1}(A)\mathbf{b}. \quad (2.4)$$

In other words, we approximate $A^{-1}\mathbf{b}$ by $p_{k-1}(A)\mathbf{b}$ for some polynomial of degree $k - 1$. In particular, when $k = n$, the Cayley-Hamilton Theorem [24] indicates that A^{-1} can be expressed as a $(n - 1)$ -degree polynomial using the characteristic polynomial of A . This fact, combined with the update formula for GMRES shown in (2.4), can be used to show that GMRES converges to the exact solution $\mathbf{x}^*$ in *at most* n iterations. While GMRES will converge in at most n iterations, there is still the possibility of it converging even sooner, and to further analyze this behaviour, the following definition is introduced.

Definition 2.1.2. The *minimal polynomial* of a vector $\mathbf{v}$ with respect to A is the nonzero monic polynomial p of lowest degree such that

$$p(A)\mathbf{v} = \mathbf{0}.$$

The degree of the minimal polynomial is called the *grade of $\mathbf{v}$ with respect to A* .

For the same reasons discussed a moment ago, GMRES will see convergence to the exact solution after μ iterations, where μ is the grade of $\mathbf{r}_0$ with respect to A . With Definition 2.1.2 established, the following property regarding Krylov subspaces emerges.

Proposition 2.1.1. ([36]) Let μ be the grade of $\mathbf{r}_0$, then $\mathcal{K}_\mu$ is *invariant* under A and $\mathcal{K}_k = \mathcal{K}_\mu$ for all $k \geq \mu$. i.e.,

$$A\mathcal{K}_\mu = \mathcal{K}_\mu.$$

[Proposition 2.1.1](#) highlights the property that the dimension of Krylov subspaces are bounded above by the grade of the initial residual $\mathbf{r}_0$. More precisely, we have

$$\dim(\mathcal{K}_k) = \min(k, \text{grade}(\mathbf{r}_0)). \quad (2.5)$$

A larger Krylov subspace will be generated at each step of the GMRES iteration until eventually $k = \mu$, at which point the GMRES update will produce the exact solution x^* to problem (2.1). This finite termination of the method results from the minimization of the residual (2.3) and the grade

We now present the basic algorithm for GMRES.

Algorithm 1 GMRES

```

1: Compute  $r_0 = b - Ax_0$ 
2: Compute  $\beta = \|r_0\|_2$ 
3: Compute  $q_1 = r_0/\beta$ 
4: for  $i = 1, \dots, m$  do
5:   Compute  $w_i = Aq_i$ 
6:   for  $j = 1, \dots, i$  do
7:      $h_{j,i} = q_j^T w_i$ 
8:      $w_i = w_i - h_{j,i}q_j$ 
9:   end for
10:   $h_{i+1,i} = \|w_i\|_2$ 
11:  if  $h_{i+1,i} = 0$  then
12:    Set  $m := i$  and go to 16
13:  end if
14:   $q_{i+1} = w_i/h_{i+1,i}$ 
15: end for
16: Find  $y_m$  that minimizes  $\|\beta e_1 - \tilde{H}_m y\|_2$ 
17: Compute  $x_m = x_0 + Q_m y_m$ 

```

A complete explanation of [Algorithm 1](#) is unnecessary for the purposes of this thesis but interested readers are referred to [36] if a deeper analysis is required. The most notable steps in [Algorithm 1](#) include the *Arnoldi procedure* in lines 4-15 to generate an orthogonal

basis for the Krylov subspace, and the minimization in line 16 for the residual that is then used to compute the update $\mathbf{x}_m$ in the following line.

GMRES is not the focus of this thesis, but as discussed in [Section 2.3.1](#), it has a deep connection with Anderson acceleration. Further, Anderson acceleration falls into the category of Krylov subspace methods [\[7\]](#) and thus it is important to understand other methods, such as GMRES, that also belong to this group of iterative methods.

2.2 General Fixed-point Iterations

We now shift focus from GMRES to a more general type of numerical method or class of numerical methods known as fixed-point iterations. Consider solving for a fixed point $\mathbf{x} \in \mathbb{R}^n$ of the function $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. In other words, solving the system

$$\mathbf{q}(\mathbf{x}) = \mathbf{x}. \quad (2.6)$$

A common approach to solve [\(2.6\)](#) is by taking a sufficiently chosen initial guess $\mathbf{x}_0$ and iterating

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k), \quad k = 0, 1, 2, \dots \quad (2.7)$$

The procedure outlined in [\(2.7\)](#) is known as a fixed-point iteration and various different numerical methods exist that follow this form. What is meant by a “sufficiently chosen initial guess” for fixed-point iterations depends on the application but generally, this condition requires that $\mathbf{x}_0$ is close enough to the solution $\mathbf{x}^*$ such that the iteration will converge. To gain a better understanding of fixed-point iterations, we now introduce some of the most popular examples.

2.2.1 Newton’s Method

For example, consider Newton’s Method, which attempts to solve for the roots of a function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by iterating

$$\mathbf{x}_{k+1} = \mathbf{x}_k - J(\mathbf{x}_k)^{-1} \mathbf{f}(\mathbf{x}_k), \quad k = 0, 1, 2, \dots, \quad (2.8)$$

where $J(\mathbf{x}_k)$ is the Jacobian matrix of $\mathbf{f}$ at the point $\mathbf{x}_k$. If we write

$$\mathbf{q}(\mathbf{x}) = \mathbf{x} - J(\mathbf{x})^{-1} \mathbf{f}(\mathbf{x}),$$

then if $\mathbf{x}^*$ is a fixed point of $\mathbf{q}$, then

$$\mathbf{x}^* = \mathbf{q}(\mathbf{x}^*) = \mathbf{x}^* - J(\mathbf{x}^*)^{-1} \mathbf{f}(\mathbf{x}^*),$$

meaning

$$-J(\mathbf{x}^*)^{-1} \mathbf{f}(\mathbf{x}^*) = \mathbf{0},$$

therefore $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. That is, $\mathbf{x}^*$ is a root of $\mathbf{f}$ as desired. Newton's method is widely used in applications as it often attains a *quadratic order of convergence*, which can be interpreted as meaning that asymptotically the number of correct digits in the approximate solution roughly doubles with each step. The topic of convergence order and convergence rate will be discussed in more depth in [Section 2.3](#) when introducing the convergence properties of Anderson acceleration.

One major drawback of Newton's method is the requirement that the Jacobian be available at each step, which may not be practical. In an attempt to overcome this problem, Quasi-Newton methods may be considered, which instead use an approximation of the Jacobian or the inverse Jacobian at each step. A simple example of a Quasi-Newton method for one-dimensional problems is given by the secant method

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \quad k = 1, 2, 3, \dots, \quad (2.9)$$

which approximates the Jacobian at each step by a secant line joining $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$, starting from initial guesses $x_0, x_1 \in \mathbb{R}$. This secant update removes the need for Jacobian computations but at the cost of slower convergence.

2.2.2 Power Iteration

Another fixed-point iteration example worth discussing is the power iteration method,

$$\mathbf{v}_{k+1} = \frac{A\mathbf{v}_k}{\|A\mathbf{v}_k\|}, \quad k = 0, 1, 2, \dots, \quad (2.10)$$

which converges to an eigenvector $\mathbf{v} \in \mathbb{R}^n$ of matrix $A \in \mathbb{R}^{n \times n}$. More specifically, the power iteration converges to the eigenvector associated with the dominant eigenvalue (largest magnitude eigenvalue) assuming such an eigenvalue exists and the initial vector $\mathbf{v}_0$ has a nonzero component in the direction of the dominant eigenvector. Furthermore, after computing a suitable approximation $\tilde{\mathbf{v}} \approx \mathbf{v}$ using the power iteration, the *Rayleigh quotient* can be used to approximate the spectral radius associated with symmetric A .

Definition 2.2.1. The Rayleigh quotient for a given matrix $B \in \mathbb{R}^{n \times n}$ and nonzero vector $\mathbf{x} \in \mathbb{R}^n$ is given by

$$R(B, \mathbf{x}) = \frac{\mathbf{x}^T B \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Remark 2.2.1. The above definition for a Rayleigh quotient makes no assumption as to whether matrix B is symmetric but the Rayleigh quotient for a nonsymmetric B evaluates to the Rayleigh quotient of the symmetric part of B , i.e.,

$$R(B, \mathbf{x}) = R((B + B^T)/2, \mathbf{x}). \quad (2.11)$$

Now, if $\tilde{\mathbf{v}}$ is an approximation found using the power iteration and $\{\lambda_j\}_{j=1}^n$ are the real eigenvalues of a symmetric matrix A , then

$$\text{spectral radius}(A) = \max(|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|) \approx R(A, \tilde{\mathbf{v}}) = \frac{\tilde{\mathbf{v}}^T A \tilde{\mathbf{v}}}{\tilde{\mathbf{v}}^T \tilde{\mathbf{v}}}.$$

In [Chapter 6](#), Rayleigh quotients are used to define some Anderson acceleration variant algorithms. Shown below are a few properties of these quotients that are needed for the analysis presented in [Chapter 6](#). Let $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ and $S \in \mathbb{R}^{n \times n}$ be a symmetric matrix, then for all $0 \neq c \in \mathbb{R}$ we have

$$R(S, \mathbf{x}) = R(S, c\mathbf{x}). \quad (2.12)$$

If $(s_1, \mathbf{u}_1)$ and $(s_2, \mathbf{u}_2)$ are two eigenpairs of S , with $\mathbf{u}_1, \mathbf{u}_2$ orthogonal, then

$$R(S, \mathbf{u}_1 + \mathbf{u}_2) = \frac{s_1 \|\mathbf{u}_1\|^2 + s_2 \|\mathbf{u}_2\|^2}{\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2}. \quad (2.13)$$

2.2.3 Contraction Mappings

Determining whether a fixed-point iteration converges or not in general can be tricky, as we will soon see for Anderson acceleration. Perhaps the most useful result concerning the convergence of general fixed-point iterations is given by the *Banach fixed-point theorem* or the *contraction mapping theorem* [[3](#), [18](#)].

Definition 2.2.2. Let $\Omega \subset \mathbb{R}^n$. We say a function $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ is *Lipschitz continuous* on Ω with *Lipschitz constant* $\sigma_{\mathbf{f}}$ if

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)\| \leq \sigma_{\mathbf{f}} \|\mathbf{x}_1 - \mathbf{x}_2\|,$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$

Definition 2.2.3. Let $\Omega \subset \mathbb{R}^n$. We say a function $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ is a *contraction mapping* on Ω if $\mathbf{f}$ is Lipschitz continuous with Lipschitz constant $\sigma_{\mathbf{f}} < 1$.

Theorem 2.2.1. Let Ω be a closed subset of $\mathbb{R}^n$ and let $\mathbf{q}$ be a contraction mapping on Ω with Lipschitz constant $\sigma_{\mathbf{q}} < 1$ such that $\mathbf{q}(\mathbf{x}) \in \Omega$ for all $\mathbf{x} \in \Omega$. Then there is a unique fixed point $\mathbf{x}^*$ of $\mathbf{q}$ and the iteration defined by (2.7) converges to $\mathbf{x}^*$ for all initial iterates $\mathbf{x}_0 \in \Omega$.

2.3 Anderson Acceleration

Iteration (1.1) shows that an AA update is comprised of a regular fixed-point iteration $\mathbf{q}(\mathbf{x}_k)$ along with a weighted sum of function evaluations to previous iterates $\sum_{j=1}^k \beta_j^{(k)}(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-j}))$. Instead of only considering the most recent iterate $\mathbf{x}_k$ in the update for $\mathbf{x}_{k+1}$, AA instead employs a number of previous iterates in an attempt to produce a more accurate approximation $\mathbf{x}_{k+1}$. One thing to notice with the form of (1.1) is that as the iteration progresses, the number of previous iterates used also increases. This can pose problems later in the method as the storage costs increase and the computational cost of optimization problem (1.2) increases. As such, many variations of (1.1) have been proposed to help alleviate this issue, the most notable example being *truncated* or *windowed* AA. As opposed to iteration (1.1), which will now be referred to *full-memory* AA, windowed AA only considers a given number of the most recent iterates to produce the next iterate. Denote by $m > 0$ the *depth* or *window size* and let $m_k = \min(k, m)$, then the windowed AA iteration is given by

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \sum_{j=1}^{m_k} \beta_j^{(k)}(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-j})) \quad k = 0, 1, 2, \dots, \quad (2.14)$$

where the coefficients follow from the associated minimization problem

$$\min_{\{\beta_j^{(k)}\}} \left\| \mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)}(\mathbf{r}_k - \mathbf{r}_{k-j}) \right\|^2. \quad (2.15)$$

For the remainder of this thesis, the windowed AA iteration shown in (2.14) will be denoted by $\text{AA}(m)$ or sometimes just AA if the context is clear. Further, the full-memory AA iteration given in (1.1) will be denoted by $\text{AA}(\infty)$.

Remark 2.3.1. Some authors further denote $\text{AA}(0)$ to be the fixed-point iteration (2.7) as it would not consider any other iterates besides $\mathbf{x}_k$. This is consistent with (2.14) with the convention that $\sum_{j=1}^0 a_j = 0$.

One thing to mention about the $AA(m)$ iteration (2.14) is the use of m_k rather than just m . Using m_k allows the user to build up to the desired window size m using a single initial iterate $\mathbf{x}_0$. That is, until the window size is reached, $AA(m)$ behaves in the same manner that $AA(\infty)$ would. Instead of using m_k as defined above, a user could instead start from a set of initial iterates $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ and immediately begin $AA(m)$ without the initial process where $m_k \rightarrow m$. This type of process is often referred to as AA with *general initial guess* [7].

As mentioned earlier, there are several other variations of AA . Another notable AA algorithm will be discussed in Chapter 6 that is generally known as *restarted* AA . Iterations of this form seek to limit the memory requirement of AA , just as $AA(m)$ does, by instead restarting the method using the most recent iterate after a given number of steps have been taken. A specific class of restarting AA methods will be analyzed in Chapter 6 that builds off the analysis presented in [8] with a comparison to made to $AA(m)$.

2.3.1 Relationship with Other Iterations

As AA has seen more application in recent years, a number of comparisons and links have been made in the literature between AA and other iterative methods. The most notable of these relations will be discussed in this section.

GMRES

Generally, AA can be applied to find fixed points of nonlinear functions $\mathbf{q}$ but it can be analyzed more specifically in the context of linear problems. In this linear context, a relationship exists between $AA(\infty)$ and GMRES that is termed “essentially equivalent” [31]. Specifically, consider affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $M \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$, and let $\mathbf{x}_k^A$, $\mathbf{x}_k^G$ denote the k -th iterate of an $AA(\infty)$ iteration and GMRES iteration respectively. Then given some basic assumptions, we have

$$\mathbf{x}_{k+1}^A = \mathbf{q}(\mathbf{x}_k^G) = M\mathbf{x}_k^G + \mathbf{b}, \quad (2.16)$$

when $\mathbf{x}_0^A = \mathbf{x}_0^G$. Therefore, $AA(\infty)$ iterates can be directly obtained from corresponding GMRES iterates. As the above relation might suggest, AA applied to linear problems can be analyzed using the framework of Krylov subspace methods [7]. To be more precise in the relationship between $AA(\infty)$ and GMRES shown in (2.16), the basic assumptions required include that the matrix $I - M$ is nonsingular and that GMRES does not exhibit any *stagnation*. This means that the GMRES residuals satisfy $\|\mathbf{r}_j^G\| < \|\mathbf{r}_{j-1}^G\|$ for each $j =$

$1, \dots, k$. The term stagnation generally refers to an iteration making no progress towards the solution of the underlying problem for a given number of iterations and stagnation for GMRES specifically has been analyzed extensively in sources such as [19, 20, 37, 41]. In Chapter 4, we explore the problem of stagnation in AA(m).

Multisecant Quasi-Newton Methods

As presented in [13], AA falls into the larger category of multisecant quasi-Newton methods. These are methods which replace the Jacobian in the Newton iteration (2.8) with an approximate Jacobian that satisfies a given number of secant conditions. In particular, AA approximates the inverse Jacobian J^{-1} at step k using a matrix G_k that satisfies the secant conditions

$$G_k(\mathbf{r}_k - \mathbf{r}_{k-j}) = \mathbf{x}_k - \mathbf{x}_{k-j}, \quad j = 1, \dots, m,$$

as well as the so-called *no change* condition given by

$$(G_k + I)\mathbf{u} = 0,$$

for all $\mathbf{u}$ orthogonal to the subspace spanned by the vectors $\{\mathbf{r}_k - \mathbf{r}_{k-j}\}_{j=1}^m$. This classification of AA identifies the method as an example of the *generalized Broyden's second method*. A detailed description of Broyden's methods and the connection with AA is given in [13, 12].

The identification of AA as a Multisecant method can be illustrated through the following simple example. Here, we show that AA(1) applied to solve for the fixed point of a function $\mathbf{q} : \mathbb{R} \rightarrow \mathbb{R}$ corresponds to applying the one-dimensional secant method to solve for the root of the function $f(x) = x - q(x)$. Recall the one-dimensional secant method presented in (2.9). Consider applying AA(1) to a function $q : \mathbb{R} \rightarrow \mathbb{R}$. Using (2.14), when $k \geq 1$ we have

$$x_{k+1} = q(x_k) + \beta_k[q(x_k) - q(x_{k-1})].$$

From (2.15) and the facts that $n = 1$ and $m = 1$, we can easily see that

$$\beta_k = \frac{-r_k}{r_k - r_{k-1}} = -\frac{x_k - q(x_k)}{x_k - x_{k-1} - q(x_k) + q(x_{k-1})}.$$

Substituting this coefficient we have

$$x_{k+1} = q(x_k) - \frac{x_k - q(x_k)}{x_k - x_{k-1} - q(x_k) + q(x_{k-1})}[q(x_k) - q(x_{k-1})].$$

If we now define $f(x) := x - q(x)$, then fixed points of q correspond to roots of f . Substituting this f above we obtain

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \quad (\text{the secant method! (2.9)})$$

meaning this AA(1) iteration applied to q corresponds to the secant method applied to f . Therefore, we can determine the convergence of AA(1) in this situation by applying well-known results concerning the convergence of the secant method, such as its superlinear order of convergence given by the golden ratio $\varphi = (1 + \sqrt{5})/2$ [9].

2.3.2 Notions of Convergence

A potential shortfall of AA in the current literature is the lack of a clear understanding of its convergence properties. In particular, there has been little success in proving that the *asymptotic convergence factor* is an improvement upon the underlying fixed-point iteration, despite the general numerical observation that this is the case. To understand this problem in more depth, the following convergence definitions are first introduced

Definition 2.3.1. A sequence $\{\xi_k\}$ converges *q-linearly* (*quotient-linearly*) to ξ^* if there exists a $\hat{C} \in (0, 1)$ such that

$$\|\xi_{k+1} - \xi^*\| \leq \hat{C} \|\xi_k - \xi^*\|$$

for all $k \geq 0$. Alternatively (according to [25]), we say a sequence converges *q-linearly* to ξ^* with *q-factor* $C \in (0, 1)$ if

$$\limsup_{k \rightarrow \infty} \frac{\|\xi_{k+1} - \xi^*\|}{\|\xi_k - \xi^*\|} = C.$$

Definition 2.3.2. A sequence $\{\xi_k\}$ converges *r-linearly* (*root-linearly*) to ξ^* with *r-linear convergence factor* $\rho_{\{\xi_k\}}$ given by

$$\rho_{\{\xi_k\}} = \limsup_{k \rightarrow \infty} \|\xi_k - \xi^*\|^{1/k},$$

if $\rho_{\{\xi_k\}} \in (0, 1)$. If $\rho_{\{\xi_k\}} = 0$, then we say $\{\xi_k\}$ is *r-superlinearly* convergent. For much of this thesis, we will also refer to the r-linear convergence factor as the *asymptotic convergence factor*.

For AA, it is important to note that, in general, the r-linear convergence factor is strongly dependent on the initial iterate $\mathbf{x}_0$ [7]. This contrasts with the underlying fixed-point iteration (2.6) whose r-linear convergence factor $\rho_{\{\mathbf{x}_k\}}^{\text{FP}}$ is largely independent of the initial iterate when $\mathbf{q}(\mathbf{x})$ is differentiable [7, 25]. Specifically, for an affine iteration function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, we have $\rho_{\{\mathbf{x}_k\}}^{\text{FP}} = \text{spectral radius}(M)$ except when the initial iterate $\mathbf{x}_0$ is parallel to a nondominant eigenvector of M .

2.3.3 Existing Convergence Literature

The first convergence results established for AA were presented in [39], which showed convergence of AA(m) in both the linear and nonlinear settings, under certain conditions. These results represented an important step in understanding whether AA will converge, but this analysis presented no insight into the acceleration observed of AA. That is, it was not shown in [39] that AA improves the convergence rate of the underlying fixed-point iteration. To aid in the discussion of AA convergence rates in the following chapters, it will be helpful to present the basic convergence theorem and proof for the linear setting as first seen in [39]. Consider using AA(m) to solve for the fixed point of $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $M \in \mathbb{R}^{n \times n}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Substituting this affine $\mathbf{q}$ into (2.14) results in the update

$$\mathbf{x}_{k+1} = M \left[\mathbf{x}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{x}_k - \mathbf{x}_{k-j}) \right] + \mathbf{b}. \quad (2.17)$$

The residuals corresponding to (2.17) are given by

$$\mathbf{r}_k = \mathbf{x}_k - \mathbf{q}(\mathbf{x}_k) = \mathbf{x}_k - M\mathbf{x}_k - \mathbf{b} = (I - M)\mathbf{x}_k - \mathbf{b}. \quad (2.18)$$

Combining (2.17) and (2.18) leads to the following residual update formula

$$\begin{aligned} \mathbf{r}_{k+1} &= (I - M)\mathbf{x}_{k+1} - \mathbf{b} \\ &= (I - M)M \left[\mathbf{x}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{x}_k - \mathbf{x}_{k-j}) \right] - M\mathbf{b} \\ &= M \left[(I - M)\mathbf{x}_k - \mathbf{b} + (I - M) \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{x}_k - \mathbf{x}_{k-j}) \right] \\ &= M \left[\mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right], \end{aligned} \quad (2.19)$$

where $\mathbf{r}_1 = M\mathbf{r}_0$ and $\mathbf{r}_0 = (I - M)\mathbf{x}_0 - \mathbf{b}$ for some initial iterate $\mathbf{x}_0$. The main tool to show convergence in this linear setting is the residual bound given by

$$\left\| \mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right\| \leq \|\mathbf{r}_k\|. \quad (2.20)$$

The bound in (2.20) is a result of the solution of the coefficients $\{\beta_j^{(k)}\}_{j=1}^{m_k}$ in optimization problem (2.15). Since these coefficients are chosen such that the left-hand side of (2.20) is minimized, and the fact taking $\beta_j^{(k)} = 0$ for all $j = 1, \dots, m_k$ results in the right-hand side, it is clear that (2.20) holds. Now, using (2.19) and (2.20) we can conclude that

$$\begin{aligned} \|\mathbf{r}_{k+1}\| &= \left\| M \left[\mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right] \right\| \\ &\leq \|M\| \left\| \mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right\| \\ &\leq \|M\| \|\mathbf{r}_k\|. \end{aligned}$$

Since $\mathbf{r}(\mathbf{x}^*) = \mathbf{x}^* - \mathbf{q}(\mathbf{x}^*) = \mathbf{0}$, the q-linear convergence according to Definition 2.3.1 of the residual iteration (2.19) is achieved if $\|M\| < 1$. As stated earlier, this convergence theory does not seem to suggest any improvement to the convergence rate of $\|M\|$ for the underlying fixed-point iteration.

