An Optimization Framework for Polynomial Zerofinders

Aaron Melman and Bill Gragg

1. INTRODUCTION. Newton’s method is a well-known iterative method for solving the equation \( f(x) = 0 \). It is defined recursively by

\[
x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}
\]

for an appropriate choice of \( x_0 \). There are two standard and equivalent ways to derive it: algebraically (via the Taylor expansion) and geometrically (via the tangent line). We show initially how it can be derived in yet another way, via a constrained optimization problem, by considering the special case of solving \( p(x) = 0 \) for its smallest zero, where \( p(x) \) is a polynomial with all real zeros. We then do the same for several more classical zerofinders and construct some apparently new ones using the same optimization framework. While some of the resulting methods are applicable only to polynomial equations, others, like Newton’s method, can be used for general nonlinear equations.

The usefulness of this approach does not limit itself to the mere construction of zerofinders. As we will see later, it also allows us to carry out a controlled enhancement of those zerofinders.

The paper is organized as follows: in section 2 we introduce our notation and state some preliminaries, in section 3 we derive Newton’s and other methods, and in section 4 we show how these methods can be enhanced.

2. PRELIMINARIES. In what follows we consider a polynomial \( p(x) \) defined by

\[
p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = (x - \xi_1)(x - \xi_2)\ldots(x - \xi_n),
\]

where the zeros \( \xi_j \) are all real and satisfy \( \xi_1 \leq \xi_2 \leq \ldots \leq \xi_n \). If \( p(x) \) is the characteristic polynomial of a matrix \( A \), then its zeros represent the eigenvalues of \( A \). The following can easily be verified:

\[
\sum_{j=1}^{n} (x - \xi_j) = nx + a_{n-1} = \text{tr}(xI - A),
\]

\[
\sum_{j=1}^{n} \frac{1}{x - \xi_j} = \frac{p'(x)}{p(x)} = \text{tr}(xI - A)^{-1},
\]

\[
\sum_{j=1}^{n} \frac{1}{(x - \xi_j)^2} = \left( \frac{p'(x)}{p(x)} \right)^2 - \frac{p''(x)}{p(x)} = \text{tr}(xI - A)^{-2}.
\]

For future ease of writing, we define:

\[
\alpha_x = \sum_{j=1}^{n} (x - \xi_j), \quad \beta_x = \sum_{j=1}^{n} \frac{1}{x - \xi_j}, \quad \gamma_x = \sum_{j=1}^{n} \frac{1}{(x - \xi_j)^2}.
\]
3. ZEROFINDERS. As we mentioned in the introduction, the general optimization framework we are about to discuss not only recovers several classical methods, but also generates new ones. We illustrate the main idea at the hand of classical methods, from Newton’s through Ostrowski’s to Laguerre’s method, before deriving an improved Newton method.

Newton’s method. Consider a polynomial \( p(x) \) of degree \( n \) with all of its zeros \( \xi_j \) real and simple and with \( \xi_1 < \xi_2 < \ldots < \xi_n \). We can compute a lower bound on the smallest zero by solving the following minimization problem for \( \bar{x} \) in \( (-\infty, \xi_1) \):

\[
\min \left\{ \xi_1 : \sum_{j=1}^{n} \frac{1}{\bar{x} - \xi_j} = \beta_{\bar{x}}; \, \xi_j > \bar{x}, \, j = 1, 2, \ldots, n \right\}.
\]

In other words, we are looking for the smallest possible zero of all polynomials that satisfy a given shifted moment of their zeros. The solution of this problem is clearly a lower bound on the smallest zero of \( p(x) \), for its zeros satisfy this constraint and many others as well. The change of variables \( \omega_j = 1/(\bar{x} - \xi_j) \) transforms the problem into

\[
\min \left\{ \omega_1 : \sum_{j=1}^{n} \omega_j = \beta_{\bar{x}}; \, \omega_j \leq 0, \, j = 1, 2, \ldots, n \right\}.
\]

The constraints in this problem represent the portion of a hyperplane in a “quadrant” of \( \mathbb{R}^n \). The optimal solution is readily seen to be given by \( \omega_1 = \beta_{\bar{x}} \) and \( \omega_j = 0 \) (\( j \neq 1 \)) and corresponds to \( \xi_1 = \bar{x} - 1/\beta_{\bar{x}} \). This bound can be iteratively refined, by now replacing \( \bar{x} \) with the newfound bound and repeating the process. From the definition of \( \beta_{\bar{x}} \) it is easy to see that this is nothing other than Newton’s method.

An upper bound on the largest zero results from the associated maximization problem. If the smallest zero is not simple, similar solutions are obtained. It is interesting to note that for Newton’s method the outcome of the optimization problem in this case is the Newton step, but with the ratio \( p(\bar{x})/p'(\bar{x}) \) multiplied by the multiplicity of the zero, precisely what is needed to maintain second-order convergence!

