SIAM REVIEW
Vol. 60, No. 1, pp. 3–55

Inverse, Shifted Inverse, and Rayleigh Quotient Iteration as Newton's Method*

Richard A. Tapia†
J. E. Dennis, Jr.‡
Jan P. Schäfermeyer§

Dedicated to Roland Glowinski on the Occasion of his 80th Birthday

Abstract. The l_2 normalized inverse, shifted inverse, and Rayleigh quotient iterations are classic algorithms for approximating an eigenvector of a symmetric matrix. This work establishes rigorously that each iterate produced by one of these three algorithms can be viewed as a Newton's method iterate followed by a normalization. The equivalences given here are not meant to suggest changes to the implementations of the classic eigenvalue algorithms. However, they add further understanding to the formal structure of these iterations, and they provide an explanation for their good behavior despite the possible need to solve systems with nearly singular coefficient matrices. A historical development of these eigenvalue algorithms is presented. Using our equivalences and traditional Newton's method theory helps to gain understanding as to why normalized Newton's method, inverse iteration, and shifted inverse iteration are only linearly convergent and not quadratically convergent, as would be expected, and why a new linear system need not be solved at each iteration. We also give an explanation as to why our normalized Newton's method equivalent of Rayleigh quotient iteration is cubically convergent and not just quadratically convergent, as would be expected.

Key words. inverse power method, inverse iteration, shifted inverse iteration, Rayleigh quotient iteration, Newton's method, multiplier substitution method

AMS subject classifications. 65F15, 49M37, 49M15, 65K05

DOI. 10.1137/15M1049956

*Received by the editors November 30, 2015; accepted for publication (in revised form) January 20, 2017; published electronically February 7, 2018. The material presented in this paper began with discussions between the first two authors in 1986. The material has been the subject of numerous presentations over these past 30 years. One such talk was given by the first author at the IPAM at UCLA in June of 2012 at the celebration of Tony Chan's 60th Birthday. Discussions with various attendees, including *SIAM Review* Editorial Board representation, motivated the writing of this paper. Moreover, it was suggested that the paper contain appropriate historical content and pursue a somewhat more informal approach than the standard terse mathematical paper. We have attempted to do this.

http://www.siam.org/journals/sirev/60-1/M104995.html

Funding: The work of the first author was supported in part by funds associated with the Maxfield-Oshman Chair in Engineering at Rice University.

†Corresponding author. Department of Computational and Applied Mathematics, Rice University, Houston, TX 77251-1892 (rat@rice.edu).

‡Department of Computational and Applied Mathematics, Rice University, Houston, TX 77251-1892, and University of Washington, Seattle, WA 98195 (dennis@rice.edu).

§Berlin, Germany (jpschaefermeyer@gmail.com).

Contents

PART I: Motivation and History

1. Introduction. The initial objective of this study was to answer the question: In what sense, if any, can Rayleigh quotient iteration (RQI) be viewed as Newton's method? A by-product of our study is that we also are able to provide an answer to the question: In what sense, if any, can inverse iteration and shifted inverse iteration be viewed as Newton's method? Since two of the authors work in the area of

nonlinear programming, it is satisfying that the equivalences we demonstrate between l_2 normalized Newton's method and l_2 normalized inverse iteration, shifted inverse iteration, and RQI are constructed by turning to three classic algorithms from that literature.

Equivalences between algorithms from different disciplines can be of value. They not only aid in understanding the algorithms under discussion, but they may also facilitate the design of improved algorithms. In this paper we work exclusively in the finite-dimensional vector space R^n. The various p norms on R^n will be referred to as l_p norms and denoted by $\| \ \|_p$. When no subscript is used we mean an arbitrary norm. All operator norms are the norms induced by the underlying normed vector spaces. By a unit vector $x \in R^n$ we mean an x such that $\|x\| = 1$. The choice of norm will be clear from the context; if not, it will be stated. In describing an iterative process we will tend to use integer subscripts on the iterates. However, we will also find it convenient to write the current iterate with no subscript and use the subscript "+" to denote the subsequent iterate. Hence the notation $x_{k+1} = S(x_k)$ or $x_+ = S(x)$ means that x_k or x is the current iterate and x_{k+1} or x_+ is the subsequent iterate. We will try to use the letter σ for eigenvalue estimates and the letter λ for multiplier estimates, though a large part of our theory concerns situations in which the two coincide.

1.1. Overview of Paper. Technical preliminaries concerning standard convergence notions in computational mathematics are given in Appendix A. These notions will be important for some distinctions in various theories or algorithms discussed in the paper. The reader not familiar with these notions will gain a heightened appreciation for the material in the paper by first becoming familiar with them in Appendix A.

Part I of this paper consists of sections 1–6 where we motivate equivalences to Newton's method and present a fairly complete historical development of three popular algorithms for the algebraic eigenvalue problem. In section 1.2 we briefly introduce Newton's method and in section 1.3 we introduce the inverse iteration, shifted inverse iteration, and RQI algorithms for the symmetric eigenvalue problem.

A basic premise of this paper is that when an algorithm for a problem requires the solution of a new square linear system at each iteration and has been "tuned and pruned," i.e., has gone through a series of modifications (improvements), so that it is effective and fast, then it is highly likely that it is a form of Newton's method applied to some form of the original problem. As the reader will soon see, RQI is a prime example of such activity. A historical development of inverse iteration is given in section 2, of shifted inverse iteration in section 3, and of RQI in section 4. Several applications of Newton's method for the algebraic eigenvalue problem from the literature are discussed in section 5, most notably those of Peters and Wilkinson [51] and Tapia and Whitley [73].

It is a common belief that an RQI iterate is identical to the iterate obtained from a form of Newton's method, yet in our opinion such an equivalence has not been properly established in the literature. We succeed because we use a different formulation of the problem as a nonlinear equation to which we apply Newton's method. Hence, in section 6 we set the stage for our complete equivalence, given in section 10, by discussing several of the better known implied equivalences and highlighting their shortcomings in terms of being complete equivalences. In this section we also show that the Tapia–Whitley understanding of Newton's method for the eigenvalue problem leads to a derivation of RQI in a short, clean, and precise alternative manner.

Part II of this paper consists of sections 7–15. The main thrust of Part II is the establishment of algorithmic equivalences between the three popular algorithms for

the algebraic eigenvalue problem and a form of Newton's method. In section 7 we give the historical development of three equality constrained optimization algorithms that form the basis of our normalized Newton's method equivalence results presented in the subsequent three sections. In sections 8, 9, and 10 we demonstrate that the iterates obtained from inverse iteration, shifted inverse iteration, and RQI, respectively, can each be viewed as an iterate obtained from a form of normalized Newton's method. In section 13 we apply standard Newton's method theory to our normalized Newton's method equivalence of RQI in an effort to explain why the algorithm gives cubic convergence and not just quadratic convergence as the standard Newton theory might suggest. The lengthy and challenging task of demonstrating the result that the constant in the cubic convergence of our normalized Newton's method equivalence of RQI is less than or equal to 1 is presented in section 14. Section 15 contains summarizing and concluding statements.

1.2. Newton's Method. Consider $F : R^n \to R^n$ and let $x_* \in R^n$ be a zero of F, i.e., $F(x_*) = 0$. Given x_0, Newton's method constructs the sequence

$$x_{k+1} = x_k + \Delta x_k, \quad k = 0, 1, \ldots, \tag{1.1}$$

by solving the square linear system

$$F'(x_k)\Delta x_k = -F(x_k) \quad \text{for} \quad \Delta x_k. \tag{1.2}$$

Constructions (1.1) and (1.2) can be concisely written as

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k) \tag{1.3}$$

when $F'(x_k)$ is invertible. Inherent in this formulation is the hope that $\{x_k\}$ will converge to a zero of F.

Normalized Newton (NN). For the nonlinear equation problem

$$F(x) = 0,$$

when it is known that x_*, the solution of interest, has the property that $\|x_*\| = 1$, we modify Newton's method as follows:

(a) Solve $\quad F'(x_k)\Delta x_k = -F(x_k).$

(b) Normalize $\quad x_{k+1} = \dfrac{x_k + \Delta x_k}{\|x_k + \Delta x_k\|}.$

We shall refer to this as the normalized Newton's method.

Newton-like Methods. If instead of solving for Δx_k from (1.2) we solve for Δx_k from

$$M_k\Delta x_k = -F(x_k), \tag{1.4}$$

where M_k is a given $n \times n$ matrix, then the iterative method (1.1) is said to be a Newton-like method. In the literature Newton-like methods are also called Newton-type or Newton-based methods. Of course, Newton-like methods may also be normalized.

1.3. Inverse, Shifted Inverse, and Rayleigh Quotient Iteration. In the algorithms considered we restrict our attention to the symmetric eigenvalue problem. Let A be a symmetric $n \times n$ matrix.

The Symmetric Eigenvalue Problem. Find $(x_* \neq 0, \sigma_*) \in R^n \times R$ such that

$$Ax_* = \sigma_* x_*.$$

The Rayleigh Quotient. For symmetric A the Rayleigh quotient

$$\sigma_R(x) = \frac{x^T A x}{x^T x} \quad (x \neq 0) \tag{1.5}$$

has the following enhanced accuracy property: If (x_*, σ_*) is an eigenpair for A, then for any x we have

$$|\sigma_R(x) - \sigma_*| = O(\|x - x_*\|^2). \tag{1.6}$$

This property follows from the well-known fact that at an eigenvector x_* the Rayleigh quotient is stationary, i.e., $\nabla \sigma_R(x_*) = 0$. This fact is established in section 9.

The following algorithmic definitions follow those given in Parlett [49].

Inverse Iteration (II). Choose a unit vector x_0 and for $k = 0, 1, \ldots$, repeat the following:

(a) Solve $\quad Ay_{k+1} = x_k \quad$ for $\quad y_{k+1}$.

(b) Normalize $\quad x_{k+1} = \dfrac{y_{k+1}}{\|y_{k+1}\|}$.

Shifted Inverse Iteration (SII). Choose a scalar σ and a unit vector x_0 and for $k = 0, 1, \ldots$, repeat the following:

(a) Solve $\quad (A - \sigma I) y_{k+1} = x_k \quad$ for $\quad y_{k+1}$.

(b) Normalize $\quad x_{k+1} = \dfrac{y_{k+1}}{\|y_{k+1}\|}$.

The scalar σ is called the *shift*.

Rayleigh Quotient Iteration. Choose a unit vector x_0 and for $k = 0, 1, \ldots$, repeat the following:

(a) Solve $\quad (A - \sigma_R(x_k) I) y_{k+1} = x_k \quad$ for $\quad y_{k+1}$.

(b) Normalize $\quad x_{k+1} = \dfrac{y_{k+1}}{\|y_{k+1}\|}$.

Most texts do not make a formal distinction between inverse iteration and shifted inverse iteration, since the former is the special case of the latter where $\sigma = 0$. We prefer to make a distinction because of the differences in the corresponding equivalent Newton's method formulations we give in sections 8 and 9.

2. The Historical Development of Inverse Iteration and the Power Method. The power method and inverse iteration are finite-dimensional examples of the so-called *method of successive approximations.* Hence, we first touch briefly on the origin of this latter method. It is often attributed to Picard, who wrote extensively about it and also used it in his existence proof for ODE initial value problems [52]. However, already in 1747 Clairaut had presented the method to the French Academy

of Sciences and, along with d'Alembert and Euler, had used it to compute the Moon's orbit around the Earth [23].

We will focus here on the application of the method of successive approximations to eigenvalue problems. In this context it was first used in 1830 by Liouville [37] as an analytical tool to prove the existence of eigenfunctions of the one-dimensional heat equation. We also briefly mention a graphical procedure for the eigenvalue computation of one-dimensional elastic continua known as Vianello's [75] or Stodola's method [71], which was shown to be equivalent to the method of successive approximations by von Mises [38].

Since the power method enters into our historical discussions, we give it a formal definition.

Power Method. Choose a unit vector x_0 and for $k = 0, 1, \ldots$, repeat the following:

(a) Form $\quad y_{k+1} = Ax_k.$

(b) Normalize $\quad x_{k+1} = \dfrac{y_{k+1}}{\|y_{k+1}\|}.$

2.1. Historical Development.

1885 – Schwarz. In the second part of his paper on minimal surfaces [62], H.A. Schwarz proved the existence of the lowest eigenfunction of the two-dimensional Laplacian by first using the method of successive approximations to construct an infinite sequence of functions w_n,

$$\Delta w_n + p \cdot w_{n-1} = 0, \quad n = 1, 2, \ldots,$$

where $p > 0$ is a bounded continuous scalar valued function. He then showed that the infinite sum of the functions w_n converges to a limiting function which is the eigenfunction. Schwarz's paper was highly influential in the development of functional analysis [15] and also led to the power method and the inverse iteration method as numerical procedures.

1905 – Schmidt. In his Ph.D. thesis [60], which has recently been translated into English by Stewart [70], Erhard Schmidt proved, among other things, the existence of eigenfunctions of integral operators with a symmetric kernel. The proof is based on successive approximations, but uses iterated kernels, the equivalent of matrix powers in a function space.

1909 – Kowalewski. In §197 of his textbook *Determinantentheorie* [36], Gerhard Kowalewski[1] gave a finite-dimensional version of Schmidt's existence proof for a symmetric matrix T. He then proposed "a *simple* infinite process to get to an invariant of T"—the power method—and proved its convergence in the case of a dominant eigenvalue [36, pp. 520–521].

1913 – Müntz. In two short notes in the 1913 *Proceedings of the French Academy of Sciences* [41, 42], without giving proofs of convergence, Chaim Müntz presented both the power method and the inverse power method for the symmetric and nonsymmetric matrix eigenvalue problems, as well as for the generalized eigenvalue problem with symmetric matrices. Müntz also hinted at the possibility of calculating simulta-

[1] Kowalewski, who had been a student of Lie and specialized in continuous groups, worked closely with Schmidt while both were in Bonn. In Schmidt's 1908 paper [61] he is given credit for the geometric interpretation of the concept of a function space.

neously all eigenvalues of a matrix by iterating on a set of linearly independent vectors and orthonormalizing them after each iteration in the manner suggested by Schmidt. In a paper in 1918 he described the latter method more thoroughly [43]. In modern textbooks it is known as orthogonal or subspace iteration and can be shown to be equivalent to the basic QR algorithm; see, e.g., Golub and Van Loan [22, Chapter 7.3]. Although it was presented in von Mises's 1929 survey paper [39], the method is far too tedious for hand computation and was taken up again with the advent of computers in the 1960s; see, e.g., Wilkinson [80, p. 607].[2]

3. The Historical Development of Shifted Inverse Iteration. The contemporary numerical linear algebra community, rather uniformly, attributes shifted inverse iteration to Helmut Wielandt. He presented it in 1944 in an unpublished report [78] written in German. Yet, Peters and Wilkinson in the introduction of their paper [51] write "[shifted] inverse iteration is usually attributed to Wielandt (1944) though a number of people seem to have had the idea independently." Indeed, Wilkinson's own version of the algorithm published in 1958 [79] bears a strong resemblance to the method put forward by Jacobi in 1846 [33].

Jacobi had adopted a two-stage process to compute the eigenvalues and eigenvectors of a symmetric matrix. In the first stage, the matrix is preconditioned by a sequence of plane rotations to make it diagonally dominant, which allows one to take the kth diagonal element and the corresponding unit vector e_k as initial estimates for an eigenpair. In the second stage (see Golub and Van Loan [22, Chapter 10.6] for a detailed description) Jacobi omitted the kth equation[3] while keeping this component of the eigenvector fixed and computed the remaining components by shifted inverse iteration using his own linear equation solver [32]. The omitted equation served as an update formula for the shift. Wilkinson's algorithm differs from Jacobi's mainly in that it uses a constant shift and has no preconditioning stage, which makes the choice of an appropriate starting vector for a nondiagonally dominant matrix difficult. For further details, see Ipsen [31].

Despite this impressive early start in eigenvalue computation, Jacobi's method did not gain widespread acceptance. Seidel, who had done the actual numerical computations for him, remembered especially the first part of the method as being very tedious [63]. Yet, the rotation algorithm was mentioned from time to time in textbooks and was known to Runge, Krylov, and Hessenberg.

The lack of matrix notation[4] makes the second part of Jacobi's paper hard to read, which might be a reason why his inverse iteration method fell into almost complete oblivion.[5] When den Boer [5] wrote her master's thesis on Jacobi's method in 1991, Sleijpen and van der Vorst [66] became aware of it and included it in their version of Davidson's method. So, without denying Jacobi the status of inventor of the method, the modern history of shifted inverse iteration started with Wielandt.

In the first half of the 20th century, the most common approach to the computation of higher eigenvectors, proposed by Pohlhausen [54], consisted in the subtrac-

[2] In 1928 Albert Einstein sought mathematical help with his unified field theory and employed Müntz and Lanczos as his coworkers for about one year [48]. It is not known whether the three of them discussed numerical linear algebra during this time, but it is interesting to speculate about the relationship between Müntz's and Lanczos's methods of eigenvalue computation.

[3] This idea had already appeared in an 1829 paper by Cauchy [6] that Jacobi was familiar with.

[4] While the concept of a matrix dates back to ancient China, formal matrix notation and matrix algebra were not introduced until some years after Jacobi's death.

[5] We could find only two applications, one by Harzer [24] and one by Ritz [57], and one mention in the textbook by Bodewig [4, p. 281].

tion of the already computed eigenvectors from the current iterates. A more refined method called deflation, which modifies the matrix such that the already computed eigenvalues are cast to zero, had been suggested by Hotelling [27].

A third approach, proposed independently by Aitken [1] and van den Dungen [17], and later transformed into the qd-method by Rutishauser, is based on the link between the spectrum of an operator and the poles of a meromorphic function and a formula by Hadamard to find all poles of such a function (see Parlett [50]).

The latter two methods were also contemplated by Wielandt in a paper that was written in 1943 [77]. One year later, however, he wrote the report [78] that contained his most lasting contribution to numerical analysis.

Wielandt Shifted Inverse Iteration. In an effort to retain the sense of Wielandt's contributions we will follow his presentation rather closely, except for notation.

For $x \in R^n$, Wielandt uses the notation x_j to denote the jth component of x. We will find it more convenient to write this quantity as $e_j^T x$, where e_j is the jth natural basis vector, i.e., the vector whose jth component is 1 and all other components are zero.

Consider a matrix $A \in R^{n\times n}$ and choose an integer $j \in \{1, \ldots, n\}$, an approximate eigenvalue $\sigma \in R$, and an approximate eigenvector x_0 satisfying $e_j^T x_0 = 1$. For $k = 0, 1, \ldots$, repeat the following:

$$\text{Solve} \quad (A - \sigma I) z_k = x_k \quad \text{for} \quad z_k. \tag{3.1}$$

$$\text{Normalize} \qquad x_{k+1} = \frac{z_k}{e_j^T z_k}. \tag{3.2}$$

Wielandt's normalization (3.2) with various choices of j is commonly used in the algorithmic eigenvalue problem literature. Peters and Wilkinson arrange for j to be the last component of x. Even if j were chosen to be the first component of z_k of largest absolute value to allow for l_∞ normalization, one can see that if x_k converges to an eigenvector, j would become fixed eventually. In the remainder of the paper we will refer to the normalization (3.2) as *the single-component normalization.*

Wielandt immediately follows his algorithmic presentation by saying that the quantity

$$\sigma + \frac{1}{e_j^T z_k} \tag{3.3}$$

gives an estimate that should converge to the true eigenvalue.

After some discussion concerning algorithmic convergence, Wielandt presents only one numerical example. In his nonsymmetric 3×3 example the eigenvalue estimate σ in (3.1) changes from iteration to iteration. His algorithm is as follows.

