1. What It's All About. In 1943 and 1944 Helmut Wielandt wrote a series of five papers, with the theme 'contributions to the mathematical treatment of complex eigenvalue problems' [20, 24, 21, 22, 23]. The papers deal with eigenvalue problems for algebraic operators, for linear Fredholm integral operators, and for linear ordinary differential operators in the form of boundary value problems. In particular, they discuss the location of eigenvalues of matrices in the complex plane [20]; as well as the computation of eigenvalues and eigenfunctions via the power method [24, 21, 22], and inverse iteration [23]. Only a single paper, [24], was published in a journal; the others exist as technical reports.

In this commentary we focus on the computational aspects of Wielandt’s work on eigenvalue problems.

2. Impressions. Helmut Wielandt is well known for his work in matrix and group theory [6]. While reading his papers I realised with surprise that he also did pioneering work in computational numerical analysis.

In fact, Wielandt was a computational numerical analyst of the same calibre as Jim Wilkinson, whose ground-breaking work on round-off error analysis [26] and the algebraic eigenvalue problem [27] continues to influence computational linear algebra to this day. Both, Wielandt and Wilkinson had a thorough understanding of the mathematical theory in addition to an extraordinary intuition regarding the effects caused by finite precision arithmetic. Both used a good measure of common sense to solve problems: aiming for the simplest possible approach; avoiding excessive formalism; and not hesitating to argue on the back of the envelope, so to speak, if that was enough to get the point across. For instance, both illustrated the numerical behaviour of their methods with plenty of well-chosen examples; and both provided estimates of the computational errors and the operation counts.

Wielandt’s criteria in 1944 for designing numerical methods agree with the principles that have been formalised over the last thirty years and are in use today. His goal was to design reliable numerical methods. By this he meant methods that are simple and insensitive to computational errors [21, Section I.3(A)]. The last feature corresponds to what we now call numerical stability. Because computations were still done by hand at times, Wielandt valued simple and repetitive algorithms, as do today’s programmers of vector and parallel machines. He was also aware of problems associated with finite precision arithmetic, such as overflow and catastrophic cancellation, which continue to cause serious concern even now. There is no doubt that Wielandt would fit right into today’s scientific computing scene and feel quite comfortable there. He was clearly ahead of his time with regard to his ideas about the design of numerical
algorithms and his intuition about the effect of computational errors. This is even more astonishing considering the scarcity of computing machinery in the 1940’s.

3. Wielandt’s Problem. Wielandt’s motivation for studying eigenvalue problems came from vibration problems in the design of aircraft wings [23, §1]. The problem (called ‘Flatterrechnung’ in German) is to determine which airstream velocities cause an aircraft wing to oscillate and become unstable [21, §1.2].

The system of differential equations that models the deformation of the wing constitutes a boundary value problem [21, §1.2]

\[ Ax(\xi) = \nu^2 Bx(\xi), \]

where \( A \equiv A(\xi) \) is a differential operator that describes the elasticity properties of the wing; and the matrix \( B \equiv B(\xi, \nu/\nu) \) describes the forces acting on the wing. The unknown quantities are the complex scalar \( \nu \) representing the frequency; the vector \( x(\xi) \) representing the displacement of the wing at point \( \xi \); and the positive scalar \( \nu \) representing the airstream velocity. One wants to determine \( \nu \) and \( \nu \) so that the boundary value problem has a non-trivial solution \( x(\xi) \).

Wielandt proposes two computational formulations for this boundary value problem. In the first formulation, he fixes the airstream velocity \( \nu \). The wing is stable at this velocity if all eigenvalues \( \nu^2 \) have positive imaginary part. This formulation is computationally difficult because \( B \) itself depends on the eigenvalue. Wielandt rejects it.

In the second formulation, Wielandt fixes \( \nu/\nu \). Since the entries of \( B \) now depend only on \( \xi \), the eigenvalue problem is linear. Then the airstream velocities that induce oscillations are those for which there exist real positive eigenvalues \( \nu^2 \). Wielandt opts in favour of this formulation.