Another paper of interest [29] establishes a per-iteration residual bound for AA that reveals the mechanism by which acceleration is achieved. It is unnecessary to present the exact bound presented in this work, but the following relationship is emphasized

$$\|\mathbf{r}_{k+1}\| \leq \mathcal{O}(\|\mathbf{r}_k\|^2) + \mathcal{O}(\|\mathbf{r}_k\| \|\mathbf{r}_{k-1}\|) + \dots + \mathcal{O}(\|\mathbf{r}_k\| \|\mathbf{r}_{k-m_k}\|).$$

One shortcoming here is that the established residual bound tells us nothing about the asymptotic behavior of AA. While the exact bound presented in [29] might suggest that AA will accelerate the convergence over the underlying fixed-point iteration, as we will see in the following chapter, there exists certain nontrivial linear maps that see no acceleration from AA.

It was shown [11, 33] that AA slows the convergence corresponding to superlinearly converging fixed-point iterations. This is a rather disappointing discovery as it implies that AA cannot be used to speed up quadratically converging Newton iterations. A further exploration of AA applied to Newtons method is studied in [6, 30].

Improved Convergence for Symmetric Maps

Recently, it was shown in [14] that $\text{AA}(m)$ has an improved asymptotic convergence factor over the underlying fixed-point iteration for affine $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ where M is a symmetric matrix. In addition, this linear result is extended to show that nonlinear functions $\mathbf{q}$ also see an improved convergence rate if they have a symmetric Jacobian at the solution. This convergence result marks the first convergence theory showing that $\text{AA}(m)$ asymptotically achieves the *acceleration* implied by its name. As such, it is important to discuss the techniques used and the applicability of the results if we wish to develop better results in the future. To begin, we first introduce the main theorem of [14].

Theorem 2.3.1. ([14], Theorem 1) Let $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $M \in \mathbb{R}^{n \times n}$ is symmetric with $\|M\| < 1$ and let $\lambda_{\max}(M)$, $\lambda_{\min}(M)$ denote the largest and smallest eigenvalue of M respectively. Then, for any initial iterate $\mathbf{x}_0$, $\text{AA}(m)$ converges to the fixed point of $\mathbf{q}$ with r -linear convergence rate bounded by $\sqrt{w_0} \|M\|$, where

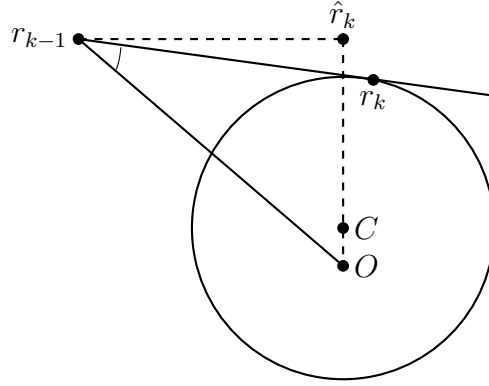
$$w_0 = \sup_{a \geq 0} \sin \left[\sin^{-1} \left(\frac{|\lambda_{\max}(M) - \lambda_{\min}(M)|a}{\sqrt{4 + (2 - \lambda_{\max}(M) - \lambda_{\min}(M))^2 a^2}} \right) + \left| \tan^{-1}(a) - \tan^{-1} \left(\frac{a}{2} (2 - \lambda_{\max}(M) - \lambda_{\min}(M)) \right) \right| \right].$$

Specifically, the residuals $\mathbf{r}_k$ generated by $\text{AA}(m)$ satisfy

$$\frac{\|\mathbf{r}_{k+1}\|}{\|\mathbf{r}_{k-1}\|} \leq w_0 \|M\|, \quad \text{for all } k \geq 2,$$

and w_0 is bounded above by $\|M\|$ with $w_0 = \|M\|$ if and only if $\lambda_{\max}(M) = -\lambda_{\min}(M)$.

The basic idea behind the proof of Theorem 2.3.1 is to take advantage of the symmetry of M by using a geometric interpretation of the $\text{AA}(m)$ iteration. More specifically, the bound shown in Theorem 2.3.1 is the maximization of the angle between $\mathbf{r}_{k-1}$ and $\mathbf{r}_k$, where $\mathbf{r}_k$ lies in the circle centered as $C = (\lambda_{\max}(M) + \lambda_{\min}(M))\hat{\mathbf{r}}_k/2$ with radius $(\lambda_{\max}(M) - \lambda_{\min}(M))\|\hat{\mathbf{r}}_k\|/2$, with $\hat{\mathbf{r}}_k$ taken to be the minimum norm affine combination of $\mathbf{r}_{k-m-1}, \dots, \mathbf{r}_{k-1}$ such that $\mathbf{r}_k = M\hat{\mathbf{r}}_k$ according to (2.19). At this point, interested readers are encouraged to visit [14] for the derivation of w_0 and the proof that $w_0 \leq \|M\|$.



In the next chapter, we explore the convergence rate of AA in more detail, with a special focus given to the linear setting. There, exact convergence rates will be given for a certain class of linear maps.

Chapter 3

Asymptotic Convergence Rates

In general, determining the r-linear convergence factor or *asymptotic convergence factor* of an AA iteration is extremely difficult. Despite the wide use of AA in practice, there have been almost no convergence results that establish an improved asymptotic convergence rate for AA over the underlying fixed-point iteration. This chapter will largely focus on analyzing this convergence factor problem, particularly in the linear setting. That is, we will focus on the asymptotic convergence factor for AA applied to solve for the fixed point $\mathbf{x}^* \in \mathbb{R}^n$ of an affine function $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Let $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $M \in \mathbb{R}^{n \times n}$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Assume $A = I - M$ is invertible so that $\mathbf{q}$ has a unique fixed point given by

$$\mathbf{x}^* = (I - M)^{-1}\mathbf{b} = A^{-1}\mathbf{b}. \quad (3.1)$$

If we consider applying AA(m) to $\mathbf{q}$ for some initial iterate $\mathbf{x}_0$ to solve for $\mathbf{x}^*$, then recall from the previous chapter that we have the iteration given by

$$\mathbf{x}_{k+1} = M \left[\mathbf{x}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{x}_k - \mathbf{x}_{k-j}) \right] + \mathbf{b}$$

and residual iteration given by

$$\mathbf{r}_{k+1} = M \left[\mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right]$$

as first shown in (2.17) and (2.19) respectively, where the acceleration coefficients are chosen to minimize the norm

$$\min_{\{\beta_j^{(k)}\}} \left\| \mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right\|^2.$$

Remark 3.0.1. Our assumption that A is invertible also allows us to express the iterates $\mathbf{x}_k$ in terms of the residuals $\mathbf{r}_k$, $k = 0, 1, 2, \dots$. Since $\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b}$, invertibility of A implies that $\mathbf{x}_k = A^{-1}(\mathbf{r}_k + \mathbf{b})$. Therefore, in this chapter and much of the rest of this thesis, we largely focus on the residual iteration as this more naturally fits with the solution of the acceleration coefficients as well as removes the need to consider the vector $\mathbf{b}$ in our iteration.

We are interested in determining the convergence factor associated with the residual sequence generated by applying AA(m) to linear problem (2.19) for some initial residual $\mathbf{r}_0 \in \mathbb{R}^n$. Recall from Definition 2.3.2 that the r -linear convergence factor or asymptotic convergence factor for a given initial residual $\mathbf{r}_0$ is given by

$$\rho(\mathbf{r}_0) := \rho_{\{\mathbf{r}_k\}} = \limsup_{k \rightarrow \infty} \|\mathbf{r}_k\|^{1/k}.$$

In the case of AA(m), this asymptotic convergence factor is strongly dependent on the initial residual. Therefore, we are often interested in the *worst-case r -linear convergence factor* given by

$$\rho^* = \sup_{\mathbf{r}_0 \in \mathbb{R}^n} \rho(\mathbf{r}_0), \quad (3.2)$$

which represents the worst-performing convergence behaviour of AA applied to the problem among all initial residuals. In this chapter, we will also frequently give special attention to AA(1). Taking $m = 1$ in (2.18) we have the residual update

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})]. \quad (3.3)$$

Note that in this setting, there is a single acceleration coefficient β_k to solve for. If $\mathbf{r}_k \neq \mathbf{r}_{k-1}$, then β_k has unique solution given by

$$\beta_k = -\frac{\mathbf{r}_k^T(\mathbf{r}_k - \mathbf{r}_{k-1})}{\|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2}. \quad (3.4)$$

If we instead have $\mathbf{r}_k = \mathbf{r}_{k-1}$, then our iteration has begun to stagnate, which we explore in more detail in Chapter 4. Notably, if stagnation occurs here, then the choice for β_k in (3.3) is irrelevant and we can thus use any value desired (throughout this thesis we often take $\beta_k := 0$ in this case according to [7]).

3.1 Investigation of a Simple Linear Setting

To begin our investigation, it may be helpful to first introduce a simple example to get a better idea of the problem of determining the asymptotic convergence factor. Consider applying AA(1) to solve for the fixed point $\mathbf{x}^*$ of

$$\mathbf{q}(\mathbf{x}) = M\mathbf{x} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \mathbf{x}, \quad \text{where } \lambda_1, \lambda_2 < 1, \quad (3.5)$$

which is obviously given by $\mathbf{x}^* = \mathbf{0}$. Given an initial iterate $\mathbf{x}_0$, our initial residual is given by

$$\mathbf{r}_0 = (I - M)\mathbf{x}_0 = \begin{bmatrix} 1 - \lambda_1 & 0 \\ 0 & 1 - \lambda_2 \end{bmatrix} \mathbf{x}_0.$$

Our next residual is then given by $\mathbf{r}_1 = M\mathbf{r}_0$, from which point we can begin iterating (3.3) using the solution for β_k given by (3.4). Even in this simplified setting, developing an exact convergence theory for AA(1) is challenging. The nonlinearity introduced by β_k leads to complicated behaviour. To see this, Fig. 3.1 highlights the oscillatory behaviour of the acceleration coefficients and approximate convergence factors as the iteration count increases for a specific example. Fig. 3.2 illustrates the dependence of the asymptotic convergence factor on the initial iterate and highlights the initial iterates that result in the worst-case asymptotic convergence of (3.5). In this case, we have $\rho^* \approx 0.5074$, which represents an improvement in the asymptotic convergence of the underlying fixed-point iteration given by $\rho^{\text{FP}} \approx \|M\| = 0.7$. Notably, Fig. 3.2 suggests there is some symmetry in the convergence of AA(1) applied to (3.5). This observation can be explained by the fact that the asymptotic convergence factor is scaling invariant, as demonstrated by the following proposition.

Proposition 3.1.1. Consider applying AA(m) for any initial residual $\mathbf{r}_0$ to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $M \in \mathbb{R}^{n \times n}$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. If the iteration converges, then

$$\rho(\mathbf{r}_0) = \rho(\alpha\mathbf{r}_0),$$

for all $\alpha \in \mathbb{R} \setminus \{0\}$.

Proof. Consider applying AA(m) to $\mathbf{q}$ with initial residual $\tilde{\mathbf{r}}_0$ where $\tilde{\mathbf{r}}_0 = \alpha\mathbf{r}_0$. Then $\tilde{\mathbf{r}}_1 = M\tilde{\mathbf{r}}_0 = \alpha M\mathbf{r}_0 = \alpha\mathbf{r}_1$. If we have $\tilde{\mathbf{r}}_j = \alpha\mathbf{r}_j$ for all $k - m_k \leq j \leq k$, then using (2.18) we see

$$\tilde{\mathbf{r}}_{k+1} = M \left[\tilde{\mathbf{r}}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\tilde{\mathbf{r}}_k - \tilde{\mathbf{r}}_{k-j}) \right] = \alpha M \left[\mathbf{r}_k + \sum_{j=1}^{m_k} \beta_j^{(k)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right] = \alpha\mathbf{r}_{k+1}.$$

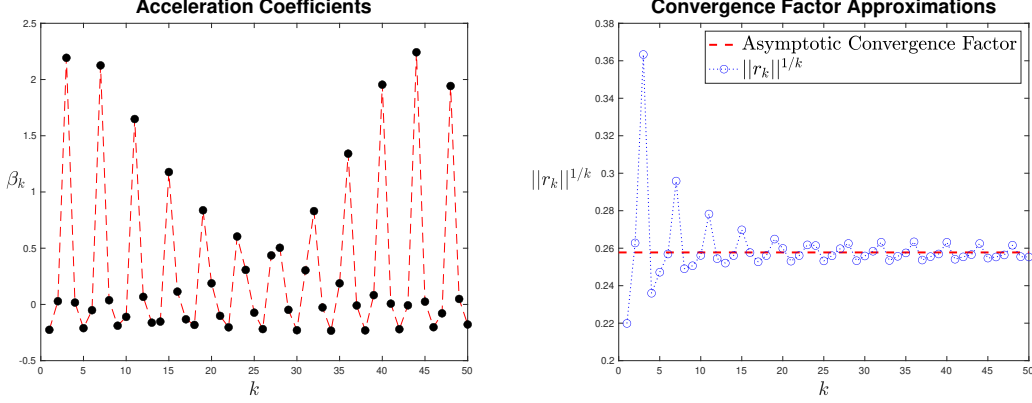


Figure 3.1: Acceleration coefficients β_k as a function of iteration number (left) and the approximate convergence factor $\|\mathbf{r}_k\|^{1/k}$ as a function of iteration number (right) for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$, and $\mathbf{r}_0 = (-0.1285, 0.6687)^T$.

It follows inductively that $\tilde{\mathbf{r}}_k = \alpha \mathbf{r}_k$ for all $k \geq 0$ and therefore

$$\rho(\tilde{\mathbf{r}}_0) = \limsup_{k \rightarrow \infty} \|\tilde{\mathbf{r}}_k\|^{1/k} = \limsup_{k \rightarrow \infty} \|\alpha \mathbf{r}_k\|^{1/k} = \limsup_{k \rightarrow \infty} |\alpha|^{1/k} \|\mathbf{r}_k\|^{1/k} = \rho(\mathbf{r}_0),$$

as desired. $\square$

Fig. 3.3 illustrates a few key areas that are interesting for analysis. Perhaps most notably, the bound given by Theorem 2.3.1 shown in yellow indeed improves upon the convergence rate of the underlying fixed-point iteration. Comparing this bound with the worst-case AA(1) iterations shown in blue highlights a sizable difference between the improved bound of [14] and the true worst-case asymptotic convergence factor. Another interesting observation is that the case where $\lambda_1 = \lambda_2$ seems to suggest that the worst-case convergence factor is zero, meaning superlinear convergence is achieved. In particular we can actually conclude that this case results in $\mathbf{r}_2 = \mathbf{0}$, meaning exact convergence is achieved in just two iterations. In fact, this can be shown to be true for scalar matrices in general as the following proposition asserts.

Proposition 3.1.2. Consider applying AA(m) for any initial residual $\mathbf{r}_0$ to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$, M is a scalar matrix (i.e., $M = \lambda I$ for some $\lambda \neq 0$), and $\mathbf{r}_1 = M\mathbf{r}_0$. Then, we have $\mathbf{r}_2 = \mathbf{0}$.

Proof. Let $\mathbf{r}_0$ be some initial residual and $\mathbf{r}_1 = M\mathbf{r}_0 = \lambda\mathbf{r}_0$. Then β_1 is chosen to minimize

$$\|\mathbf{r}_1 + \beta_1(\mathbf{r}_1 - \mathbf{r}_0)\|^2 = \|\lambda\mathbf{r}_0 + \beta_1(\lambda - 1)\mathbf{r}_0\|^2 = \|(\lambda + \beta_1(\lambda - 1))\mathbf{r}_0\|^2.$$

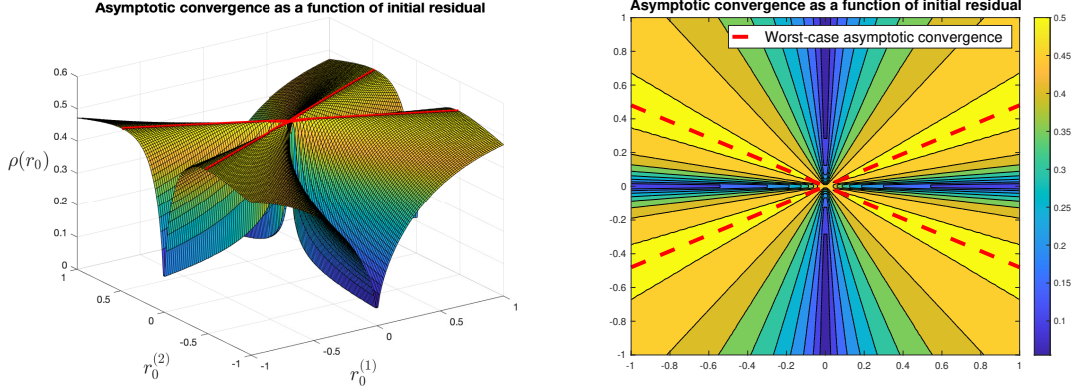


Figure 3.2: Asymptotic convergence factor as a function of initial residual $\mathbf{r}_0 \in [-1, 1]^2$ for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$. A 128×128 grid was used for the initial residuals and convergence factors were numerically found after generating the residuals according to (3.3) for each grid point.

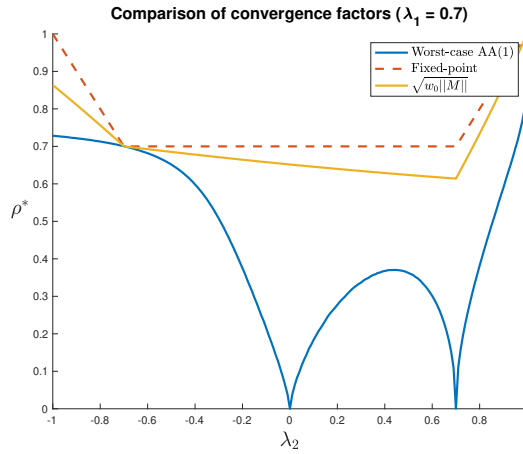


Figure 3.3: Worst-case asymptotic convergence factor for AA(1) (blue) applied to (3.5) as a function of λ_2 with λ_1 fixed compared with the associated convergence factor for the underlying fixed-point iteration (red) and the bound for the convergence factor (yellow) given by Theorem 2.3.1.

The above expression is minimized when

$$\beta_1 = -\frac{\lambda}{\lambda - 1},$$

resulting in the next residual satisfying

$$\begin{aligned} \mathbf{r}_2 &= M[\mathbf{r}_1 + \beta_1(\mathbf{r}_1 - \mathbf{r}_0)] \\ &= \lambda(\lambda + \beta_1(\lambda - 1))\mathbf{r}_0 \\ &= \lambda(0)\mathbf{r}_0 = \mathbf{0}, \end{aligned}$$

as desired. □

Another interesting thing to note about [Fig. 3.3](#) is the worst-case convergence factor when $\lambda_1 = -\lambda_2$, as it appears that AA(1) does not improve the convergence of the associated fixed-point iteration. As we will see in the next section, this case is especially interesting because AA(1) applied to this affine map exhibits a few properties that allow us to *exactly quantify the asymptotic convergence factor for any initial iterate*.