Wielandt Iteration. Consider a matrix $A \in R^{n\times n}$, an approximate eigenvalue σ_0, an integer $j \in \{1, \ldots, n\}$, and an approximate eigenvector x_0 satisfying $e_j^T x_0 = 1$. For $k = 0, 1, \ldots$, repeat the following:

$$\text{Solve} \qquad (A - \sigma_k I) y_{k+1} = x_k \quad \text{for} \quad y_{k+1}. \tag{3.4}$$

$$\text{Normalize} \qquad x_{k+1} = \frac{y_{k+1}}{e_j^T y_{k+1}}. \tag{3.5}$$

$$\text{Update the eigenvalue estimate} \qquad \sigma_{k+1} = \sigma_k + \frac{1}{e_j^T y_{k+1}}. \tag{3.6}$$

Wielandt iteration (3.4)–(3.6) for the example given in the paper converges to an eigenpair extremely fast. The convergence is clearly quadratic and probably cubic. In an attempt to better understand Wielandt iteration, then Rice University graduate student Josef Sifuentes performed a numerical study comparing Wielandt iteration with RQI. Sifuentes conducted an extensive study working with matrices of varying sizes. He considered both symmetric and nonsymmetric matrices, but in the latter case he made sure that the matrices had real eigenvalues. In all cases, except in the nonsymmetric 3×3 example presented in Wielandt's paper, RQI outperformed Wielandt iteration. Moreover, in all cases except the one from his paper, Wielandt iteration gave quadratic convergence. In the case of the example from the paper, the convergence appeared cubic and was essentially the same as that given by RQI. In several cases for nonsymmetric matrices the convergence of RQI deteriorated to quadratic, but it still performed as well as Wielandt iteration.

This numerical study motivated us to conjecture and prove the following theorem.

THEOREM 3.1. *Wielandt iteration* (3.4)–(3.6) *is equivalent to Newton's method on the nonlinear system*

$$(A - \sigma I)x = 0, \tag{3.7}$$

$$e_j^T x = 1. \tag{3.8}$$

Proof. Given the pair (x, σ) with $e_j^T x = 1$, Newton's method on (3.7)–(3.8) consists of constructing $(x + \Delta x, \sigma + \Delta\sigma)$ by solving the linear system

$$\begin{pmatrix} A - \sigma I & -x \\ e_j^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\sigma \end{pmatrix} = - \begin{pmatrix} (A - \sigma I)x \\ 0 \end{pmatrix}. \tag{3.9}$$

The system (3.9) tells us that

$$x + \Delta x = \Delta\sigma (A - \sigma I)^{-1} x, \tag{3.10}$$

$$e_j^T \Delta x = 0. \tag{3.11}$$

From (3.10), using (3.11) and the fact that $e_j^T x = 1$, we obtain

$$\Delta\sigma = \frac{1}{e_j^T (A - \sigma I)^{-1} x}. \tag{3.12}$$

Now, (3.5) follows from (3.10) and (3.12) and (3.6) follows from (3.12). ∎

Since in our numerical study RQI performed so much better than Wielandt iteration, we must ask if Wielandt was familiar with the Rayleigh quotient. If he had used it, instead of (3.6), to update the eigenvalue estimate in his algorithm, then he would have discovered RQI some 13 years prior to Ostrowski, as described in the next section.

The literature is rich with applications of various forms of Newton's method to the algebraic eigenvalue problem. We discuss some of these in section 5. However, we believe that Wielandt iteration (3.4)–(3.6) should be viewed as the first Newton's method application in history to the eigenvalue problem, with some consideration given to Jacobi [33] (see section 5).

For historical reasons we will spend considerable time attempting to answer the following two questions:

Question 1: Was Wielandt familiar with the Rayleigh quotient? If he was, then why did he not use it in his algorithm?

Question 2: Did Wielandt know that his second iterative procedure was Newton's method?

Helmut Wielandt was a pure mathematician specializing in group theory and, in particular, permutation groups. In 1942 Wielandt was attached to the Kaiser Wilhelm Institute and the Aerodynamics Research Institute at Göttingen, a situation not uncommon during those times for mathematicians in Germany, the United Kingdom, and the United States. Concerning that period Wielandt writes,

> I had to work on vibration problems. I am indebted to that time for valuable discoveries: on the one hand, the applicability of abstract tools to the solution of concrete problems, on the other hand, the—for a pure mathematician—unexpected difficulty and unaccustomed responsibility of numerical evaluation. It was a matter of estimating eigenvalues of non-self-adjoint differential equations and matrices. [6]

While there is much interest in this passage, for our purposes, it is significant that Wielandt was concerned primarily with nonsymmetric problems. However, we mention in passing that Wielandt's statement leads us to believe that he was somewhat of a stranger concerning numerical methods, yet Ipsen [29, 30], an expert on Wielandt, compares him favorably to Jim Wilkinson.

Wielandt's respect for the difficulty of nonsymmetric problems was heightened by his observation that for nonnormal matrices the iterates of the power method behave in an erratic way, especially with a starting vector pointing approximately in the direction of the second eigenvector [77, p. 123]. On pages 124–125 of this paper he analyzes the behavior of the power method for a class of lower triangular 2×2 matrices with complex coefficients and obtains inequalities that define the set of starting vectors in the complex plane showing nonmonotonic convergence toward the dominant eigenvector. This anomaly was significant to Wielandt.

From 1942 onwards Wielandt kept mathematical notebooks. About ten years ago they were scanned, transcribed, and made accessible to the public.[7]

On pages 64–65 of Volume 3 of Wielandt's notes, which covers the years 1943 and 1944, he proves an inclusion theorem for the eigenvalues of a symmetric matrix with prescribed first row stated in terms of the Rayleigh quotient. So at the time he proposed his algorithm he was aware of the Rayleigh quotient. We believe that he did not use it in the definition of his algorithm because he deemed it unsuitable for the computation of eigenvalues of non-self-adjoint (nonsymmetric) problems. He did not even consider numerical experiments with the Rayleigh quotient to see if it outperformed his updating procedure. This opinion was not unique, as can be seen from Ostrowski's introductory remarks in his 1959 paper on the RQI method [47].

We now turn to Question 2 stated above. Again from Volume 3 of Wielandt's mathematical notes we learn that indeed it was not Newton's method but an algebraic procedure that he had in mind when he devised his two shifted inverse iteration methods.

[6] From the website http://www-history.mcs.st-andrews.ac.uk/Biographies/Wielandt.html.

[7] http://www3.math.tu-berlin.de/numerik/Wielandt/.

After having discovered (pages 32–33) that the eigenvalues of a simple ODE can be expressed as a continued fraction, he introduces on page 41 the pseudocode of the inverse iteration method with the following words:

> It is a continued-fraction-like method that can be used to improve the eigenvector and eigenvalue of a matrix in the following way:

On page 49 of these notes Wielandt refers to this procedure as "fractional iteration."

At a conference in Darmstadt, Germany, in 1950, Heinz Unger gave a talk on "The Nonlinear Treatment of Eigenvalue Problems," in which he presented the Newton's method formulation (3.9) for the algebraic eigenvalue problem given by (3.7)–(3.8). To his conference presentation [74] Unger added the following interesting footnote:

> A detailed publication about 'Newton's method for eigenvalue problems' is in preparation where, among other things, the connection with 'fractional iteration' will be addressed. I am grateful to Prof. Wielandt for bringing this connection to my attention during the conference.

Unger's paper was well received in Germany. In his much-used textbook [83] Rudolf Zurmühl gave a unified presentation of Unger's and Wielandt's methods and presented an algorithm that avoids the increasing ill-conditioning of Wielandt iteration as the shift approaches the true eigenvalue. Other variants of Unger's method are due to Wittmeyer [81] and to Collatz in his 1966 monograph [9]. As all three authors quote Unger's 1950 paper, it seems that the promised detailed publication was never written. Much of the latter material that we have discussed and an English translation of Unger's note can be found on the webpage of Hubert Schwetlick.[8]

In referring to Unger's Newton's method formulation of the algebraic eigenvalue problem, Schwetlick makes the following statement:

> Unfortunately, the English speaking numerical linear algebra community did not recognize Unger's basic paper. For instance, Peters and Wilkinson did not mention it in their famous survey [51].

This statement seems to be somewhat valid, and we will discuss this and related Peters–Wilkinson oversights in section 5. However, we should add that Householder in his well-known 1964 text [28] does mention Unger, as did Anselone and Rall in their 1968 paper [2].

So from the Unger footnote we know that by 1950 Wielandt was aware that Wielandt iteration, as we describe it in (3.4)–(3.6), was related to Newton's method. The question then is: Did Wielandt posses this awareness at the time of the 1944 writing of his legendary paper? We think not. But on page 41 in Volume 4 of his notebooks, written in February 1945, there is an entry in which Wielandt states that Newton's method is equivalent to the fractional iteration method, giving exactly the same formulae as in Theorem 3.1. So, if he was not aware that this algorithm was Newton's method at the time of writing of the paper, he became so very soon after. Moreover, we acknowledge Wielandt's intuition in that in tuning inverse iteration he actually presented an algorithm that is equivalent to a form of Newton's method.

4. The Historical Development of Rayleigh Quotient Iteration. We present several contributions from the literature that we believe played a role in the development of RQI. In what follows, $\sigma_R(x)$ represents the Rayleigh quotient defined by (1.5).

[8] http://www.math.tu-dresden.de/~schwetli/Unger.html.

1894 – Lord Rayleigh (John William Strutt). In the second edition of his pioneering and critically important text titled the *The Theory of Sound* [56, p. 110], Lord Rayleigh, while studying the natural frequencies of vibrating systems, proposed the following algorithm for approximating an eigenpair:

$$(A - \sigma_R(x)B)x_+ = e_1, \tag{4.1}$$

where e_1 denotes the first natural coordinate vector $(1, 0, \ldots, 0)^T$, x is the current eigenvector estimate, and x_+ is the next estimate.[9]

The algorithm went unnoticed until Kohn rediscovered it in 1949 (see below).

1940 – Southwell. In 1940 Sir Richard Southwell's [68] influential text titled *Relaxation Methods in Engineering Science* was published. Southwell developed and promoted a school of thought referred to as relaxation methods. The relaxation method philosophy influenced the design of algorithms for algebraic eigenvalue problems, throughout the 1940s and well into the 1950s. Following Ipsen [29] we extract Crandall's view of relaxation methods from the introduction of his 1951 paper [11]:

> A unique feature of relaxation methods is that they are approximate procedures which are not rigidly prescribed in advance. The precise procedure to be followed is not dictated but is left to the intuition and accumulated experience of the computer. The computer's intelligence is thus an active or dynamic link in the computational chain. It is this fact which has made relaxation so attractive to many computers.

In those times a computer was a human being, hopefully a creative one.

It is interesting then that disciples of the Southwell relaxation methods school, in formulating numerical procedures, often intentionally left sufficient flexibility in the formulation for the computer to include numerical techniques and experimentation in the implementation. Thus, relaxation methods were essentially a general framework which allowed flexibility and promoted creativity on the part of both the applied science modeler and the human computer. It is remarkable to see how much change occurred in the ensuing decades concerning the notion of a numerical procedure. Of course, the impetus for this change has been the fact that the human computer, with considerable responsibility for making decisions, has been replaced with a digital computer and a programmer who is not only not encouraged, but is actually not allowed, to make delicate decisions or changes in the implementation (programming) of the scientist's numerical algorithm.

While we do not present a specific algorithm for Southwell, we felt that it was important to introduce his relaxation methods philosophy since it had a significant influence on the design of many algorithms from this time period in England and in the United States.

1944 – Wielandt. We have discussed Wielandt's contribution in detail in the previous section. Clearly, for transparent reasons Wielandt was not influenced by the relaxation method philosophy.

1946 – Todd. After the war, in 1945, John Todd returned to King's College in London and gave his first course in numerical mathematics. In 1957 he and his wife, the well-respected linear algebraist Olga Taussky Todd, moved to Caltech to start a program in numerical mathematics. Todd's lecture notes from this extended period

[9] Actually Lord Rayleigh considered the symmetric generalized eigenproblem $Ax = \lambda Bx$, where the Rayleigh quotient becomes $x^T Ax/x^T Bx$.

were widely circulated. In these notes he presented the following algorithm for the algebraic eigenvalue problem:

$$\text{Approximately solve} \quad (A - \sigma_R(x)I)x_+ = 0 \quad \text{for} \quad x_+. \tag{4.2}$$

It is abundantly clear that Todd's algorithm (4.2) is given in the context of the Southwell relaxation school. In 1992 at a conference at UCLA the first author had the opportunity to ask Todd what he meant by "approximately solve," since the clear and unique solution is the zero vector. He replied in a pleasant and straightforward manner, "approximately solve is to be interpreted by the individual that is designing the algorithm."

1949 – Kohn. In a letter to the editor, Kohn [35] suggests the following algorithm for the algebraic eigenproblem:

$$(A - \sigma_R(x)I)x_+ = e_i, \tag{4.3}$$

where e_i, $i = 1, \ldots, n$, is any natural basis vector for R^n. He claims that if x_0 is sufficiently close to an eigenvector, then the sequence $\{\sigma_R(x_k)\}$ converges to the associated eigenvalue and the convergence is quadratic. This quadratic convergence is stated as q-quadratic convergence; see Appendix A.

Observe that Kohn's algorithm (4.3) is a slight modification of (4.1), the algorithm originally proposed by Lord Rayleigh, a fact that was pointed out by Crandall in his 1951 paper discussed below. Kohn did not mention Lord Rayleigh's algorithm or any other algorithm in his brief letter.

1951 – Crandall. In 1951 in an interesting paper [11] titled "Iterative Procedures Related to Relaxation Methods for Eigenvalue Problems" and communicated to the Royal Society of London by Sir Richard Southwell himself, Crandall proposes the following algorithm for the algebraic eigenvalue problem. Given x, solve for x_+ from

$$(A - \sigma_R(x)I)x_+ = x. \tag{4.4}$$

Actually Crandall considered the generalized eigenvalue problem $Ax = \lambda Bx$, however, for our comparative purposes it is sufficient to take B equal to the identity matrix.

We observe that Crandall's algorithm is unnormalized RQI; see section 1.3. Crandall believed that he had established (what we would call) r-cubic convergence in both the approximate eigenvector sequence and the approximate eigenvalue sequence; see Appendix A. This issue will be further addressed in the following section.

In further reinforcing Southwell's influence we point out that the final paragraph of Crandall's paper includes the following statement:

> From the preceding analysis it is possible to infer the following guiding principles to aid the relaxation computer in choosing his tactics.

1957–1958 – Ostrowski. In 1957 and 1958 Ostrowski authored the first two parts of a six-part paper titled "On the Convergence of the Rayleigh Quotient Iteration for the Computation of the Characteristic Roots and Vectors." We first observe that at the time of Ostrowski's paper RQI had not been defined. Hence, we must conclude that in this paper he intended to define such an algorithm. Part I [45] of Ostrowski's paper appeared in 1957, and Part II [46] appeared in 1958. We will see that, motivated by interesting circumstances, Ostrowski defined RQI as we know it today in Part II of his paper. We now discuss Part I of the paper.

Ostrowski – Part I. Ostrowski begins his paper by writing down (4.2) and saying that his attention was drawn to this method by John Todd, who used it in his lectures as long ago as 1945. He followed this by saying,

> The crucial point in the discussion above is of course a suitable rule for the computation of the "approximate solution" x_k in (4.2). The rule that I will use in the first part of this discussion consists in taking an arbitrary vector $\eta \neq 0$ and in putting

$$x_{k+1} = (A - \sigma_R(x_k)I)^{-1}\eta. \tag{4.5}$$

> The theoretical arguments in support of this rule are given in another paper.

He then establishes local and q-quadratic convergence for the approximate eigenvalue sequence. In a footnote he mentions that this result is consistent with the note of Kohn.

In section 9 of the paper Ostrowski expresses concern that the vector η on the right-hand side of (4.5) is fixed. This leads him to combine his idea with the idea of Wielandt's shifted inverse iteration (3.1)–(3.2) to obtain the algorithm

$$(A - \sigma_R(x)I)x_+ = x. \tag{4.6}$$

Let (x_*, σ_*) be an eigenpair of A. Ostrowski proves local (x_0 near x_*) convergence for the approximate eigenvalue sequence $\{\sigma_k = \sigma_R(x_k)\}$ to σ_* with the property

$$\frac{(\sigma_{k+1} - \sigma_*)}{(\sigma_k - \sigma_*)^3} \to \gamma \quad \text{as} \quad k \to \infty \tag{4.7}$$

for some positive γ. Observe that (4.7) reflects q-cubic convergence.

Now, Ostrowski was aware of Kohn [35], but he was not aware of Crandall [11]. Notice that Crandall's algorithm (4.4) is exactly Ostrowski's algorithm (4.6). While Ostrowski's paper was in print, George Forsythe, founder of the Computer Science Department at Stanford and a close colleague, directed him to Crandall's work. Ostrowski added a note to the paper acknowledging this fact and then stated, "Professor Crandall establishes the cubic character of convergence of x_k in the rule (4.6). However, he does not arrive at our asymptotic formula (4.7), which is the principal result of our paper."

Ostrowski – Part II. Ostrowski begins Part II of his paper by reflecting back on Crandall's paper. He writes,

> In the paper by Crandall the cubic character of the convergence is proved in the following sense for the sequence of vectors x_k obtained by the rule (4.6): If $x_k \to \eta$, and if for a sequence of numbers ϵ_k with $\epsilon_k \to 0$,
>
> $$x_k - \eta = O(\epsilon_k),$$
>
> then
>
> $$x_{k+1} - \eta = O(\epsilon_k^3).$$
>
> However, from this result it does not follow that $\sigma_{k+1} - \sigma = O((\sigma_k - \sigma)^3)$, and still less that the asymptotic formula (4.7) holds.

What Ostrowski seems to be implying is that the r-cubic convergence of the approximate eigenvector sequence $\{x_k\}$ does not imply the q-cubic convergence of the

approximate eigenvalue sequence $\{\sigma_k\}$. Ostrowski seems to have missed the fact that Crandall claimed r-cubic convergence of the approximate eigenvalue sequence $\{\sigma_k\}$; but this also would not imply the q-cubic convergence of this sequence that Ostrowski established. See Appendix A for the definition of these convergence notions.

We now express a major concern with Crandall's conclusions.

THEOREM 4.1. *The iteration sequence $\{x_k\}$ generated by Crandall–Ostrowski iteration* (4.6)*, i.e., unnormalized RQI, cannot converge.*

Proof. Suppose $x_k \to \eta$. Then from (4.6) we have

$$A\eta = (1 + \sigma_R(\eta))\eta.$$

Hence, η is an eigenvector for A with corresponding eigenvalue $1 + \sigma_R(\eta)$, but if η is an eigenvector, then its corresponding eigenvalue is $\sigma_R(\eta)$; hence

$$1 + \sigma_R(\eta) = \sigma_R(\eta),$$

which is a contradiction. ∎

While the approximate eigenvector sequence will not converge, it is certainly possible that the approximate eigenvalue sequence will converge. In section 21 of his paper Ostrowski promotes the idea of normalizing the iterates defined by (4.5) with the following powerful result.[10]

THEOREM 4.2 (Ostrowski (1958)). *Suppose that the RQI sequence has been defined using an l_2 norm normalization. Then, if the Rayleigh quotient sequence converges to an eigenvalue, the RQI sequence converges to a corresponding unit eigenvector.*

We have finally reached our destination having started with Lord Rayleigh in 1894 and ending with Ostrowski in 1958. We have completed a 64-year journey of tuning and pruning that has brought us to RQI as we know it today. Notice that the Rayleigh quotient sequence is not changed by normalization, so it is not surprising that it converges without normalization.

5. The Historical Development of Newton's Method for the Eigenvalue Problem. The literature abounds with applications of Newton's method to the algebraic eigenvalue problem. In this section, following introductory statements, we will highlight several of the applications which we view as appropriate background for our current study. We will pay close attention to what we consider to be the earlier applications. Applications which have the flavor of both Newton's method and RQI are reviewed in section 6.

To begin with, we see that the literature contains attempts to obtain an eigenvalue of a real symmetric matrix by minimizing the Rayleigh quotient. However, such an activity is not productive when employing algorithms related to Newton's method. This is because Newton's method does not behave well when the sought solution is not isolated. We say that a solution is isolated if there exists an open set in which this solution is the only solution contained in the set. An attempt at correcting this is to minimize the Rayleigh quotient subject to the constraint that a solution must have its l_2 norm equal to 1. In this case the solutions will be isolated for simple eigenvalues. This leads to the following problem.

[10] Ostrowski did not formally state his result as a theorem, and we paraphrase his statement. This theorem is undoubtedly the first result in history that concerns the RQI method.