Notation. We formulate eigenvalue problems as

\[ Ax = \lambda x, \]

where the complex scalar \( \lambda \) denotes an eigenvalue. In case of the algebraic eigenvalue problem, \( A \) is a complex square matrix and the non-zero complex vector \( x \) is an eigenvector. If \( A \) is an integral or differential operator then \( x \) is a complex valued eigenfunction. We refer to the eigenvalues of maximal modulus as dominant eigenvalues.

4. Locating Eigenvalues. At the time when Wielandt wrote the papers, algorithms for determining eigenvalues mainly involved explicitly computing the characteristic polynomial and then finding the zeros of this polynomial [29, Section 6]. However, Wilkinson showed that this strategy is numerically most unreliable [25, 29], [27, Section 7.5]. This is because finite precision arithmetic introduces errors in the polynomial coefficients; and the zeros of the polynomial are extremely sensitive to these errors. In contrast, the sensitivity of eigenvalues to changes in the original matrix elements is much lower. To avoid amplifying the sensitivity of eigenvalues, many state-of-the-art numerical methods use unitary similarity transformations to simplify the matrix, and then compute the eigenvalues of the simplified matrix [3, 4, 5, 9, 14, 17, 18, 27].

In [20] Wielandt proposes two methods: one to compute the characteristic polynomial of a complex matrix \( A \) and a second one to locate the roots of a polynomial with complex coefficients. Wielandt’s objective was to develop methods for locating eigenvalues that are simpler, more efficient, and more suitable for mechanical calculations than other methods in use at the time [20, pp 1, 2].
The method for computing the coefficients of the characteristic polynomial is based on the calculation of powers of A and their traces. Unbeknownst to Wielandt, he rediscovered Leverrier's method. This method was first published in 1840 [7, §25], [11, §6.7]; and Householder calls it 'the earliest practical method of finding the characteristic polynomial' [11, p 172].

Surprisingly, Wielandt already realises the danger of catastrophic cancellation, which happens when the difference of two nearly equal numbers is of the same magnitude as the uncertainty in the numbers, hence the computed difference is totally inaccurate. To avoid subtraction of large numbers, he suggests computing the characteristic polynomial of a shifted matrix A - µI. The difficulty then is to find a suitable shift µ.

Leverrier's method requires $O(n^4)$ operations when A is of order n. A simplified version of the method for n = 4 carried out with five to seven decimal digits required at the time 3 1/2 to 4 1/2 hours of computation 'for various calculators and electronic calculators' [20, § II.5].

Wielandt's method for locating the roots of a polynomial with complex coefficients amounts to counting the number of roots in the upper half-plane. He refers to the 1898 German edition of a book by Routh that contains a method for counting the roots in the left half-plane of a polynomial with real coefficients [16, §§296-8]. In his extension to polynomials with complex coefficients, Wielandt proposes a method to count the number of roots with positive imaginary part. Locating the roots of a fourth degree polynomial when carrying five decimal digits required 20-30 minutes at the time [20, Section III.4]. It is not clear, though, whether these measurements refer to man time or machine time.

Because this technical report was apparently not published, Wielandt's method must have gone unnoticed. The same method is described, for instance, by Barnett [1, Theorem 3.15], yet Barnett attributes its origin [1, p 253] to a 1963 paper by Parks [13].

5. The Power Method. The papers [21, 22, 24] are devoted to the power method for computing eigenfunctions of linear operators, to which Wielandt simply refers as 'the iteration process'. He argues [21, §1] that in the context of vibration problems for aircraft wings, this kind of iterative method provides more reliable results because it obviates the need for additional conditions that describe the vibrations. The particular choice of such conditions, which are required by other methods, strongly affects the computational results and therefore introduces an element of arbitrariness into these methods.