3.2 Asymptotic Convergence Factors for Certain Linear Maps

In this section, we will present two classes of affine maps, a symmetric and skew-symmetric class, for which we can determine the asymptotic convergence factor for any initial residual $\mathbf{r}_0$. Further, we will consider how the convergence analysis corresponding to these maps can be extended to larger systems and similar iterations. Specifically, consider using AA(1) to solve for the fixed point $\mathbf{x}^* \in \mathbb{R}^n$ of the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^2$ and $M = M_\lambda$ or $M = W_\lambda$ with

$$M_\lambda = \begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix} \tag{3.6a}$$

and

$$W_\lambda = \begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix}, \tag{3.6b}$$

where $\lambda \in \mathbb{R} \setminus \{0, 1\}$. The symmetric matrix M_λ in [\(3.6a\)](#) represents the case identified in [Fig. 3.3](#) where the worst-case convergence factor for AA(1) appears to coincide with the convergence rate of the underlying fixed-point iteration. The skew-symmetric matrix W_λ

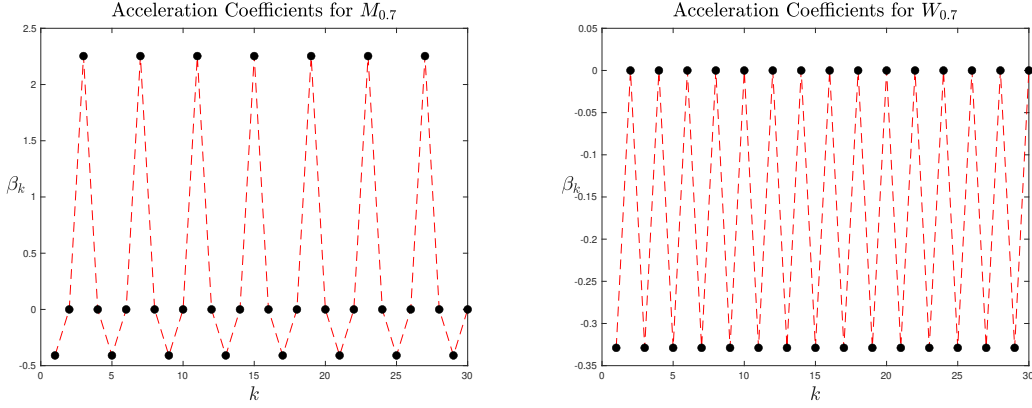


Figure 3.4: Acceleration coefficients β_k as a function of iteration number for AA(1) applied to affine function with $M = M_{0.7}$ (left) and $M = W_{0.7}$ (right).

in (3.6b) contains the same values along the skew-diagonal as the diagonal in M_λ . It is also important to notice that the eigenvalues of M_λ are λ and $-\lambda$ while the eigenvalues of W_λ are $i\lambda$ and $-i\lambda$. A useful property that M_λ and W_λ share is that

$$\|M_\lambda \mathbf{v}\|^2 = \lambda^2 v_1^2 + (-\lambda)^2 v_2^2 = \lambda^2 \|\mathbf{v}\|^2 = \lambda^2 v_2^2 + (-\lambda)^2 v_1^2 = \|W_\lambda \mathbf{v}\|^2, \quad (3.7)$$

for all $\mathbf{v} = (v_1, v_2)^T \in \mathbb{R}^2$. We may occasionally need the above property since it allows us to split up the norm of the following residual

$$\|\mathbf{r}_{k+1}\|^2 = \|M\|^2 \|\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})\|^2 = \lambda^2 \|\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})\|^2. \quad (3.8)$$

After applying AA(1) to affine $\mathbf{q}$ with $M = M_\lambda$ or $M = W_\lambda$, a few key observations can be made about the iteration. In particular, the acceleration coefficients appear to follow a specific pattern with zero acceleration being provided every other iteration ($\beta_{2j} = 0$ for all $j \geq 1$). Fig. 3.4 highlights this observation for each map by taking $\lambda = 0.7$ and displaying the acceleration coefficients corresponding to a random initial residual. One visible difference that Fig. 3.4 illustrates is that the coefficients corresponding to W_λ appear to alternate between two distinct values while the coefficients corresponding to M_λ cycle between 4 points. Additionally these plots clearly show that the coefficients for both maps take the value zero every other iteration. This behaviour can be proved theoretically. To see this, we first introduce the following lemma.

Lemma 3.2.1. The AA(1) iteration (3.3) applied to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^2$ and $M = M_\lambda$ (3.6a) or $M = W_\lambda$ (3.6b) obeys the residual relation

$$\|\mathbf{r}_{k+1}\|^2 = \mathbf{r}_{k+1}^T M \mathbf{r}_{k-1}, \quad (3.9)$$

for all $k \geq 1$.

Proof. Let $k \geq 1$. From the unique solution (3.4) for β_k , we have

$$(\mathbf{r}_k - \mathbf{r}_{k-1})^T [\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] = 0. \quad (3.10)$$

Taking M to be M_λ or W_λ from (3.6a) or (3.6b) respectively, it is clear that $M^T M = \lambda^2 I_2$, where I_2 is the 2×2 identity matrix. Thus, using (3.10), (3.8) and residual update (3.3), we have

$$\begin{aligned} \|\mathbf{r}_{k+1}\|^2 &= \lambda^2 [\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})]^T [\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] \\ &= \lambda^2 \mathbf{r}_k^T [\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] = \mathbf{r}_{k+1}^T M \mathbf{r}_k. \end{aligned}$$

Now, since

$$\mathbf{r}_{k+1}^T M \mathbf{r}_k = \|\mathbf{r}_{k+1}\|^2 - \beta_k \mathbf{r}_{k+1}^T M (\mathbf{r}_k - \mathbf{r}_{k-1}),$$

it follows that $\beta_k \mathbf{r}_{k+1}^T M (\mathbf{r}_k - \mathbf{r}_{k-1}) = 0$. Therefore

$$\mathbf{r}_{k+1}^T M \mathbf{r}_{k-1} = \mathbf{r}_{k+1}^T M \mathbf{r}_k = \|\mathbf{r}_{k+1}\|^2$$

as desired. $\square$

The relation in Lemma 3.2.1 now allows us to prove the peculiar property of these iterations naturally providing zero acceleration on alternating steps. That is, we now have the following result.

Proposition 3.2.1. The AA(1) iteration (3.3) applied to linear functions (3.6a) or (3.6b) for any initial residual $\mathbf{r}_0$ has $\beta_{2k} = 0$ for all $k \geq 1$.

Proof. Let M be given by M_λ or W_λ from (3.6a) or (3.6b) respectively. We will proceed by induction. Start by assuming that $\mathbf{r}_k = M \mathbf{r}_{k-1}$, as this corresponds to $\beta_{k-1} = 0$. In the base case ($k = 1$) however, this β_0 coefficient is not present but can analogously taken to be 0. From Lemma 3.2.1, it follows that

$$\|\mathbf{r}_{k+1}\|^2 = \mathbf{r}_{k+1}^T M \mathbf{r}_{k-1} = \mathbf{r}_{k+1}^T \mathbf{r}_k,$$

meaning,

$$-\mathbf{r}_{k+1}^T (\mathbf{r}_{k+1} - \mathbf{r}_k) = 0. \quad (3.11)$$

Then, if $\mathbf{r}_{k+1} \neq \mathbf{r}_k$, using the unique solution for our coefficients given by (3.4) with (3.11) results in $\beta_{k+1} = 0$. If $\mathbf{r}_{k+1} = \mathbf{r}_k$, recall that we set $\beta_{k+1} = 0$ (we can choose anything for β_{k+1} since the vector it is multiplied by will be the zero vector). Consequently, the next iteration is determined by the rule $\mathbf{r}_{k+2} = M \mathbf{r}_{k+1}$, and thus we have our result. $\square$

3.2.1 Convergence for W_λ

From this point, we will inspect the convergence of M_λ and W_λ under AA(1) separately as they require more individual analysis. To start, we will introduce the convergence rate for W_λ .

Proposition 3.2.2. Consider AA(1) applied to $\mathbf{q}(\mathbf{x}) = W_\lambda \mathbf{x} + \mathbf{b}$ for some $\mathbf{b} \in \mathbb{R}^2$ and any initial residual $\mathbf{r}_0 \in \mathbb{R}^2$. Then, the coefficients β_k follow the 2-periodic behaviour

$$\begin{cases} \beta_{2k} &= 0 \\ \beta_{2k+1} &= -\frac{\|W_\lambda\|^2}{1+\|W_\lambda\|^2} \end{cases} \quad (3.12)$$

for all $k \geq 0$, and the r-linear convergence factor is given by

$$\rho_{W_\lambda} = \frac{\|W_\lambda\|}{(1 + \|W_\lambda\|^2)^{1/4}}.$$

Proof. As shown in [Proposition 3.2.1](#), applying AA(1) to (3.6b) for any initial iterate results in $\beta_{2k} = 0$ and thus $\mathbf{r}_{2k+1} = W_\lambda \mathbf{r}_{2k}$ for all $k \geq 0$. Since

$$\mathbf{r}_{2k+1}^T \mathbf{r}_{2k} = \mathbf{r}_{2k}^T W_\lambda^T \mathbf{r}_{2k} = -\lambda \mathbf{r}_{2k}^{(2)} \mathbf{r}_{2k}^{(1)} + \lambda \mathbf{r}_{2k}^{(2)} \mathbf{r}_{2k}^{(1)} = 0, \quad (3.13)$$

it follows that

$$\beta_{2k+1} = -\frac{\|\mathbf{r}_{2k+1}\|^2}{\|\mathbf{r}_{2k+1}\|^2 + \|\mathbf{r}_{2k}\|^2} = -\frac{\|W_\lambda\|^2 \|\mathbf{r}_{2k}\|^2}{\|W_\lambda\|^2 \|\mathbf{r}_{2k}\|^2 + \|\mathbf{r}_{2k}\|^2} = -\frac{\|W_\lambda\|^2}{\|W_\lambda\|^2 + 1}, \quad (3.14)$$

as shown in (3.12). The r-linear convergence factor ρ_{W_λ} will be uncovered by considering the relative magnitudes of consecutive residuals which we denote by $y_k := \|\mathbf{r}_k\| / \|\mathbf{r}_{k-1}\|$. Since $\mathbf{r}_{2k+1} = W_\lambda \mathbf{r}_{2k}$, we see $y_{2k+1} = \|W_\lambda\|$ for all $k \geq 0$. Using (3.14), (3.10), and (3.13), we have

$$\begin{aligned} y_{2k+2}^2 &= \frac{\|\mathbf{r}_{2k+2}\|^2}{\|\mathbf{r}_{2k+1}\|^2} = \frac{\|W_\lambda\|^2 \|\mathbf{r}_{2k+1} + \beta_{2k+1}(\mathbf{r}_{2k+1} - \mathbf{r}_{2k})\|^2}{\|\mathbf{r}_{2k+1}\|^2} \\ &= \frac{\|W_\lambda\|^2 (\|\mathbf{r}_{2k+1}\|^2 + \beta_{2k+1}^2 \|\mathbf{r}_{2k+1} - \mathbf{r}_{2k}\|^2)}{\|\mathbf{r}_{2k+1}\|^2} \\ &= \frac{\|W_\lambda\|^2 (\|\mathbf{r}_{2k+1}\|^2 + \beta_{2k+1}^2 \|\mathbf{r}_{2k+1}\|^2)}{\|\mathbf{r}_{2k+1}\|^2} = \frac{\|W_\lambda\|^2}{1 + \|W_\lambda\|^2} \end{aligned}$$

From the periodic behaviour present in the acceleration coefficients we then have

$$\begin{cases} y_{2k-1} &= \|W_\lambda\| \\ y_{2k} &= \frac{\|W_\lambda\|}{\sqrt{1+\|W_\lambda\|^2}}, \end{cases} \quad \text{for all } k \geq 1. \quad (3.15)$$

Now, applying a telescoping procedure we get the desired convergence rate.

$$\begin{aligned} \rho_{W_\lambda} &= \lim_{k \rightarrow \infty} \|\mathbf{r}_k\|^{1/k} = \lim_{k \rightarrow \infty} \frac{\|\mathbf{r}_k\|^{1/k}}{\|\mathbf{r}_0\|^{1/k}} = \lim_{k \rightarrow \infty} \left(\prod_{j=0}^{k/2-1} \frac{\|\mathbf{r}_{2j+2}\|}{\|\mathbf{r}_{2j}\|} \right)^{1/k} \\ &= \lim_{k \rightarrow \infty} \left(\prod_{j=0}^{k/2-1} y_{2j+1} y_{2j+2} \right)^{1/k} = \lim_{k \rightarrow \infty} \left(\prod_{j=0}^{k/2-1} \frac{\|W_\lambda\|^2}{\sqrt{1+\|W_\lambda\|^2}} \right)^{1/k} \\ &= \lim_{k \rightarrow \infty} \left(\left[\frac{\|W_\lambda\|^2}{\sqrt{1+\|W_\lambda\|^2}} \right]^{k/2} \right)^{1/k} = \frac{\|W_\lambda\|}{(1+\|W_\lambda\|^2)^{1/4}}. \end{aligned}$$

□

The above result indicates the peculiar property of these skew symmetric matrices possessing AA(1) convergence factors that are independent of the initial iterate. This is a rather unusual property as typical applications of AA show convergence rates that are highly dependent on the initial residual, as illustrated in examples such as [Fig. 3.1](#). An interesting consequence of this independence on the initial residual is that the skew symmetric matrix of [\(3.6b\)](#) *always* improves upon the fixed-point iteration since

$$\rho_{W_\lambda} = \frac{\|W_\lambda\|}{(1+\|W_\lambda\|^2)^{1/4}} < \|W_\lambda\| = \lambda.$$

Furthermore, the fixed-point iteration converges only if the iteration function $\mathbf{q}(\mathbf{x})$ is contractive which implies $\text{spectral radius}(M) < 1$ in the linear case but AA can still see convergence for non-contractive maps as discussed in [\[29\]](#). For maps of the form [\(3.6b\)](#), we can have non-contractive maps with $\|W_\lambda\| > 1$ and still see convergence of AA(1) despite the fixed-point iteration failing to converge. Specifically, by solving for when $\rho_{W_\lambda} < 1$ we see that we instead require

$$\lambda = \|W_\lambda\| < \sqrt{\frac{1+\sqrt{5}}{2}} \approx 1.272$$

for convergence of AA(1).

3.2.2 Convergence of M_λ

The pattern in the coefficients described by [Proposition 3.2.1](#) results in the iteration being identical to the “restarted” Anderson acceleration variation analyzed in [\[?\]](#). Denoted by rAA(1), this restarted variant *enforces* no acceleration every other iteration by using the two-step update

$$\mathbf{x}_k = \mathbf{q}(\mathbf{x}_{k-1}), \quad (3.16)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})), \quad (3.17)$$

where the coefficient β_k is determined by the same optimization problem [\(2.15\)](#) with unique solution [\(3.4\)](#). In fact, the following proposition outlines how we can use the results of [\[8\]](#) to help establish the convergence of AA(1) applied to linear iteration [\(3.6a\)](#).

Proposition 3.2.3. Consider AA(1) applied to M_λ as given by [\(3.6a\)](#) and set $\varepsilon := r_0^{(2)}/r_0^{(1)}$ for initial residual $\mathbf{r}_0 = (r_0^{(1)}, r_0^{(2)})^T$. Then, the r-linear convergence factor is given by

$$\rho_{M_\lambda}(\mathbf{r}_0) = \left[\frac{4\varepsilon^2\lambda^4}{\varepsilon^4(1+\lambda)^2 + 2\varepsilon^2(\lambda^2+1) + (1-\lambda)^2} \right]^{1/4}. \quad (3.18)$$

Furthermore, the *worst-case* asymptotic convergence factor among all initial residuals $\mathbf{r}_0$ is given by $\rho_{M_\lambda}^* = \|M_\lambda\| = |\lambda|$.

Proof. For AA(1) applied to [\(3.6a\)](#), since according to [Proposition 3.2.1](#) every other acceleration weight is zero, we can write the iteration in the two-step form of [\(3.16\)](#), [\(3.17\)](#). That is, the AA(1) and rAA(1) iterations follow the same update rule when applied to [\(3.6a\)](#) and thus generate the same iterates when beginning from the same initial iterate. As a consequence, we can simply apply the r-linear convergence result of [\[8, Theorem 4.3\]](#) for rAA(1) applied to 2×2 symmetric linear systems to our problem leading to [\(3.18\)](#) as stated. Further, as established in [\[8, Lemma 4.4\]](#), we can maximize this asymptotic convergence factor with respect to ε to obtain

$$\varepsilon^* = \pm \sqrt{\frac{|-\lambda(1-\lambda)|}{|\lambda(1+\lambda)|}} = \pm \sqrt{\frac{|1-\lambda|}{|1+\lambda|}}, \quad (3.19)$$

which gives $\rho_{M_\lambda}^* = |\lambda|$ when substituted into [\(3.18\)](#). □

The above proposition reveals an interesting quality of map (3.6a) when AA(1) is applied. The worst-performing initial iterate leads to no improvement or acceleration over the underlying fixed-point iteration. In fact, taking $\varepsilon = \varepsilon^*$ according to (3.19), it can be seen that $\beta_k = 0$ for all $k \geq 1$ and thus the regular fixed-point iteration (2.7) emerges. With this discovery in mind, we now have a class of maps that *do not* see any improvement in the worst case when AA(1) is applied. This is undesirable of an *acceleration* technique but the asymptotic convergence factor (3.18) here also highlights calculable acceleration for all other initiations. The next result highlights a relationship between AA(m) iterations that will allow us to generalize the convergence given in Proposition 3.2.3 to a larger class of symmetric matrices.

Proposition 3.2.4. Let $S \in \mathbb{R}^{n \times n}$ be a symmetric matrix with eigendecomposition

$$S = Q\Sigma Q^T,$$

for diagonal $\Sigma \in \mathbb{R}^{n \times n}$, orthogonal $Q \in \mathbb{R}^{n \times n}$, and let ρ_S, ρ_Σ be the r-linear convergence factors associated with AA(m) applied to linear iterations $\mathbf{q}_1(\mathbf{x}) = S\mathbf{x} + \mathbf{b}_1$, $\mathbf{q}_2(\mathbf{x}) = \Sigma\mathbf{x} + \mathbf{b}_2$ for some $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^n$ respectively. Then, ρ_S satisfies

$$\rho_S(Q\mathbf{r}_0) = \rho_\Sigma(\mathbf{r}_0)$$

for initial residual $\mathbf{r}_0 \in \mathbb{R}^n$ assuming the AA(m) iteration converges for Σ .

Proof. To begin, it is helpful to distinguish between the residuals produced by AA(m) applied to Σ and S by $\{\mathbf{r}_k^\Sigma\}$ and $\{\mathbf{r}_k^S\}$ respectively. Setting $\mathbf{r}_0^S := Q\mathbf{r}_0^\Sigma$, we have

$$\mathbf{r}_1^S = S\mathbf{r}_0^S = Q\Sigma Q^T(\mathbf{r}_0^S) = Q\Sigma Q^T(Q\mathbf{r}_0^\Sigma) = Q\Sigma\mathbf{r}_0^\Sigma = Q\mathbf{r}_1^\Sigma.$$

We now proceed by induction. Assume $\mathbf{r}_k^S = Q\mathbf{r}_k^\Sigma$ for $0 \leq k \leq K$. Now, using the fact that $\mathbf{r}_k^S = Q\mathbf{r}_k^\Sigma$ for $K - m_K \leq k \leq K$ and plugging these residuals into (2.15), by the norm-preserving property of Q , the acceleration coefficients $\{\beta_j^{(K)}\}$ used to compute $\mathbf{r}_{K+1}^S$ and $\mathbf{r}_{K+1}^\Sigma$ are the same. Our next residual is thus

$$\begin{aligned} \mathbf{r}_{K+1}^S &= S\mathbf{r}_K^S + \sum_{j=1}^{m_K} \beta_j^{(K)} S(\mathbf{r}_K^S - \mathbf{r}_{K-j}^S) \\ &= Q\Sigma Q^T(Q\mathbf{r}_K^\Sigma) + \sum_{j=1}^{m_K} \beta_j^{(K)} Q\Sigma Q^T(Q\mathbf{r}_K^\Sigma - Q\mathbf{r}_{K-j}^\Sigma) \\ &= Q \left(\Sigma\mathbf{r}_K^\Sigma + \sum_{j=1}^{m_K} \beta_j^{(K)} \Sigma(\mathbf{r}_K^\Sigma - \mathbf{r}_{K-j}^\Sigma) \right) \\ &= Q\mathbf{r}_{K+1}^\Sigma. \end{aligned}$$

Now, since

$$\|\mathbf{r}_k^S\| = \|Q\mathbf{r}_k^\Sigma\| = \|\mathbf{r}_k^\Sigma\|, \quad \text{for all } k \geq 0,$$

and by the asymptotic convergence factor definition [Definition 2.3.2](#), we have our result. $\square$

It should be noted that [Proposition 3.2.4](#) can be stated in a more general form for any two similar matrices $A, B \in \mathbb{R}^{n \times n}$ related by the similarity transform $A = QBQ^T$ for orthogonal $Q \in \mathbb{R}^{n \times n}$. That is, it is not required for this result to strictly relate the asymptotic convergence for a symmetric matrix with the diagonal matrix of its eigenvalues.

The following Corollary uses [Proposition 3.2.4](#) to expand this *no-acceleration* observation to a larger class of symmetric matrices, further cementing the idea that AA(1) does not always result in acceleration.

Corollary 3.2.1. Let $S \in \mathbb{R}^{n \times n}$ be symmetric with eigenvalues

$$\lambda_1 = -\lambda_2 \text{ with } 1 > |\lambda_1| = |\lambda_2| > |\lambda_3| \geq |\lambda_4| \geq \dots \geq |\lambda_n|,$$

and let $\rho_S(\mathbf{r}_0)$ be the AA(1) asymptotic convergence factor associated with linear map $\mathbf{q}_1(\mathbf{x}) = S\mathbf{x} + \mathbf{b}_1$ for some $\mathbf{r}_0, \mathbf{b}_1 \in \mathbb{R}^n$. Then, the worst-case AA(1) asymptotic convergence factor over all initial residuals is given by $\rho_S^* = \|S\|$.

Proof. The eigendecomposition of S may be written as

$$S = Q\Sigma Q^T, \quad \text{where } \Sigma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

for some orthogonal $Q \in \mathbb{R}^{n \times n}$. From [Proposition 3.2.4](#), we have $\rho_S(Q\mathbf{r}_0) = \rho_\Sigma(\mathbf{r}_0)$ so it is sufficient to shift our focus to the worst asymptotic convergence of AA(1) applied to linear map $\mathbf{q}_2(\mathbf{x}) = \Sigma\mathbf{x} + \mathbf{b}_2$ for some $\mathbf{b}_2 \in \mathbb{R}^n$. Now, by “zeroing out” components of our initial residual that do not correspond to the M_λ block of Σ we can recover the asymptotic convergence associated with M_λ given by [\(3.18\)](#). In particular, choosing

$$\mathbf{r}_0^* = \alpha \begin{bmatrix} 1 \\ \pm \sqrt{\frac{1-\lambda_1}{1-\lambda_2}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{for some } \alpha \in \mathbb{R} \setminus \{0\}, \quad (3.20)$$

in accordance with [\(3.19\)](#) leads to $\rho_\Sigma(\mathbf{r}_0^*) = \|\Sigma\|$. From [Proposition 3.2.4](#), we then see $\rho_S(Q\mathbf{r}_0^*) = \|\Sigma\| = \|S\|$. According to [\[39\]](#), AA does not perform worse than the underlying fixed-point iteration and thus since $\rho_{\text{FP}} = \|S\|$, we have our result. $\square$

Chapter 4

Stagnation

The goal of this chapter is to analyze the presence of *stagnation* in AA iterations with a finite window size ($m < \infty$). The term stagnation describes the scenario in which consecutive iterates are equal, showing no advancement of the iteration to the solution. The problem of stagnation for AA(∞) with affine $\mathbf{q}$ has been analyzed in [31] using the relationship with GMRES discussed in Chapter 2. Specifically, it was found that stagnation of GMRES corresponds to stagnation of AA. One key difference between these methods is exposed with the presence of stagnation however. If stagnation occurs, GMRES will continue to generate larger Krylov subspaces until it eventually converges exactly to the solution of (2.6), while AA (with $m = \infty$) will continue to stagnate indefinitely, making no progress towards the solution of (2.6).