The l_2 Norm Eigenvector Optimization Problem. Given a real symmetric matrix $A \in \mathcal{R}^{n\times n}$,

$$\begin{aligned} &\text{minimize} && \frac{1}{2}x^T Ax \\ &\text{subject to} && \frac{1}{2}(x^T x - 1) = 0. \end{aligned} \tag{5.1}$$

In deriving the first-order necessary conditions for problem (5.1) we first consider the Lagrangian

$$\ell(x,\lambda) = \frac{1}{2}x^T Ax - \frac{1}{2}\lambda(x^T x - 1) \tag{5.2}$$

and then write

$$\nabla \ell(x,\lambda) = 0,$$

which gives us our first-order necessary conditions as

$$(A - \lambda I)x = 0, \tag{5.3}$$

$$\frac{1}{2}(x^T x - 1) = 0. \tag{5.4}$$

Observe that the unit eigenvectors of A are stationary points of the Lagrangian and the corresponding eigenvalues are the Lagrange multipliers. Also observe that the nonlinear system (5.3)–(5.4) is square with isolated solutions for simple eigenvalues; hence it is amenable to Newton's method. Moreover, it is known that the Jacobian matrix for this system (see (6.7)) is nonsingular at the eigenpair (λ_*, x_*) provided λ_k is simple. For a proof see Peters and Wilkinson [51, p. 345].

Anselone and Rall [2] in 1968 suggest applying Newton's method to the algebraic eigenvalue problem by applying Newton's method to the square nonlinear system

$$(A - \lambda I)x = 0, \tag{5.5}$$

$$\phi(x) - 1 = 0. \tag{5.6}$$

They call (5.6) the normalization equation. The choice of $\phi(x) = x^T x$, i.e., normalizing with the l_2 norm squared, is a very popular choice.

Anselone and Rall list a host of authors who have employed (5.5)–(5.6) with the l_2 norm squared normalization dating back to L.V. Kantorovich as early as 1948. Many of these applications were in infinite-dimensional Hilbert spaces.

1846 – Jacobi. In Chapter 10 of the fourth edition of their book [22], Golub and Van Loan effectively argue that the two-stage Jacobi process can be viewed as an inexact formulation of Newton's method. While we are confident that Jacobi was not aware of this phenomenon,[11] we feel that it is appropriate to list his contribution in our Newton's method timeline.

1944 – Wielandt. Even though Wielandt may not have been aware at the time of presentation of his report that his proposed iteration (3.4)–(3.6) was Newton's method on the system (3.7)–(3.8), it is clear that it should be considered the first example in modern times of Newton's method applied to the eigenvalue problem.

[11] Certainly Jacobi was aware of Newton's method in one dimension. However, the lack of matrix notation would have made Newton's method difficult to formulate or recognize in multiple dimensions. Indeed, Simpson [65], who is credited with extending Newton's method from one dimension to two dimensions in 1740, stated the method in an awkward fashion using Cramer's rule since matrix notation and matrix algebra were not available at the time.

1948 – Kantorovich. In his 1948 text [34], Nobel laureate L.V. Kantorovich began his well-received theory of Newton's method leading to the popular Kantorovich theorem for Newton's method, the topic of the Anselone and Rall [2] paper. He considered the algebraic eigenvalue problem in a Hilbert function space setting and suggested the use of Newton's method applied to (5.3)–(5.4) as a solution technique.

1950 – Unger. Unger in 1950 in a direct fashion suggested using Newton's method on the system (3.7)–(3.8) to solve the eigenvalue problem. Some consider this to be the first application of Newton's method to the eigenvalue problem. However, as we discussed in previous sections, it was later determined that this Newton's method application is equivalent to the iteration method given by Wielandt in 1944.

1953 – Hestenes. In a 1953 National Bureau of Standards report [25], Hestenes considered an application of Newton's method to the algebraic eigenvalue problem. Certainly aware of the shortcomings of using Newton's method to minimize the Rayleigh quotient which we described above, he constructed iterative methods to generate the characteristic polynomial and then brought in Newton's method to find the roots of this polynomial.

1979 – Peters and Wilkinson. Concern for the fact that shifted inverse iteration may require the solution of nearly singular linear systems led Peters and Wilkinson [51], in a paper titled "Inverse Iteration, Ill-Conditioned Equations and Newton's Method," to study inverse iteration in the context of the normalized Newton's method. They naturally first turned to Newton's method on the nonlinear system (5.3)–(5.4) followed by an l_2 normalization. However, they quickly realized that if instead they used the linear normalization constraint, $e_j^T x = 1$ in (5.6) for some j, then, since Newton iterates satisfy linear constraints, normalized Newton's method would reduce to Newton's method. Hence there would be no need for the normalization step and the quadratic convergence of the algorithm would be guaranteed by standard Newton's method theory, unlike the l_2 norm case. In comparing Newton's method on (5.3)–(5.4) to shifted inverse iteration, Peters and Wilkinson made the following statement:

> One step of inverse iteration gives the same result to first order as one Newton step. It is interesting though, that inverse iteration is sometimes the exact equivalent of the classical Newton's method. This is true, for example for the system

$$Ax - \sigma x = 0, \tag{5.7}$$

$$e_j^T x = 1. \tag{5.8}$$

Often in the literature we see authors state that Peters and Wilkinson [51] demonstrated that shifted inverse iteration is Newton's method to first order. To make this statement more precise, we give a formal presentation of the Peters–Wilkinson understanding.

THEOREM 5.1.

(a) *The eigenvector estimate obtained from one step of Newton's method on the system* (5.3)–(5.4) *with initial estimates* (x_0, σ_0) *is a scalar multiple of the eigenvector estimate obtained from one step of shifted inverse iteration with initial estimate* (x_0, σ_0) *using any normalization factor.*

(b) *The eigenvector estimate obtained from one step of Newton's method on the system* (5.7)–(5.8) *with initial estimates* (x_0, σ_0) *is identical to the eigenvector estimate obtained from one step of shifted inverse iteration with initial estimate* (x_0, σ_0) *provided that the normalization is the Wielandt normalization*

given in (3.2) *and* x_0 *has been normalized according to this normalization, i.e.,* $e_j^T x_0 = 1$.

Proof. The Newton equations for the system (5.3)–(5.4) are given by (6.7). From the first equation it follows that

$$x_+ = -\Delta\lambda(A - \lambda I)^{-1}x. \tag{5.9}$$

Part (a) of the theorem follows directly from (5.9) and part (b) follows directly from (3.10) and (3.12). ∎

Remark 5.2. The understanding given in part (b) of Theorem 5.1 is clearly inherent in Theorem 3.1.

Remark 5.3. Seemingly unaware of Unger's proposal of Newton's method on (3.7)–(3.8), some 30 years later, Peters and Wilkinson [51] proposed the identical Newton's method formulation for the algebraic eigenvalue problem. Moreover, it is somewhat ironic that, unaware of Wielandt iteration, they then manipulated the Newton's method formulation to come up with expressions for the new eigenvector and eigenvalue approximations and some 35 years later arrived at exactly Wielandt iteration (3.4)–(3.6). See [51, p. 347]. They concluded, as Wielandt proposed, that the eigenvalue correction is the normalization factor. In essence, they established Theorem 5.1.

Remark 5.4. We see from Schwetlick's[12] description of Unger's 1950 paper that Unger not only proposed Newton's method on the system (5.7)–(5.8) nearly 30 years before Peters and Wilkinson [51], but he also stated a version of the Peters–Wilkinson insight that inverse iteration is Newton's method to first order. Specifically, Unger derived (3.10), but he did not derive (3.12), so he stated part (b) of Theorem 5.1 with the Newton iterate being a scalar multiple of the shifted inverse iterate, not realizing that the scalar was 1.

1988 – Tapia–Whitley. In researching material for the present paper, the first author became intrigued with the Peters–Wilkinson quick dismissal of the l_2 normalized Newton's method in favor of the single-component normalized Newton's method, in large part because the latter algorithm reduced to Newton's method, and as such quadratic convergence was guaranteed, while in the former case the retention of quadratic convergence was not immediate. Hence, he recruited then Rice Computational and Applied Mathematics graduate student David Whitley to do a numerical study of these two approaches. The study showed that while the single-component formulation gave quadratic convergence as expected, the l_2 formulations demonstrated a convergence rate that was q-superquadratic, but not yet cubic. Whitley did a superb job on this task and played a major role in conjecturing and establishing the following theorem; see Tapia and Whitley [73].

THEOREM 5.5 (Tapia–Whitley). *Let* A *be a symmetric matrix,* x_0 *a vector with* $\|x_0\|_2 = 1$, *and* σ_0 *a scalar that is not an eigenvalue of* A. *Also, let the sequence* $\{x_k, \sigma_k\}$ *be generated by the* l_2 *normalized Newton's method on the nonlinear system* (5.3)–(5.4) *initial data* $\{x_0, \sigma_0\}$. *If* $\{\sigma_k\}$ *converges to an eigenvalue* σ_* *and* $\sigma_k \neq \sigma_*$ *for all* k, *then* $\{\sigma_k\}$ *converges with* q*-order* $1 + \sqrt{2}$, *and* $\{x_k\}$ *converges with the same* q*-order to a corresponding eigenvector* x_*.

Tapia and Whitley conclude their study with the following remarks:

[12] http://www.math.tu-dresden.de/~schwetli/Unger.html.

> The motivation for this work was the intriguing and, to us, surprising nonintegral superquadratic convergence rate indicated by our numerical experiments. We were also quite pleased that we were able to establish a q-rate of convergence of $1 + \sqrt{2}$ both for the x-iterate alone and for the σ-iterate alone. The classical Newton's method theory gives q-quadratic convergence in the pair (x, σ) for the Peters–Wilkinson algorithm, which implies an r-quadratic rate in each of the x and σ variables. The direct proof given by Dongarra, Moler, and Wilkinson [16] also only establishes r-quadratic convergence in x for the same algorithm.
>
> We experimented with several problem formulations that used p-norms other than the l_2 norm; for no value of p except $p = 2$ did our estimated convergence rates exceed 2. On the basis of these experiments, we consider it extremely unlikely that any p-norm formulation other than the one presented here produces superquadratic convergence.

Remark 5.6. Upon reflection, it seems appropriate to refer to Newton's method on the system (3.7)–(3.8) (single-component Newton's method) and Newton's method on the system (5.3)–(5.4) (l_2 norm Newton's method) as the classic Newton's method formulations for the algebraic eigenvalue problem. It is of interest to point out that Kantorovich's 1948 formulation of the l_2 norm Newton's method predates Unger's 1950 formulation of the single-component Newton's method. However, in a strict sense the single-component Newton's method should be credited to Wielandt from 1944; hence it technically predates Kantorovich.

6. Newton-RQI Equivalences from the Literature. The existing literature contains many applications to the algebraic eigenvalue problem of algorithms with a Newton's method flavor which also have an RQI flavor. In this section, as background for our Newton-RQI equivalence given in section 11, we investigate several implied claims of equivalences between a form of a normalized Newton's method and the RQI algorithm and demonstrate the various deficiencies of these implied equivalences, as they concern a true Newton's method equivalence to RQI.

We remind the reader that the primary objective of the present study is to find an equation for which the normalized pure Newton's method iterate obtained from this equation and the Rayleigh quotient iterate coincide. As we shall see, equivalences between RQI and a normalized Newton-like method are not that hard to construct. Indeed, RQI itself can be viewed as a normalized Newton-like method according to the definition given in section 1.2. In many cases the main disqualifier from a true Newton's method equivalence is the replacement of a Newton multiplier estimate in the algorithm with the Rayleigh quotient multiplier estimate. As such, the Newton's method becomes a Newton-like method.

An additional subtlety, which is often ignored in the literature, is that demonstrating that an iterate in question is a nonzero scalar multiple of the Rayleigh quotient iterate implies only that this iterate when normalized coincides with either the Rayleigh quotient iterate or its negative. It is surprising to see how many authors simply gloss over this aspect of an equivalence.

In an interesting paper, Auchmuty [3] proposes a Newton's method formulation for solving the algebraic eigenvalue problem. He demonstrates that he can obtain cubic convergence from a modification of his basic algorithm by implementing two linear solves with the same coefficient matrix per iteration. Such an algorithm has the flavor of Newton's method and the flavor of the cubic convergence of RQI.

Some statements in the literature concerning the equivalence of Newton's method and RQI are excessively general. For example, Stewart [69, p. 424] highlights the Peters–Wilkinson formulation of Newton's method for the eigenvalue problem discussed in the previous section and for its equivalence with RQI he turns to a theorem from Dennis and Schnabel [14] that says that any two superlinearly converging methods launched from the same point must up to higher-order terms produce the Newton direction. This is true, but not as descriptive as it might be here since one of our algorithms is quadratically convergent and the other is cubically convergent.

1986 – Chatelin. In the middle 1980s, Chatelin [7, 8] produced excellent work on Newton's method and the algebraic eigenvalue problem. She stated an equivalence between RQI and the l_2 normalized Newton's method applied to

$$F(x) = (A - \lambda(x)I)x, \tag{6.1}$$

with the choice for $\lambda(x)$ given by

$$\lambda_C(x) = x^T Ax. \tag{6.2}$$

There are a couple of ways this approach fails to provide a complete equivalence between the normalized Newton's method and RQI. First, if A is nonsingular and x of unit norm satisfies $\lambda_C(x) = 0$, then the Newton iterate is zero and the normalized Newton iterate is not defined. To see this, observe that $\nabla\lambda_C(x)^T x = 2\lambda_C(x)$ and $F'(x) = (I - 2xx^T)A - \lambda_C(x)I$. When $\lambda_C(x) = 0$, the Newton equation becomes

$$(I - 2xx^T)Ax_+^N = 0,$$

and since $(I - 2xx^T)^2 = I$ and A is also nonsingular we must have $x_+^N = 0$. Clearly $x = (1, 0)^T$ and

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are such that $\lambda_C(x) = 0$ and A is nonsingular.

The second shortcoming is hard to avoid. A straightforward calculation shows that the Newton iterate for (6.1) can be written as

$$x_+ = 2(x^T A\Delta x)(A - \lambda(x)I)^{-1}x. \tag{6.3}$$

It follows from (6.3) that the normalized Newton iterate, assuming the Newton iterate is not zero, will be equal to the RQI iterate *or* its negative. We can not guarantee equality to the RQI iterate, so while the Chatelin equivalence is not quite complete, we add that of all the implied equivalences between a form of normalized Newton's method and RQI in the literature, we view this Chatelin equivalence as the closest to a true equivalence. Indeed, it served to motivate the complete equivalence that we give in section 10 using a different F.

1988 – Tapia–Whitley. In this subsection, we use the approach of Tapia and Whitley [73] in the construction of their convergence rate theory for the l_2 normalized Newton's method applied to the eigenvalue problem to give a concise derivation of the classic RQI algorithm. We will refer often to the various convergence notions described in section A1 of Appendix A. Also we use the notation $x = \pm y$ to mean $x = y$ or $x = -y$. We make explicit the possibly different sign between Newton and RQI.

In the introductory section of their paper, Tapia and Whitley give an informal statement of the following equivalence between a Newton-like method and RQI.

PROPOSITION 6.1. *Consider an l_2 unit vector x. Let x_{RQI} denote the RQI iterate obtained from x and let x_{NN} denote the l_2 normalized Newton-like iterate obtained from the Newton equations* (6.7) *for this x and let λ be replaced by the Rayleigh quotient for x. Then*

$$x_{NN} = \pm x_{RQI}. \tag{6.4}$$

Proof. By way of background recall that the l_2 normalized eigenvector optimization problem (5.1) has as first-order necessary conditions (5.3)–(5.4). Moreover, Newton's method on the system (5.3)–(5.4) consists of the iterative process

$$x_+ = x + \Delta x, \tag{6.5}$$

$$\lambda_+ = \lambda + \Delta\lambda, \tag{6.6}$$

where

$$\begin{pmatrix} A - \lambda I & -x \\ x^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} (A - \lambda I)x \\ \frac{1}{2}(x^T x - 1) \end{pmatrix}. \tag{6.7}$$

The first equation in (6.7) tells us that

$$x + \Delta x = \Delta\lambda (A - \lambda I)^{-1} x. \tag{6.8}$$

Expression (6.8) is sufficient to prove the proposition. However, for the sake of completeness we note that $x^T x = 1$ by hypothesis, and $x^T \Delta x = 0$ by the second equation of (6.7), so straightforward algebra gives

$$\Delta\lambda = \frac{1}{x^T (A - \lambda I)^{-1} x}; \tag{6.9}$$

hence (6.8) becomes

$$x + \Delta x = \frac{(A - \lambda I)^{-1} x}{x^T (A - \lambda I)^{-1} x}. \tag{6.10}$$

If we replace λ in (6.10) with $\lambda_R(x)$, the Rayleigh quotient at x, then (6.10) leads to (6.4) and establishes the proposition. ∎

In addition to the Newton iteration process (6.5)–(6.6), we are interested in the l_2 normalized Newton iteration process

$$x_+ = \frac{(x + \Delta x)}{\|x + \Delta x\|_2}, \tag{6.11}$$

$$\lambda_+ = \lambda + \Delta\lambda, \tag{6.12}$$

where $(\Delta x, \Delta y)$ is obtained from (6.7).

Motivating the RQI Algorithm.

- From standard Newton's method theory we know that the sequence $\{(x_k, \lambda_k)\}$ defined by (6.5)–(6.6) has a q-convergence rate of 2; hence each of the component sequences has an r-convergence rate of 2.
- Tapia–Whitley [73] demonstrated that the q-convergence rate of both sequences $\{x_k\}$ and $\{\lambda_k\}$ defined by (6.11)–(6.12) is $1 + \sqrt{2}$. Hence "closing" the norm constraint improved the convergence rate from 2 to $1 + \sqrt{2}$ and gave us q-rates instead of just r-rates for the $\{x_k\}, \{\lambda_k\}$ sequences. (Recall that Peters and Wilkinson [51] were concerned that this l_2 norm normalization might destroy quadratic convergence. Indeed, it improves the rate.)

- Based on the success described in the previous bullet, it seems that a good idea would be to choose λ_+ to "close" the equation

$$(A - \lambda I)x_+ = 0.$$

But we can't do this, because for fixed x_+ this is an overdetermined system in λ. However, we can "close" in the least-squares sense; i.e., let λ_+ be the solution of

$$\min_{\lambda} \|(A - \lambda I)x_+\|_2.$$

This gives

$$\lambda_+ = \frac{x_+^T A x_+}{x_+^T x_+} = \sigma_R(x_+).$$

According to Proposition 6.1, our closed-in-x and least-squares closed-in-λ Newton-like method is RQI. As such it will be r-cubically convergent in x; see Wilkinson [80, p. 636]. See also Corollary 14.2, which tells us that in the case of a simple eigenvalue we will have q-cubic convergence in x, and it then follows that we will have an r-convergence rate of 6 in λ (recall (1.6)).

We summarize what we have learned as a meaningful derivation of the RQI algorithm:

- Newton's method on problem (6.7) gives q-quadratic convergence.
- Newton's method followed by a (normalization) closing of the norm constraint gives a q-convergence rate of $1 + \sqrt{2}$.
- Newton's method followed by a (normalization) closing of the norm constraint, followed by a least-squares closing of the gradient of the Lagrangian-equal-to-zero equation, leads to a convergence rate of 3 in x. It is also RQI modulo the sign of the iterate.

Finally, we present the following additional motivation for the choice of RQI over the Newton's method formulations (3.9) and (6.7). Writing λ for σ, we see that the first equations in (3.9) and (6.7) lead to an equation of the form (6.4). Hence each of the three methods, single-component normalized Newton's method (3.9), l_2 normalized Newton's method (6.7), and RQI, produces a new eigenvector estimate that is a scalar multiple of that given by shifted inverse iteration. The new eigenvalue estimates for these three methods are related to each other in an interesting way. Straightforward algebra can be used to obtain the following expressions.

The Single-Component Normalized Newton Method.

$$\lambda_+ = \frac{e_j^T A x_+}{e_j^T x_+}. \tag{6.13}$$

The l_2 Normalized Newton Method.

$$\lambda_+ = \frac{x^T A x_+}{x^T x_+}. \tag{6.14}$$

The RQI Algorithm.

$$\lambda_+ = \frac{{x_+}^T A x_+}{x_+^T x_+}. \tag{6.15}$$

We know that the choice of (6.13) leads to an algorithm that is quadratically convergent, the choice (6.14) leads to an algorithm that exhibits a convergence rate of $1 + \sqrt{2}$, and the choice of (6.15) leads to an algorithm that is cubically convergent.

Our theory tells us that the RQI formula should be preferred to the l_2 normalized Newton method, which in turn should be preferred to the single-component normalized Newton method. The following reflection on the topic should lead one to conjecture this ordering.