The power method applied to a complex square matrix A generates a sequence

---

1 Wielandt emphasizes that computing the necessary matrix powers via inner products constitutes a repetitive process with a simple control flow that can be easily checked for errors [20, §II.1, p 4]. In fact, he suggests the use of templates to facilitate the computation of inner products, an idea that has recently resurfaced on a grander scale [2]. He checks for errors in the matrix multiplication $C = AB$ by comparing the column sums of $C$ to column sums determined from linear combinations of columns of A. This approach is also recommended by Faddeeva [7, §25]. Wielandt argues that the number of operations due to error checking is of lower order than the number of operations required to redo the whole process from scratch [20, §II.2].

2 In comparison, a state of the art method for finding eigenvalues requires an order of magnitude fewer operations; e.g. to find the eigenvalues of a real symmetric matrix via Householder tridiagonalisation followed by the implicitly shifted QR algorithm costs $\frac{1}{2}n^3 + O(n^2)$ arithmetic operations [9, §§8.2.3].
of vectors $y_k$ from a given starting vector $y_0$ according to
\[ y_k = s_k A y_{k-1}, \quad k \geq 1. \]
Here $s_k$ represents a scalar responsible for normalising $y_k$, either in some norm or with regard to a particular component. If everything goes well, the iterates $y_k$ converge to an eigenvector associated with a dominant eigenvalue $\lambda$; and the ratio $y_{k,i} / y_{k-1,i}$ of two corresponding components of $y_k$ and $y_{k-1}$ converges to $\lambda$. At Wielandt’s time the power method was known for the algebraic eigenvalue problem, e.g. [8, §4.17], [15, §V]; and for integral equations with Hermitian kernel [19], [24, p.95].

In the journal paper [24] Wielandt analyses the power method for algebraic operators, for linear Fredholm integral operators, and for linear ordinary differential operators in the form of boundary value problems. His primary interest in [24], though, was the eigenvalue problem for integral operators [24, §4A]. Wielandt’s train of thought for analysing the power method [24, §1A] is essentially the following: Start out with the algebraic eigenvalue problem, since, among the three types of operators, it is best understood. The results for matrices then carry over, more or less naturally, to integral operators. This in turn opens the door to boundary value problems with an explicit, known Green’s function because they can be expressed as Fredholm integral operators of the second kind [24, §5B].

6. Algebraic Operators. In his ‘convergence proof’ [24, §4A], as he optimistically calls it, Wielandt analyses the composition of the iterates $y_k$ as the power method progresses [24, §(4Ad)]: If the starting vector $y_0$ contains contributions of eigenvectors associated with dominant eigenvalues then the contributions of these eigenvectors in the $y_k$ increase as $k \to \infty$.

In particular when the matrix $A$ has only a single dominant eigenvalue $\lambda$, the $y_k$ converge to an eigenvector associated with $\lambda$, provided the starting vector $y_0$ contains a contribution of this eigenvector. This is because $y_k$ is a multiple of the vector $A^k y_0$. Hence, as $k$ increases, the contribution in $y_k$ of the eigendirection associated with $\lambda$ increases faster than the contributions associated with the smaller eigenvalues.

Judging by the way Wielandt formulated his convergence result, he realised very well that it does not always imply convergence. About thirty years later, Wilkinson shows that the power method does in general not converge when $A$ has several dominant but unequal eigenvalues, e.g. complex conjugate eigenvalues [28, p 363]. Although the $y_k$ are composed to an ever larger extent of eigenvectors associated with the dominant eigenvalues, they do not converge to any particular direction. Alternatively, if $A$ has a dominant eigenvalue that is defective then convergence of the iterates to an associated eigenvector is very slow because the contributions of the associated generalised eigenvectors diminish only slowly [28, p. 362].