Several authors have considered the problem of stagnation for GMRES [41, 19, 37, 20]. A deeper analysis of stagnation for GMRES is presented in [20], where necessary and sufficient conditions are established for *complete stagnation* and *partial stagnation*. The work of [20] is built off of the more general work of [15] which characterizes the matrices A and initial residuals $\mathbf{r}_0$ resulting in any desired nonincreasing convergence curve for the normed residuals $\|\mathbf{r}_k\|$, $k = 0, 1, 2, \dots$. To our knowledge, general results concerning the problem of stagnation for AA(m) with $m < \infty$ in the linear and nonlinear setting have not been considered and so the goal of this work is to explore this problem and any threat it may pose to AA iterations in practice.

One term we use frequently in this chapter is defined below.

Definition 4.0.1. Given some fixed-point function $\mathbf{q}$ and initial iterate $\mathbf{x}_0$, we say AA displays *immediate stagnation* if $\mathbf{x}_1 = \mathbf{x}_2$.

Additionally, it will be occasionally helpful to consider the following matrix-vector reformulation of linear residual iteration (2.18). We may define

$$\boldsymbol{\beta}_k = \begin{bmatrix} \beta_k^{(1)} \\ \vdots \\ \beta_k^{(m_k)} \end{bmatrix} \quad \text{and} \quad R_k = \begin{bmatrix} \mathbf{r}_k - \mathbf{r}_{k-1} & \mathbf{r}_k - \mathbf{r}_{k-2} & \cdots & \mathbf{r}_k - \mathbf{r}_{k-m_k} \end{bmatrix}, \quad (4.1)$$

so that least squares problem (2.15) can take the equivalent form

$$\min_{\boldsymbol{\beta}} \|\mathbf{r}_k + R_k \boldsymbol{\beta}\|^2, \quad (4.2)$$

with solution satisfying the normal equations

$$R_k^T R_k \boldsymbol{\beta}_k = -R_k^T \mathbf{r}_k. \quad (4.3)$$

Meaning (2.18) takes the equivalent form

$$\mathbf{r}_{k+1} = M(\mathbf{r}_k + R_k \boldsymbol{\beta}_k).$$

Recall from Chapter 2, for a given $\mathbf{x}_0$ we have the following relationship between the iterates $\{\mathbf{x}_k^A\}_{k=0}^\infty$, $\{\mathbf{x}_k^G\}_{k=0}^\infty$ produced by AA and GMRES respectively:

$$\mathbf{x}_{k+1}^A = \mathbf{q}(\mathbf{x}_k^G) = M\mathbf{x}_k^G + \mathbf{b}$$

In [31], an emphasis is placed on the stagnation and convergence properties of AA(∞) with respect to the stagnation of GMRES and the grade of A with respect to $\mathbf{r}_0$. While the results of this chapter are instead concerned with the stagnation of AA(m) for some depth $0 < m < \infty$, the stagnation results of [31] are still of great importance in our analysis as the iterates produced by AA(m) and AA(∞) are the same during the initial period where $m_k = k$. Specifically, the authors define the number

$$\mu(A, \mathbf{x}_0) = \max\{k \in \mathbb{N} : \mathbf{x}_1^{AA} - \mathbf{x}_0, \mathbf{x}_2^{AA} - \mathbf{x}_0, \dots, \mathbf{x}_k^{AA} - \mathbf{x}_0 \text{ are linearly independent}\},$$

and compare it with the grade of A with respect to $\mathbf{r}_0$

$$\nu(A, \mathbf{r}_0) = \min\{k \in \mathbb{N} : \mathbf{r}_0, A\mathbf{r}_0, \dots, A^k \mathbf{r}_0 \text{ are linearly dependent}\}.$$

If $\mu < \nu$, then GMRES begins stagnating at index μ until it eventually converges exactly to $\mathbf{x}^*$, while AA(∞) will stagnate forever. A consequence of the relationship between AA and GMRES is that initial iterates that result in the immediate stagnation of GMRES

$(\mathbf{x}_0^G = \mathbf{x}_1^G \neq \mathbf{x}^*)$ also correspond to the immediate stagnation of AA ($\mathbf{x}_2^A = \mathbf{x}_1^A$). To see this, recall that GMRES updates according to the rule

$$\mathbf{x}_k^G = \mathbf{x}_0 + \boldsymbol{\gamma}_k, \quad \text{with } \boldsymbol{\gamma}_k = \arg \min_{\boldsymbol{\gamma} \in \mathcal{K}_k} \|A(\mathbf{x}_0 + \boldsymbol{\gamma}) - \mathbf{b}\|. \quad (4.4)$$

In the above update, the vector $\boldsymbol{\gamma}_k$ is chosen from the Krylov subspace

$$\mathcal{K}_k = \mathcal{K}_k(A, \mathbf{r}_0) = \text{span}\{\mathbf{r}_0, A\mathbf{r}_0, \dots, A^{k-1}\mathbf{r}_0\}, \quad (4.5)$$

where the initial residual is given by $\mathbf{r}_0 = A\mathbf{x}_0 - \mathbf{b}$. Using (4.4) and (4.5), the first GMRES iterate takes the form

$$\mathbf{x}_1^{\text{GMRES}} = \mathbf{x}_0 + \alpha \mathbf{r}_0, \quad \text{where } \alpha = -\frac{\mathbf{r}_0^T A \mathbf{r}_0}{\|A \mathbf{r}_0\|^2}. \quad (4.6)$$

It is clear from the above GMRES update that immediate stagnation is achieved when $\alpha = 0$, or alternatively, when the initial residual satisfies the orthogonality condition $\mathbf{r}_0 \perp A\mathbf{r}_0$. Now consider the first and second AA(m) iterates according to (2.17):

$$\begin{aligned} \mathbf{x}_1^A &= M\mathbf{x}_0 + \mathbf{b} \\ \mathbf{x}_2^A &= M[\mathbf{x}_1^A + \beta_1(\mathbf{x}_1^A - \mathbf{x}_0)] + \mathbf{b} \end{aligned}$$

It is then clear that we have immediate stagnation $\mathbf{x}_1^A = \mathbf{x}_2^A$ if $\beta_1 = -1$. Using (3.4) and the fact that $\mathbf{r}_1^A = M\mathbf{r}_0$, this value for β_1 can be achieved if

$$0 = \frac{\mathbf{r}_0^T (\mathbf{r}_1^A - \mathbf{r}_0)}{\|\mathbf{r}_1^A - \mathbf{r}_0\|^2} \implies \mathbf{r}_0^T A \mathbf{r}_0 = 0.$$

This once again highlights the relationship between GMRES and AA, confirming that linear AA iterations will see immediate stagnation if the same is true of GMRES. In particular, these iterations immediately stagnate if $\mathbf{r}_0 \perp A\mathbf{r}_0$. In the linear setting, we will use the residual update formula of (2.19) to present our results as our assumption that A is invertible implies the stagnation of residuals corresponds to the stagnation of iterates. The following corollary is a natural consequence of the stagnation results of [31] and the equivalence of AA(m) and AA(∞) for iterates $\mathbf{x}_k$, where $k \leq m + 1$.

Corollary 4.0.1. Consider applying AA(m) to affine fixed-point function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, where $M \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$, $A = I - M$ is invertible, and $\mathbf{r}_1 = M\mathbf{r}_0 \neq \mathbf{0}$. Then the residuals satisfy $\mathbf{r}_1 = \mathbf{r}_2 = \dots = \mathbf{r}_{m+1}$ if and only if $\mathbf{r}_0 \perp A\mathbf{r}_0$.

Proof. Theorem 1 of [31] is sufficient in proving the above corollary. $\square$

An important remark is that immediate stagnation as described by Corollary 4.0.1 is only possible if A is an indefinite matrix since the requirement that $\mathbf{r}_0 \perp A\mathbf{r}_0$ implies $\mathbf{r}_0^T A \mathbf{r}_0 = 0$.

4.1 Stagnation for General Problems

The problem of stagnation becomes more complicated when considering nonlinear iteration functions $\mathbf{q}(\mathbf{x})$ but there are still a few general conclusions we can make. To begin, we will state a property of the vector

$$\hat{\mathbf{r}}_{k+1} = \mathbf{r}_k + R_k \boldsymbol{\beta}_k, \quad (4.7)$$

which has minimum norm with respect to $\boldsymbol{\beta}_k$ as (4.2) shows.

Lemma 4.1.1. Let $\hat{\mathbf{r}}_{k+1}$ be defined as in (4.7) with general iteration function $\mathbf{q}$, then

$$\hat{\mathbf{r}}_{k+1} \perp \hat{\mathbf{r}}_{k+1} - \mathbf{r}_k.$$

Proof. From definition (4.7) and solution for $\boldsymbol{\beta}_k$ as given by (4.3), we have

$$[\hat{\mathbf{r}}_{k+1} - \mathbf{r}_k]^T \hat{\mathbf{r}}_{k+1} = [R_k \boldsymbol{\beta}_k]^T [\mathbf{r}_k + R_k \boldsymbol{\beta}_k] = \boldsymbol{\beta}_k^T [R_k^T \mathbf{r}_k + R_k^T R_k \boldsymbol{\beta}_k] = 0,$$

confirming the orthogonality property desired. $\square$

We note that [Lemma 4.1.1](#) is actually a special case of Proposition 5 of [\[7\]](#). We will need [Lemma 4.1.1](#) repeatedly in the analysis of stagnation for AA. Our next result is concerned with the immediate stagnation of nonlinear problems. We denote $\mathbf{q}^2(\mathbf{x})$ as two applications of $\mathbf{q}$, i.e. $\mathbf{q}^2(\mathbf{x}) = \mathbf{q}(\mathbf{q}(\mathbf{x}))$.

Theorem 4.1.1. Consider applying AA(m) to nonlinear fixed-point iteration function $\mathbf{q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ where $\mathbf{x}_1 = \mathbf{q}(\mathbf{x}_0) \neq \mathbf{x}^*$. Define $\mathbf{w}_0 := \mathbf{q}^2(\mathbf{x}_0) - 2\mathbf{q}(\mathbf{x}_0) + \mathbf{x}_0$ and assume $\mathbf{w}_0 \neq \mathbf{0}$. Then, we have immediate stagnation $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_{m+1}$ if $\mathbf{x}_0$ satisfies

$$\mathbf{x}_0 - \mathbf{q}(\mathbf{x}_0) \perp \mathbf{w}_0 \quad (4.8)$$

Proof. To begin, it is important to notice that if $\mathbf{x}_1 = \mathbf{q}(\mathbf{x}_0) \neq \mathbf{x}^*$, then we must have $\mathbf{x}_1 \neq \mathbf{x}_0$. Since $\mathbf{0} \neq \mathbf{w}_0 = \mathbf{q}^2(\mathbf{x}_0) - 2\mathbf{q}(\mathbf{x}_0) + \mathbf{x}_0 = \mathbf{r}_1 - \mathbf{r}_0$, we can use the unique solution for coefficient β_1 , as given by (3.4), as we satisfy the conditions that $m_k = 1$ and $\mathbf{r}_1 \neq \mathbf{r}_0$. From here it is easy to verify that $\beta_1 = -1$ if $\mathbf{x}_0$ satisfies (4.8). As a result, our next iterate is given by

$$\mathbf{x}_2 = (1 + \beta_1)\mathbf{q}(\mathbf{x}_1) - \beta_1\mathbf{q}(\mathbf{x}_0) = \mathbf{q}(\mathbf{x}_0) = \mathbf{x}_1.$$

If $\mathbf{x}_2 = \mathbf{x}_1$, then we also have $\mathbf{r}_2 = \mathbf{r}_1$ and the matrix R_2 takes the form

$$R_2 = \begin{bmatrix} \mathbf{0} & \mathbf{r}_2 - \mathbf{r}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{r}_1 - \mathbf{r}_0 \end{bmatrix}.$$

Since β_2 solves the normal equations [Eq. \(4.3\)](#), we have

$$\begin{bmatrix} 0 & 0 \\ 0 & \|\mathbf{r}_1 - \mathbf{r}_0\|^2 \end{bmatrix} \beta_2 = - [\mathbf{0} \quad \mathbf{r}_1 - \mathbf{r}_0]^T \mathbf{r}_1,$$

meaning

$$\beta_2 = \begin{bmatrix} \xi \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \xi \\ -1 \end{bmatrix}, \quad \text{where } \xi \in \mathbb{R} \text{ is a free parameter.}$$

Since $\mathbf{q}(\mathbf{x}_2) = \mathbf{q}(\mathbf{x}_1)$, the next iterate is given by

$$\mathbf{x}_3 = \mathbf{q}(\mathbf{x}_2) + \beta_1[\mathbf{q}(\mathbf{x}_2) - \mathbf{q}(\mathbf{x}_0)] = \mathbf{q}(\mathbf{x}_1) - [\mathbf{q}(\mathbf{x}_1) - \mathbf{q}(\mathbf{x}_0)] = \mathbf{q}(\mathbf{x}_0) = \mathbf{x}_1$$

independently of ξ . This stagnation update procedure continues as

$$R_k = [\mathbf{0} \quad \dots \quad \mathbf{0} \quad \mathbf{r}_1 - \mathbf{r}_0], \quad \text{for } k \leq m.$$

Repeating the process described above reveals that $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_{m+1}$ which is then halted as the size of R_k is constrained by the depth m . Specifically, this stagnation period finishes with R_{m+1} being the zero-matrix and the following iterate $\mathbf{x}_{m+2}$ being determined by a regular fixed-point update $\mathbf{x}_{m+2} = \mathbf{q}(\mathbf{x}_{m+1}) = \mathbf{q}(\mathbf{x}_1)$. $\square$

An interesting point to mention of [Theorem 4.1.1](#) is that the assumption that $\mathbf{w}_0 \neq 0$ can be satisfied if the operator $\mathcal{A}(\mathbf{x}) = \mathbf{x} - \mathbf{q}(\mathbf{x})$ has an inverse. In this situation, we have $\mathbf{r}_1 \neq \mathbf{r}_0$ since $\mathbf{x}_1 \neq \mathbf{x}_0 \implies \mathcal{A}^{-1}(\mathbf{r}_1) \neq \mathcal{A}^{-1}(\mathbf{r}_0)$ by the definition of our residuals. It is easy to check that condition [\(4.8\)](#) with affine $\mathbf{q}$ and residuals given by [\(2.19\)](#) results in $\mathbf{r}_0^T A \mathbf{r}_0 = 0$ as in [Corollary 4.0.1](#). [Theorem 4.1.1](#) identifies the possibility of poorly chosen initial iterates immediately falling into a period where no progress is made towards the solution of the underlying fixed-point problem. The duration of this period of stagnation is determined by the depth m , with larger m resulting in longer periods. If one wishes to solve a fixed-point problem using a small fixed number of iterations of AA(m), then [Theorem 4.1.1](#) suggests that a larger m may lead to no improvement towards the solution. [\(4.8\)](#) may seem unlikely to occur with a given iteration function $\mathbf{q}$ for initial iterate $\mathbf{x}_0$ chosen uniformly at random from the problem domain, but as we shall see in [Section 4.2](#), there exists some examples where stagnation is guaranteed. We now present a necessary condition of iteration functions $\mathbf{q}$ to see stagnation for AA(1).

Proposition 4.1.1. Consider AA(1) applied to a nonlinear iteration function $\mathbf{q}$. If $\mathbf{x}_{k+1} = \mathbf{x}_k$ for some $k \geq 1$, then $\mathbf{q}$ is noncontractive.

Proof. Suppose $\mathbf{x}_{k+1} = \mathbf{x}_k$ for some $k \geq 1$, then from ?? we see

$$\mathbf{x}_k = \mathbf{q}(\mathbf{x}_k) + \beta_k[\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})],$$

or alternatively,

$$\mathbf{r}_k = \beta_k[\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})]. \quad (4.9)$$

Recalling that the vector

$$\hat{\mathbf{r}}_{k+1} := \mathbf{r}_k + \beta_k[\mathbf{r}_k - \mathbf{r}_{k-1}]$$

has minimized norm with respect to β_k , and thus satisfies

$$\hat{\mathbf{r}}_{k+1} \perp \hat{\mathbf{r}}_{k+1} - \mathbf{r}_k,$$

we can substitute (4.9) to result in

$$(\beta_k[\mathbf{x}_k - \mathbf{x}_{k-1}])^T(\beta_k[\mathbf{r}_k - \mathbf{r}_{k-1}]) = 0.$$

Since $\mathbf{r}_k = \mathbf{x}_k - \mathbf{q}(\mathbf{x}_k)$, we have

$$[\mathbf{x}_k - \mathbf{x}_{k-1}]^T[(\mathbf{x}_k - \mathbf{x}_{k-1}) - (\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1}))] = 0.$$

Applying Cauchy-Schwarz results in

$$\|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \leq \|\mathbf{x}_k - \mathbf{x}_{k-1}\| \|\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})\|.$$

Now, dividing out the common factor $\|\mathbf{x}_k - \mathbf{x}_{k-1}\|$ shows $\mathbf{q}$ must be noncontractive. A related result is detailed in [39] that establishes that AA(m) applied to linear iteration functions $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ sees q-linear convergence of the residuals with a factor no worse than the underlying fixed-point iteration, that is, $\|\mathbf{r}_{k+1}\| / \|\mathbf{r}_k\| \leq \|M\|$ for matrices M with $\|M\| < 1$. In other words, it was shown that $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ cannot see stagnation for any depth m if $\|M\| < 1$. $\square$

Proposition 4.1.1 shows that AA(1) cannot stagnate if $\mathbf{q}$ is contractive. This result also informs us that the immediate stagnation of AA(m), as described by **Theorem 4.1.1**, cannot be achieved for contractive $\mathbf{q}$ since $m_k = 1$ in the update for $\mathbf{x}_2$. That is, condition (4.8) can only be achieved for some $\mathbf{x}_0$ if $\mathbf{q}$ is noncontractive.

4.2 Stagnation for Linear Problems

There are several additional conclusions we can make about the stagnation behaviour of AA applied to linear problems. In particular, the linearity provides a nice framework to analyze the problem of *consecutive stagnation periods*. That is, iterations that stagnate according to the results described in [Theorem 4.1.1](#) can, under the right conditions, immediately fall into another period of stagnation. In fact, several of these stagnation periods can be chained together to produce the following residual pattern:

$$\left\{ \begin{array}{ll} \mathbf{r}_0 & = (I - M)\mathbf{x}_0 \\ \mathbf{r}_1 & = M\mathbf{r}_0 = \mathbf{r}_2 = \cdots = \mathbf{r}_{m+1} \\ \mathbf{r}_{m+2} & = M\mathbf{r}_{m+1} = \mathbf{r}_{m+3} = \cdots = \mathbf{r}_{2m+2} \\ \mathbf{r}_{2m+3} & = M\mathbf{r}_{2m+2} = \mathbf{r}_{2m+4} = \cdots = \mathbf{r}_{3m+3} \\ \vdots & \vdots \\ \mathbf{r}_{(j-1)(m+1)+1} & = M\mathbf{r}_{(j-1)(m+1)} = \mathbf{r}_{(j-1)(m+1)+2} = \cdots = \mathbf{r}_{j(m+1)} \\ \vdots & \vdots \end{array} \right.$$

The key to understanding this behaviour lies in the AA update following an immediate stagnation period as detailed in [Theorem 4.1.1](#). As explained in the proof, R_k continues to grow by a zero column as $k \rightarrow m$, resulting in the same update for each iterate $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{m+1}$. After this period of stagnation, we have a zero-matrix for R_{m+1} and the resulting iterate $\mathbf{x}_{m+2}$ is determined by a simple fixed-point update $\mathbf{x}_{m+2} = \mathbf{q}(\mathbf{x}_{m+1}) = \mathbf{q}(\mathbf{x}_1)$. In terms of the residuals of a linear $\mathbf{q}$, this means that $\mathbf{r}_1 = \mathbf{r}_2 = \cdots = \mathbf{r}_{m+1}$ and the following residual is determined by $\mathbf{r}_{m+2} = M\mathbf{r}_1$. This means that the residual update following a period of stagnation is analagous to starting AA(m) with initial residual $\mathbf{r}_1$, so we may again encounter a period of stagnation if $\mathbf{r}_1$ satisfies the conditions of [Theorem 4.1.1](#), or more specifically, the linear conditions given by [Corollary 4.0.1](#). Recall from [Corollary 4.0.1](#) that immediate stagnation for linear iteration functions $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ is achieved if the initial residual satisfies $\mathbf{r}_0 \perp A\mathbf{r}_0$. Therefore, we can have an additional stagnation period following an immediate stagnation period if $\mathbf{r}_0^T A\mathbf{r}_0 = 0$ and

$$0 = \mathbf{r}_1^T A\mathbf{r}_1 = \mathbf{r}_0^T M^T A M\mathbf{r}_0.$$

If we let μ denote the number of stagnation periods, where $\mu = 1$ corresponds to the immediate stagnation result of [Corollary 4.0.1](#), we can extend this idea further to conclude

that if $\mathbf{r}_0$ satisfies the following system of quadratic forms

$$\begin{cases} \mathbf{r}_0^T A \mathbf{r}_0 & = 0 \\ (M\mathbf{r}_0)^T A (M\mathbf{r}_0) & = 0 \\ (M^2\mathbf{r}_0)^T A (M^2\mathbf{r}_0) & = 0 \\ \vdots & \vdots \\ (M^{\mu-1}\mathbf{r}_0)^T A M^{\mu-1}\mathbf{r}_0 & = 0 \end{cases} \quad (4.10)$$

then we see the residual pattern

$$\begin{cases} \mathbf{r}_0 & = (I - M)\mathbf{x}_0, \\ \mathbf{r}_1 & = M\mathbf{r}_0 = \mathbf{r}_2 = \cdots = \mathbf{r}_{m+1}, \\ \mathbf{r}_{m+2} & = M\mathbf{r}_{m+1} = \mathbf{r}_{m+3} = \cdots = \mathbf{r}_{2m+2}, \\ \vdots & \vdots \\ \mathbf{r}_{(\mu-1)(m+1)+1} & = M\mathbf{r}_{(\mu-1)(m+1)} = \mathbf{r}_{(\mu-1)(m+1)+2} = \cdots = \mathbf{r}_{\mu(m+1)}. \end{cases} \quad (4.11)$$

At this point, it is worth mentioning that the above residual behaviour is not necessarily restricted to the linear setting. That is, we could observe the same residual behaviour of a nonlinear $\mathbf{q}$ but analysis of this process is made more tractable under this linear setting.