Let (x_*, λ_*) be an eigenpair of the matrix A. Since the Rayleigh quotient at x_* gives λ_*, the Rayleigh quotient at points near x_* should serve as a good estimate for λ_*, and the nearer the better. Now observe that (6.13) can be viewed as an approximation to λ_*, the Rayleigh quotient at x_*, where e_j and x serve as approximations to x_*. Similar observations hold for (6.14) with x and x_+ serving as approximations to x_*, and for (6.15) with x_+, the most recent information, serving as an approximation to x_*. Hence, one would expect (6.15) to be preferred to (6.14), which in turn would be preferred to (6.13), as our theory tells us.

1996 – Sleijpen–van der Vorst. In this subsection we discuss the popular and highly successful Jacobi–Davidson algorithm proposed by Sleijpen and van der Vorst in 1996 and its relationship to Newton's method. By way of motivation and background, we first consider Davidson's 1975 method and the Morgan and Scott 1986 extension of Davidson's method. All these methods are so-called subspace methods and, except for Davidson's method, the correction equation is intended to be solved inexactly. For our purposes it is sufficient to consider the most basic formulations of these algorithms. Excellent presentations and discussions of these methods can be found in Sorensen [67] and in Chapter 10 of Golub and Van Loan [22].

In 1975 Davidson [12] suggested the following approach to the algebraic eigenvalue problem. Given (σ, x) as an estimate to an eigenpair of A, consider $x_+ = x + \Delta x$ as an improved eigenvector estimate, where Δx is obtained by solving the correction equation

$$\operatorname{diag}(A - \sigma I)\Delta x = -r, \tag{6.16}$$

where

$$r = (A - \sigma I)x \tag{6.17}$$

and $\operatorname{diag}(B)$ denotes the diagonal matrix consisting of the diagonal of the matrix B. See section 6.1 of Sorensen [67].

In 1986 Morgan and Scott [40] generalized Davidson's approach by replacing the correction equation (6.16) with the full correction equation

$$(A - \sigma I)\Delta x = -r. \tag{6.18}$$

Numerical analysts expressed concern over the formulation (6.18) because an exact solution would cause the algorithm to "stagnate," i.e., Δx would be equal to $-x$ and the subspace would not be expanded. For our purposes it suffices to say that $x_+ = x + \Delta x = 0$. Van der Vorst [76, p. 116] expressed concern over the formulation (6.16) for similar reasons, because for a diagonal matrix A it will stagnate. Yet all acknowledge that Davidson's method is very effective for the class of chemistry problems that he was considering, a class where the matrix A is strongly diagonally dominant.

Our explanation of this perplexing situation is that the correction equation (6.17) must be viewed in the sense of the Southwell relaxation method described in section 4. Indeed, (6.17) is exactly the equation (4.2) posed by John Todd in 1946. Recall that

it was the Southwell relaxation formulation suggested by John Todd that motivated Ostrowski in the construction of the RQI method. Now, Todd's instruction to approximately solve the equation is in alignment with Davidson solving (6.16). The fact that Davidson was so successful adds modern-day credibility to the age-old Southwell relaxation methods philosophy.

In 1996 Sleijpen and van der Vorst suggested improving the algorithm and avoiding stagnation by imposing on (6.18) the orthogonality constraint

$$\Delta x^T x = 0. \tag{6.19}$$

They proposed a new approach which they viewed as a combination of Jacobi's approach described in the beginning of section 3 and Davidson's method. Their correction equation is

$$(I - xx^T)(A - \sigma I)(I - xx^T)\Delta x = -r, \quad \Delta x^T x = 0, \tag{6.20}$$

where r is given by (6.17) and σ is the Rayleigh quotient for A at x. The authors were aware that this basic formulation was related to Newton's method, i.e., (6.7) with λ replaced by σ. To see this, merely substitute Δx from (6.10) into the first equation in (6.20), realize that σ is the Rayleigh quotient, and recall that Δx from (6.10) satisfies the second equation in (6.20). Of course, in a strict sense their algorithm, in this simple form, is the Newton-like method discussed in Proposition 6.1, and it is strictly not Newton's method.

It is interesting to point out that for symmetric problems, the defining characteristic (6.20) of the Jacobi–Davidson method and its relationship to Newton's method appeared in the 1982 paper by Sameh and Wisniewski [58] on the tracemin method. See (2.4)–(2.9) in Sameh and Tong [59]. The former paper also included the important step of constructing an expanding subspace.

1998 – Wu–Saad–Stathopoulos. In [82] Wu, Saad, and Stathopoulos study several Newton's method formulations for the algebraic eigenvalue problem. They first consider the equation (6.1)–(6.2) that we credited to Chatelin earlier in this section. The normalized Newton's method on this equation they call the constrained Newton recurrence. Various properties of the method are established. In the l_2 norm normalized Newton's method on the system (5.3)–(5.4), which entails solving the linear system (6.7), the authors consider replacing λ in (6.7) with $\lambda_R(x)$, the Rayleigh quotient at x, and essentially establish Proposition 6.1. They call this latter algorithm the normalized augmented Newton recurrence. Finally, the authors suggest the algorithm that consists of replacing the Jacobian of the Chatelin function (6.1)

$$F'(x) = A - \lambda(x)I - xx^T(A + A^T)$$

in the normalized Newton's method on $F(x) = 0$ given by (6.1) (their so-called constraint Newton recurrence) with the matrix

$$J_I(x) = A - \lambda(x)I + \alpha xx^T, \tag{6.21}$$

where $\lambda(x) = x^T Ax$ and α is a free parameter chosen at each step. The matrix $J_I(x)$ in (6.21) was previously used by Galick [21] in an algorithm called inflated inverse iteration; hence these authors call this modified version of their constrained Newton recurrence inflated Newton recurrence.

1998 – Edelman–Arias–Smith. In an effort to gain understanding of why RQI is cubically convergent, Edelman, Arias, and Smith [18] consider a Newton-like method which they claim to be the special case of their Newton–Grassmann manifold method when the manifold is a sphere. The algorithm is as follows: Given x with $\|x\|_2 = 1$, consider the projector

$$\Pi = I - xx^T$$

and the Rayleigh quotient

$$\lambda = x^T Ax.$$

$$\text{Solve} \quad \Pi(A - \lambda I)\Delta x = -\Pi Ax \quad \text{for} \quad \Delta x. \tag{6.22}$$

$$\text{Let} \qquad x_+ = x + \Delta x. \tag{6.23}$$

They state that the solution to (6.22) is

$$\Delta x = -x + \frac{(A - \lambda I)^{-1}x}{x^T(A - \lambda I)^{-1}x} \tag{6.24}$$

and that $x + \Delta x$ when normalized by the l_2 norm is the RQI iterate normalized by the l_2 norm.

A direct substitution, observing that $\Pi x = 0$, shows that the solutions of (6.22) are of the form

$$\Delta x = -x + \alpha(A - \lambda I)^{-1}x \tag{6.25}$$

for any scalar α. The choice $\alpha = 1$, after normalization, gives the RQI iterate directly; but in this case the second equation in (6.7) would not be satisfied, i.e., Δx would not satisfy the linearized constraint. So the choice (6.24) is the choice of (6.10) which we know is the solution of (6.7) and (6.8) and is the unique solution from (6.19) that satisfies linearized constraints; equivalently, it is the unique choice that makes Δx orthogonal to x, and that is what Edelman, Arias, and Smith suggest. The price that they pay for making this choice is that after normalizing they obtain the RQI iterate modulo its sign.

Finally, it should be clear that the formulation (6.22) can be viewed as a Newton-like method, but not as Newton's method.

2002 – Simoncini–Eldén. In [64] Simoncini and Eldén replace the system (6.22) with the system

$$\Pi(A - \lambda(x)I)\Pi\Delta x = -\Pi Ax. \tag{6.26}$$

Observe that Δx from (6.26) will satisfy (6.21) if $\Pi\Delta x = \Delta x$, equivalently, $(\Delta x)^T x = 0$, i.e., if Δx satisfies linearized constraints. So, in order to obtain from (6.26) a solution that is also a solution to (6.21), we must choose Δx given by (6.24) before this choice will lead us to the RQI iterate modulo the sign of the iterate.

PART II: Algorithmic Equivalences

7. Historical Development of Four Equality Constrained Optimization Algorithms. For the sake of convenience we first restate the unconstrained optimization problem and the equality constrained optimization problem in R^n.

Unconstrained Optimization. The unconstrained optimization problem in R^n is

$$\min_{x \in R^n} f(x), \tag{7.1}$$

where $f : R^n \to R$.

Equality Constrained Optimization. The equality constrained minimization problem in R^n is

$$\begin{aligned} \min_{x \in R^n} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \end{aligned} \tag{7.2}$$

where $f : R^n \to R$ and $h : R^n \to R^m$ with $m \leq n$.

Throughout the history of optimization, unconstrained optimization has been more receptive to effective theory and algorithmic construction than has equality constrained optimization. Hence, since the 1940s or so, there has been considerable activity in addressing the equality constrained optimization problem by instead considering an unconstrained optimization problem or a sequence of such problems. This latter approach carries the flavor of Newton's method in that a nonlinear problem is addressed by considering a sequence of more accessible problems, in the Newton case, linear problems. The approach of reducing the solution of a constrained minimization problem to solving a sequence of unconstrained minimization problems became known as sequential unconstrained minimization techniques or SUMT, as was highlighted by the award-winning 1968 text *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* [19] by Fiacco and McCormick.

Recall that the first-order necessary conditions for the equality constrained optimization problem (7.2) are

$$\begin{aligned} \nabla_x \ell(x, \lambda) &= 0, \\ h(x) \quad &= 0, \end{aligned} \tag{7.3}$$

where the Lagrangian $\ell(x, \lambda)$ is given by

$$\ell(x, \lambda) = f(x) + \lambda^T h(x). \tag{7.4}$$

Observe that in (5.2) for compatibility of notation we attached a minus sign to the multiplier λ. Here we use a plus sign, the usual approach. It should be clear that the formulations are equivalent, leading only to a change in the sign of the multiplier.

7.1. The Courant l_2 Penalty Function Method. In 1943 Courant [10] proposed what today is called the l_2 penalty function for problem (7.2),

$$P(x; c) = f(x) + \frac{c}{2} h(x)^T h(x) \quad (c > 0). \tag{7.5}$$

The constant c is called the penalty constant. We first observe that

$$\nabla P(x; c) = \nabla f(x) + \nabla h(x) c\, h(x), \tag{7.6}$$

where $\nabla h(x) = (\nabla h_1(x), \ldots, \nabla h_m(x))$.

We are interested in obtaining solutions of problem (7.2) by minimizing the penalty function (7.5). Toward that end we observe that a solution x_* of problem (7.2)

with associated multiplier λ_* must satisfy the necessary conditions (7.3). Hence, if x_* is also a minimizer of the penalty function (7.5), we must have that $\nabla P(x_*, c) = 0$. This implies from (7.6) that if $\nabla h(x_*)$ has full rank, then

$$c\, h(x_*) = \lambda_*. \tag{7.7}$$

However, in general $\lambda_* \neq 0$ and since $h(x_*) = 0$, formally we must have $c = +\infty$. Hence, the l_2 penalty function method consists of sequentially minimizing the penalty function (7.5) for a sequence of penalty constants $\{c_k\}$ such that $c_k \to +\infty$ and $c_k > c_{k-1}$. In spite of this shortcoming, this technique has seen fairly wide use in optimization applications throughout the years.

7.2. The Hestenes–Powell Augmented Lagrangian Multiplier Method. We consider replacing the l_2 penalty function (7.5) with the augmented Lagrangian for problem (7.2),

$$L(x, \lambda; c) = f(x) + \lambda^T h(x) + \frac{c}{2} h(x)^T h(x) \quad (c > 0), \tag{7.8}$$

to obtain the augmented Lagrangian multiplier method.

Given (x, λ, c), obtain x_+ as the solution of

$$\min_x L(x, \lambda; c).$$

Update (λ, c):

$$\lambda \to \lambda_+,$$
$$c \to c_+.$$

Observe that a minimizer in x of the augmented Lagrangian (7.8) must satisfy the necessary condition $\nabla_x L(x, \lambda; c) = 0$, where

$$\nabla_x L(x, \lambda; c) = \nabla f(x) + \nabla h(x)(\lambda + c\, h(x)). \tag{7.9}$$

Hence, in contrast to the penalty function, minimizers of the Lagrangian are not precluded from being solutions of the constrained problem (7.2) for finite values of c, so in the multiplier method it is not necessary to choose $c \to +\infty$. However, the rub is that if we obtain convergence in the multiplier method, say, $(x_k, \lambda_k; c_k) \to (x_*, \lambda_*; c_*)$, it does not follow from (7.9) alone that $h(x_*) = 0$. In other words, stationary points of the augmented Lagrangian are not necessarily stationary points of problem (7.2), let alone solutions of problem (7.2).

Hestenes [26] and Powell [55] are credited independently with formulating the multiplier method in 1969. Hestenes' motivation came from the desire to avoid the need to use an unbounded sequence of penalty constants as is required in the penalty function method.

7.3. The Fletcher Exact-Penalty Function Method. In the interests of efficiency, it was natural for researchers in optimization to ask whether it is possible to move away from minimizing a sequence of penalty functions to minimizing just one penalty function in order to obtain the solution of the constrained optimization problem (7.2). Such a function has historically been called an exact-penalty function. In 1970 Fletcher [20] proposed the exact-penalty function

$$F(x; c) = L(x, \lambda(x); c), \tag{7.10}$$

where $\lambda(x)$ is the well-known multiplier approximation formula

$$\lambda(x) = -(\nabla h(x)^T \nabla h(x))^{-1} \nabla h(x)^T \nabla f(x). \tag{7.11}$$

The multiplier approximation formula is often called the least-squares multiplier approximation formula because it arises as the least-squares solution for λ of the necessary condition

$$\nabla f(x) + \nabla h(x)\lambda = 0.$$

Fletcher demonstrated that for c sufficiently large solutions of problem (7.2) are minimizers of his penalty function (7.10), i.e., $F(x;c)$ is an exact-penalty function.

Issues with Fletcher's approach are the following:

- How large is sufficiently large for the penalty constant c?
- The derivatives of f and h are out of phase, e.g., $\nabla F(x;c)$ involves $\nabla^2 f(x)$ and $\nabla^2 h_i(x)$.

7.4. The Multiplier Substitution Equation. Tapia, in a series of papers, developed what he called the multiplier substitution quasi-Newton methods. For a brief introduction, see section 8 of Tapia [72]. The multiplier substitution approach avoids dealing with the issues described above associated with the Fletcher exact-penalty function by, instead of minimizing an unconstrained function, working with the nonlinear equation system

$$\nabla_x L(x, \lambda(x); c) = 0, \tag{7.12}$$

where $\lambda(x)$ is given by (7.11). A strong justification of this approach is the following theorem; appropriate counterparts do not hold for the previous two penalty function methods. As is standard in the literature we call a vector x a *regular* point of problem (7.2) if the set $\{\nabla h_1(x), \ldots, \nabla h_m(x)\}$ is linearly independent.

THEOREM 7.1. *Let the vector x be a regular point of problem* (7.2)*. Then x is a zero of the multiplier substitution equation* (7.12) *if and only if it is a stationary point of problem* (7.2)*, i.e., solution of equations* (7.3)*.*

Proof. We first observe that

$$\nabla_x L(x, \lambda(x); c) = \nabla f(x) + \nabla h(x)(\lambda(x) + c\, h(x)). \tag{7.13}$$

Let (x_*, λ_*) be a solution of (7.3). Then $h(x_*) = 0$ and $\lambda(x_*) = \lambda_*$, so x_* is a zero of the multiplier substitution equation (7.12). Now let x_* satisfy (7.12), substitute $\lambda(x_*)$ into (7.12), and left multiply the expression by $\nabla h(x_*)^T$ to obtain

$$\nabla h(x_*)^T \nabla h(x_*)\, c\, h(x_*) = 0. \tag{7.14}$$

Since we are assuming that x_* is a regular point, the matrix in (7.14) is invertible. Hence $c\, h(x_*) = 0$, and since we are assuming that $c \neq 0$ we must have $h(x_*) = 0$. So (x_*, λ_*) is a solution of (7.3) and x_* is a constrained stationary point. ∎

8. Inverse Iteration as Newton's Method. At the beginning of our journey we offer the reader the following road sign. We will derive theory establishing the equivalence between the l_2 normalized Newton's method iterate on an appropriate equation described in the previous four subsections and an iterate obtained from inverse iteration, shifted inverse iteration, or RQI. Each of the three equations involves a penalty constant. As the reader will see, it is a straightforward matter to show that the Newton iterate is a scalar multiple of the iterate in question. The challenge will be to

show that it is possible to choose the penalty constant so that the l_2 normalization will reduce the scalar to 1, so that the iterates in question will coincide.

We now present the following theorem showing that l_2 normalized Newton's method applied to the problem of minimizing the penalty function (7.5) for the eigenvalue optimization problem (5.1), i.e.,

$$\min_x P(x;c) = \frac{1}{2}x^T Ax + \frac{c}{8}(x^T x - 1)^2 \quad \text{for} \quad c \neq 0 \tag{8.1}$$

for appropriate choices of the penalty constant c, generates an iterate that is identical to the iterate obtained from l_2 normalized inverse iteration as defined in section 1.3. Recall the notation introduced in Proposition 6.1 that $x = \pm y$ means either $x = y$ or $x = -y$.

THEOREM 8.1. *Consider a symmetric invertible matrix A and an l_2 unit vector x_0. Let x_{II} be the l_2 normalized inverse iteration iterate obtained from x_0, and let x_{NN} be the l_2 normalized Newton iterate obtained from the problem of minimizing the penalty function* (8.1) *with penalty constant c and initial iterate x_0, i.e., from applying the l_2 normalized Newton's method to the nonlinear equation*

$$\nabla P(x;c) = Ax + \frac{c}{2}(x^T x - 1)x = 0. \tag{8.2}$$

Then x_{II} is well defined and x_{NN} is well defined if and only if $c \neq 0$ and $1 + cx_0^T A^{-1} x_0 \neq 0$. Moreover, for such choices of c

$$x_{NN} = \pm x_{II}. \tag{8.3}$$

Furthermore, for the following choices of penalty constant, both iterates are well defined and identical:

$$\begin{aligned} c > 0 \quad \text{or} \quad c < -\frac{1}{x_0^T A^{-1} x_0} \quad &\text{if} \quad x_0^T A^{-1} x_0 > 0, \\ c > 0 \qquad\qquad\qquad\quad &\text{if} \quad x_0^T A^{-1} x_0 = 0, \\ \text{and} \quad 0 < c < -\frac{1}{x_0^T A^{-1} x_0} \quad &\text{if} \quad x_0^T A^{-1} x_0 < 0. \end{aligned} \tag{8.4}$$

Proof. We have for inverse iteration

$$x_{II} = \frac{A^{-1}x_0}{\|A^{-1}x_0\|_2}. \tag{8.5}$$

Now, for Newton's method we need the following quantity derived from (8.2):

$$\nabla^2 P(x;c) = A + c\,xx^T + \frac{c}{2}\,(x^T x - 1)I. \tag{8.6}$$

So the Newton correction Δx for the l_2 unit vector x_0 satisfies

$$\Delta x = -(A + c\,x_0 x_0^T)^{-1} A x_0. \tag{8.7}$$

We next call on the Sherman–Morrison formula (see page 50 of Ortega and Rheinboldt [44]) to expand the inverse term in (8.7) and obtain

$$\Delta x = -x_0 + \frac{c}{1 + cx_0^T A^{-1} x_0} A^{-1} x_0, \tag{8.8}$$

with the additional information that the inverse exists (the Newton iterate is well defined) if and only if $1 + cx_0^T A^{-1}x_0 \neq 0$. This latter condition comes from the Sherman–Morrison formula when we assume that A is invertible. The Newton correction (8.7) could be well defined even for A which is not invertible; see (i) in Theorem 11.1. Furthermore, (8.8) tells us that the Newton iterate is not zero and the normalized Newton iterate is well defined if and only if $c \neq 0$. Equation (8.8) leads directly to (8.3). Moreover, reflection on the scalar multiplying $A^{-1}x_0$ in (8.8) tells us that the scalar will be positive and the iterates in question will agree under the conditions stated in (8.4). ∎

It is well known that inverse iteration is linearly convergent and converges to a unit eigenvector corresponding to the smallest (in absolute value) eigenvalue, yet we have just demonstrated that it could be the l_2 normalized Newton's method, an algorithm that for a fixed penalty constant c is quadratically convergent under standard assumptions (see Theorem A.1 of Appendix A). Hence we must address this seeming inconsistency. The algorithm in question is the l_2 normalized Newton's method on the nonlinear equation

$$\nabla P(x; c) = Ax + \frac{c}{2}(x^T x - 1)x = 0. \tag{8.9}$$

In most Newton's method applications when quadratic convergence fails, it is because the Jacobian matrix at the limit of the Newton sequence is not invertible. However, in our current application this is not the case. Theorem 11.1 shows that under mild conditions the Jacobian matrix at the limit point, an l_2 unit eigenvector, is invertible.