7. Integral and Differential Operators. The power method $y_k \equiv A y_{k-1}$ for Fredholm integral operators with symmetric kernel was already known in 1930 [19, §II.7]. Wielandt extends it to nonsymmetric kernels [24, §§4B, C]. For a linear Fredholm integral operator of the second kind, the operator $A$ is defined by [24, §4C]
\[ Ay = \int_0^1 K(\xi, \eta) y(\eta)d\eta + k_1(\xi) y(\xi_1) + \cdots + k_m(\xi) y(\xi_m), \]
where the given complex-valued functions $K(\xi, \eta)$, $k_1(\xi), \ldots, k_m(\xi)$ are continuous for $\xi, \eta \in [0, 1]$; and $\xi_1, \ldots, \xi_m \in [0, 1]$. Such operators occur in vibration problems with continuously distributed masses as well as point masses [24, p 120].
The proof for non-Hermitian integral operators is not as straightforward as the one for algebraic and Hermitian integral operators because the iterates admit, in general, no representation in terms of eigenfunctions [24, p 116], [22, p 4]. The difficulty is to express what it means for a starting vector to contain a contribution with regard to a certain eigenvalue. However, with the proper identifications the statement of the ‘convergence’ result for integral operators looks like the one for algebraic operators [24, §§4B, C].

In order to transfer the analysis from integral to differential operators, Wielandt assumes that an explicit Green’s function for the differential equation is known [24, §5B]. One of the examples used to illustrate the application of the power method to differential operators is the boundary value problem [24, §1A]

\[ y''(\xi) = \lambda g(\xi)y(\xi), \quad y(0) = y(1) = 0, \]

where the complex-valued function \( g(\xi) \) is continuous on \([0, 1]\). Wielandt rejects what he calls the obvious identification

\[ Ay \equiv \frac{1}{y(\xi)}y''(\xi), \]

because, although \( y(\xi) \) may be twice differentiable, \( Ay \) is in general not [24, p 133].

Instead, he thinks of the differential operator as an integral operator and defines \( A \) by [24, Beispiel (2Bc)]

\[ Ay \equiv z, \]

where \( z \equiv z(\xi) \) is the unique solution to

\[ z''(\xi) = g(\xi)y(\xi), \quad z(0) = z(1) = 0. \]

If \( y \) is an eigenfunction of \( A \) corresponding to eigenvalue \( \kappa \) then \( Ay = \kappa y \). The definition of \( A \) implies that \( \kappa y \) solves

\[ \kappa y''(\xi) = g(\xi)y(\xi). \]

Thus the eigenvalues of \( A \) are the reciprocals of the ones we want, \( \kappa = 1/\lambda \).

One iteration of the power method therefore involves the solution of the boundary value problem

\[ y''_k(\xi) = g(\xi)y_{k-1}(\xi), \quad y_k(0) = y_k(1) = 0. \]

Hence these iterates are mathematically identical\(^3\) to the iterates one gets when applying the power method to the corresponding integral operator. Since the power method tries to converge to an eigenfunction associated with a dominant eigenvalue \( \kappa \) of \( A \), we get an eigenfunction associated with the eigenvalue \( \lambda \) of smallest modulus [24, §1B].

For general boundary value problems we set [24, §§(5Bb)],

\[ Ay \equiv \frac{d^m y}{d\xi^m} + a_m \frac{1}{d\xi^m} \frac{1}{y} + \ldots + a_0(\xi)y, \]

\[ By \equiv b_m \frac{1}{d\xi^m} \frac{1}{y} + \ldots + b_0(\xi)y, \]

---

\(^3\) ‘Mathematically identical’ means that equality holds in exact arithmetic but not necessarily in finite precision arithmetic because the operations required to compute the two sets of iterates are different.
where the coefficients are continuous on \([0,1]\), and the boundary operators \(r_i y\) and \(t_i y\) at the interval endpoints 0 and 1 have constant coefficients, \(1 \leq i \leq m\). To compute an eigenfunction associated with an eigenvalue of smallest modulus of 

\[ Ax = \lambda B x, \quad r_i x = \lambda t_i x, \quad 1 \leq i \leq m, \]

the power method produces iterates from a starting function \(y_0\) via 

\[ A y_k = B y_{k-1}, \quad r_i y_k = t_i y_{k-1}, \quad 1 \leq i \leq m. \]

8. Deflation Et Cetera. Eigenfunctions of smaller eigenvalues can be determined by removing from \(A\) the contribution of the maximal eigenvalue without affecting the other eigenvalues. This is what we now call deflation. The power method can then be applied to the deflated matrix to compute an eigenfunction associated with the next largest eigenvalue. One repeats this process as often as necessary, going through the eigenvalues in order of decreasing magnitude.