4.2.1 Skew-symmetric Matrices

One observation we can immediately make is the effect that a skew-symmetric matrix $A = I - M$ would have when considering system (4.10). Recall that if a matrix $A \in \mathbb{R}^{n \times n}$ is skew-symmetric (i.e., $A^T = -A$), we have

$$\mathbf{v}^T A \mathbf{v} = 0 \quad \text{for all } \mathbf{v} \in \mathbb{R}^n. \quad (4.12)$$

Considering system (4.10), it is clear that if A is skew-symmetric then we will have immediate stagnation for any given initial residual. Further, after this immediate stagnation period, $AA(m)$ will fall into another period of stagnation (of the same length according to m). Moreover, each stagnation period will be followed by another stagnation period, continuing this procedure forever. This leads us to the following result.

Theorem 4.2.1. If $AA(m)$ is applied to solve for the fixed point of a linear function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, where $M \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$, and $M = I - A$ with $A \in \mathbb{R}^{n \times n}$ skew-symmetric, then the iteration will diverge for all initial iterates $\mathbf{x}_0$.

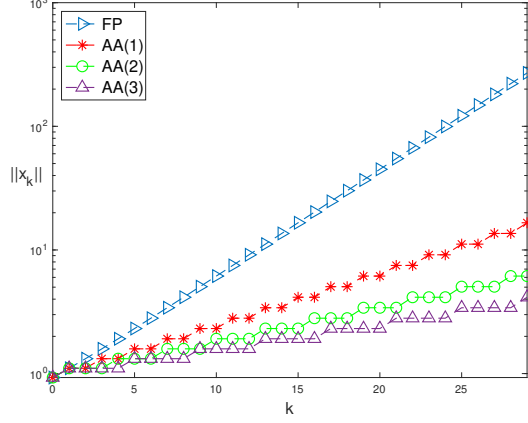


Figure 4.1: Normed iterates $\|\mathbf{x}_k\|$ as a function of iteration number k for $\text{AA}(m)$ and FP applied to a random skew-symmetric A with random initial iterate, comparing $m = 1, 2, 3$. Stagnation periods of length $m + 1$, separated by fixed-point iterations, can be observed for each choice of m .

Proof. The proof follows from the fact that $\text{AA}(m)$ repeatedly falls into stagnation periods as a result of (4.10), (4.11), the fact of skew-symmetric matrices shown in (4.12), and the natural divergence of the underlying fixed-point iteration with M defined as in the theorem statement. Since $M = I - A$ consists only of eigenvalues with modulus greater than 1, we have $\|M\| > 1$, and thus the underlying fixed-point iteration must diverge. The residual behaviour of (4.11) identifies that $\text{AA}(m)$ contains the same residuals of the underlying fixed-point iteration with m additional stagnating residuals between each fixed-point update. $\square$

An important note of Theorem 4.2.1 is that the assumption of invertibility of A implies that n must be even. In some sense, we might say that $\text{AA}(m)$ in the problem setting of Theorem 4.2.1 slows down the divergence of the underlying fixed-point iteration by a factor depending on the depth parameter m , as highlighted in Fig. 4.1. This figure shows that AA sees repeated stagnation periods with the length of each corresponding to the depth parameter m .

4.2.2 Symmetric Matrices

If we consider symmetric $A \in \mathbb{R}^n$, implying $M = I - A$ is also symmetric, there are a few more results we can explore concerning the discussed repeated stagnation period behaviour. To begin, it is first beneficial to reform system (4.10). Let $M = Q\Sigma Q^T$ be the eigendecomposition of M with diagonal matrix $\Sigma \in \mathbb{R}^n$ containing its eigenvalues $\{\lambda_j\}_{j=1}^n$. Now, if there exists an initial residual $\mathbf{r}_0$ satisfying (4.10) given some $\mu \in \mathbb{N}$, then the following linear system

$$G_\mu(\lambda_1, \dots, \lambda_n)\mathbf{v} = \mathbf{0}, \quad (4.13)$$

has a solution where matrix $G_\mu \in \mathbb{R}^{\mu \times n}$ is given by

$$G_\mu(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} 1 - \lambda_1 & 1 - \lambda_2 & \dots & 1 - \lambda_n \\ \lambda_1^2(1 - \lambda_1) & \lambda_2^2(1 - \lambda_2) & \dots & \lambda_n^2(1 - \lambda_n) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{2(\mu-1)}(1 - \lambda_1) & \lambda_2^{2(\mu-1)}(1 - \lambda_2) & \dots & \lambda_n^{2(\mu-1)}(1 - \lambda_n) \end{bmatrix}, \quad (4.14)$$

and

$$\mathbf{v} = \begin{bmatrix} (\mathbf{w}_0^{(1)})^2 \\ (\mathbf{w}_0^{(2)})^2 \\ \vdots \\ (\mathbf{w}_0^{(n)})^2 \end{bmatrix} \quad \text{with } \mathbf{w}_0 = Q^T \mathbf{r}_0. \quad (4.15)$$

If we wish to find a solution to the system of quadratic forms in (4.10), we can attempt to solve (4.13) for a *nonnegative vector* $\mathbf{v}$ ($\mathbf{v}^{(j)} \geq 0$ for all $1 \leq j \leq n$). This formulation leads to the following two results.

Theorem 4.2.2. Consider applying AA(m) to linear fixed-point function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, where $M \in \mathbb{R}^{n \times n}$ is symmetric, $\mathbf{b} \in \mathbb{R}^n$, $A = I - M$ is invertible, and $\mathbf{r}_1 = M\mathbf{r}_0 \neq \mathbf{0}$. If all eigenvalues $\{\lambda_j\}_{j=1}^n$ of M have distinct magnitude ($|\lambda_i| \neq |\lambda_j|$ for all $i \neq j$) and $\mu \geq n$, then the only solution to $G_\mu(\lambda_1, \dots, \lambda_n)\mathbf{v} = \mathbf{0}$ is $\mathbf{v} = \mathbf{0}$. As a consequence, matrices M with distinct eigenvalue magnitudes cannot see more than $n - 1$ consecutive stagnation periods.

Proof. If all eigenvalues $\{\lambda_j\}_{j=1}^n$ have distinct magnitude, then the columns of $G_\mu(\lambda_1, \dots, \lambda_n)$ are linearly independent and therefore deny any nonzero solution $\mathbf{v}$ to (4.13) when $\mu \geq n$. From the connection established between (4.10) and (4.13), there does not exist a nonzero residual $\mathbf{r}_0$ that produces $\geq n$ stagnation periods after application of AA(m). $\square$

Theorem 4.2.3. Consider applying AA(m) to linear fixed-point function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$, where $M \in \mathbb{R}^{n \times n}$ is symmetric, $\mathbf{b} \in \mathbb{R}^n$, $A = I - M$ is invertible, and $\mathbf{r}_1 = M\mathbf{r}_0 \neq \mathbf{0}$. If there exists a pair of eigenvalues of M such that $\lambda_i = -\lambda_j$ for $i \neq j$ where $|\lambda_i| > 1$, then there exists a nonnegative solution to $G_\mu(\lambda_1, \dots, \lambda_n)\mathbf{v} = \mathbf{0}$ for all $\mu \in \mathbb{N}$. As a consequence, matrices M that feature an eigenvalue pair with these properties can see consecutive stagnation periods forever.

Proof. We will prove the above result by constructing such a vector $\mathbf{v}^* \neq \mathbf{0}$ that satisfies our system $G_\mu \mathbf{v}^* = \mathbf{0}$. Let $i^* < j^*$ denote the indices where $\lambda_{i^*} = -\lambda_{j^*}$ and set $\mathbf{v}_d^* = 0 \ \forall d \neq i^*, j^*$. Now, consider the submatrix of M given by the 2×2 diagonal matrix $\text{diag}(\lambda_{i^*}, \lambda_{j^*})$ with corresponding subvector $(\mathbf{v}_{i^*}^*, \mathbf{v}_{j^*}^*)$. The corresponding columns of $G_\mu(\lambda_{i^*}, \lambda_{j^*})$ are linearly dependent and thus $G_\mu(\lambda_{i^*}, \lambda_{j^*})(\mathbf{v}_{i^*}^*, \mathbf{v}_{j^*}^*) = \mathbf{0}$ by taking $\mathbf{v}_{i^*}^* = \frac{1-\lambda_{j^*}}{\lambda_{i^*}-1} \mathbf{v}_{j^*}^*$. In summary, we have

$$\mathbf{v}_d^* = \begin{cases} 0, & \text{for all } d \neq i^*, j^*, \\ \xi, & \text{for } d = j^*, \\ \frac{1-\lambda_{j^*}}{\lambda_{i^*}-1} \xi, & \text{for } d = i^*, \end{cases}, \quad \text{where } \xi \text{ is a free parameter.}$$

With this construction of $\mathbf{v}^*$, we can then find an initial residual that results in repeatedly falling into consecutive stagnation periods by using (4.15). $\square$

4.3 Nonlinear Example

Consider applying AA to solve

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad \text{where } \mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 + 10x_2 \\ \sqrt{5}(x_3 - x_4) \\ (x_2 - 2x_3)^2 \\ \sqrt{10}(x_1 - x_4)^2 \end{bmatrix}. \quad (4.16)$$

Note that x_j is denoting the j -th component of $\mathbf{x}$ here. $\mathbf{f}$ is known as the *Powell singular function* [32, 21]. In this experiment, we will consider solving the more general system

$$\mathbf{F}^{[n]}(\mathbf{x}) = \mathbf{0},$$

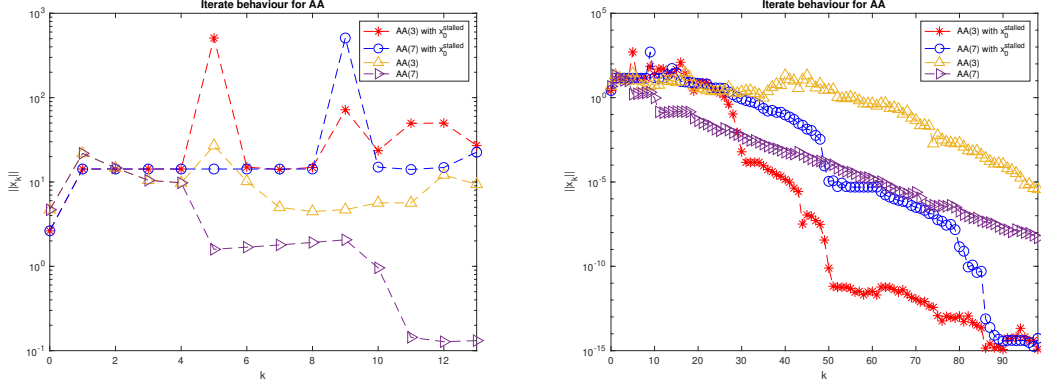


Figure 4.2: Normed iterates $\|\mathbf{x}_k\|$ as a function of iteration number k ($0 \leq k \leq 13$ in left plot and $0 \leq k \leq 100$ in right plot) for AA applied to $\mathbf{q}(\mathbf{x})$ of (4.18). Yellow and purple plots take initial iterate according to [21] ($\mathbf{x}_0 = (3, -1, 0, 1)^T$).

where $\mathbf{F}^{[n]} : \mathbb{R}^{4n} \rightarrow \mathbb{R}^{4n}$ is the *extended Powell singular function* [21] and

$$\begin{cases} \mathbf{F}_j^{[n]} &= x_j + 10x_{j+1}, \\ \mathbf{F}_{j+1}^{[n]} &= \sqrt{5}(x_{j+2} - x_{j+3}), \\ \mathbf{F}_{j+2}^{[n]} &= (x_{j+1} - 2x_{j+2})^2, \\ \mathbf{F}_{j+3}^{[n]} &= \sqrt{10}(x_j - x_{j+3})^2, \end{cases} \quad \text{with } j = 1, 5, 9, \dots, n-3. \quad (4.17)$$

The solution to each of these nonlinear problems, is given by $\mathbf{x}^* = \mathbf{0}$. We can translate these problems into an equivalent fixed-point problem:

$$\mathbf{x} = \mathbf{q}(\mathbf{x}), \quad \text{where } \mathbf{q}(\mathbf{x}) = \mathbf{x} - \mathbf{F}^{[n]}(\mathbf{x}). \quad (4.18)$$

Informed by Theorem 4.1.1, we can solve for an initial iterate $\mathbf{x}_0$ such that AA(m) exhibits immediate stagnation. Specifically, we will denote one such solution by $\mathbf{x}_0^{\text{stalled}}$ where

$$\mathbf{x}_0^{\text{stalled}} \approx (0.6836, 1, 1, 1)^T \quad (4.19)$$

Fig. 4.2 highlights the immediate stagnation property that AA(m) is capable of, given the correct conditions of the initial iterate $\mathbf{x}_0$ outlined in Theorem 4.1.1. Considering the iterates generated by AA(3) with $\mathbf{x}_0^{\text{stalled}}$ (shown in red), the left plot illustrates the 4 stalled iterates $x_1 = x_2 = x_3 = x_4$ predicted by Theorem 4.1.1 while AA(7) with $\mathbf{x}_0^{\text{stalled}}$

(shown in blue) features 8 stalled iterates. The left plot indicates that initial iterates that result in immediate stalled iterates (blue and red plots) may perform worse than other initializations in the preasymptotic phase while the right plot suggests that these stalled iterations may still perform better asymptotically.

Chapter 5

Acceleration Coefficients for Linear Problems

A source of much of the difficulty surrounding the asymptotic convergence of AA is the nonlinearity introduced by the acceleration coefficients. In papers concerning AA's convergence, there is often bounds or assumptions placed on the coefficients to ensure convergence. In this section we will derive a few expressions for the acceleration coefficients associated with affine functions $\mathbf{q}$.

5.0.1 Coefficients for AA(1)

To begin, we will first consider the case where $m = 1$ and we only have one coefficient to solve for at each step. Here, we introduce the following quantities

$$y_k := \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_{k-1}\|}, \quad \text{and} \quad \phi_k := \angle(\mathbf{r}_k, \mathbf{r}_{k-1}). \quad (5.1)$$

The ϕ_k quantity represents the angle between $\mathbf{r}_k$ and $\mathbf{r}_{k-1}$ and its calculation is through the formula

$$\cos(\phi_k) = \frac{\mathbf{r}_k^T \mathbf{r}_{k-1}}{\|\mathbf{r}_k\| \|\mathbf{r}_{k-1}\|}.$$

The purpose of introducing these values is to rewrite the solution for β_k , as given by (3.4). This technique can also be seen in [7]. Specifically, after dividing the numerator and denominator by $\|\mathbf{r}_{k-1}\|^2$ we now have

$$\beta_k = \frac{-\mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1})}{\|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2} = \frac{-\|\mathbf{r}_k\|^2 + \mathbf{r}_k^T \mathbf{r}_{k-1}}{\|\mathbf{r}_k\|^2 - 2\mathbf{r}_k^T \mathbf{r}_{k-1} + \|\mathbf{r}_{k-1}\|^2} = \frac{-y_k^2 + y_k \cos(\phi_k)}{y_k^2 - 2y_k \cos(\phi_k) + 1}. \quad (5.2)$$

From the linear convergence result in [39], we have $\|\mathbf{r}_k\| \leq \|M\| \|\mathbf{r}_{k-1}\|$, and therefore we have $0 \leq y_k \leq \|M\|$, and from the properties of the cosine function, we have $-1 \leq \cos(\phi_k) \leq 1$. This realization allows us to consider the acceleration coefficient as a function of $\cos(\phi_k)$, y_k on the domain $[-1, 1] \times [0, \|M\|]$. Let $\Omega = [-1, 1] \times [0, \|M\|]$ and consider the function $\beta : \Omega \rightarrow \mathbb{R}$ given by

$$\beta(x, y) = \frac{-y^2 + xy}{y^2 - 2xy + 1}. \quad (5.3)$$

Then, it is clear from (5.2) that we have

$$\beta_k = \beta(\cos(\phi_k), y_k) = \frac{-y_k^2 + y_k \cos(\phi_k)}{y_k^2 - 2y_k \cos(\phi_k) + 1}. \quad (5.4)$$

From the above expression, we can then attempt to understand the qualitative behaviour of the acceleration coefficients β_k through the analysis of the function β on the domain Ω . For example, since

$$y_k^2 - 2y_k \cos(\phi_k) + 1 = (y_k - \cos(\phi_k))^2 + 1 - \cos^2(\phi_k) = (y_k - \cos(\phi_k))^2 + \sin^2(\phi_k),$$

the denominator in (5.4) is zero if and only if $y_k = \cos(\phi_k)$ and $\sin(\phi_k) = 0$. This corresponds to the point $(\phi_k, y_k) = (0, 1)$, or alternatively, the situation in which we have stagnation $\mathbf{r}_k = \mathbf{r}_{k-1}$. As explained in Chapter 4, in the instance that $\mathbf{r}_k = \mathbf{r}_{k-1}$, the optimization problem given for our acceleration coefficients (3.4) becomes rank-deficient. It is important to note that the point $(\phi_k, y_k) = (0, 1)$ can only be achieved in AA iterations where $\|M\| \geq 1$ since otherwise $y_k \leq \|M\| < 1$. Fig. 5.1 showcases the surface described by the function β in the domain Ω for a specific value of $\|M\|$. This highlights the extreme values of β , allowing us to prove the following theorem.

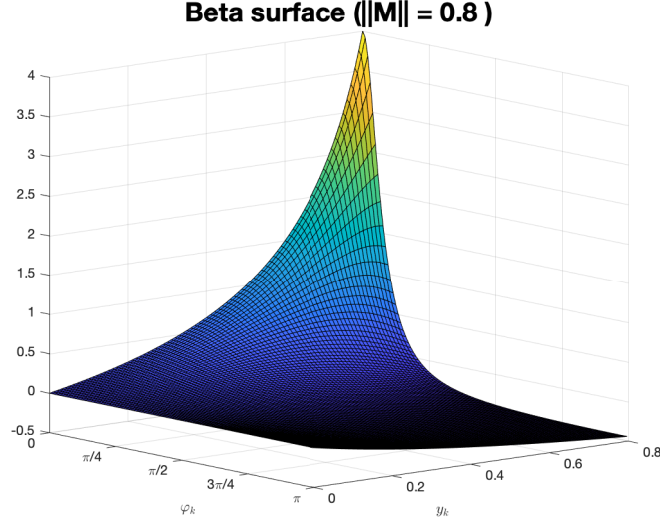


Figure 5.1: Acceleration coefficients β_k as a function of $y_k \in [0, \|M\|]$ and $\phi_k \in [0, \pi]$ according to (5.4)

Theorem 5.0.1. Consider applying AA(1) for any initial residual $\mathbf{r}_0$ to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\|M\| < 1$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Then,

$$-\frac{\|M\|}{1 + \|M\|} \leq \beta_k \leq \frac{\|M\|}{1 - \|M\|},$$

for all $k \geq 1$.

Proof. This result follows from the properties of the function β defined in (5.3). It is not difficult to show using typical calculus techniques that the function $\beta(x, y)$ on the domain $\Omega = [-1, 1] \times [0, \|M\|]$, with $\|M\| < 1$, achieves its maximum at the point $(x, y) = (1, \|M\|)$ and minimum at the point $(x, y) = (-1, \|M\|)$. At these points, we have

$$\beta(-1, \|M\|) = -\frac{\|M\|}{1 + \|M\|}, \quad \text{and} \quad \beta(1, \|M\|) = \frac{\|M\|}{1 - \|M\|},$$

therefore

$$-\frac{\|M\|}{1 + \|M\|} \leq \beta(x, y) \leq \frac{\|M\|}{1 - \|M\|}, \quad \text{for all } (x, y) \in \Omega. \quad (5.5)$$

If we define ϕ_k and y_k as in (5.1), then $\beta_k = \beta(\cos(\phi_k), y_k)$ according to (5.4). When AA(1) is applied to an affine function with $\|M\| < 1$, then [39] shows that $y_k \leq \|M\|$ for all $k \geq 1$.