Suppose that our algorithm converges to an l_2 unit eigenvector, say, x_*, with associated eigenvalue λ_*. From (8.9) we have

$$\nabla P(x_*; c) = \lambda_* x_*. \tag{8.10}$$

In general, λ_* will not be zero, so the rub is that while the l_2 normalization undoubtedly enhances global convergence, it precludes convergence to a zero of the equation on which we are performing Newton's method; thus it is not surprising that we only have linear convergence. Now, if we removed the normalization step, then under the mild conditions of (i) of Theorem 11.1, we would retain quadratic convergence, but at the expense of having to solve a different linear system at each iteration. This balance is mathematical poetic justice.

It is of interest to point out the following. Suppose that A has zero as a simple eigenvalue. Then A is not invertible and inverse iteration is not defined. However, according to Theorem 11.1 if we choose $c \neq 0$, then locally the normalized Newton's method is still well defined. Moreover, the l_2 unit eigenvector corresponding to this zero eigenvalue is a zero of the nonlinear equation (8.9). Hence standard Newton's method theory tells us that our normalized Newton's method will be locally and quadratically convergent to this l_2 normalized eigenvector. However, a new linear system will have to be solved at each iteration.

It is also worth stressing that the l_2 normalized Newton iterates, except for sign, are independent of the penalty constant c, as long as $c \neq 0$.

9. Shifted Inverse Iteration as Newton's Method. We present the following theorem showing that the l_2 normalized Newton's method applied to the problem of minimizing the augmented Lagrangian (7.8) for the eigenvalue optimization problem (7.1), i.e.,

$$\min_x L(x, \lambda; c) = \frac{1}{2}x^T Ax - \frac{\lambda}{2}(x^T x - 1) + \frac{c}{8}(x^T x - 1)^2 \quad \text{for} \quad c \neq 0, \tag{9.1}$$

for appropriate choices of the penalty constant c, generates an iterate that is identical to the iterate obtained from l_2 normalized shifted inverse iteration with shift λ, as defined in section 1.3.

THEOREM 9.1. *Consider a symmetric matrix A, a scalar λ which is not an eigenvalue of A, and an l_2 unit vector x_0. Let x_{SII} be the l_2 normalized shifted inverse iteration iterate obtained from x_0 with shift λ, and let x_{NN} be the l_2 normalized Newton iterate obtained for* (9.1)*, the problem of minimizing the augmented Lagrangian for the eigenvalue optimization problem* (7.1) *with penalty constant c, multiplier estimate λ, and initial iterate x_0, i.e., from applying the normalized Newton's method to the nonlinear equation*

$$\nabla L(x, \lambda; c) = (A - \lambda I)x + \frac{c}{2}(x^T x - 1)x. \tag{9.2}$$

Then $B = (A - \lambda I)$ is invertible, and x_{SII} is well defined. Also, x_{NN} is well defined if and only if $c \neq 0$ and $1 + cx_0^T B^{-1} x_0 \neq 0$. Moreover, for such choices of c,

$$x_{NN} = \pm x_{SII}. \tag{9.3}$$

Furthermore, the two iterates are well defined and identical for the following choices of penalty constant:

$$\begin{aligned} & c > 0 \quad \text{or} \quad c < -\frac{1}{x_0^T B^{-1} x_0} && \text{if} \quad x_0^T B^{-1} x_0 > 0, \\ & c > 0 && \text{if} \quad x_0^T B^{-1} x_0 = 0, \\ \text{and} \qquad & 0 < c < -\frac{1}{x_0^T B^{-1} x_0} && \text{if} \quad x_0^T B^{-1} x_0 < 0. \end{aligned} \tag{9.4}$$

Proof. We observe that shifted inverse iteration can be viewed as inverse iteration with the matrix A replaced by the matrix $B = A - \lambda I$. This latter matrix is invertible since λ is assumed to not be an eigenvalue of A. Moreover, Newton's method on the augmented Lagrangian can be viewed as Newton's method on the penalty function with the matrix A replaced by the matrix B. Hence our theorem follows from Theorem 8.1 by making this substitution. ∎

All the comments made at the end of section 8 concerning inverse iteration apply here to shifted inverse iteration. Reiterating, the l_2 normalization undoubtedly enhances global convergence and allows for an implementation that does not require us to solve a new linear system at each iteration, but precludes quadratic convergence. This latter statement follows from the fact that while convergence is obtained, it is not to a zero of the governing Newton's method equation.

10. Rayleigh Quotient Iteration as Newton's Method. As argued in our introductory sections, in attempting to demonstrate that a given algorithm is Newton's method, the challenge most often is the identification of the underlying Newton's method equation. Historically, this has certainly been the case concerning the interpretation of RQI as Newton's method. In our current search it is reasonable to consider a function of the form

$$F(x) = (A - \lambda(x) I)x, \tag{10.1}$$

where $\lambda(x)$ is a C^1 functional defined on R^n. If x_* is a zero of F, then x_* will be an eigenvector of A with corresponding eigenvalue $\lambda(x_*)$. On the other hand, we

certainly want l_2 unit eigenvectors, i.e., solutions of (7.3)–(7.4), to be zeros of F in (10.1). An immediate way of accomplishing this objective is to require that $\lambda(x)$ in (10.1) agree with the Rayleigh quotient for l_2 unit vectors.

The present study began sometime in the late 1980s when the second author, after a linear algebra class lecture, approached the first author asking if RQI is normalized Newton's method on the equation $F(x) = 0$ where F is given by (10.1) and $\lambda(x)$ given by the Rayleigh quotient. A quick calculation showed that for every initial estimate x, the Newton iterate was the zero vector; hence the normalized Newton iterate was never defined. Motivated by the fact that we seemed so far from a complete equivalence, we began the present study.

It follows then that for various choices of the functional λ in (10.1) we will have various levels of equivalence between RQI and normalized Newton's method on (10.1). A necessary condition for the equivalence to be complete is that the set of exceptional points, points where the Newton iterate is zero, is empty.

Recall that in section 6 we discussed Chatelin's choice for $\lambda(x)$ in (10.1) given by

$$\lambda_C(x) = x^T Ax.$$

We saw that Chatelin's choice does not give a complete equivalence because the Newton step can be zero and the sign can be wrong, and for our first attempt using the Rayleigh quotient for $\lambda(x)$ in (10.1) the Newton step was always zero. Neither happens in the multiplier substitution formulation we analyze now.

The multiplier substitution equation (7.12) for the eigenvalue optimization problem (7.1) is written

$$F(x;c) = (A - \lambda_c(x)I)x = 0 \quad \text{for} \quad c \neq 0, \tag{10.2}$$

where

$$\lambda_c(x) = \lambda_R(x) - \frac{c}{2}(x^T x - 1) \tag{10.3}$$

with

$$\lambda_R(x) = \frac{x^T Ax}{x^T x}, \quad x \neq 0. \tag{10.4}$$

It is worth noting that for the eigenvalue optimization problem (7.1) the least-squares multiplier approximation formula (7.11) becomes the Rayleigh quotient. We also note that the multiplier substitution function $F(x;c)$ given in (10.2) has the form (10.1), as expected. We now present our equivalence theorem.

THEOREM 10.1. *Consider a symmetric matrix A and an l_2 unit vector x_0 such that $\lambda_R(x_0)$, the Rayleigh quotient at x_0, is not an eigenvalue of A. Let x_{RQI} be the l_2 Rayleigh quotient iterate obtained from x_0 and let x_{NN} be the l_2 normalized Newton iterate obtained from x_0 on* (10.2)*, the multiplier substitution equation for the eigenvalue optimization problem* (8.1) *with penalty constant c. Then $B = (A - \lambda(x_0)I)$ is invertible and x_{RQI} is well defined. Also, x_{NN} is well defined if and only if $c \neq 0$ and $1 - cx_0^T B^{-1} x_0 \neq 0$. Moreover, for such choices of c we have*

$$x_{NN} = \pm x_{RQI}. \tag{10.5}$$

Furthermore, the two iterates are well defined and identical for the following choices

of penalty constant:

$$\begin{aligned} & c<0 \quad \textit{or} \quad \frac{1}{x_0^T B^{-1} x_0} < c \quad \textit{if} \quad x_0^T B^{-1} x_0 > 0, \\ & c<0 \qquad\qquad\qquad\quad \textit{if} \quad x_0^T B^{-1} x_0 = 0, \\ \textit{and} \quad & \frac{1}{x_0^T B^{-1} x_0} < c < 0 \qquad \textit{if} \quad x_0^T B^{-1} x_0 < 0. \end{aligned} \tag{10.6}$$

Proof. We have for l_2 normalized RQI,

$$x_{RQI} = \frac{(A-\lambda_R(x_0)I)^{-1}x_0}{\|(A-\lambda_R(x_0)I)^{-1}x_0\|_2}. \tag{10.7}$$

Moreover, this iterate is well defined since $\lambda_R(x)$ is not an eigenvalue of A. Now, for Newton's method we need the following quantities derived from (10.2)–(10.4) using x for x_0:

$$\nabla\lambda_R(x) = \frac{2(A-\lambda_R(x)I)x}{x^T x}, \quad x \neq 0, \tag{10.8}$$

$$\nabla\lambda_c(x) = \nabla\lambda_R(x) - c\,x, \tag{10.9}$$

$$F'(x;c) = (I-2xx^T)(A-\lambda_R(x)I) + c\,xx^T. \tag{10.10}$$

We will use the notation

$$B = A - \lambda_R(x)I \tag{10.11}$$

and

$$D = (I-2xx^T)B. \tag{10.12}$$

Observe that if B is invertible, then $D^{-1} = B^{-1}(I-2xx^T)$.

We will demonstrate that

$$x + \Delta x = \alpha(c)B^{-1}x, \tag{10.13}$$

where

$$\alpha(c) = \frac{-c}{1-c\,x^T B^{-1}x}. \tag{10.14}$$

Toward this end we first observe that

$$\Delta x = -(D + c\,x^T x)^{-1}Bx. \tag{10.15}$$

As before, turning to the Sherman–Morrison formula (see page 50 of [44]) leads to

$$\Delta x = -x + \alpha(c)B^{-1}x. \tag{10.16}$$

In this derivation we used the fact that $x^T Bx = 0$. The Sherman–Morrison formula tells us the additional information that the inverse in (10.15) exists, i.e., the Newton iterate is well defined, if and only if $1 - cx^T B^{-1}x \neq 0$.

Furthermore, (10.13) and (10.14) tell us that the Newton iterate is not zero and the normalized Newton iterate is well defined if and only if $c \neq 0$. Reflection on $\alpha(c)$ given by (10.14) leads us directly to (10.5). Finally, reflection on (10.16) establishes the last part of the theorem involving the inequalities in (10.6). ∎

COROLLARY 10.2. *Let $\{x^k_{RQI}\}$ denote the RQI sequence beginning at x_0 and let $\{x^k_{NN}\}$ denote the normalized Newton iteration sequence also beginning at x_0. Then, if both sequences are well defined, we have*

$$x^k_{NN} = \pm x^k_{RQI} \quad \forall\ k. \tag{10.17}$$

Proof. We use NOR(x) to denote the normalized Newton iterate at x and use RQI(x) to denote the Rayleigh quotient iterate at x. It follows that for $k \geq 0$,

$$x^{k+1}_{NN} = NOR(x^k_{NN}) = \pm RQI(x^k_{NN}) = \pm RQI(\pm x^k_{RQI}) = \pm x^{k+1}_{RQI}. \tag{10.18}$$

The first equality in (10.18) follows by definition, the second and third follow from Theorem 10.1, and the last follows from the fact that the Rayleigh quotient is homogeneous of degree zero. ∎

Remark 10.3. It is important to realize that the Newton iterate at x may exist without satisfying $1 - cx^TB^{-1}x \neq 0$. This assumption is the price that we pay for using the Sherman–Morrison formula under the assumption that B, hence D, is invertible. This remark is reinforced by part (iii) of Theorem 11.1.

Remark 10.4. For B given by (10.11) the eigenvalues of B^{-1} are $(\lambda_1 - \lambda(x))^{-1}, \ldots, (\lambda_n - \lambda(x))^{-1}$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Hence, for x near x_*, a unit eigenvector of A corresponding to the eigenvalue λ_* of A, the quantity $x^TB^{-1}x$ will be approximately $(\lambda_* - \lambda(x))^{-1}$, a quantity that will be arbitrarily large for x arbitrarily close to x_*. Hence, for any c the condition $1 - cx^TB^{-1}x \neq 0$ will always be satisfied locally, i.e., for x near x_*.

11. Invertibility of the Jacobian Matrix at a Unit Eigenvector. The advisability of considering a function of the form (10.1) for our normalized Newton's method equivalences of inverse iteration, shifted inverse iteration, and RQI has been validated. More specifically, if we consider an equation of the form

$$F(x; c) = (A - \lambda_c(x)I)x = 0, \tag{11.1}$$

then Theorems 8.1, 9.1, and 10.1 tell us that inverse iteration, shifted inverse iteration, and RQI, respectively, can be viewed as the normalized Newton's method applied to the nonlinear equation (11.1) with the choices for $\lambda_c(x)$ given, respectively, by

$$\lambda_c(x) = -\frac{c}{2}(x^Tx - 1), \tag{11.2}$$

$$\lambda_c(x) = \lambda - \frac{c}{2}(x^Tx - 1), \tag{11.3}$$

and

$$\lambda_c(x) = \lambda_R(x) - \frac{c}{2}(x^Tx - 1), \tag{11.4}$$

where $\lambda_R(x)$ is given by (10.4).

In the proof of Theorem 10.1 we established the fact that under mild conditions the Jacobian matrix (10.10) is invertible if the l_2 unit vector x is not an eigenvector. In this section we are interested in determining the invertibility of $F'(x_*; c)$ when x_* is an l_2 unit eigenvector of the matrix A for our three choices of $\lambda_c(x)$, (11.2), (11.3), and (11.4).

THEOREM 11.1. *Let $F_{II}(x;c)$, $F_{SII}(x;c)$, and $F_{RQI}(x;c)$ denote the function $F(x;c)$ described in* (11.1) *with the choices of $\lambda_c(x)$ given by* (11.2), (11.3), *and* (11.4), *respectively. Also, let x_1 be an l_2 unit eigenvector of the symmetric matrix A with corresponding eigenvalue λ_1. Finally, let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of A. Then the following hold:*

(i) *The eigenvalues of $F'_{II}(x_1;c)$ are $\lambda_1 + c, \lambda_2, \ldots, \lambda_n$. Hence $F'_{II}(x_*;c)$ is invertible if $\lambda_1 + c \neq 0$, and $\lambda_2, \ldots, \lambda_n$ are not zero.*

(ii) *The eigenvalues of $F'_{SII}(x_1;c)$ are $\lambda_1 - \lambda + c, \lambda_2 - \lambda, \ldots, \lambda_n - \lambda$. Hence $F'_{SII}(x_*;c)$ is invertible if $\lambda_1 - \lambda + c \neq 0$, and $\lambda_2, \ldots, \lambda_n$ are not equal to λ.*

(iii) *The eigenvalues of $F'_{RQI}(x_1;c)$ are $c, \lambda_2 - \lambda_1, \ldots, \lambda_n - \lambda_1$. Hence $F'_{RQI}(x_*;c)$ is invertible if $c \neq 0$, and λ_1 is a simple eigenvalue.*

Proof. Straightforward calculations, recalling (10.8), show that

(i)′ $F'_{II}(x_*;c) = A + cx_*x_*^T$,

(ii)′ $F'_{SII}(x_*;c) = (A - \lambda I) + cx_*x_*^T$, and

(iii)′ $F'_{RQI}(x_*;c) = (A - \lambda_1 I) + cx_*x_*^T$.

Since A is a real $n \times n$ symmetric matrix, it has orthonormal eigenvectors $x_1, \ldots, x_n$. Applying the Jacobian matrices in (i)′–(iii)′ to these n eigenvectors leads immediately to statements (i)–(iii). ∎

12. Derivatives of the Normalized Newton Operator. It was originally our intention at this point in our study to apply standard Newton's method theory to study the local convergence properties of inverse iteration, shifted inverse iteration, and RQI via our normalized Newton's method algebraic equivalences. However, the fact that for inverse iteration and shifted inverse iteration the equivalent normalized Newton's method algorithms do not generate sequences that converge to a zero of the equation in question precludes us from carrying out the proposed study for these latter two algorithms, i.e., standard Newton's method theory does not apply in these two cases; see the remarks at the end of section 6. So we limit our proposed study to the q-cubic convergence of RQI via our normalized Newton's method equivalence, which was always our primary concern. Toward this end we develop several differentiation formulae in the remainder of this section and then construct the Newton's method theory in the following two sections.

Consider the nonlinear equation

$$F(x) = 0. \tag{12.1}$$

Recall that by the Newton operator for (12.1) we mean

$$N(x) = x - F'(x)^{-1}F(x), \tag{12.2}$$

and by the normalized Newton operator for (12.1) we mean

$$S(x) = N(x)/\|N(x)\|_2. \tag{12.3}$$

In the following propositions we assume that F has as many continuous derivatives as are indicated.

PROPOSITION 12.1. *Consider $x_* \in R^n$ such that $F'(x_*)$ is nonsingular and $F(x_*) = 0$. Then*

$$\|N(x_*)\|_2 S'(x_*)(\cdot) = (I - x_*x_*^T)F'(x_*)^{-1}F''(x_*)(F'(x_*)^{-1}F(x_*), \cdot). \tag{12.4}$$

So, for $x_* \neq 0$ we have

$$S'(x_*) = 0. \tag{12.5}$$

Moreover,

$$S''(x_*) = [I - x_* x_*^T]F'(x_*)^{-1}F''(x_*) \tag{12.6}$$

and

$$S'''(x_*) = [I - x_* x_*^T]F'(x_*)^{-1}[2F'''(x_*) - 3F''(x_*)(F'(x_*)^{-1}F''(x_*), \cdot)]. \tag{12.7}$$

Proof. Writing (12.2) as

$$F'(x)(x - N(x)) = F(x) \tag{12.8}$$

and differentiating implicitly gives

$$F''(x)(x - N(x), \cdot) = F'(x)N'(x)(\cdot). \tag{12.9}$$

Recall that if $P(x) = \|x\|_2$, then $P(x)^2 = x^T x$; hence $P'(x)(\eta) = x^T\eta/\|x\|_2$. Now write (12.2) as

$$\|N(x)\|_2 S(x) = N(x) \tag{12.10}$$

and differentiate implicitly to obtain

$$\|N(x)\|_2 S'(x) = [I - S(x)S(x)^T]N'(x). \tag{12.11}$$

Differentiating $N(x)$ as given by (12.2) implicitly and combining (12.8), (12.10), and (12.11), and the fact that $S(x_*) = x_*$, we obtain (12.4) and (12.5) readily follows. Since $F(x_*) = 0$, we have $N(x_*) = x_*$ and $N'(x_*) = 0$ from (12.9). By combining this information with two successive implicit differentiations on (12.9) and (12.11), we obtain

$$F''(x_*) = F'(x_*)N''(x_*), \tag{12.12}$$

$$2F'''(x_*) - 3F''(x_*)(N''(x_*), \cdot) = F'(x_*)N'''(x_*), \tag{12.13}$$

$$S''(x_*) = [I - x_* x_*^T]N''(x_*), \tag{12.14}$$

and

$$S'''(x_*) = [I - x_* x_*^T]N'''(x_*). \tag{12.15}$$

The combination (12.12) and (12.14) gives (12.6), while the combination of (12.13) and (12.15) gives (12.7). ∎

13. Local Cubic Convergence of the Normalized Newton Method. We are interested in the multiplier substitution equation for the eigenvalue optimization problem (7.1) given by (10.2), (10.3), and (10.4) and denote it as

$$F(x; c) = (A - \lambda_c(x)I)x = 0, \tag{13.1}$$

where $\lambda_c(x)$ is given by (10.3) with $\lambda_R(x)$ given by (10.4). We first collect several derivative formulae that will be useful in our analysis. Let x_* be an l_2 unit eigenvector of the symmetric matrix A and corresponding to the eigenvalue λ_1. We have

$$\nabla\lambda_R(x) = 2[A - \lambda_R(x)I]x/x^Tx. \tag{13.2}$$

Our first observation is that

$$\nabla\lambda_R(x_*) = 0. \tag{13.3}$$

Next, we see that straightforward differentiation leads to

$$F''(x_*;c)(u,u) = 2c(x_*^Tu)u - 2(\nabla\lambda_R(x_*)^Tu)u - (u^T\nabla^2\lambda_c(x_*)u)x_*. \tag{13.4}$$

Now, taking (13.3) into account we have

$$F''(x_*;c)(u,u) = 2c(x_*^Tu)u - (u^T\nabla^2\lambda_c(x_*)u)x_*. \tag{13.5}$$

THEOREM 13.1. *Let $\lambda_1,\dots,\lambda_n$ denote the eigenvalues of the matrix A. Consider x_* to be an l_2 unit eigenvector of A corresponding to a simple eigenvalue λ_1. Also, let x be an arbitrary vector in R^n.*

(i) *For $c \neq 0$, $F(x;c) = 0$ if and only if x is an l_2 unit eigenvector of the matrix A in the symmetric eigenvalue problem described in section* 1.3.