It was known at Wielandt’s time how to deflate a matrix whose maximal eigenvalue is simple [8, §4.18], [10], [15, §VI]. Wielandt extends this deflation mechanism and presents a whole class of projectors for deflation of linear operators [24, §3]. They are constructed from the previously computed eigenfunctions and work even when the eigenvalues are defective. A particular deflation operator is the Gram-Schmidt orthogonalisation method [24, §§3Dc, 5C]. In this case the projector is orthogonal, and the deflation process amounts to orthogonalising each iterate against the already computed eigenfunctions. Wielandt’s class of deflation operators is still being used in the solution of non-Hermitian algebraic eigenvalue problems [17, §V.2.1]. As for today’s state of affairs, Saad bemoans that, although deflation procedures are ‘essential in the design of [numerically] robust eigenvalue algorithms’, the literature on deflation procedures for non-Hermitian matrices is scarce [17, p 150].

The report [21] is a less mathematical description of the results from the publication [24] in the context of the aircraft wing design. In addition to computational examples that illustrate the theoretical results in [24], Wielandt comments on the choice of dependent and independent variables, the appropriate size of the discretisation, and the interpretation of the solutions.

In [22] Wielandt simplifies the convergence proof from [24, §4B] for the power method applied to Fredholm integral operators with non-Hermitian kernel.

9. Inverse Iteration. Inverse iteration is mathematically identical to the power method applied to \((A - \bar{\lambda} I)^{-1}\), where \(\bar{\lambda}\) is an approximation to some eigenvalue of \(A\) and \(I\) is the identity.

In particular, if \(A\) is a real or complex square matrix \(A\) and \(\bar{\lambda}\) is given, then inverse iteration generates a sequence of vectors \(x_k\) from a given starting vector \(x_0\) by solving the linear systems 

\[ (A - \bar{\lambda} I)x_k = s_k x_{k-1}, \quad k \geq 1. \]

Here \(s_k\) represents a scalar responsible for normalising \(x_k\), either in some norm or with regard to a particular component. Under the same conditions as for the power method, the sequence of iterates \(x_k\) converges to an eigenvector associated with the eigenvalue \(\lambda\) closest to \(\bar{\lambda}\).

Nowadays inverse iteration is the method of choice when one has at one’s disposal approximations to a specified subset of eigenvalues for which one wants to compute eigenvectors. The development of inverse iteration from Wielandt’s time to the present day (October 1994) is reviewed in [12].
In [23] Wielandt introduces inverse iteration. He refers to inverse iteration as \emph{fractional iteration} (‘gebrochene Iteration’ in German) because the matrix \((A - \lambda I)^{-1}\) is a fractional linear function of \(A\) [23, p 3]. He points out that the use of inverse iteration is beneficial in the stability analysis of vibrating systems that are small perturbations of systems whose behaviour is known [23, §II]. In this case good approximations to the eigenvalues of the perturbed system are available. He briefly discusses the application of inverse iteration to integral and differential operators [23, §§III(b), (c)], but emphasizes its use for the solution of the generalised algebraic eigenvalue problem [23, §V].