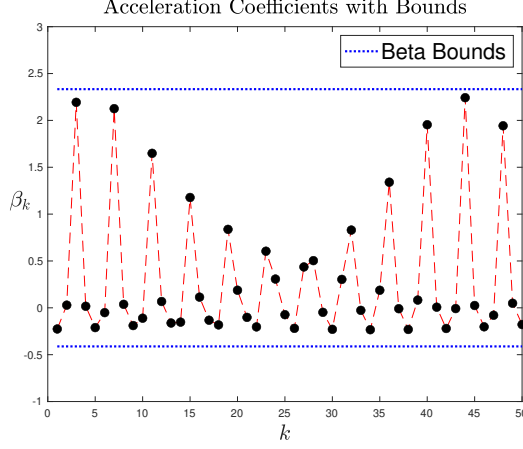


Figure 5.2: Acceleration coefficients β_k as a function of iteration number for AA(1) applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$, and $\mathbf{r}_0 = (-0.1285, 0.6687)^T$. The dotted blue lines correspond to the upper and lower bounds for β_k as determined by Theorem 5.0.1 with upper bound given by $7/3$ and lower bound given by $-7/17$

Since $(\cos(\phi_k), y_k) \in \Omega$ for all $k \geq 1$, using the bounds in (5.5) results in the desired bound for the acceleration coefficients. $\square$

Remark 5.0.1. The proof of Theorem 5.0.1 asserts that the function $\beta(x, y)$ in the domain Ω is maximized when $(x, y) = (1, \|M\|)$ and minimized when $(x, y) = (-1, \|M\|)$. In an AA(1) iteration, the point $(\cos(\phi_k), y_k) = (1, \|M\|)$ corresponds to parallel residuals $\mathbf{r}_k, \mathbf{r}_{k-1}$, with $\mathbf{r}_k = \|M\| \mathbf{r}_{k-1}$, while the point $(\cos(\phi_k), y_k) = (-1, \|M\|)$ corresponds to parallel residuals $\mathbf{r}_k, \mathbf{r}_{k-1}$, with $\mathbf{r}_k = -\|M\| \mathbf{r}_{k-1}$.

If we consider the 2×2 example shown in Fig. 3.1 where $\|M\| = \lambda_1 = 0.7$, we have

$$-\frac{7}{17} = -\frac{0.7}{1+0.7} \leq \beta_k \leq \frac{0.7}{1-0.7} = \frac{7}{3}, \quad \text{for all } k \geq 1,$$

which is highlighted in Fig. 5.2. It's important to note that for AA(1) with a given initial residual $\mathbf{r}_0$, there may not be a $K \geq 1$ such that β_K achieves the lower or upper bounds detailed in Theorem 5.0.1. Instead, the bound shown in Theorem 5.0.1 holds over all possible initial residuals.

The expression for the acceleration coefficients shown in (5.4) is useful for understanding the possible values these coefficients can take for a given AA(1) iteration, as described by

[Theorem 5.0.1](#). Another use of (5.4) is for bounding the norm of the residual

$$\|\mathbf{r}_{k+1}\| = \|M(\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}))\|.$$

As defined in the previous section, let

$$\hat{\mathbf{r}}_{k+1} := \mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}), \quad (5.6)$$

such that $\|\mathbf{r}_{k+1}\| = \|M\hat{\mathbf{r}}_{k+1}\|$. Now, we have the following theorem

Theorem 5.0.2. Consider applying AA(1) for any initial residual $\mathbf{r}_0$ to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\|M\| < 1$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. If $y_k = \|\mathbf{r}_k\| / \|\mathbf{r}_{k-1}\|$ and $\phi_k = \angle(\mathbf{r}_k, \mathbf{r}_{k-1})$, then

$$\|\mathbf{r}_{k+1}\| \leq \|\mathbf{r}_k\| \frac{\|M\| \sin(\phi_k)}{\sqrt{y_k^2 - 2y_k \cos(\phi_k) + 1}}$$

for all $k \geq 1$.

Proof. Since $\|\mathbf{r}_{k+1}\|^2 = \|M(\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}))\|^2$ for all $k \geq 1$, we have

$$\|\mathbf{r}_{k+1}\|^2 \leq \|M\|^2 \|\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})\|^2.$$

Now, using the solution for β_k when $m = 1$, as given by (3.4), we have

$$\begin{aligned} \|\mathbf{r}_{k+1}\|^2 &\leq \|M\|^2 \|\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})\|^2 \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + 2\beta_k \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1}) + \beta_k^2 \|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2) \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + 2\beta_k \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1}) - \beta_k \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1})) \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + \beta_k \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1})) \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + \beta_k (\|\mathbf{r}_k\|^2 - \mathbf{r}_k^T \mathbf{r}_{k-1})) \\ &= \|\mathbf{r}_k\|^2 \|M\|^2 \left(1 + \beta_k \left(1 - \frac{\mathbf{r}_k^T \mathbf{r}_{k-1}}{\|\mathbf{r}_k\|^2}\right)\right). \end{aligned}$$

Substituting in y_k and ϕ_k we have

$$\begin{aligned} \|\mathbf{r}_{k+1}\|^2 &\leq \|\mathbf{r}_k\|^2 \|M\|^2 \left(1 + \left(\frac{-y_k^2 + y_k \cos(\phi_k)}{y_k^2 - 2y_k \cos(\phi_k) + 1}\right) \left(\frac{y_k - \cos(\phi_k)}{y_k}\right)\right) \\ &= \|\mathbf{r}_k\|^2 \|M\|^2 \left(\frac{1 - \cos^2(\phi_k)}{y_k^2 - 2y_k \cos(\phi_k) + 1}\right). \end{aligned}$$

Taking the square root of the above leads to the desired bound. $\square$

Remark 5.0.2. The bound in [Theorem 5.0.2](#) is also proved in [29, 7] using slightly different approaches.

5.0.2 Coefficients for AA(2)

The previous section developed an expression for the acceleration coefficients associated with AA(1) by introducing the numbers y_k and ϕ_k that represent the relative length and relative angle of the previous two residuals respectively. In this section, we generalize this approach slightly to the case where $m = 2$. To do this, an additional quantity ψ_k is introduced, which represents the angle made between $\mathbf{r}_k$ and $\mathbf{r}_{k-2}$. Specifically,

$$\cos(\psi_k) = \frac{\mathbf{r}_k^T \mathbf{r}_{k-2}}{\|\mathbf{r}_k\| \|\mathbf{r}_{k-2}\|}.$$

While the introduction of ψ_k may suggest an associated quantity for $\|\mathbf{r}_k\| / \|\mathbf{r}_{k-2}\|$, note that

$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_{k-2}\|} = \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_{k-1}\|} \frac{\|\mathbf{r}_{k-1}\|}{\|\mathbf{r}_{k-2}\|} = y_k y_{k-1},$$

meaning another variable is unnecessary.

Recall that the AA(2) residual update formula takes the form

$$\mathbf{r}_{k+1} = M \left(\mathbf{r}_k + \sum_{j=1}^m \beta_k^{(j)} (\mathbf{r}_k - \mathbf{r}_{k-j}) \right) = M(\mathbf{r}_k + R_k \boldsymbol{\beta}_k),$$

where

$$R_k = [\mathbf{r}_k - \mathbf{r}_{k-1} \quad \mathbf{r}_k - \mathbf{r}_{k-2}], \quad \text{and} \quad \boldsymbol{\beta}_k = \begin{bmatrix} \beta_k^{(1)} \\ \beta_k^{(2)} \end{bmatrix}. \quad (5.7)$$

When $R_k^T R_k$ has an inverse, the solution to $\boldsymbol{\beta}_k$ is given by

$$\boldsymbol{\beta}_k = -(R_k^T R_k)^{-1} R_k^T \mathbf{r}_k. \quad (5.8)$$

The situation in which $R_k^T R_k$ does not have an inverse corresponds to the case in which the columns of R_k are linearly dependent, i.e. $\mathbf{r}_k - \mathbf{r}_{k-2} = \alpha(\mathbf{r}_k - \mathbf{r}_{k-1})$ for some $\alpha \in \mathbb{R}$. For this section, we will ignore this case and strictly consider the formula for $\boldsymbol{\beta}_k$ given in (5.8), since linear dependence of the columns of R_k implies the solution for the acceleration coefficients are given by minimizing the norm of the vector

$$\mathbf{r}_k + \boldsymbol{\beta}_k^{(1)} (\mathbf{r}_k - \mathbf{r}_{k-1}) + \boldsymbol{\beta}_k^{(2)} (\mathbf{r}_k - \mathbf{r}_{k-2}) = \mathbf{r}_k + (\boldsymbol{\beta}_k^{(1)} + \alpha \boldsymbol{\beta}_k^{(2)}) (\mathbf{r}_k - \mathbf{r}_{k-1}),$$

for some $\alpha \in \mathbb{R}$. This reduces to the $m = 1$ case with $\beta_k = \beta_k^{(1)} + \alpha \beta_k^{(2)}$ found using (5.4). It is therefore not of interest to consider the case where $R_k^T R_k$ does not have an inverse

as this reduces to the analysis for the acceleration coefficients presented in the previous section. Now, since

$$R_k^T R_k = \begin{bmatrix} \|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2 & (\mathbf{r}_k - \mathbf{r}_{k-1})^T (\mathbf{r}_k - \mathbf{r}_{k-2}) \\ (\mathbf{r}_k - \mathbf{r}_{k-1})^T (\mathbf{r}_k - \mathbf{r}_{k-2}) & \|\mathbf{r}_k - \mathbf{r}_{k-2}\|^2 \end{bmatrix},$$

our assumption that the columns of R_k are linearly independent implies $R_k^T R_k$ has an inverse and

$$(R_k^T R_k)^{-1} = \frac{1}{\det(R_k^T R_k)} \begin{bmatrix} \|\mathbf{r}_k - \mathbf{r}_{k-2}\|^2 & -(\mathbf{r}_k - \mathbf{r}_{k-1})^T (\mathbf{r}_k - \mathbf{r}_{k-2}) \\ -(\mathbf{r}_k - \mathbf{r}_{k-1})^T (\mathbf{r}_k - \mathbf{r}_{k-2}) & \|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2 \end{bmatrix} \quad (5.9)$$

where

$$\det(R_k^T R_k) = \|\mathbf{r}_k - \mathbf{r}_{k-1}\|^2 \|\mathbf{r}_k - \mathbf{r}_{k-2}\|^2 - ((\mathbf{r}_k - \mathbf{r}_{k-1})^T (\mathbf{r}_k - \mathbf{r}_{k-2}))^2. \quad (5.10)$$

Additionally, we have

$$R_k^T \mathbf{r}_k = \begin{bmatrix} \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1}) \\ \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-2}) \end{bmatrix}. \quad (5.11)$$

For the k -th step of a given AA(2) iteration, define the numbers

$$h_1 := y_k - \cos(\phi_k), \quad (5.12)$$

$$h_2 := y_k y_{k-1} - \cos(\psi_k), \quad (5.13)$$

$$h_3 := y_k^2 - 2y_k \cos(\phi_k) + 1, \quad (5.14)$$

$$h_4 := y_k^2 y_{k-1}^2 - 2y_k y_{k-1} \cos(\psi_k) + 1, \quad (5.15)$$

$$h_5 := y_k^2 y_{k-1} - y_k y_{k-1} \cos(\phi_k) - y_k \cos(\psi_k) + \cos(\phi_{k-1}). \quad (5.16)$$

Now, dividing by $\|\mathbf{r}_{k-1}\|^2 \|\mathbf{r}_{k-2}\|^2$ in a manner similar to (5.2), using the expressions (5.8), (5.9), (5.11) we have

$$\beta_k^{(1)} = y_k \times \left(\frac{h_2 h_5 - h_1 h_4}{h_3 h_4 - h_5^2} \right) \quad (5.17)$$

$$\beta_k^{(2)} = y_k y_{k-1} \times \left(\frac{h_1 h_5 - h_2 h_3}{h_3 h_4 - h_5^2} \right) \quad (5.18)$$

Equipped with the above expression for β_k , we can prove the following theorem which provides an explicit AA(2) residual bound with respect to the relative quantities y_k , y_{k-1} , ϕ_k , ϕ_{k-1} , and ψ_k .

Theorem 5.0.3. Consider applying AA(2) for any initial residual $\mathbf{r}_0$ to the affine function $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ where $\|M\| < 1$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. If $y_k = \|\mathbf{r}_k\| / \|\mathbf{r}_{k-1}\|$, $\phi_k = \angle(\mathbf{r}_k, \mathbf{r}_{k-1})$, $\psi_k = \angle(\mathbf{r}_k, \mathbf{r}_{k-2})$, and h_1, h_2, h_3, h_4, h_5 are defined as in (5.12)-(5.16), then

$$\|\mathbf{r}_{k+1}\| \leq \|\mathbf{r}_k\| \|M\| \sqrt{\frac{h_3 h_4 - h_5^2 - h_1^2 h_4 + 2h_1 h_2 h_5 - h_2^2 h_3}{h_3 h_4 - h_5^2}}.$$

Proof. Since $\|\mathbf{r}_{k+1}\|^2 = \|M(\mathbf{r}_k + R_k \boldsymbol{\beta}_k)\|^2$ for all $k \geq 2$, we have

$$\|\mathbf{r}_{k+1}\|^2 \leq \|M\|^2 \|\mathbf{r}_k + R_k \boldsymbol{\beta}_k\|^2.$$

Now, using the solution for $\boldsymbol{\beta}_k$ when $m = 2$, as given by (5.8), we have

$$\begin{aligned} \|\mathbf{r}_{k+1}\|^2 &\leq \|M\|^2 \|\mathbf{r}_k + R_k \boldsymbol{\beta}_k\|^2 \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + 2\boldsymbol{\beta}_k^T R_k^T \mathbf{r}_k + \boldsymbol{\beta}_k^T R_k^T R_k \boldsymbol{\beta}_k) \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + 2\boldsymbol{\beta}_k^T R_k^T \mathbf{r}_k - \boldsymbol{\beta}_k^T R_k^T \mathbf{r}_k) \\ &= \|M\|^2 (\|\mathbf{r}_k\|^2 + \boldsymbol{\beta}_k^T R_k^T \mathbf{r}_k) \end{aligned}$$

Using our expression for $\boldsymbol{\beta}_k$ shown in (5.17) and (5.18), we have

$$\begin{aligned} \boldsymbol{\beta}_k^T R_k^T \mathbf{r}_k &= y_k \left(\frac{h_2 h_5 - h_1 h_4}{h_3 h_4 - h_5^2} \right) \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-1}) + y_k y_{k-1} \left(\frac{h_1 h_5 - h_2 h_3}{h_3 h_4 - h_5^2} \right) \mathbf{r}_k^T (\mathbf{r}_k - \mathbf{r}_{k-2}) \\ &= \|\mathbf{r}_k\|^2 \left(\frac{h_2 h_5 - h_1 h_4}{h_3 h_4 - h_5^2} \right) (y_k - \cos(\phi_k)) + \|\mathbf{r}_k\|^2 \left(\frac{h_1 h_5 - h_2 h_3}{h_3 h_4 - h_5^2} \right) (y_k y_{k-1} - \cos(\psi_k)) \\ &= \|\mathbf{r}_k\|^2 \left[h_1 \left(\frac{h_2 h_5 - h_1 h_4}{h_3 h_4 - h_5^2} \right) + h_2 \left(\frac{h_1 h_5 - h_2 h_3}{h_3 h_4 - h_5^2} \right) \right]. \end{aligned}$$

Finally, this results in

$$\begin{aligned} \|\mathbf{r}_{k+1}\|^2 &\leq \|M\|^2 \left(\|\mathbf{r}_k\|^2 + \|\mathbf{r}_k\|^2 \left[h_1 \left(\frac{h_2 h_5 - h_1 h_4}{h_3 h_4 - h_5^2} \right) + h_2 \left(\frac{h_1 h_5 - h_2 h_3}{h_3 h_4 - h_5^2} \right) \right] \right) \\ &= \|\mathbf{r}_k\|^2 \|M\|^2 \left(\frac{h_3 h_4 - h_5^2 - h_1^2 h_4 + 2h_1 h_2 h_5 - h_2^2 h_3}{h_3 h_4 - h_5^2} \right). \end{aligned}$$

Now, taking the root of each side gives us the desired residual bound. $\square$

To our knowledge, the residual bound for AA(2) shown in [Theorem 5.0.3](#) is the first of its kind. The general residual bound shown in [\[29\]](#) takes a similar form in the linear

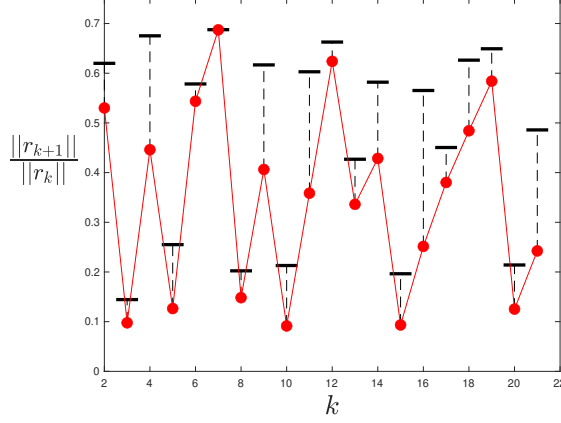


Figure 5.3: Illustration of AA(2) residual bounds from [Theorem 5.0.3](#) applied to affine $\mathbf{q}$ with $M = \text{diag}(0.7, -0.5, 0.3)$ and $\mathbf{r}_0$ chosen randomly. The red points correspond to the values of $y_{k+1} = \|\mathbf{r}_{k+1}\| / \|\mathbf{r}_k\|$ generated by AA(2) and the black bars correspond to the bound given for each y_{k+1} in [Theorem 5.0.3](#). The iterations considered in the above plot range from $k = 2$ to $k = 21$.

setting when $m = 2$, except that the bound is in terms of a variable that represents the optimization gain at step k , which in general is unknown. The bound in [Theorem 5.0.3](#) writes this optimization gain explicitly in terms of the angles between residual vectors $\phi_k, \phi_{k-1}, \psi_k$ and the relative lengths of residual vectors y_k, y_{k-1} .

To see the residual bound given by [Theorem 5.0.3](#) in action, consider the following example. Let $\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b}$ with $\mathbf{x}, \mathbf{b} \in \mathbb{R}^3$ and $M = \text{diag}(0.7, -0.5, 0.3)$ and consider the residuals generated by AA(2) applied to $\mathbf{q}$ for some initial residual $\mathbf{r}_0$. By dividing each side of the bound in [Theorem 5.0.3](#) by $\|\mathbf{r}_k\|$, we obtain a bound for y_{k+1} . In [Fig. 5.3](#), we plot the actual values for y_{k+1} versus the bound obtained in [Theorem 5.0.3](#) as a function of the iteration number k .

In theory, the approach used in [Theorem 5.0.3](#) to find a bound for the normed residuals can be extended to $m > 2$ by introducing more quantities that represent relative angles and relative lengths of previous residuals. That is, if we define the numbers

$$y_k = \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_{k-1}\|}, \quad (5.19)$$

$$\phi_k^{[j]} = \angle(\mathbf{r}_k, \mathbf{r}_{k-j}), \quad (5.20)$$

then we could express β_k as a function of m relative lengths

$$y_j = \frac{\|\mathbf{r}_j\|}{\|\mathbf{r}_{j-1}\|}, \quad j = k - m + 1, \dots, k,$$

and $m(m + 1)/2$ relative angles

$$\begin{aligned} \phi_j^{[1]} &= \angle(\mathbf{r}_j, \mathbf{r}_{j-1}), \quad j = k - m + 1, \dots, k, \\ \phi_j^{[2]} &= \angle(\mathbf{r}_j, \mathbf{r}_{j-2}), \quad j = k - m + 2, \dots, k, \\ &\vdots \\ \phi_j^{[m-1]} &= \angle(\mathbf{r}_j, \mathbf{r}_{j-m+1}), \quad j = k - 1, k, \\ \phi_j^{[m]} &= \angle(\mathbf{r}_k, \mathbf{r}_{k-m}), \end{aligned}$$

resulting in $m + m(m + 1)/2 = m(m + 3)/2$ total variables. With respect to [Theorem 5.0.3](#), the above formulation implies $\phi_k = \phi_k^{[1]}$, $\phi_{k-1} = \phi_{k-1}^{[1]}$, and $\psi_k = \phi_k^{[2]}$. In the future, this idea of expressing β_k in these terms will be explored in more detail in the hopes of extracting a residual bound that shows improved convergence rates over the underlying fixed-point iteration.

Chapter 6

Restarted Anderson Acceleration

As discussed in [Chapter 2](#), various modifications have been proposed for the AA algorithm. In particular, several authors have explored the use of *restarts* during AA iterations [[16](#), [26](#), [4](#), [8](#)]. This technique suggests applying AA(m) for some number of iterations K , at which point the iteration is restarted by taking the last iterate $\mathbf{x}_K$ and using it as an initial iterate for AA(m) while discarding all previous information. This restarting strategy is similar to restarted GMRES, often denoted by GMRES(m), which is analyzed in various sources such as [[10](#), [2](#), [22](#), [17](#)]. For example, consider the restarted AA(1) algorithm studied in [[4](#), [8](#)] which enforces a restart every other iteration. This algorithm updates according to the rule

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) \quad k = 0, 2, 4, \dots, \quad (6.1)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 1, 3, 5, \dots, \quad (6.2)$$

where β_k is found in the usual manner ([3.4](#)). Here, we adopt the notation used in [[8](#)] and refer to the above algorithm as rAA(1). It is clear from the above formulas that restarts correspond to the even iterations k where $\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k)$, as only information from the most previous iterate is used in the next update. In this chapter, we develop exact convergence rates associated with various restarted AA(1) algorithms applied to the affine function

$$\mathbf{q}(\mathbf{x}) = M\mathbf{x} + \mathbf{b},$$

where $M \in \mathbb{R}^{n \times n}$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Additionally, we make the assumption that $A := I - M$ is invertible such that a unique solution for the fixed-point problem is guaranteed. While these results do not directly inform us of the convergence behaviour of AA(1) in general, these algorithms are similar enough that the analysis of these methods may have implications for

the analysis of AA(1). Further, numerical experiments will be investigated that compare the convergence of these restarted methods with the typical windowed AA that this thesis has largely been concerned with.

6.1 Convergence of rAA(2)

In this section, we discuss a new restarted AA(1) algorithm which will be denoted by rAA(2). This method differs from the method analyzed in [8] by including an additional AA(1) update between each restart. That is, rAA(2) updates according to the rule

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) \quad k = 0, 3, 6, \dots, \quad (6.3a)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 1, 4, 7, \dots, \quad (6.3b)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 2, 5, 8, \dots, \quad (6.3c)$$

where the coefficients β_k are determined in the usual manner according to (3.4). Here, restarts occur every three iterations, where the iteration number satisfies $k \equiv 0 \pmod{3}$. Another way we might chose to think about this formulation is given as follows. If we follow [39] and denote AA(0) as the fixed-point iteration, then our proposed restarted AA(1) algorithm can be found by taking $m_k = \min(1, k \pmod{3})$ in (2.14).