(ii) *For $c \neq 0$, $F'(x_*;c)^{-1}$ exists.*

(iii) *For a choice of $c \neq 0$ the normalized Newton's method sequence* (12.3) *obtained from* (13.1)*, and denoted by*

$$x_{k+1} = S(x_k), \quad k = 0,1,\dots, \tag{13.6}$$

is locally well defined, i.e., well defined in a neighborhood of x_*.

Moreover,

(a) $S(x_*) = x_*$,

(b) $S'(x_*) = 0$, *and*

(c) $\|S''(x_*)(x_k - x_*, x_k - x_*)\| \leq \frac{c}{\min_{j\neq 1}|\lambda_j - \lambda_1|}(\|x_k - x_*\|_2)^3$.

Hence, the l_2 normalized Newton's method applied to (13.1) *with $c \neq 0$ is locally q-cubically convergent to each unit eigenvector of A corresponding to a simple eigenvalue.*

Proof of (i). This statement follows from Theorem 7.1 and the fact that a unit eigenvector cannot be the zero vector, i.e., it must be a regular point.

Proof of (ii). The statement follows from part (iii) of Theorem 11.1.

Proof of (iii). Statement (a) follows from part (i) of the theorem; statement (b) follows from part (i) of the theorem and (12.5). We now prove statement (*c*). Let x_* be a unit eigenvector of A corresponding to a simple eigenvalue, say, λ_1. From (b) above we know that $S'(x_*) = 0$. This is sufficient to give local q-quadratic convergence. Our task is now to establish q-cubic convergence. We will use (12.6) to evaluate $S''(x_*)$.

From part (iii) of Theorem 11.1 we see that the eigenvalues of

$$V = (I - x_*x_*^T)F'(x_*;c)^{-1} \tag{13.7}$$

are zero and $(\lambda_j - \lambda_1)^{-1}$, $j \neq 1$. Also, V is symmetric since A is symmetric. Hence $\|V\| = 1/\min_{j\neq 1}|\lambda_j - \lambda_1|$. Observe from part (iii) of Theorem 11.1 that x_* as an

eigenvector of A is also an eigenvector of $F'(x_*, c)$ and therefore of $F'(x_*, c)^{-1}$. Hence $Vx_* = 0$. It follows from (12.6), (13.5), and the material presented above that

$$\|S''(x_*)(u,u)\| = 2c\|u\|_2|x_*^T u|/\min_{j\neq 1}|\lambda_j - \lambda_1|. \tag{13.8}$$

Now, let $\{x_k\}$ denote the normalized Newton iteration sequence (13.6). Since $\|x_k\|_2 = \|x_*\|_2 = 1$, we have

$$x_*^T(x_* - x_k) = 1 - x_*^T x_k = \frac{1}{2}(x_k - x_*)^T(x_k - x_*). \tag{13.9}$$

Combining (13.8) and (13.9) with $u = x_k - x_*$ gives

$$\|S''(x_*)(x_k - x_*, x_k - x_*)\| = c\|x_k - x_*\|_2^3/\min_{j\neq 1}|\lambda_j - \lambda_1|. \tag{13.10}$$

From Taylor's theorem, recalling that $S'(x_*) = 0$, we can write

$$x_{k+1} - x_* = S(x_k) - S(x_*) = \frac{1}{2}S''(x_*)(x_k - x_*, x_k - x_*) + O(\|x_k - x_*\|_2^3). \tag{13.11}$$

It now follows from (13.10) and (13.11) that

$$\overline{\lim_k}\, \|x_{k+1} - x_*\|_2/\|x_k - x_*\|_2^3 < +\infty.$$

This establishes the cubic convergence and the theorem. ∎

COROLLARY 13.2. *Let x_* be a unit eigenvector of A corresponding to a simple eigenvalue λ_*. Then, for x_0 sufficiently close to x_*, the RQI sequence beginning with x_0 converges q-cubically to either x_* or $-x_*$.*

Proof. From the theorem consider x_0 in a neighborhood of x_*, such that the normalized Newton's method $\{x_{NN}^k\}$ obtained from (13.1) for a fixed penalty constant c beginning at x_0 is q-cubically convergent to x_*. Since $\lambda_R(x_{NN}^k) \to \lambda_*$ and $x_{NN}^k = \pm x_{RQI}^k \,\forall\, k$ (see Corollary 10.2 and Remark 10.4), it follows that the RQI sequence $\{x_{RQI}^k\}$ converges (see Theorem 4.2). However, it can only converge to x_* or $-x_*$. Moreover, since $\| - x_{NN}^k - (-x_*)\| = \|x_{NN}^k - x_*\|$, in either case the convergence is q-cubic. ∎

Remark 13.3. It is now well established that the RQI iterates and the normalized Newton iterates obtained from our approaches to problem (5.1) can be negatives of each other. Contrasting part (iii) of Theorem 13.1 to Corollary 13.2 leads us to believe that, at least locally, normalized Newton's method makes the more robust choice of sign. We are not able to ascertain that an RQI iterate will not jump out of a small neighborhood of x_*. This brings the simplicity of the formulation of RQI into question.

In Theorems 8.1, 9.1, and 10.1 we demonstrated the powerful results that each iterate produced by either inverse iteration, shifted inverse iteration, or RQI can be viewed as a normalized Newton's method iterate obtained from a given penalty function equation for specified choices of the penalty constant. However, for complete agreement of the entire sequences the penalty constant may have to be changed from iteration to iteration. Hence, technically we do not have equivalence with normalized Newton's method on a single equation. Moreover, it is our opinion that our theorems are sharp and we do not have equivalences that involve only a single equation, in our case for a fixed penalty constant. However, the contemporary literature abounds with statements along the lines "it is well known that RQI is equivalent to normalized

Newton's method." As mentioned earlier, the implied equivalences fail because either the method is not Newton's method, but a Newton-like method, or the sign of the iterates is not taken into account in the equivalence, or the equivalence does not hold for all starting points, or a combination of all three deficiencies. However, there are times when the RQI sequence does agree with the normalized Newton's method on a single equation, as the following corollary demonstrates.

COROLLARY 13.4. *Consider an RQI sequence* $\{x_{RQI}^k\}$ *starting at* x_0. *Suppose that the sequence* $\{\lambda_R(x_{RQI}^k)\}$ *converges to an eigenvalue* λ_*. *Then the following hold.*

(i) *The sequence* $\{x_{RQI}^k\}$ *converges r-cubically to a corresponding unit eigenvector* x_*.

(ii) *Moreover, for a* $c \neq 0$, *suppose that the normalized Newton sequence on* (10.2), *also starting at* x_0, *is well defined. Then this sequence converges r-cubically to* x_* *or* $-x_*$. *In the case of convergence to* x_*, *there exists an integer* $K(c)$ *such that for* $k \geq K(c)$ *the RQI sequence and the normalized Newton sequence are identical.*

(iii) *In addition, if* λ_* *is simple, then for each* $c \neq 0$ *there exists an integer* $K(c)$ *such that the RQI sequence and the normalized Newton sequence starting at* $x_{RQI}^{K(c)}$ *are identical. Hence, in this case, the convergence of* $\{x_{RQI}^k\}$ *to* x_* *is also q-cubic.*

Proof of (i). This is a standard result and the proof follows from Theorem 4.1 and Wilkinson [80, p. 636].

Proof of (ii). The task here is to show that for a given $x_0 \in R^n$ and a given penalty constant $c \neq 0$, the RQI iterates x_{RQI}^k starting with x_0 and the normalized Newton iterates x_{NN}^k starting with x_0 cannot switch in agreement and disagreement of sign an infinite number of times. We use the notation B_k to denote B given by (10.11) with $x = x_{NN}^k$, and $\alpha_k(c)$ to denote $\alpha(c)$ given by (10.14) with x given by x_{NN}^k and B given by B_k. If at the kth step we have $\alpha_k(c) > 0$ and the two iterates agree, then they will agree at the $(k+1)$th step, while if they disagree, they will disagree also at the $(k+1)$th step. On the other hand, if $\alpha_k(c) < 0$, then the opposite is true: we will see a switching behavior. Observe that an infinite number of switches implies that $\alpha_k(c) < 0$ an infinite number of times. This means from (10.14) that $(x_{NN}^k)^T B_k^{-1} x_{NN}^k$ would be negative an infinite number of times. But we know from Remark 10.4 that this quantity is arbitrarily large for large k. Hence, the last inequality in (10.5) tells us that there can be no fixed penalty constant c such that the iterates switch an infinite number of times. The r-cubic convergence of the normalized Newton sequence follows directly from the r-cubic convergence of the RQI sequence, observing that $\| - x_{NN}^k - (-x_*)\| = \|x_{NN}^k - x_*\|$.

Proof of (iii). The proof follows from Remark 10.3, Corollary 10.2, and part (iii) of Theorem 13.1. Observe that by the local convergence of the normalized Newton's method (13.6) for any choice of $c \neq 0$ there exists an integer $K(c)$ such that for $k \geq K(c)$, the normalized Newton iterate at x_{RQI}^k is well defined and equal to the Rayleigh quotient iterate at x_{RQI}^k since both sequences converge to x_*. ∎

Part (ii) of Corollary 13.4 actually implies the following global convergence result.

COROLLARY 13.5. *For a fixed penalty constant c, the normalized Newton sequence, when it is well defined, is globally convergent in the sense of RQI, i.e., if for a given* x_0 *the RQI sequence converges, then so does the normalized Newton sequence for this* x_0.

Remark 13.6. According to Corollary 13.4 it could happen that the tail of the normalized Newton's method sequence coincides with the tail of the RQI sequence. Hence it would be r-cubically convergent, even though the Jacobian at the limit point may be singular, i.e., λ_* may not be simple. This would be a surprising phenomenon in standard Newton's method theory. On the other hand, if λ_* is simple, then Corollary 13.2 tells us that both sequences would be q-cubically convergent and this is a surprising phenomenon in standard RQI theory.

Remark 13.7. If in Theorem 13.1 we used Newton's method instead of the normalized Newton's method, then the convergence would be q-quadratic. Thus the use of the normalization improves the convergence from quadratic to cubic. This contrasts with inverse iteration and shifted inverse iteration where the normalization causes the convergence rate to deteriorate from quadratic to linear because the zeros of the governing equation are not unit vectors.

In establishing q-cubic convergence to zero of the error angle sequence in RQI, Parlett [49, p. 72] states, "In part the phenomenal convergence rate (of RQI) can be attributed to the stationarity of the Rayleigh quotient at eigenvectors," i.e., to the fact that $\nabla\lambda_R(x_*) = 0$; see (13.3).

A similar derivation to that used for (13.8) using (13.4) instead of (13.5) leads to

$$\|S''(x_*)(u,u)\| = 2\|Vu\|_2|\lambda_R(x_*)^T u - cx_*^T u|, \tag{13.12}$$

where $\lambda_R(x)$ is the Rayleigh quotient and V is given by (13.7).

Knowing that $\nabla\lambda_R(x_*) = 0$ (see (12.3)), we will have q-cubic convergence from (13.12) if and only if $x_*^T u = O(\|u\|_2^2)$ for $u = x_* - x_k$. This observation completes the picture in the sense that q-cubic convergence of the iteration sequence comes from the fact that the Rayleigh quotient is stationary at x_* and that $\|x_*\|_2 = \|x_k\|_2 = 1$ implies

$$x_*^T(x_* - x_k) = \frac{1}{2}\|x_* - x_k\|_2^2. \tag{13.13}$$

Neither fact alone accounts for the q-cubic convergence. Thus, our version of Parlett's statement is that the q-cubic convergence rate (of RQI) is due to the stationarity of the Rayleigh quotient at eigenvectors and the interesting property of the Euclidean norm (13.13).

14. The Convergence Constant in the Normalized Newton's Method. Our next task will be to demonstrate the striking result that the convergence rate constant in the normalized Newton method equivalent of RQI is less than or equal to 1. We first establish a lemma that is based on the material contained on pages 73–74 of Parlett [49].

LEMMA 14.1. *Let $\{x_k\}$ denote the normalized Newton sequence* (13.6) *for a given choice of penalty c. Assume that for all k the penalty constant c is such that the conditions of Theorem* 10.1 *are satisfied, and therefore x_k is well defined. Finally, assume that $\{x_k\}$ converges to x_* (necessarily a unit eigenvector of A) with associated eigenvalue λ_*. If*

$$\eta_k = (x_k - x_*)/\|x_k - x_*\|_2, \quad k = 0, 1, \ldots, \tag{14.1}$$

then either η_k converges to a unit eigenvector orthogonal to x_*, or there are two eigenvalues of A equidistant from λ_*, say, λ_p and λ_q, and the accumulation points of

$\{\eta_k\}$ are of the form $\alpha v_p + \beta v_q$, where $\alpha^2 + \beta^2 = 1$ and v_p and v_q are unit eigenvectors associated with λ_p and λ_q, respectively, and orthogonal to x_*.

Proof. Let $\{v_1, \ldots, v_n\}$ denote an orthonormal set of eigenvectors of A satisfying $Av_j = \lambda_j v_j$, $j = 1, \ldots, n$. We lose no generality by assuming that $v_1 = x_*$, and hence $\lambda_* = \lambda_1$. For $k = 0, 1, \ldots$, write

$$x_k = x_* \cos\phi_k + u_k \sin\phi_k, \tag{14.2}$$

where u_k is a unit vector orthogonal to x_* and ϕ_k denotes the angle between x_k and x_*, i.e., $\cos\phi_k = x_*^T x_k$. Applying $(A - \lambda_R(x_k)I)^{-1}$ to (14.2) gives

$$(A-\lambda_R(x_k)\cdot I)^{-1}x_k = x_* \cos\phi_k/(\lambda_1 - \sigma(x_k)) + u_{k+1}\sin\phi_k \|A-\lambda_R(x_k)\cdot I)^{-1}u_k\|, \tag{14.3}$$

where

$$u_{k+1} = (A - \lambda_R(x_k)\cdot I)^{-1}u_k / \|A - \lambda_R(x_k)\cdot I)^{-1}u_k\|_2 \tag{14.4}$$

and is orthogonal to x_*. According to Theorem 9.1 and the definition of RQI we see that $(A - \lambda_R(x_k)\cdot I)^{-1}x_k$ is a multiple of x_{k+1}; hence $\{u_k\}$ defined by (14.2) can be generated by the formula (14.4). Suppose that

$$|\lambda_i - \lambda_1| = \min_{j\neq 1} |\lambda_j - \lambda_1|. \tag{14.5}$$

Recall that

$$(A - \lambda_R(x_k)\cdot I)^{-1}v_j = (\lambda_j - \lambda_R(x_k))^{-1}v_j. \tag{14.6}$$

Suppose for the moment that $\lambda = \lambda_R(x_k)$ for all k. Since $x_*^T u_0 = 0$, we can write

$$u_0 = \alpha_2 v_2 + \cdots + \alpha_n v_n. \tag{14.7}$$

From (14.4), (14.6), and (14.7), we have

$$u_k = \left(\sum_{j\neq 1}(\lambda_j - \lambda)^{-k}\alpha_j v_j\right) \Big/ \left(\sum_{j\neq 1}(\lambda_i - \lambda)^{-2k}\alpha_j^2\right)^{\frac{1}{2}}. \tag{14.8}$$

Now, multiplying the top and bottom of (14.8) by $(\lambda_i - \lambda)^k$, where λ_i is a member of $\{\lambda_2, \ldots, \lambda_n\}$ which is closest to λ_*, we see that the coefficients of v_j for any j such that λ_j is not closest to λ converge to zero as $k \to +\infty$. Therefore, if λ_i is uniquely closest to λ, then $u_k \to v_i$. However, if $\lambda_p = \lambda_1 + \delta$ and $\lambda_q = \lambda_1 - \delta$, then any accumulation point of $\{u_k\}$ will be of the form $\alpha v_p + \beta v_q$. Since $v_p^T v_q = 0$, we have $\alpha^2 + \beta^2 = 1$. The same arguments hold if we replace λ with $\lambda_R(x_k)$.