10. The Idea. Wielandt gives a simple justification for inverse iteration in case of differential operators [21, §III(c)]. His argument applies to a larger class of linear operators, though. Consider the generalised eigenvalue problem

\[ Ax = \lambda B x, \]

which reverts to the ordinary eigenvalue problem when \(B\) is the identity. Suppose we have an approximation \(\bar{\lambda}\) to \(\lambda\). Choose \(\epsilon\) so that \(\lambda = \bar{\lambda} + \epsilon\). Then we can write the eigenvalue problem as

\[ (A - \bar{\lambda} B) x = \epsilon B x. \]

Now replace \(x\) on the left by \(x_k\) and on the right by \(x_{k-1}\); and write \(s_k\) instead of \(\epsilon\). The resulting iterative method is inverse iteration [23, §§III(a), (c)]

\[ (A - \bar{\lambda} B) x_k = s_k B x_{k-1}, \quad k \geq 1. \]

When \(B\) is the identity, this is just the formulation of inverse iteration from §9.

In the special case when \(B\) is invertible, the generalised eigenvalue problem is mathematically identical to the ordinary eigenvalue problem \(B^{-1} A x = \lambda x\). Wielandt points out that, unlike most other methods, inverse iteration avoids the costly, explicit formation of \(B^{-1}\) [23, §III(a)]. Moreover, it can even be used when \(B\) is not invertible.

The power method for differential operators described in §7 is just a special case of inverse iteration with \(\bar{\lambda} = 0\). With regard to differential operators, therefore, the only difference between the power method and inverse iteration is the shift.

The advantage of inverse iteration over the power method is, loosely speaking, its random access to the eigenvalues compared to the power method’s sequential access. This was also Wielandt’s motivation for introducing inverse iteration [23, §II]: it can compute eigenfunctions for any specified subset of eigenvalues, provided that approximations to these eigenvalues are available. The power method, in contrast, has to walk through the eigenvalues one by one in descending order, and deflate, until it hits the eigenvalues of interest.

The weakness of the power method compared to inverse iteration is caused by the shift. Of course, nothing keeps us from investing a shift in the power method, like so:

\[ y_k \equiv s_k (A - \bar{\lambda} I) y_{k-1}. \]

Besides accelerating the convergence, this isn’t going to make any difference, though. For instance in the Hermitian case, where all eigenvalues lie on the real line,

\[ \lambda_1 \leq \ldots \leq \lambda_n, \]

the eigenvalue of maximal modulus, \(\max\{|\lambda_1 - \bar{\lambda}|, |\lambda_n - \bar{\lambda}|\}\), still comes from one of the two ends of the spectrum of \(A\), but never from an interior eigenvalue [28, p 363], [14, §4-2-3]. The power method uses the shift only in a linear fashion, while inverse iteration uses the shift non-linearly.
Wielandt also recognised correctly that the quality of the shift can be improved and the convergence accelerated if the shift is updated during each iteration [23, §IV]. The most successful incarnation of this strategy is Rayleigh quotient iteration, where the shift $\lambda$ equals the Rayleigh quotient; i.e.

$$\tilde{\lambda} = x_k^T A x_k / x_k^T x_k$$

when $A$ is a real symmetric matrix [14, §4.6]. An extension of Rayleigh quotient iteration to non-Hermitian matrices is possible [27, §9.62].

11. A Final Look Back. Helmut Wielandt’s ‘contributions to the mathematical treatment of complex eigenvalue problems’ [20, 24, 21, 22, 23] have definitely not received the recognition and attention they deserve. Perhaps this is so because only one of the papers ever got published in a journal. But even that one does not seem to have been widely read.

Like Wilkinson, Wielandt possessed a thorough mathematical understanding and an extraordinary intuition regarding computational phenomena; as well as a good measure of common sense in his approach to problem solving. What makes Wielandt’s contributions even more remarkable is the fact that he preceded Wilkinson by twenty years, working at a time when computing machinery was more or less non-existent and numerical analysis was even less fashionable. Because of their wealth of information and numerical examples, we benefit from reading Wielandt’s papers – even now, fifty years after they were written.

Acknowledgement. I thank Hans Schneider for inviting me to write this contribution, and for his generous and lucid advice. Stan Eisenstat, Tim Kelley, Carl Meyer, and especially Fernando Reitich provided many suggestions that greatly improved the quality of the paper.

REFERENCES