To analyze the convergence of rAA(2) we investigate the associated residual iterations. Since $\mathbf{r}_k = \mathbf{x}_k - \mathbf{q}(\mathbf{x}_k) = A\mathbf{x} - \mathbf{b}$, our assumption that A has an inverse implies that $\mathbf{x}_k = A^{-1}(\mathbf{r}_k + \mathbf{b})$. Substituting this into (6.3), we have the rAA(2) residual iteration

$$\mathbf{r}_{k+1} = M\mathbf{r}_k \quad k = 0, 3, 6, \dots, \quad (6.4a)$$

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] \quad k = 1, 4, 7, \dots, \quad (6.4b)$$

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] \quad k = 2, 5, 8, \dots \quad (6.4c)$$

To aid in the analysis of this method, we introduce a reformulation of (6.4) with the following proposition.

Proposition 6.1.1. Let $\{\mathbf{r}_k\}$ be the residuals generated by applying the restarted AA iteration shown in (6.4). Then,

$$\mathbf{r}_{k+1} = \mathcal{F}(\mathbf{r}_{k-2}), \quad k = 2, 5, 8, \dots, \quad (6.5)$$

where $\mathcal{F} : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$ is the three-step residual operator given by

$$\mathcal{F}(\mathbf{u}) := M^2[I - R(A^{-1}, A(I - A)\mathbf{u})A]\mathbf{u}, \quad (6.6)$$

and $R(A^{-1}, A(I - A)\mathbf{u})$ is the Rayleigh quotient seen in Definition 2.2.1 for the matrix A^{-1} with the nonzero vector $A(I - A)\mathbf{u}$.

Proof. Consider $\mathbf{r}_{k-2}$ where $k \in \{2, 5, 8, \dots\}$. From (6.4), we then have $\mathbf{r}_{k-1} = M\mathbf{r}_{k-2}$ and therefore

$$\mathbf{r}_k = M[\mathbf{r}_{k-1} + \beta_{k-1}(\mathbf{r}_{k-1} - \mathbf{r}_{k-2})] = M[I - (1 + \beta_{k-1})A]\mathbf{r}_{k-2},$$

where

$$1 + \beta_{k-1} = 1 + \frac{\mathbf{r}_{k-2}^T M^T A \mathbf{r}_{k-2}}{\mathbf{r}_{k-2}^T A^T A \mathbf{r}_{k-2}} = \frac{\mathbf{r}_{k-2}^T A \mathbf{r}_{k-2}}{\mathbf{r}_{k-2}^T A^T A \mathbf{r}_{k-2}} = \frac{\mathbf{r}_{k-2}^T A^T A^{-1} A \mathbf{r}_{k-2}}{\mathbf{r}_{k-2}^T A^T A \mathbf{r}_{k-2}}.$$

Putting these equations together, we have

$$\mathbf{r}_k = M[I - R(A^{-1}, A\mathbf{r}_{k-2})A]\mathbf{r}_{k-2}. \quad (6.7)$$

Now, our next residual is determined by the update

$$\begin{aligned} \mathbf{r}_{k+1} &= M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1})] \\ &= M[M(I - R(A^{-1}, A\mathbf{r}_{k-2})A)\mathbf{r}_{k-2} - \beta_k R(A^{-1}, A\mathbf{r}_{k-2})M A \mathbf{r}_{k-2}] \\ &= M^2 [I - (1 + \beta_k)R(A^{-1}, A\mathbf{r}_{k-2})A] \mathbf{r}_{k-2}, \end{aligned}$$

where

$$\begin{aligned} 1 + \beta_k &= 1 + \frac{(M[I - R(A^{-1}, A\mathbf{r}_{k-2})A]\mathbf{r}_{k-2})^T (R(A^{-1}, A\mathbf{r}_{k-2})M A \mathbf{r}_{k-2})}{(R(A^{-1}, A\mathbf{r}_{k-2})M A \mathbf{r}_{k-2})^T (R(A^{-1}, A\mathbf{r}_{k-2})M A \mathbf{r}_{k-2})} \\ &= \frac{1}{R(A^{-1}, A\mathbf{r}_{k-2})} \frac{(M\mathbf{r}_{k-2})^T (M A \mathbf{r}_{k-2})}{(M A \mathbf{r}_{k-2})^T (M A \mathbf{r}_{k-2})} \\ &= \frac{R(A^{-1}, A(I - A)\mathbf{r}_{k-2})}{R(A^{-1}, A\mathbf{r}_{k-2})}. \end{aligned}$$

Finally, substituting this in for the expression we found earlier for $\mathbf{r}_{k+1}$, we now have

$$\mathbf{r}_{k+1} = M^2 [I - R(A^{-1}, A(I - A)\mathbf{r}_{k-2})A] \mathbf{r}_{k-2} = \mathcal{F}(\mathbf{r}_{k-2}), \quad (6.8)$$

as desired. $\square$

The following theorem highlights an interesting property of initial residuals that consist of a sum of two eigenvectors of the matrix M . From this point on, we make the assumption that M is a symmetric matrix, which ensures that M has real eigenvalues with orthogonal eigenvectors.

Theorem 6.1.1. Let $\mathbf{u}_i, \mathbf{u}_j, i \neq j$, denote two orthogonal eigenvectors of M with eigenvalues m_i and m_j respectively. Define

$$\mathbf{z}(\varepsilon) := c_i \mathbf{u}_i + c_j \mathbf{u}_j = c_i (\mathbf{u}_i + \varepsilon \mathbf{u}_j), \quad \text{with } \varepsilon = \frac{c_j}{c_i}, \quad (6.9)$$

where $c_i, c_j \neq 0$ are real constants. Then, if $m_i = m_j$,

$$\mathcal{F}(\mathbf{z}(\varepsilon)) = 0. \quad (6.10)$$

Otherwise, if $m_i \neq m_j$, then $\mathbf{z}(\varepsilon)$ solves the following nonlinear eigenvalue problem

$$\mathcal{F}(\mathcal{F}(\mathbf{z}(\varepsilon))) = \lambda(\varepsilon)\mathbf{z}(\varepsilon), \quad (6.11)$$

where

$$\lambda(\varepsilon) = \frac{(m_2 - m_1)^2 m_1^4 m_2^4}{[m_1^2(m_1 - 1)^2 + m_2^2(m_2 - 1)^2 \varepsilon^2][m_1^2 + m_2^2/\varepsilon^2]}. \quad (6.12)$$

Proof. Suppose $m_1 = m_2$. Then, since rAA(2) follows the same residual updates as AA(1) until a restart occurs after the third iteration, we can apply [Proposition 3.1.2](#) to conclude that $\mathbf{r}_2 = \mathbf{0}$. The following residual $\mathbf{r}_3$ will also be the zero vector as we have reached that exact solution. Therefore, $\mathbf{0} = \mathbf{r}_3 = \mathcal{F}(\mathbf{r}_0)$. If we instead have distinct m_i, m_j , then we can analyze the effect of $\mathcal{F}$ on the vector $\mathbf{z}$. For the sake of readability, set $i = 1, j = 2$, and let $a_1 = 1 - m_1, a_2 = 1 - m_2$ be the eigenvalues of A . Substituting $\mathbf{z}(\varepsilon) = c_1(\mathbf{u}_1 + \varepsilon \mathbf{u}_2)$ into $\mathcal{F}$, we have

$$\mathcal{F}(\mathbf{z}(\varepsilon)) = c_1[m_1^2(1 - R(A^{-1}, A(A - I)\mathbf{z}(\varepsilon))a_1)\mathbf{u}_1 + m_2^2(1 - R(A^{-1}, A(A - I)\mathbf{z}(\varepsilon))a_2)\varepsilon \mathbf{u}_2].$$

Using properties [\(2.12\)](#) and [\(2.13\)](#) of Rayleigh quotients outlined in [Chapter 2](#), we have

$$\begin{aligned} R(A^{-1}, A(I - A)\mathbf{z}(\varepsilon)) &= R(A^{-1}, c_1 A(I - A)(\mathbf{u}_1 + \varepsilon \mathbf{u}_2)) \\ &= R(A^{-1}, a_1(1 - a_1)\mathbf{u}_1 + a_2(1 - a_2)\varepsilon \mathbf{u}_2) \\ &= \frac{\frac{1}{a_1} \|a_1(1 - a_1)\mathbf{u}_1\|^2 + \frac{1}{a_2} \|a_2(1 - a_2)\varepsilon \mathbf{u}_2\|^2}{\|a_1(1 - a_1)\mathbf{u}_1\|^2 + \|a_2(1 - a_2)\varepsilon \mathbf{u}_2\|^2} \\ &= \frac{a_1(1 - a_1)^2 + a_2(1 - a_2)^2 \varepsilon^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2 \varepsilon^2}. \end{aligned}$$

Now,

$$\begin{aligned} \mathcal{F}(\mathbf{z}(\varepsilon)) &= c_1 \left[\frac{a_2(1 - a_2)^2(a_2 - a_1)m_1^2 \varepsilon^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2 \varepsilon^2} \mathbf{u}_1 - \frac{a_1(1 - a_1)^2(a_2 - a_1)m_2^2 \varepsilon}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2 \varepsilon^2} \mathbf{u}_2 \right] \\ &= \underbrace{\frac{c_1(a_2 - a_1)a_2(1 - a_2)^2 m_1^2 \varepsilon^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2 \varepsilon^2}}_{=: \tilde{c}_1} \left[\mathbf{u}_1 + \underbrace{\left(\frac{-a_1(1 - a_1)^2 m_2^2}{a_2(1 - a_2)^2 m_1^2 \varepsilon} \right)}_{=: \tilde{\varepsilon}} \mathbf{u}_2 \right]. \quad (6.13) \end{aligned}$$

Therefore, another application of $\mathcal{F}$ results in

$$\mathcal{F}(\mathcal{F}(\mathbf{z}(\varepsilon))) = \frac{\tilde{c}_1(a_2 - a_1)a_2(1 - a_2)^2m_1^2\tilde{\varepsilon}^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2\tilde{\varepsilon}^2} \left[\mathbf{u}_1 + \left(\frac{-a_1(1 - a_1)^2m_2^2}{a_2(1 - a_2)^2m_1^2\tilde{\varepsilon}} \right) \mathbf{u}_2 \right].$$

Since $\frac{-a_1(1-a_1)^2m_2^2}{a_2(1-a_2)^2m_1^2\tilde{\varepsilon}} = \varepsilon$ by definition, we have

$$\mathcal{F}(\mathcal{F}(\mathbf{z}(\varepsilon))) = \frac{(a_2 - a_1)a_2(1 - a_2)^2m_1^2\varepsilon^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2\varepsilon^2} \frac{(a_2 - a_1)a_2(1 - a_2)^2m_1^2\tilde{\varepsilon}^2}{a_1^2(1 - a_1)^2 + a_2^2(1 - a_2)^2\tilde{\varepsilon}^2} [c_1(\mathbf{u}_1 + \varepsilon\mathbf{u}_2)].$$

Finally, substituting in the expression for $\tilde{\varepsilon}$ and $a_i = 1 - m_i$, after simplifying we get

$$\mathcal{F}(\mathcal{F}(\mathbf{z}(\varepsilon))) = \underbrace{\left(\frac{(m_2 - m_1)^2m_1^4m_2^4}{[m_1^2(m_1 - 1)^2 + m_2^2(m_2 - 1)^2\varepsilon^2][m_1^2 + m_2^2/\varepsilon^2]} \right)}_{\lambda(\varepsilon)} \mathbf{z}(\varepsilon),$$

proving our claim. $\square$

Now, we can exactly quantify the asymptotic convergence factor associated with initial residuals of the form presented in [Theorem 6.1.1](#), as the following theorem claims.

Theorem 6.1.2. Consider rAA(2) applied to a symmetric matrix $M \in \mathbb{R}^{n \times n}$. Let the initial residual have the form $\mathbf{r}_0 = c_i\mathbf{u}_i + c_j\mathbf{u}_j$, $i \neq j$, where $c_i, c_j \neq 0$ are real constants and $\mathbf{u}_i, \mathbf{u}_j$ are orthogonal eigenvectors of M with eigenvalues m_i and m_j respectively. If we denote $\varepsilon := c_j/c_i$, such that $\mathbf{r}_0 = c_i(\mathbf{u}_i + \varepsilon\mathbf{u}_j)$, then the r-linear convergence factor of rAA(2) applied to $\mathbf{r}_0$ is given by

$$\rho(\mathbf{r}_0) = \left(\frac{(m_j - m_i)^2m_i^4m_j^4}{[m_i^2(1 - m_i)^2 + m_j^2(1 - m_j)^2\varepsilon^2][m_i^2 + m_j^2/\varepsilon^2]} \right)^{\frac{1}{6}}. \quad (6.14)$$

Proof. The proof follows as a result of [Theorem 6.1.1](#) and a similar telescoping process to what was seen in the proof of [Proposition 3.2.2](#). Using [Proposition 6.1.1](#) and [Theorem 6.1.1](#), we have

$$\mathbf{r}_{k+6} = \mathcal{F}(\mathcal{F}(\mathbf{r}_k)) = \lambda(\varepsilon)\mathbf{r}_k, \quad k = 0, 6, 12, \dots$$

This implies that

$$\frac{\|\mathbf{r}_{k+6}\|}{\|\mathbf{r}_k\|} = \lambda(\varepsilon).$$

Therefore, we have the following r-linear convergence factor

$$\begin{aligned}\rho(\mathbf{r}_0) &= \lim_{k \rightarrow \infty} \|\mathbf{r}_k\|^{1/k} = \lim_{k \rightarrow \infty} \frac{\|\mathbf{r}_k\|^{1/k}}{\|\mathbf{r}_0\|^{1/k}} = \lim_{k \rightarrow \infty} \left(\prod_{j=0}^{k/6-1} \frac{\|\mathbf{r}_{6(j+1)}\|}{\|\mathbf{r}_{6j}\|} \right)^{1/k} \\ &= \lim_{k \rightarrow \infty} [\lambda(\varepsilon)^{k/6}]^{1/k} = \lambda(\varepsilon)^{1/6}.\end{aligned}$$

From the definition of $\lambda(\varepsilon)$ given in (6.12), we have our result. $\square$

While [Theorem 6.1.2](#) only holds for residuals that are defined by a linear combination of two eigenvectors, this theorem allows us to fully classify the convergence corresponding to 2×2 symmetric matrices, giving us the following corollary.

Corollary 6.1.1. Consider rAA(2) applied to a symmetric matrix $M \in \mathbb{R}^{2 \times 2}$. Let the initial residual have the form $\mathbf{r}_0 = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2$, where $c_1, c_2 \neq 0$ are real constants and $\mathbf{u}_1, \mathbf{u}_2$ are the orthogonal eigenvectors of M with eigenvalues m_1 and m_2 respectively. If we denote $\varepsilon := c_2/c_1$, such that $\mathbf{r}_0 = c_1(\mathbf{u}_1 + \varepsilon \mathbf{u}_2)$, then the r-linear convergence factor of rAA(2) applied to $\mathbf{r}_0$ is given by

$$\rho(\mathbf{r}_0) = \left(\frac{(m_2 - m_1)^2 m_1^4 m_2^4}{[m_1^2(1 - m_1)^2 + m_2^2(1 - m_2)^2 \varepsilon^2][m_1^2 + m_2^2/\varepsilon^2]} \right)^{\frac{1}{6}}. \quad (6.15)$$

Furthermore, the worst-case r-linear convergence factor among all initial residuals is given by

$$\rho^* = \sup_{\mathbf{r}_0 \in \mathbb{R}^2} \rho(\mathbf{r}_0) = \left[\frac{|m_2 - m_1| m_1^2 m_2^2}{m_1^2 |1 - m_1| + m_2^2 |1 - m_2|} \right]^{\frac{1}{3}}.$$

Proof. It is clear from [Theorem 6.1.2](#) that (6.15) holds where $\mathbf{r}_0$ is defined as a linear combination of the two eigenvectors of M . It can be found using typical calculus techniques that

$$\arg \max_{\varepsilon \in \mathbb{R}} \lambda(\varepsilon) = \pm \sqrt{\left| \frac{1 - m_1}{1 - m_2} \right|}.$$

Substituting the above value for ε in (6.15) results in the given value for ρ^* shown above. $\square$

In [Fig. 6.1](#) and [Fig. 6.2](#) a few illustrations are given that show off the convergence of rAA(2) for a 2×2 linear setting. In particular, [Fig. 6.2](#) appears to suggest that the

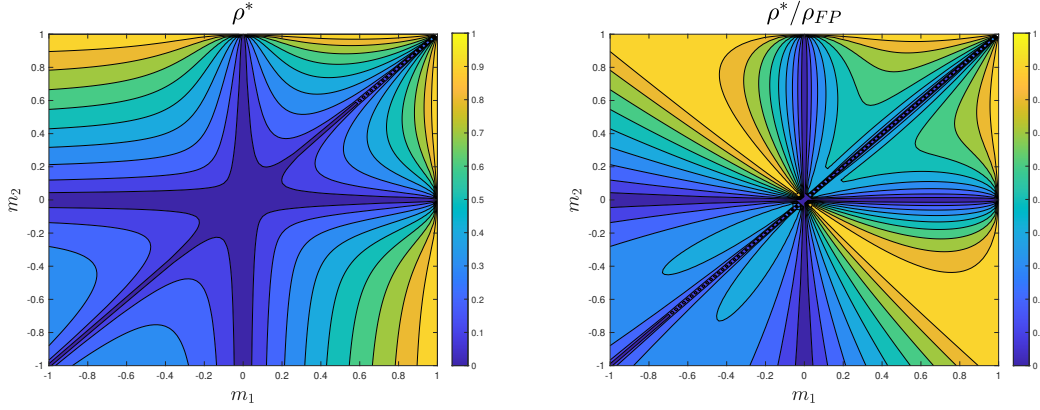


Figure 6.1: Convergence of rAA(2) for symmetric $M \in \mathbb{R}^{2 \times 2}$ with eigenvalues m_1 and m_2 . (Left) Worst-case asymptotic convergence factor of rAA(2) according to [Corollary 6.1.1](#). (Right) Ratio of worst-case asymptotic convergence factor of rAA(2) and convergence factor of fixed-point iteration.

convergence of AA(1) and rAA(2) are quite similar when applied to problem (3.5). In fact, it appears that the worst-case asymptotic convergence factor for rAA(2) may be an upper bound on the worst-case asymptotic convergence factor of AA(1). Proving this observation would lead to a more accurate convergence factor bound than that which was proved in [14]. A similar observation can be made for the restarted iteration presented in [8] and in the future, this idea will be explored in more detail.

6.2 Convergence of arAA

In this section, another restarted variant of AA(1) is presented in which restarts occur in an alternating fashion between every two and three iterations. We term this method *Alternating Restarted Anderson Acceleration* (arAA) and iterates are produced according to the rule

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) \quad k = 0, 5, 10, \dots, \quad (6.16a)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 1, 6, 11, \dots, \quad (6.16b)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) \quad k = 2, 7, 12, \dots, \quad (6.16c)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 3, 8, 13, \dots, \quad (6.16d)$$

$$\mathbf{x}_{k+1} = \mathbf{q}(\mathbf{x}_k) + \beta_k(\mathbf{q}(\mathbf{x}_k) - \mathbf{q}(\mathbf{x}_{k-1})) \quad k = 4, 9, 14, \dots \quad (6.16e)$$

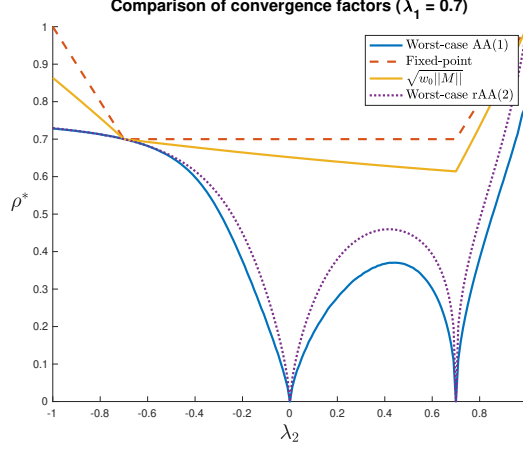


Figure 6.2: Worst-case asymptotic convergence factor for AA(1) (blue) applied to (3.5) as a function of λ_2 with λ_1 fixed compared with the associated convergence factor for the underlying fixed-point iteration (red), the bound for the convergence factor (yellow) given by Theorem 2.3.1, and the worst-case asymptotic convergence factor for rAA(2) (purple) according to Corollary 6.1.1.

In our linear setting, with A invertible, we have the associated residual iteration

$$\mathbf{r}_{k+1} = M\mathbf{r}_k \quad k = 0, 5, 10, \dots, \quad (6.17a)$$

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}))] \quad k = 1, 6, 11, \dots, \quad (6.17b)$$

$$\mathbf{r}_{k+1} = M\mathbf{r}_k \quad k = 2, 7, 12, \dots, \quad (6.17c)$$

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}))] \quad k = 3, 8, 13, \dots, \quad (6.17d)$$

$$\mathbf{r}_{k+1} = M[\mathbf{r}_k + \beta_k(\mathbf{r}_k - \mathbf{r}_{k-1}))] \quad k = 4, 9, 14, \dots \quad (6.17e)$$

In essence, the iteration described above is an alternating combination of the rAA(2) method presented in the previous section and the method presented in [8]. We analyze this method in a similar fashion to the previous section. To begin, suppose $k \in \{0, 5, 10, \dots\}$ and consider the residual vector $\mathbf{r}_k$ generated according to (6.17). Then, following (6.7), we have

$$\mathbf{r}_{k+2} = \mathcal{R}(\mathbf{r}_k) := M[I - R(A^{-1}, A\mathbf{r}_k)A]\mathbf{r}_k.$$

Here we denote by $\mathcal{R} : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^n$ the *two-step residual propagation operator* as first seen in [8]. Following this update, we restart the method and perform a fixed-point iteration to find the next residual. Then, using the notation introduced in the previous section, we have the residual

$$\mathbf{r}_{k+5} = \mathcal{F}(\mathbf{r}_{k+2}) = M^2 [I - R(A^{-1}, A(I - A)\mathbf{r}_{k+2})A] \mathbf{r}_{k+2},$$

according to (6.17). Putting these updates together, we can express (6.17) using the five-step update

$$\mathbf{r}_{k+5} = (\mathcal{F} \circ \mathcal{R})(\mathbf{r}_k), \quad k = 0, 5, 10, \dots \quad (6.18)$$

The following proposition is similar to Theorem 6.1.1 in the sense that it characterizes the behaviour of the operator featured in (6.18) on vectors that are linear combinations of two eigenvectors of M .