Our proof will be completed once we show that the accumulation points of $\{\eta_k\}$ are the same as the accumulation points of $\{u_k\}$ or of $\{-u_k\}$ or of both sequences. In the former case the α and β in the statement of the lemma will be the same as the α and β described in the paragraph above. In the second case they will be the negatives of these α and β, and in the third case they will be the above α and β or their negatives. From (14.2), we can write

$$x_k - x_* = x_*(1 - \cos\phi_k) - u_k \sin\phi_k. \tag{14.9}$$

Since $x_*^T u_k = 0$, from (14.9) we have

$$\|x_* - x_k\|_2^2 = 2(1 - \cos\phi_k). \tag{14.10}$$

Combining (14.1), (14.9), and (14.10) gives

$$\eta_k = \alpha_k x_* + \beta_k u_k, \tag{14.11}$$

where

$$\alpha_k = (1 - \cos\phi_k)^{\frac{1}{2}}/\sqrt{2} \tag{14.12}$$

and

$$\beta_k = -\sin\phi_k/(2 - 2\cos\phi_k)^{\frac{1}{2}}. \tag{14.13}$$

Since $\phi_k \to 0$, it follows that $\alpha_k \to 0$. Also, it is easy to show that $\beta_k^2 \to 1$; hence, the accumulation points of $\{\beta_k\}$ are either 1, -1, or both. It follows that the accumulation points of $\{\eta_k\}$ are the same as those of $\{u_k\}$ or $\{-u_k\}$, or both. □

THEOREM 14.1. *Let $\{x_k\}$ denote the normalized Newton sequence* (13.6) *for a given choice of penalty constant $c \neq 0$. Assume that for all k the penalty constant c is such that the conditions of Theorem* 10.1 *are satisfied and therefore x_k is well defined. Finally, assume that $\{x_k\}$ converges to x_* (necessarily a unit eigenvector of A) with associated simple eigenvalue λ_1. Then*

$$\overline{\lim_k} \|x_{k+1} - x_*\|_2/\|x_k - x_*\|_2^3 \leq 1. \tag{14.14}$$

Proof. Let us begin by collecting the formulae that will be needed. If

$$\lambda_R(x) = x^T Ax/x^T x \tag{14.15}$$

and

$$\lambda_c(x) = \lambda_R(x) - \frac{c}{2}(x^T x - 1), \tag{14.16}$$

then from (10.8) and (10.9) we have

$$\nabla\lambda_c(x_*) = -cx_*. \tag{14.17}$$

Differentiation in (10.9) leads to

$$\nabla^2\lambda_c(x_*) = 2(A - \lambda_1 I) - cI. \tag{14.18}$$

Moreover, if

$$F(x; c) = (A - \lambda_c(x) \cdot I)x,$$

then $F'(x; c)$ is given by (10.10), $F''(x_*; c)$ is given by (13.5), and

$$F'''(x; c)(\eta, \eta, \eta) = -3(\eta^T \nabla^2\lambda_c(x_*)\eta)\eta - \lambda_c'''(x_*)(\eta, \eta, \eta)x_*. \tag{14.19}$$

Let us write

$$F'(x_*; c)^{-1} = V + c^{-1}x_* x_*^T, \tag{14.20}$$

where

$$V = \sum_{j \neq 1} (\lambda_j - \lambda_1)^{-1} v_j v_j^T. \tag{14.21}$$

The correctness of (14.20) follows from the fact that both sides of (14.20) agree on $v_1, \ldots, v_n$. Observe that

$$(I - x_* x_*^T) F'(x_*; c)^{-1} = V. \tag{14.22}$$

Let us use the notation

$$e_+ = x_{k+1} - x_* \quad \text{and} \quad e = x_k - x_*. \tag{14.23}$$

From Taylor's theorem, recalling that $S'(x_*) = 0$, we can write

$$e_+ = S(x_k) - S(x_*) = \frac{1}{2} S''(x_*)(e, e) + \frac{1}{6} S'''(e, e, e) + O(\|e\|^4). \tag{14.24}$$

Substituting from (12.6) and (12.7) into (14.24) and recalling (14.22), we obtain

$$e_+ = \frac{1}{2} V F''(x_*) \left(e - F'(x_*)^{-1} F''(x_*)(e, e), e \right) + \frac{1}{3} V F'''(x_*)(e, e, e). \tag{14.25}$$

Recall from (13.9) the important fact that for e given by (14.23),

$$-x_*^T e = \frac{1}{2} \|e\|_2^2. \tag{14.26}$$

A straightforward computation using (14.19) and the fact that $V x_* = 0$ leads to

$$V F'''(x_*; c)(e, e, e) = -3 (e^T \nabla^2 \lambda_c(x_*) e) V e. \tag{14.27}$$

Another straightforward, but this time quite lengthy, calculation referring often to (14.26) gives

$$\begin{aligned} V F'''(x_*; c) \left(e - F'(x_*; c)^{-1} F''(x_*; c)(e, e), e \right) \\ = \left(e^T \nabla^2 \lambda_c(x_*) e \right) V e + c\, O(\|e\|^3). \end{aligned} \tag{14.28}$$

Combining (14.25), (14.27), and (14.28) leads to

$$e_+ = -\frac{1}{2} \left(e^T \nabla^2 \lambda_c(x_*) e \right) V e + c\, O(\|e\|^3) + O(\|e\|^4). \tag{14.29}$$

Substituting (14.18) into (14.29) gives

$$e_+ = -\left(e^T (A - \lambda_1 I) e \right) V e + c\, O(\|e\|^3) + O(\|e\|^4). \tag{14.30}$$

Letting $\eta_k = (x_k - x_*)/\|x_k - x_*\|$, we can write (14.30) as

$$(x_{k+1} - x_*)/\|x_k - x_*\|_2^3 = -\eta_k^T (A - \lambda_1 I) \eta_k V \eta_k + cO(1) + O(\|x_k - x_*\|_2). \tag{14.31}$$

Our objective is to show that the first two terms in (14.31) have the same accumulation points and to then apply Lemma 14.1. For any infinite sequence of indices, the fourth term in (14.31) converges to zero. However, the $c\,O(1)$ term is problematic since it is not necessarily zero. We circumvent this problem by demonstrating

that it is possible to change $c \neq 0$ in the normalized Newton construction process with arbitrarily small $c \neq 0$ without changing the iterates x_k for sufficiently large k. Toward this end, for comparison purposes, we introduce the RQI iterates beginning with x_0 and denote them by x^k_{RQI} for $k = 0, 1, \ldots$. Our assumptions imply that these iterates are well defined. For clarity of presentation we will also denote the normalized Newton iterate x_k by x^k_{NN}. From (10.18) we have

$$x^k_{NN} = \pm x^k_{RQI} \ \ \forall \ \ k. \tag{14.32}$$

Now from (10.13) and (10.14) we have, recalling the notation introduced in the proof of Corollary 10.2,

$$x^{k+1}_{NN} = NOR(x^k_{NN}) = \frac{\alpha_k(c)}{|\alpha_k(c)|}\, RQI(x^k_{NN}), \tag{14.33}$$

where

$$\alpha_k(c) = \frac{-c}{1 - cx^T_k B^{-1}_k x_k} \tag{14.34}$$

with $x_k = x^k_{NN}$, $B_k = [A - \lambda_R(x_k)I]$, and $\lambda_R(x_k)$ is the Rayleigh quotient at x_k.

Now, observe that we can write (14.34) as

$$\alpha_k(c) = \frac{-c}{1 - c\|B^{-1}_k x_k\|(x_k)^T RQI(x_k)}. \tag{14.35}$$

We have two cases to investigate:

Case 1: $\{x^k_{RQI}\}$ converges to x_*.

In this case, for k large $x^k_{NN} = x^{k+1}_{RQI}$ and $\alpha_k(c) > 0$. Also, $(x^k_{NN})^T x^{k+1}_{RQI} \to x^T_* x_* (= 1)$ as $k \to \infty$. From Remark 10.4 we see that the norms of the unnormalized Rayleigh quotient iterates must form an unbounded sequence, i.e., $\|B^{-1}_k x^k_{RQI}\|_2 \to \infty$ as $k \to \infty$. Hence, from (14.35), if $c > 0$, then for each $0 < \hat{c} < c$ there exists $K(\hat{c})$ such that $\alpha_k(\hat{c}) > 0 \ \forall \ k \geq K(\hat{c})$. For these values of k, the normalized Newton iterates generated using the penalty constants c or $\hat{c}$ are identical. The same result follows even more directly from (14.35) in the case that $c < 0$ for $c < \hat{c} < 0$.

Case 2: $\{x^k_{RQI}\}$ converges to $-x_*$.

In this case we will compare $\{-x^k_{NN}\}$ to $\{x^k_{RQI}\}$. For k large $-x^k_{NN} = x^k_{RQI}$. Since $\lambda_R(x) = \lambda_R(-x)$ it follows that $\alpha_k(c)$ is the same for x^k_{NN} or $-x^k_{NN}$. From (14.33) we have

$$-x^{k+1}_{NN} = \frac{\alpha_k(c)}{|\alpha_k(c)|} x^{k+1}_{RQI}.$$

Hence $\alpha_k(c) > 0$. Observe that

$$RQI(x^k_{NN}) = -RQI(-x^k_{NN}) = -x^{k+1}_{RQI} = x^{k+1}_{NN};$$

so $(x^k_{NN})^T RQI(x^k_{NN}) \to x^T_* x_* = 1$. As in case 1, if $c > 0$, then for each $0 < \hat{c} < c$ or if $c < 0$, then for each $c < \hat{c} < 0$, there exists $K(\hat{c})$ such that $\alpha_k(\hat{c}) > 0 \ \forall \ k \geq K(c)$. For these values of k, the negatives of the normalized Newton iterates, hence the normalized Newton iterates, generated using penalty constant c or $\hat{c}$ are identical.

This construction allows us to conclude that the first two terms in (14.31) must have the same accumulation points. We now investigate the two situations depicted in Lemma 14.1.

Case 1: η_k converges to v_i. Then

$$\begin{aligned}\lim_k \|x_{k+1} - x_*\|_2/\|x_k - x_*\|_2^3 =& |v_i^T(A - \lambda_1 I)v_i\|Vv_i\| \\ =& \ |\lambda_i - \lambda_1|/|\lambda_i - \lambda_1| \\ =& \ 1.\end{aligned}$$

Case 2: A subsequence of $\{\eta_k\}$ converges to $\alpha v_p + \beta v_q$, where $\alpha^2 + \beta^2 = 1$, $\lambda_p = \lambda_1 + \delta$, and $\lambda_q = \lambda_1 - \delta$. For the sake of notation let us assume that the subsequence is $\{\eta_k\}$ itself. No generality is lost with this assumption. We have

$$\begin{aligned}\lim_k \|x_{k+1} - x_*\|_2/\|x_k - x_*\|_2^3 &= (|\alpha^2\delta + \beta^2\delta|(\alpha^2 + \beta^2)/\delta^2)^{\frac{1}{2}} \\ &= |\alpha^2 + \beta^2| \leq 1.\end{aligned}$$

This finally proves the theorem. ∎

COROLLARY 14.2. *The q-cubic convergence discussed in Corollary* 13.4 *has a convergence constant less than or equal to* 1.

15. Summary and Concluding Remarks. In this work, we have related several well-known algorithms for the symmetric eigenvalue problem to well-known algorithms from nonlinear programming applied to a carefully chosen nonlinear program. We have given rigorous demonstrations that an iterate generated by inverse iteration, shifted inverse iteration, or RQI can be viewed as an iterate generated by the normalized Newton's method applied to minimizing the l_2 penalty function, the l_2 augmented Lagrangian, or what Tapia [72] calls the multiplier substitution equation, respectively, for appropriately chosen penalty constants. We argue that while RQI and the normalized Newton's method are very closely related, they are not equivalent, as many authors claim. It was demonstrated that for simple eigenvalues the normalized Newton's method is always well defined locally, whereas inverse iteration, shifted inverse iteration, and RQI are not. It is known that RQI gives satisfactory performance near the solution (eigenvector), even though the defining matrix is nearly singular. From our point of view, this algorithm is calculating iterates that are the same as the iterates given by an algorithm (normalized Newton's method) which is well defined at the solution and in a neighborhood of the solution. Hence, we may consider RQI as having a removable singularity at the solution. Similar statements hold for inverse iteration and shifted inverse iteration. We have demonstrated that if the RQI sequence converges to an eigenvector corresponding to a simple eigenvalue, then eventually, i.e., for large iteration index k, the sequence is identical to a normalized Newton sequence and the convergence is q-cubic and not just r-cubic. Moreover, we have demonstrated that this q-cubic convergence of RQI is a direct result of the stationarity of the Rayleigh quotient at an eigenvector and the Euclidean norm equality $2x^T(x - y) = (x - y)^T(x - y) = \|x - y\|_2^2$ for x and y with Euclidean norm 1. No doubt the Euclidean norm is special, as many authors have argued. Finally, we have established the striking result that the constant in our cubic convergence result is less than or equal to 1. This study promotes the sense that the strength of the normalized Newton's method lies in its powerful local convergence properties and its q-cubic behavior, while the strength of RQI is its global behavior and its formulation

simplicity. However, it is this formulation simplicity that compromises RQI's local behavior relative to the normalized Newton's method; see Remark 13.3.

Appendix A. A Brief Introduction to Convergence Rate and Algorithmic Convergence. When studying iterative algorithms, much can be learned by determining the convergence rate of the sequences that they generate. However, the convergence rate is only a part of the story because it is an asymptotic property. It guarantees the behavior of the iteration sequence after some fixed index of the sequence. In practice, whether this fixed index appears relatively early or relatively late will be more important in determining the effectiveness of the algorithm than the convergence rate.

A.1. Convergence Rate. Consider a sequence $\{x_k\} \subset R^n$ converging to a point $x_* \in R^n$. Assume that $\forall k$, $x_k \neq x_*$. Given $p \geq 1$, by the q_p *factor* we mean

$$q_p\{x_k\} = \overline{\lim_k} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^p}.$$

Observe that the q_p *factor* is norm dependent if $q_p \in (0, \infty)$, while it is norm independent if it is equal to 0 or ∞. We say that the *q-convergence rate* of the sequence is *at least* p if

$$q_p\{x_k\} < +\infty. \tag{A.1}$$

Moreover, we say that the *q-convergence rate* or *order* of the sequence is (*exactly*) p_*, where p_* is the supremum of all p such that the q-convergence order is at least p. In the case that $q_1\{x_k\} = +\infty$ we will say that the q-convergence rate is 1.

Notice that when we say that the q-convergence rate is p, then p is uniquely determined; however, when we say that the q-convergence rate is at least p, then p is not uniquely determined.

Observe that (A.1) is equivalent to the statement that there exist $c \geq 0$ and k_0 such that

$$\|x_{k+1} - x_*\| \leq c\|x_k - x_*\|^p \quad \forall\, k \geq k_0.$$

The following convergence notions will be of particular interest to us in the current study. Again consider a sequence $\{x_k\} \subset R^n$ converging to a point $x_* \in R^n$. Assume that $\forall\, k$, $x_k \neq x_*$. We say that the convergence of $\{x_k\}$ to x_* is

(i) q-linear if $q_1\{x_k\} < 1$,
(ii) q-quadratic if $q_2\{x_k\} < +\infty$,
(iii) q-cubic if $q_3\{x_k\} < +\infty$.

Observe that a sequence which is q-cubically convergent is also q-quadratically convergent and q-linearly convergent. This slight abuse of terminology will not create problems and is greatly preferred to the alternative of defining these notions so that they are mutually exclusive.

In addition to the q-convergence notions, we have the following r-convergence notions. As before, consider a sequence $\{x_k\} \subset R^n$ converging to a point $x_* \in R^n$. Assume that $\forall\, k$, $x_k \neq x_*$. We say that the convergence of $\{x_k\}$ to x_* has a particular *r-convergence property* if there exists a sequence of scalars (error bounds) $\{a_k\}$ converging to zero satisfying

$$\|x_{k+1} - x_*\| \leq a_k \;\forall\, k$$

and the convergence of $\{a_k\}$ to zero has the corresponding q-convergence property.

It follows that q-convergence rate is a stronger notion of convergence than is r-convergence rate. The primary distinction between the two is essentially one of consistent behavior of successive sequence elements. We see that q-convergence guarantees the particular behavior at each element of the sequence, whereas r-convergence guarantees only that the particular property will be demonstrated on average.

The following facts are well known; moreover, their proofs are reasonably straightforward:

(i) q-linear convergence is norm dependent; all other q- and r-convergence notions are norm independent.
(ii) The r-order is greater than or equal to the q-order and may be larger.
(iii) If a sequence has a particular q-convergence property, then it has the same r-convergence property.
(iv) If a sequence of pairs $\{(x_k, y_k)\}$ has a particular q-convergence property, then the sequences $\{x_k\}$ and $\{y_k\}$ have the same r-convergence property, but not necessarily the same q-convergence property. Conversely, if the sequences $\{x_k\}$ and $\{y_k\}$ have a given q-convergence rate, then the sequence $\{(x_k, y_k)\}$ has the given q-convergence rate.

The definitive study of the convergence notion we have described can be found in Chapter 9 of Ortega and Rheinboldt [44]. The qualifiers q and r are due to Ortega and Rheinboldt; they use q to signify quotient factors and r to signify root factors. However, their development of r-notions is not what has been presented here. Our development of r-notions has followed Dennis who in 1968 [13] introduced the notion of quasi-order, which is now known to be equivalent to Ortega and Rheinboldt's notion of r-order. The distinction between q- and r-convergence is quite meaningful and useful and is essentially sacred to workers in the area of computational optimization. However, for reasons not well understood, computational scientists who are not computational optimizers seem to be at best only tangentially aware of the distinction.

A.2. Algorithm Convergence. For the sake of directness and simplicity, we will pose and study our various Newton's method formulations in the general framework of functional iteration. It is well known that for Newton's method and related algorithms, such a general approach requires stronger assumptions on differentiation than does the standard Newton's method analysis. However, this is not an issue in our eigenvalue problem study, since all our functions are infinitely smooth. The general framework of functional iteration offers us flexibility and has a certain degree of pedagogical convenience and value.

Consider $S: R^n \to R^n$. We say that x_* is a fixed point of S if $x_* = S(x_*)$. Fixed point problems motivate the iterative procedure

$$x_{k+1} = S(x_k), \quad k = 0, 1, \ldots. \tag{A.2}$$

The functional iteration procedure is said to be locally convergent to a fixed point x_* if the sequence generated according to (A.2) converges to x_* for all x_0 in an open set $D(x_*)$ containing x_*. It is said to be linearly, quadratically, or cubically convergent to x_* if, under standard assumptions, all sequences which converge to x_* converge linearly, quadratically, or cubically, respectively, to x_* as defined in section A.1. The following convergence theorem will play a basic role in our analysis.

THEOREM A.1. *Consider $S: R^n \to R$ with fixed point x_*. Assume that S has two continuous derivatives in D, a neighborhood of x_*, and*

(i) $S'(x_*) = 0$.

Then functional iteration (A.2) *is locally q-quadratically convergent to* x_*. *If* S *has three continuous derivatives in* D *and, in addition to* (i),

(ii) $\|S''(x_*)(x-x_*, x-x_*)\| \leq K\|x-x_*\|^3$

for some constant K *and for all* $x \in D$, *then functional iteration* (A.2) *is locally q-cubically convergent to* x_*.

Proof. The proofs of both parts are direct applications of Taylor's theorem with remainder. For the first part we write

$$\begin{aligned}\|x_+ - x_*\| &= \|S(x) - S(x_*)\| \\ &= \left\| S'(x_*)(x-x_*) + \frac{1}{2}\int_0^1 S''(x+\theta(x-x_*))(x-x_*, x-x_*)d\theta \right\| \\ &\leq \frac{1}{2}\left\| \int_0^1 S''(x_*+\theta(x-x_*))(x-x_*, x-x_*)d\theta \right\| \\ &\leq M\|x-x_*\|^2.\end{aligned}$$

The continuity of the second derivative allows us to choose such a constant M.

For the second part we obtain in an analogous fashion

$$\|x_+ - x_*\| \leq (K+M)\|x-x_*\|^3. \qquad \square$$

If we consider Newton's method for the nonlinear equation $F(x) = 0$, then we see that we can write it as a functional iteration for the fixed point problem $x_* = S(x_*)$, where

$$S(x) = x - F'(x)^{-1}F(x).$$

For the nonlinear equation $F(x) = 0$, the standard Newton's method theory assumptions for local and q-quadratic convergence of Newton's method are as follows:

B1. There exists $x_* \in R^n$ such that $F(x_*) = 0$.

B2. The Jacobian operator F' exists and is Lipschitz continuous in an open neighborhood of x_*.

B3. The Jacobian matrix $F'(x_*)$ is nonsingular.

For a convergence proof and further detail, see page 90 of Dennis and Schnabel [14].

Our functional iteration approach to the local q-quadratic convergence of Newton's method uses B1 and B3 and replaces B2 with a stronger smoothness requirement. In what follows we shall assume that we always have as many derivatives as the theory demands. As previously stated, this is not an issue in our eigenvalue application.

We now show that appropriate smoothness, existence of a solution, and invertibility of the Jacobian matrix at the solution are sufficient conditions for the local q-quadratic convergence of Newton's method. If the solution also has l_2 norm 1, then these conditions are also sufficient for the local q-quadratic convergence of the l_2 normalized Newton's method.

Our point here is that, while it may not be obvious, the l_2 normalization does not hurt the rapid convergence of Newton's method under the standard assumptions, provided, of course, that the solution has l_2 norm equal to 1.

THEOREM A.2. *Consider* $F: R^n \to R^n$. *Assume that the following hold:*

(i) *There exists* x_* *such that* $F(x_*) = 0$.

(ii) F *is sufficiently smooth in a neighborhood of* x_*.

(iii) *The Jacobian matrix* $F'(x_*)$ *is nonsingular.*

Let

$$N(x) = x - F'(x)^{-1}F(x). \tag{A.3}$$

Then $N'(x_*) = 0$ *and Newton's method is locally q-quadratically convergent to* x_*.

If, in addition,

(iv) $\|x_*\| = 1$

and

$$S(x) = \frac{N(x)}{\|N(x)\|}, \tag{A.4}$$

then $S'(x_*) = 0$ and the normalized Newton's method is locally q-quadratically convergent to x_*.

Proof. From (A.3) we have

$$F'(x)N(x) = F'(x)(x) - F(x).$$

Differentiation at x^* leads to

$$F'(x_*)N'(x_*) + F''(x_*)(N(x_*) - x_*, \cdot) = 0.$$

Recall that $x_* = N(x_*)$ to obtain

$$F'(x_*)N'(x_*) = 0.$$

It follows by nonsingularity of $F'(x_*)$ that $N'(x_*) = 0$. For the second part write (A.4) as

$$\|N(x)\|S(x) = N(x). \tag{A.5}$$

Now observe that if $P(x) = \|x\|_2$, then $P(x)^2 = x^T x$; hence

$$P'(x)(\eta) = \frac{x^T \eta}{\|x\|_2}. \tag{A.6}$$

Differentiation at x_* of (A.5) using (A.6) leads to

$$\|N(x_*)\|S'(x_*) = (I - S(x_*)^T S(x_*))N'(x_*).$$

It follows that $S'(x_*) = 0$, since $N'(x_*) = 0$ and $\|N(x_*)\| = 1$. Recall that $N(x_*) = x_*$ and $\|x_*\| = 1$. ∎

Acknowledgments. Special thanks go to Leticia Velázquez for mathematical and technical support, to Mark Embree for encouraging us to finish the paper and for making numerous suggestions that improved the quality of the presentation, to Josef Sifuentes for his numerical assistance, and to Pete Stewart, who shed light on Schmidt's intricate existence proof. Valuable historical information on the method of successive approximations was given to us by Hubert Kalf and Martin Costabel, and on Jacobi's method by Anjet den Boer. Discussions with Roland Glowinski, Beresford Parlett, Florian Potra, and Dan Sorensen are much appreciated. We thank several excellent referees. One in particular went far beyond the call of duty in detailed reading and positive suggestions, leading to improvements in the presentation and several proofs. Finally, we are indebted to the associate editor Juan Meza for his choice of referees and his numerous suggestions for improving the presentation and to J. M. Sanz-Serna for several key corrections.

REFERENCES

[1] A.C. Aitken, *On Bernoulli's numerical solution of algebraic equations*, Proc. Roy. Soc. Edinburgh, 46 (1926), pp. 289–305, https://doi.org/10.1017/S0370164600022070. (Cited on p. 10)

[2] P.M. Anselone and L.B. Rall, *The solution of characteristic value-vector problems by Newton's method*, Numer. Math., 11 (1968), pp. 38–45, https://doi.org/10.1007/BF02165469. (Cited on pp. 13, 18, 19)

[3] G. Auchmuty, *Globally and rapidly convergent algorithms for symmetric eigenproblems*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 690–706, https://doi.org/10.1137/0612053. (Cited on p. 21)

[4] E. Bodewig, *Matrix Calculus*, North-Holland, Amsterdam, 1956, https://doi.org/10.1016/B978-1-4832-3214-0.50032-5. (Cited on p. 9)

[5] A. den Boer, *De Jacobi-methode van* 1845 *tot* 1990, Masters thesis, Department of Mathematics, Utrecht University, The Netherlands, 1991. (Cited on p. 9)

[6] A. Cauchy, *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires du mouvement des planètes*, Œuvres Complètes, Gauthier-Villars, Paris, Ser. 2, 9 (1891), pp. 174–195, https://archive.org/stream/uvrescompltesda00fragoog#page/n184/mode/2up. (Cited on p. 9)

[7] F. Chatelin, *Simultaneous Newton's iteration for the eigenproblem*, in Defect Correction Methods, Comput. Suppl. 5, K. Böhmer and H.J. Stetter, eds., Springer, Vienna, 1984, pp. 67–74, https://doi.org/10.1007/978-3-7091-7023-6_4. (Cited on p. 22)

[8] F. Chatelin, *Ill-conditioned eigenproblems*, in Large Scale Eigenvalue Problems, J. Cullum and R.A. Willoughby, eds., North-Holland Math. Stud. 127, North-Holland, Amsterdam, 1986, pp. 267–282, https://doi.org/10.1016/S0304-0208%2808%2972649-5. (Cited on p. 22)

[9] L. Collatz, *Functional Analysis and Numerical Mathematics*, Academic Press, New York, 1966, https://doi.org/10.1016/B978-1-4832-3183-9.50007-8, https://doi.org/10.1007/978-3-642-95028-5. (Cited on p. 13)

[10] R. Courant, *Variational methods for the solution of problems of equilibrium and vibrations*, Bull. Amer. Math. Soc., 49 (1943), pp. 1–23, https://doi.org/10.1090/S0002-9904-1943-07818-4. (Cited on p. 28)

[11] S.H. Crandall, *Iterative procedures related to relaxation methods for eigenvalue problems*, Proc. Roy. Soc. London Ser. A, 207 (1951), pp. 416–423, https://doi.org/10.1098/rspa.1951.0129. (Cited on pp. 14, 15, 16)

[12] E.R. Davidson, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices*, J. Comput. Phys., 17 (1975), pp. 87–94, https://doi.org/10.1016/0021-9991%2875%2990065-0. (Cited on p. 25)

[13] J.E. Dennis, Jr., *On Newton-like methods*, Numer. Math., 11 (1968), pp. 324–330, https://doi.org/10.1007/BF02166685. (Cited on p. 49)

[14] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, New York, 1983. Reprinted in the SIAM Classics in Applied Mathematics series, SIAM, Philadelphia, 1996, https://doi.org/10.1137/1.9781611971200. (Cited on pp. 22, 50)

[15] J. Dieudonné, *History of Functional Analysis*, North Holland, Amsterdam, 1981, https://doi.org/10.1016/S0304-0208%2808%2971850-4. (Cited on p. 8)

[16] J.J. Dongarra, C.B. Moler, and J.H. Wilkinson, *Improving the accuracy of computed eigenvalues and eigenvectors*, SIAM J. Numer. Anal., 20 (1983), pp. 23–45, https://doi.org/10.1137/0720002. (Cited on p. 21)

[17] F.H. van den Dungen, *Calcul des pôles simples d'une fonction méromorphe*, C.R. Acad. Sci. Paris, 176 (1923), pp. 1864–1865, http://gallica.bnf.fr/ark:/12148/bpt6k31295/f1864. (Cited on p. 10)

[18] A. Edelman, T.A. Arias, and S.T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353, https://doi.org/10.1137/S0895479895290954. (Cited on p. 27)

[19] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968. Reprinted in the SIAM Classics in Applied Mathematics series, SIAM, Philadelphia, 1990, https://doi.org/10.1137/1.9781611971316. (Cited on p. 28)

[20] R. Fletcher, *A new approach to variable metric algorithms*, Comput. J., 13 (1970), pp. 317–322, https://doi.org/10.1093/comjnl/13.3.317. (Cited on p. 29)

[21] A.T. Galick, *Efficient Solution of Large Sparse Eigenvalue Problems in Microelectronic Simulation*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1993, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.9179. (Cited on p. 26)

[22] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 4th ed., The Johns Hopkins University Press, Baltimore, 2013. (Cited on pp. 9, 18, 25)
[23] M.C. Gutzwiller, *Moon-Earth-Sun: The oldest three-body problem*, Rev. Modern Phys., 70 (1998), pp. 589–639, https://doi.org/10.1103/RevModPhys.70.589. (Cited on p. 8)
[24] P. Harzer, *Die säcularen Veränderungen der Bahnen der grossen Planeten*, Hirzel, Leipzig, 1895, https://goobi.tib.eu/viewer/image/829337687/112/. (Cited on p. 9)
[25] M.R. Hestenes, *Determination of eigenvalues and eigenvectors of matrices*, in Simultaneous Linear Equations and the Determination of Eigenvalues, L.J. Paige and O. Taussky, eds., Natl. Bur. Stand. Appl. Math. Ser. 29, U.S. Government Printing Office, Washington, D.C., 1953, pp. 89–94, https://archive.org/stream/simultaneousline00paig#page/88/mode/2up. (Cited on p. 19)
[26] M.R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory Appl., 4 (1969), pp. 303–320, https://doi.org/10.1007/BF00927673. (Cited on p. 29)
[27] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educ. Psychol., 24 (1933), pp. 417–441, https://doi.org/10.1037/h0071325. (Cited on p. 10)
[28] A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964, http://store.doverpublications.com/0486449726.html. (Cited on p. 13)
[29] I.C.F. Ipsen, *Helmut Wielandt's contributions to the numerical solution of complex eigenvalue problems*, in Helmut Wielandt: Mathematische Werke, Mathematical Works, Volume 2: Linear Algebra and Analysis, B. Huppert and H. Schneider, eds., Walter de Gruyter, Berlin, 1996, pp. 453–463, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.45.2671. (Cited on pp. 12, 14)
[30] I.C.F. Ipsen, *A history of inverse iteration*, in Helmut Wielandt: Mathematische Werke, Mathematical Works, Volume 2: Linear Algebra and Analysis, B. Huppert, H. Schneider, eds., Walter de Gruyter, Berlin, 1996, pp. 464–472, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.45.1134. (Cited on p. 12)
[31] I.C.F. Ipsen, *Computing an eigenvector with inverse iteration*, SIAM Rev., 39 (1997), pp. 254–291, https://doi.org/10.1137/S0036144596300773. (Cited on p. 9)
[32] C.G.J. Jacobi, *Über eine neue Auflösungsart der bei der Methode der kleinsten Quadrate vorkommenden linearen Gleichungen*, Astron. Nachr., 22 (1845), pp. 297–306, https://doi.org/10.1002/asna.18450222002. (Cited on p. 9)
[33] C.G.J. Jacobi, *Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, J. Reine Angew. Math., 30 (1846), pp. 51–94, https://gdz.sub.uni-goettingen.de/id/PPN243919689_0030?tify="pages":[55]. (Cited on pp. 9, 11)
[34] L.V. Kantorovich, *Functional analysis and applied mathematics*, Uspekhi Mat. Nauk, 3 (1948), pp. 89–185, http://mi.mathnet.ru/eng/umn8775. (Cited on p. 19)
[35] W. Kohn, *A variational iteration method for solving secular equations*, J. Chem. Phys., 17 (1949), p. 670, https://doi.org/10.1063/1.1747366. (Cited on pp. 15, 16)
[36] G. Kowalewski, *Einführung in die Determinantentheorie, einschließlich der unendlichen und der Fredholmschen Determinanten*, Veit, Leipzig, 1909, http://www.dbc.wroc.pl/dlibra/docmetadata?id=1705. (Cited on p. 8)
[37] J. Liouville, *Mémoire sur la théorie analytique de la chaleur*, Ann. Math. Pures Appl., 21 (1830), pp. 133–181, http://www.numdam.org/item?id=AMPA_1830-1831__21__133_0. (Cited on p. 8)
[38] R. von Mises, *Über die Stabilität rotierender Wellen*, Monatsh. Math. Phys., 22 (1911), pp. 33–52, https://doi.org/10.1007/BF01742790. (Cited on p. 8)
[39] R. von Mises and H. Pollaczek-Geiringer, *Praktische Verfahren der Gleichungsauflösung*, Z. Angew. Math. Mech., 9 (1929), pp. 152–164, https://doi.org/10.1002/zamm.19290090206. (Cited on p. 9)
[40] R.B. Morgan and D.S. Scott, *Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 817–825, https://doi.org/10.1137/0907054. (Cited on p. 25)
[41] C.H. Müntz, *Solution directe de l'équation séculaire et de quelques problèmes analogues transcendants*, C. R. Acad. Sci. Paris, 156 (1913), pp. 43–46, http://gallica.bnf.fr/ark:/12148/bpt6k3109m/f43. (Cited on p. 8)
[42] C.H. Müntz, *Sur la solution des équations séculaires et des équations intégrales*, C. R. Acad. Sci. Paris, 156 (1913), pp. 860–862, http://gallica.bnf.fr/ark:/12148/bpt6k3109m/f860. (Cited on p. 8)
[43] C.H. Müntz, *Das Hauptaxenproblem der quadratischen Formen und der symmetrischen Integralglei-chungen*, Prace Matematyczno-Fizyczne, 29 (1918), pp. 109–177, https://eudml.org/doc/215381. (Cited on p. 9)

[44] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, San Francisco, London, 1970, https://doi.org/10.1137/1.9780898719468. (Cited on pp. 31, 35, 49)
[45] A.M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors*, I, Arch. Rational Mech. Anal., 1 (1957), pp. 233–241, https://doi.org/10.1007/BF00298007. (Cited on p. 15)
[46] A.M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors*, II, Arch. Rational Mech. Anal., 2 (1958), pp. 423–428, https://doi.org/10.1007/BF00277939. (Cited on p. 15)
[47] A.M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors*, V, Arch. Rational Mech. Anal., 3 (1959), pp. 472–481, https://doi.org/10.1007/BF00284194. (Cited on p. 12)
[48] A. Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*, Oxford University Press, New York, 1982. (Cited on p. 9)
[49] B.N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980, https://doi.org/10.1137/1.9781611971163. (Cited on pp. 7, 42)
[50] B.N. Parlett, *What Hadamard Missed*, Internat. J. Appl. Sci. Comput., 14 (2007), pp. 116–136, https://doi.org/10.21236/ada327603. (Cited on p. 10)
[51] G. Peters and J.H. Wilkinson, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Rev., 21 (1979), pp. 339–360, https://doi.org/10.1137/1021052. (Cited on pp. 5, 9, 13, 18, 19, 20, 23)
[52] E. Picard, *Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires*, J. Math. Pures Appl. Ser. 4, 9 (1893), pp. 217–271, http://gallica.bnf.fr/ark:/12148/bpt6k107457q/f217. (Cited on p. 7)
[53] E. Picard, *Sur les équations linéaires du second ordre renfermant un paramètre arbitraire*, C. R. Acad. Sci. Paris, 118 (1894), pp. 379–383. (Not cited)
[54] E. Pohlhausen, *Berechnung der Eigenschwingungen statisch-bestimmter Fachwerke*, Z. Angew. Math. Mech., 1 (1921), pp. 28–42, https://doi.org/10.1002/zamm.19210010104. (Cited on p. 9)
[55] M.J.D. Powell, *A method for nonlinear constraints in minimization problems*, in Optimization, R. Fletcher, ed., Academic Press, New York, 1969, pp. 283–298. (Cited on p. 29)
[56] J.W.S.B. Rayleigh, *The Theory of Sound*, Vol. 1, 2nd ed., MacMillan and Co., New York, 1894, https://archive.org/stream/theorysound09raylgoog#page/n127/mode/2up. (Cited on p. 14)
[57] W. Ritz, *Theorie der Transversalschwingungen einer quadratischen Platte mit freien Rändern*, Ann. Phys., 333 (1909), pp. 737–786, https://doi.org/10.1002/andp.19093330403. (Cited on p. 9)
[58] A.H. Sameh and J.H. Wisniewski, *A trace minimization algorithm for the generalized eigenvalue problem*, SIAM J. Numer. Anal., 19 (1982), pp. 1243–1259, https://doi.org/10.1137/0719089. (Cited on p. 26)
[59] A.H. Sameh and Z. Tong, *The trace minimization method for the symmetric generalized eigenvalue problem*, J. Comput. Appl. Math., 123 (2000), pp. 155–175, https://doi.org/10.1016/S0377-0427%2800%2900391-5. (Cited on p. 26)
[60] E. Schmidt, *Entwickelung willkürlicher Functionen nach Systemen vorgeschriebener*, Kaestner, Göttingen, 1905. Reprinted in Math. Ann., 63 (1907), pp. 433–476, https://gdz.sub.uni-goettingen.de/id/PPN235181684_0063?tify="pages":[457]. (Cited on p. 8)
[61] E. Schmidt, *Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten*, Rend. Circ. Mat. Palermo, 25 (1908), pp. 53–77, https://doi.org/10.1007/BF03029116. (Cited on p. 8)
[62] H.A. Schwarz, *Ueber ein die Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung*, in Gesammelte Mathematische Abhandlungen, Vol. 1, Springer, Berlin, 1890, pp. 241–265, https://archive.org/stream/gesammeltemathem01schwuoft#page/240/mode/2up. (Cited on p. 8)
[63] L. Seidel, *Über ein Verfahren, die Gleichungen, auf welche die Methode der kleinsten Quadrate führt, sowie lineäre Gleichungen überhaupt, durch successive Annäherung aufzulösen*, Abh. Math.-Phys. Kl. Bayr. Akad. Wiss., 11 (1874), pp. 81–108, http://www.zobodat.at/publikation_articles.php?id=206736. (Cited on p. 9)
[64] V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159–182, https://doi.org/10.1023/A:1021930421106. (Cited on p. 27)
[65] T. Simpson, *Essays on Several Curious and Useful Subjects in Speculative and Mix'd Mathematicks: Illustrated by a Variety of Examples*, printed by H. Woodfall, Jun. For J. Nourse,

London, 1740, https://archive.org/stream/essaysonseveralc00simp#page/n95/mode/2up. (Cited on p. 18)

[66] G.L.G. Sleijpen and H.A. van der Vorst, *A Jacobi–Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425, https://doi.org/10.1137/S0895479894270427. (Cited on p. 9)

[67] D.C. Sorensen, *Numerical methods for large eigenvalue problems*, Acta Numer., 11 (2002), pp. 519–584, https://doi.org/10.1017/S0962492902000089. (Cited on p. 25)

[68] R.V. Southwell, *Relaxation Methods in Engineering Science*, Oxford University Press, London, 1940. (Cited on p. 14)

[69] G.W. Stewart, *Matrix Algorithms, Vol.* II: *Eigensystems*, SIAM, Philadelphia, 2001, https://doi.org/10.1137/1.9780898718058. (Cited on p. 22)

[70] G.W. Stewart, *Fredholm, Hilbert, Schmidt: Three Fundamental Papers on Integral Equations*, manuscript, translated with commentary by G. W. Stewart, 2011, http://www.umiacs.umd.edu/~stewart/FHS.pdf. (Cited on p. 8)

[71] A. Stodola, *Steam Turbines: With an Appendix on Gas Turbines and the Future of Heat Engines*, Van Nostrand, New York, 1905, https://archive.org/stream/steamturbineswit00stod#page/184/mode/2up. (Cited on p. 8)

[72] R.A. Tapia, *Quasi-Newton methods for equality constrained optimization: Equivalence of existing methods and a new implementation*, in Nonlinear Programming 3, O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, eds., Academic Press, New York, 1978, pp. 125–164, https://doi.org/10.1016/B978-0-12-468660-1.50010-4. (Cited on pp. 30, 47)

[73] R.A. Tapia and D.L. Whitley, *The projected Newton method has order* $1 + \sqrt{2}$ *for the symmetric eigenvalue problem*, SIAM J. Numer. Anal., 25 (1988), pp. 1376–1382, https://doi.org/10.1137/0725079. (Cited on pp. 5, 20, 22, 23)

[74] H. Unger, *Nichtlineare Behandlung von Eigenwertaufgaben*, Z. Angew. Math. Mech., 30 (1950), pp. 281–282, https://doi.org/10.1002/zamm.19500300839. (Cited on p. 13)

[75] L. Vianello, *Graphische Untersuchung der Knickfestigkeit gerader Stäbe*, Zeitschrift des Vereines deutscher Ingenieure, 52 (1898), pp. 1436–1443. (Cited on p. 8)

[76] H.A. van der Vorst, *Computational methods for large scale eigenvalue problems*, in Handbook of Numerical Analysis, Vol. VIII, P.G. Ciarlet and J.L. Lions, eds., Elsevier, Amsterdam, 2002, pp. 3–179, https://doi.org/10.1016/S1570-8659%2802%2908003-1. (Cited on p. 25)

[77] H. Wielandt, *Das Iterationsverfahren bei nicht selbstadjungierten linearen Eigenwertaufgaben*, Math. Z., 50 (1944), pp. 93–143, https://gdz.sub.uni-goettingen.de/id/PPN266833020_0050?tify="pages":[95]. (Cited on pp. 10, 12)

[78] H. Wielandt, *Beiträge zur mathematischen Behandlung komplexer Eigenwertprobleme, Teil* V: *Bestimmung höherer Eigenwerte durch gebrochene Iteration*, Bericht B 44/J/37, Aerodynamische Versuchsanstalt Göttingen, Germany, 1944. (Cited on pp. 9, 10)

[79] J.H. Wilkinson, *The calculation of the eigenvectors of codiagonal matrices*, Computer J., 1 (1958), pp. 90–96, https://doi.org/10.1093/comjnl/1.2.90. (Cited on p. 9)

[80] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. (Cited on pp. 9, 24, 41)

[81] H. Wittmeyer, *Berechnung einzelner Eigenwerte eines algebraischen linearen Eigenwertproblems durch Störiteration*, Z. Angew. Math. Mech., 35 (1955), pp. 441–452, https://doi.org/10.1002/zamm.19550351202. (Cited on p. 13)

[82] K. Wu, Y. Saad and A. Stathopoulos, *Inexact Newton preconditioning techniques for large symmetric eigenvalue problems*, Electron. Trans. Numer. Anal., 7 (1998), pp. 202–214, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.278. (Cited on p. 26)

[83] R. Zurmühl, *Praktische Mathematik für Ingenieure und Physiker*, Springer, Berlin, 1953, https://doi.org//10.1007/978-3-642-53283-2. (Cited on p. 13)