Proposition 6.2.1. Let $c_i, c_j \in \mathbb{R}$ be two nonzero constants and $\mathbf{u}_i, \mathbf{u}_j$, $i \neq j$, denote two orthogonal eigenvectors of M with eigenvalues m_i and m_j respectively. Setting $\varepsilon = c_j/c_i$, we have

$$(\mathcal{F} \circ \mathcal{R})(c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)) = \hat{C} \left[\mathbf{u}_i + \frac{m_i}{m_j} \varepsilon \mathbf{u}_j \right], \quad (6.19)$$

where the constant $\hat{C}$ is given by

$$\hat{C} = \frac{c_i(m_j - m_i)^2 m_i^3 m_j^4 \varepsilon^2}{[(1 - m_i)^2 + (1 - m_j)^2 \varepsilon^2][m_i^4 \varepsilon^2 + m_j^4]}. \quad (6.20)$$

Proof. Consider applying $\mathcal{R}$ to the vector $c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)$ first. As shown in [8], this gives us

$$\mathcal{R}(c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)) = \underbrace{\frac{c_i(m_j - m_i)(1 - m_j)m_i \varepsilon^2}{(1 - m_i)^2 + (1 - m_j)^2 \varepsilon^2}}_{=: \tilde{c}} \left[\mathbf{u}_i + \underbrace{\left(\frac{-m_j(1 - m_i)}{m_i(1 - m_j)\varepsilon} \right)}_{=: \tilde{\varepsilon}} \mathbf{u}_j \right]. \quad (6.21)$$

By substituting $a_i = 1 - m_i$ in (6.13), we then have

$$(\mathcal{F} \circ \mathcal{R})(c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)) = \frac{\tilde{c}(m_j - m_i)(1 - m_j)m_j^2 m_i^2 \tilde{\varepsilon}^2}{m_i^2(1 - m_i)^2 + m_j^2(1 - m_j)^2 \tilde{\varepsilon}^2} \left[\mathbf{u}_1 + \left(-\frac{1 - m_i}{1 - m_j} \frac{1}{\tilde{\varepsilon}} \right) \mathbf{u}_2 \right].$$

Now, plugging in $\tilde{c}$ and $\tilde{\varepsilon}$ according to (6.21) and simplifying, we get the desired result. $\square$

A critical observation to make of the expression in (6.19) is the multiplication by m_i/m_j in the second eigenvector. A by-product of this fact is that repeated applications of the operator $\mathcal{F} \circ \mathcal{R}$ result in more multiplications by m_i/m_j . Specifically, K applications of this operator takes the form

$$(\mathcal{F} \circ \mathcal{R})^K(c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)) = C \left[\mathbf{u}_i + \left(\frac{m_i}{m_j} \right)^K \varepsilon \mathbf{u}_j \right], \quad (6.22)$$

for some constant $C \in \mathbb{R}$. As the iteration progresses, this suggests that if $|m_i| > |m_j|$, then the coefficient of $\mathbf{u}_j$ explodes and if $|m_i| < |m_j|$, then the coefficient vanishes. In other words, arAA converges in the direction associated with the smallest eigenvalue. This fact also leads to the following convergence behaviour.

Theorem 6.2.1. Consider arAA applied to a symmetric matrix $M \in \mathbb{R}^{n \times n}$. Let the initial residual have the form $\mathbf{r}_0 = c_i \mathbf{u}_i + c_j \mathbf{u}_j$, $i \neq j$, where $c_i, c_j \neq 0$ are real constants and $\mathbf{u}_i, \mathbf{u}_j$ are orthogonal eigenvectors of M with eigenvalues m_i and m_j respectively. Denote $\varepsilon := c_j/c_i$, such that $\mathbf{r}_0 = c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)$. If $m_i = -m_j$, then the r-linear convergence factor of arAA applied to $\mathbf{r}_0$ is given by

$$\rho(\mathbf{r}_0) = \left(\frac{4m_i^5 \varepsilon^2}{[(1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2][\varepsilon^2 + 1]} \right)^{1/5}, \quad (6.23)$$

otherwise, the residuals are r-superlinearly convergent.

Proof. Suppose $m_i = -m_j$ and $\mathbf{r}_k = c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)$ for some $k \in \{0, 5, 10, \dots\}$. Then, using [Proposition 6.2.1](#), we have

$$\begin{aligned} \mathbf{r}_{k+5} &= (\mathcal{F} \circ \mathcal{R})\mathbf{r}_k = (\mathcal{F} \circ \mathcal{R})[c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)] \\ &= \frac{4m_i^5 \varepsilon^2}{((1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2)(\varepsilon^2 + 1)} [c_i(\mathbf{u}_i - \varepsilon \mathbf{u}_j)]. \end{aligned}$$

Another application of $\mathcal{F} \circ \mathcal{R}$ results in

$$\begin{aligned} \mathbf{r}_{k+10} &= \left(\frac{4m_i^5 \varepsilon^2}{((1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2)(\varepsilon^2 + 1)} \right)^2 [c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)] \\ &= \left(\frac{4m_i^5 \varepsilon^2}{((1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2)(\varepsilon^2 + 1)} \right)^2 \mathbf{r}_k \end{aligned}$$

Therefore, we have

$$\frac{\|\mathbf{r}_{k+10}\|}{\|\mathbf{r}_k\|} = \left(\frac{4m_i^5 \varepsilon^2}{((1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2)(\varepsilon^2 + 1)} \right)^2.$$

Thus, using a similar process to [Theorem 6.1.2](#), we have the following r-linear convergence factor

$$\begin{aligned} \rho(\mathbf{r}_0) &= \lim_{k \rightarrow \infty} \|\mathbf{r}_k\|^{1/k} = \lim_{k \rightarrow \infty} \frac{\|\mathbf{r}_k\|^{1/k}}{\|\mathbf{r}_0\|^{1/k}} = \lim_{k \rightarrow \infty} \left(\prod_{j=0}^{k/10-1} \frac{\|\mathbf{r}_{10(j+1)}\|}{\|\mathbf{r}_{10j}\|} \right)^{1/k} \\ &= \left(\frac{4m_i^5 \varepsilon^2}{[(1 - m_i)^2 + (1 + m_i)^2 \varepsilon^2][\varepsilon^2 + 1]} \right)^{1/5}. \end{aligned}$$

We now handle the case in which $|m_i| \neq |m_j|$. If $\mathbf{r}_k = c_i(\mathbf{u}_i + \varepsilon \mathbf{u}_j)$ for some $k \in \{0, 5, 10, \dots\}$, then using [Proposition 6.2.1](#) gives

$$\begin{aligned} \frac{\|\mathbf{r}_{k+5}\|^2}{\|\mathbf{r}_k\|^2} &= \frac{\hat{C}^2 \left[1 + \left(\frac{m_i}{m_j} \right)^2 \varepsilon^2 \right]}{c_i^2 [1 + \varepsilon^2]} \\ &= \underbrace{\left[\frac{(m_j - m_i)^4 m_j^8 m_i^6 \varepsilon^4}{((1 - m_i)^2 + (1 - m_j)^2 \varepsilon^2)^2 (m_i^4 \varepsilon^2 + m_j^4)^2} \right]}_{f(\varepsilon)} \underbrace{\left[\frac{1 + \left(\frac{m_i}{m_j} \right)^2 \varepsilon^2}{1 + \varepsilon^2} \right]}_{g(\varepsilon)}. \end{aligned}$$

Recall from (6.22) that repeated applications of $\mathcal{F} \circ \mathcal{R}$ scales the second eigenvector according to m_i/m_j . This means that the ε used in the above equation either goes to infinity or zero as k increases. Therefore, it is sufficient to understand the behaviour of $\|\mathbf{r}_{k+5}\| / \|\mathbf{r}_k\|$ as $k \rightarrow \infty$ by analyzing the functions $f(\varepsilon)$ and $g(\varepsilon)$ as $\varepsilon \rightarrow \infty$ and $\varepsilon \rightarrow 0$. We can then recognize that $\lim_{\varepsilon \rightarrow \infty} f(\varepsilon) = 0 = \lim_{\varepsilon \rightarrow 0} f(\varepsilon)$, and $\lim_{\varepsilon \rightarrow \infty} g(\varepsilon) = (m_1/m_2)^2$ and $\lim_{\varepsilon \rightarrow 0} g(\varepsilon) = 1$. Therefore, we can say

$$\lim_{k \rightarrow \infty} \frac{\|(\mathcal{F} \circ \mathcal{R})(\mathbf{r}_k)\|}{\|\mathbf{r}_k\|} = 0,$$

showing the r-superlinear convergence of the method according to (2.3.2). $\square$

The following corollary now follows directly as a result of the above theorem.

Corollary 6.2.1. Consider arAA applied to a symmetric matrix $M \in \mathbb{R}^{2 \times 2}$. Let the initial residual have the form $\mathbf{r}_0 = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2$, where $c_1, c_2 \neq 0$ are real constants and $\mathbf{u}_1, \mathbf{u}_2$ are the orthogonal eigenvectors of M with eigenvalues m_1 and m_2 respectively. Denote $\varepsilon := c_2/c_1$, such that $\mathbf{r}_0 = c_1(\mathbf{u}_1 + \varepsilon \mathbf{u}_2)$. If $m_1 = -m_2$, then the r-linear convergence factor of arAA applied to $\mathbf{r}_0$ is given by

$$\rho(\mathbf{r}_0) = \left(\frac{4m_1^5 \varepsilon^2}{[(1 - m_1)^2 + (1 + m_1)^2 \varepsilon^2][\varepsilon^2 + 1]} \right)^{1/5}, \quad (6.24)$$

otherwise, the residuals are r-superlinearly convergent.

In the future, more investigation into the convergence of arAA on higher-dimensional nonlinear problems is of interest. The superlinear convergence described above is appealing for an iteration method and this faster convergence is rather unusual considering AA(1)

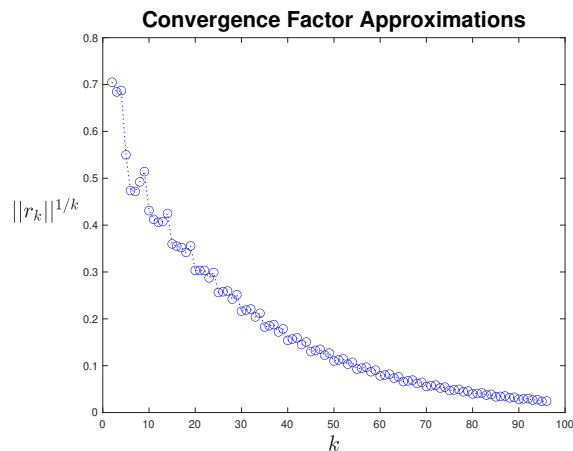


Figure 6.3: Approximate r-linear convergence factor $\|\mathbf{r}_k\|^{1/k}$ as a function of iteration number k for arAA applied to (3.5) with $\lambda_1 = 0.7$, $\lambda_2 = -0.3$ for a random initial residual $\mathbf{r}_0$.

takes more efficient steps on a per-iteration basis. In general, exploring different restarted AA methods is enticing as convergence is more easily quantified and there are clearly some examples where asymptotic convergence of restarted iterations is superior to that of the traditional AA method.

Chapter 7

Conclusion

Quantifying the r -linear convergence factor for AA iterations is a challenging problem, even in simple linear settings. While it is frequently observed that AA improves the convergence rate over the underlying fixed-point iteration, a complete theory outlining this improvement and its dependence on the initial iterate is unavailable in the current literature.

In [Chapter 3](#), we analyzed the r -linear convergence factor corresponding to AA(1) applied to a symmetric and skew-symmetric class of linear iteration functions $\mathbf{q}$. In the symmetric case, it was discovered that the worst r -linear convergence factor among all initial iterates did not improve upon the convergence rate of the fixed-point iteration. This highlighted a class of iterations where AA does not achieve any acceleration in the worst case but has quantifiable convergence improvement for all other initial iterates. In the skew-symmetric case, it was discovered that the r -linear convergence factor was independent of the initial iterate. This independence on the initial iterate results in AA(1) always improving upon the fixed-point iteration and achieving convergence for maps where the regular fixed-point iteration fails. The analysis of these simple linear maps provides a glimpse into the complicated convergence behaviour possible of AA iterations. In the future, developing more general results for the asymptotic convergence of AA applied to linear settings will be explored.

The occurrence of stagnation in AA(m) iterations is studied in [Chapter 4](#) in both linear and nonlinear settings. There, conditions are presented that are necessary for an iteration to see stagnation. These results can be used to inform users of AA(m) when stagnation may occur but the given analysis does not say anything about the asymptotic repercussions of stagnation. As highlighted in [Fig. 4.2](#), AA iterations that stagnate can still achieve better r -linear convergence factors than nonstagnating iterations. A future investigation of the

asymptotic behaviour of iterates where $AA(m)$ stagnates would be interesting.

In [Chapter 5](#), expressions are found for the acceleration coefficients in terms of the relative angles and relative lengths of previous residuals. Further, residual bounds are derived using these expressions that make the optimization gain at each step explicit in terms of previous residuals. These one-step residual bounds are insightful for understanding the acceleration AA provides at each iteration but these bounds do not tell us anything about the asymptotic behaviour. In the future, it would be interesting to explore how these bounds might be used to show an improved r -linear convergence factor over the underlying fixed-point iteration. Bounding the angle between previous residuals (such as what was shown in [\[14\]](#)) may be a reasonable avenue to further strengthen these residual bounds.

Lastly, we explored a few different restarted variations of AA in [Chapter 6](#). As highlighted by the convergence result given for $arAA$, restarting AA in a certain manner may result in improved asymptotic convergence over the regular windowed AA algorithm. As a result, it may be interesting to further investigate restarted algorithms as they may perform better. Additionally, these algorithms may provide a means to understand the convergence of AA . As seen in [Fig. 6.2](#), the worst-case r -linear convergence factor of $rAA(2)$ appears to bound the worst-case r -linear convergence of $AA(1)$. Proving this bound exists would allow us to state a bound for $AA(1)$ that better represents the acceleration provided by AA than the bound shown in [\[14\]](#). Overall, despite the difficulty in quantifying the asymptotic convergence of AA , there are several approaches one might take to show improved convergence.

References

- [1] Donald G Anderson. Iterative procedures for nonlinear integral equations. *Journal of the ACM (JACM)*, 12(4):547–560, 1965.
- [2] Allison H Baker, Elizabeth R Jessup, and Thomas Manteuffel. A technique for accelerating the convergence of restarted GMRES. *SIAM Journal on Matrix Analysis and Applications*, 26(4):962–984, 2005.
- [3] Stefan Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta mathematicae*, 3(1):133–181, 1922.
- [4] Jakub Wiktor Both, Kundan Kumar, Jan Martin Nordbotten, and Florin Adrian Radu. Anderson accelerated fixed-stress splitting schemes for consolidation of unsaturated porous media. *Computers & Mathematics with Applications*, 77(6):1479–1502, 2019.
- [5] Augustin Cauchy et al. Méthode générale pour la résolution des systemes d’équations simultanées. *Comp. Rend. Sci. Paris*, 25(1847):536–538, 1847.
- [6] Matt Dallas and Sara Pollock. Newton-Anderson at singular points. *arXiv preprint arXiv:2207.12334*, 2022.
- [7] Hans De Sterck, Yunhui He, and Oliver A Krzysik. Anderson acceleration as a Krylov method with application to convergence analysis. *Journal of Scientific Computing*, 99(1):12, 2024.
- [8] Hans De Sterck, Oliver A Krzysik, and Adam Smith. On the convergence of restarted Anderson acceleration for symmetric linear systems. *arXiv preprint arXiv:2312.04776*, 2023.

- [9] JC Ehiwario and SO Aghamie. Comparative study of bisection, Newton-Raphson and secant methods of root-finding problems. *IOSR Journal of Engineering*, 4(4):01–07, 2014.
- [10] Mark Embree. The tortoise and the hare restart GMRES. *SIAM review*, 45(2):259–266, 2003.
- [11] Claire Evans, Sara Pollock, Leo G Rebholz, and Mengying Xiao. A proof that Anderson acceleration improves the convergence rate in linearly converging fixed-point methods (but not in those converging quadratically). *SIAM Journal on Numerical Analysis*, 58(1):788–810, 2020.
- [12] Volker Eyert. A comparative study on methods for convergence acceleration of iterative vector sequences. *Journal of Computational Physics*, 124(2):271–285, 1996.
- [13] Haw-ren Fang and Yousef Saad. Two classes of multiseant methods for nonlinear acceleration. *Numerical linear algebra with applications*, 16(3):197–221, 2009.
- [14] Casey Garner, Gilad Lerman, and Teng Zhang. Improved convergence rates of Anderson acceleration for a large class of fixed-point iterations. *arXiv preprint arXiv:2311.02490*, 2023.
- [15] Anne Greenbaum, Vlastimil Pták, and Zdeněk Strakoš. Any nonincreasing convergence curve is possible for GMRES. *Siam journal on matrix analysis and applications*, 17(3):465–469, 1996.
- [16] Nicholas C Henderson and Ravi Varadhan. Damped Anderson acceleration with restarts and monotonicity control for accelerating em and em-like algorithms. *Journal of Computational and Graphical Statistics*, 28(4):834–846, 2019.
- [17] Wayne Joubert. On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems. *Numerical linear algebra with applications*, 1(5):427–447, 1994.
- [18] Carl T Kelley. *Iterative methods for linear and nonlinear equations*. SIAM, 1995.
- [19] Gerard Meurant. The complete stagnation of GMRES for $n \geq 4$. *Electron. Trans. Numer. Anal.*, 39:75–101, 2012.
- [20] Gérard Meurant. Necessary and sufficient conditions for GMRES complete and partial stagnation. *Applied Numerical Mathematics*, 75:100–107, 2014.

- [21] Jorge J Moré, Burton S Garbow, and Kenneth E Hillstom. Testing unconstrained optimization software. *ACM transactions on mathematical software (TOMS)*, 7(1):17–41, 1981.
- [22] Ronald B Morgan. A restarted GMRES method augmented with eigenvectors. *SIAM Journal on Matrix Analysis and Applications*, 16(4):1154–1171, 1995.
- [23] Peng Ni. *Anderson acceleration of fixed-point iteration with applications to electronic structure computations*. PhD thesis, Worcester Polytechnic Institute Worcester, MA, USA, 2009.
- [24] Ben Noble, James W Daniel, et al. *Applied linear algebra*, volume 477. Prentice-Hall Englewood Cliffs, NJ, 1977.
- [25] James M Ortega and Werner C Rheinboldt. *Iterative solution of nonlinear equations in several variables*. SIAM, 2000.
- [26] Wenqing Ouyang, Yang Liu, and Andre Milzarek. Descent properties of an Anderson accelerated gradient method with restarting. *SIAM Journal on Optimization*, 34(1):336–365, 2024.
- [27] Christopher C Paige and Michael A Saunders. Solution of sparse indefinite systems of linear equations. *SIAM journal on numerical analysis*, 12(4):617–629, 1975.
- [28] Yue Peng, Bailin Deng, Juyong Zhang, Fanyu Geng, Wenjie Qin, and Ligang Liu. Anderson acceleration for geometry optimization and physics simulation. *ACM Transactions on Graphics (TOG)*, 37(4):1–14, 2018.
- [29] Sara Pollock and Leo G Rebholz. Anderson acceleration for contractive and noncontractive operators. *IMA Journal of Numerical Analysis*, 41(4):2841–2872, 2021.
- [30] Sara Pollock and Hunter Schwartz. Benchmarking results for the Newton–Anderson method. *Results in Applied Mathematics*, 8:100095, 2020.
- [31] Florian A Potra and Hans Engler. A characterization of the behavior of the Anderson acceleration on linear problems. *linear Algebra and its Applications*, 438(3):1002–1011, 2013.
- [32] Michael JD Powell. An iterative method for finding stationary values of a function of several variables. *The Computer Journal*, 5(2):147–151, 1962.

- [33] Leo G Rebholz and Mengying Xiao. The effect of Anderson acceleration on superlinear and sublinear convergence. *Journal of Scientific Computing*, 96(2):34, 2023.
- [34] Youcef Saad and Martin H Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3):856–869, 1986.
- [35] Yousef Saad. Krylov subspace methods for solving large unsymmetric linear systems. *Mathematics of computation*, 37(155):105–126, 1981.
- [36] Yousef Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.
- [37] Valeria Simoncini et al. On a non-stagnation condition for GMRES and application to saddle point matrices. *Electronic Transactions on Numerical Analysis*, 37(202-213):25, 2010.
- [38] Ke Sun, Yafei Wang, Yi Liu, Bo Pan, Shangling Jui, Bei Jiang, Linglong Kong, et al. Damped Anderson mixing for deep reinforcement learning: Acceleration, convergence, and stabilization. *Advances in Neural Information Processing Systems*, 34:3732–3743, 2021.
- [39] Alex Toth and Carl T Kelley. Convergence analysis for Anderson acceleration. *SIAM Journal on Numerical Analysis*, 53(2):805–819, 2015.
- [40] Alex Toth, CT Kelley, Stuart R Slattery, Steven P Hamilton, Kevin T Clarno, and RPP Pawlowski. Analysis of Anderson acceleration on a simplified neutronics/thermal hydraulics system. Technical report, Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States). Consortium for . . . , 2015.
- [41] Ilya Zavorin, DianneP O’Leary, and Howard Elman. Complete stagnation of GMRES. *Linear Algebra and its Applications*, 367:165–183, 2003.