It is unclear where this idea first originated. An optimization idea in the derivation of Laguerre’s method can be discerned in [7], [9], and [13]. It was later used in [5] and [6] to construct both lower and higher order methods, and it appears still later in [15]. We have not seen Newton’s method derived in this manner.

This characterization of Newton’s method for polynomials as the result of an optimization problem suggests that other methods could likewise be characterized by changing or adding appropriate constraints. This is indeed the case, as we see next. But before we do, we consider a slight modification of Newton’s method. If an upper bound on the largest eigenvalue is available, then Newton’s method can be improved, at least for the computation of the smallest zero. Although from the few numerical experiments we have conducted this seems to be only a small improvement, it serves to illustrate the optimization framework. The modification arises from adding an upper bound \( b \) on the largest zero as a constraint for each of the variables (i.e., \( \bar{x} < \xi_j \leq b \)).

This gives for the optimal solution

\[
\xi_1 = \bar{x} - \left( \frac{1}{\bar{x} - b} \right)^{-1}, \quad \xi_j = b \quad (j \neq 1).
\]

To emphasize the improvement over Newton’s method, the corresponding iterative method can be written as

November 2006] OPTIMIZATION FOR POLYNOMIAL ZEROFINDERS 795
\[ x_{k+1} = x_k - \frac{1}{1 - \frac{1}{p'(x_k)}}, \frac{p(x_k)}{p'(x_k)}. \]

**Ostrowski’s method.** Consider the following problem for \( \bar{x} \) in \( (-\infty, \xi) \):

\[ \min \left\{ \zeta_1 : \sum_{j=1}^{n} \frac{1}{(\bar{x} - \xi_j)^2} = \gamma; \; \zeta_j > \bar{x}, \; j = 1, 2, \ldots, n \right\}. \]

The same change of variables as before (namely, \( \omega_j = 1/(\bar{x} - \xi_j) \)) transforms the problem into

\[ \min \left\{ \omega_1 : \sum_{j=1}^{n} \omega_j^2 = \gamma; \; \omega_j \leq 0, \; j = 1, 2, \ldots, n \right\}. \]

This time the constraints represent the portion of a sphere in a “quadrant” of \( \mathbb{R}^n \). As in Newton’s method, the optimal solution is easily obtained and is given by \( \omega_1 = \sqrt{\gamma} \) and \( \omega_j = 0 \) \( (j \neq 1) \), which corresponds to \( \zeta_1 = \bar{x} - 1/\sqrt{\gamma} \). However, here

\[ \frac{1}{\sqrt{\gamma}} = -\frac{1}{\sqrt{1 - \frac{p''(\bar{x})}{p'(\bar{x})}}}, \frac{p(\bar{x})}{p'(\bar{x})}, \]

which shows that we have obtained what is known as Ostrowski’s method (see, for example, [12], where it is called the “square root iteration”).

**Laguerre’s method.** Each different minimization problem like the ones we have been considering leads to a different method. Ideally, one would like to add as many constraints as possible. So far, we have considered only one constraint. The following problem contains two:

\[ \min \left\{ \zeta_1 : \sum_{j=1}^{n} \frac{1}{(\bar{x} - \xi_j)^2} = \beta; \sum_{j=1}^{n} \frac{1}{(\bar{x} - \xi_j)^2} = \gamma; \; \zeta_j > \bar{x}, \; j = 1, 2, \ldots, n \right\}. \]

Once again, it is convenient to introduce the change of variables \( \omega_j = 1/(\bar{x} - \xi_j) \), which transforms the problem into

\[ \min \left\{ \omega_1 : \sum_{j=1}^{n} \omega_j = \beta; \sum_{j=1}^{n} \omega_j^2 = \gamma; \; \omega_j \leq 0, \; j = 1, 2, \ldots, n \right\}. \]

For this problem, the constraints represent the portion of the intersection of a hyperplane and a sphere in a “quadrant” of \( \mathbb{R}^n \).

Under the assumption that \( \omega_j \neq 0 \) for all \( j \), the optimality conditions (see [4, sec. 2.1]) are given by

\[ 1 + \mu + 2v\omega_1 = 0, \quad \mu + 2v\omega_j = 0 \; (j \neq 1), \]

where \( \mu \) and \( v \) are the Lagrange multipliers associated with the equality constraints. The optimality conditions imply that, at the solution, all \( \omega_j \) \( (j = 2, \ldots, n) \) are equal. Setting \( y = \omega_j \) \( (j \neq 1) \) and substituting this into the constraints yields

\[ \omega_1 + (n - 1)y = \beta, \quad \omega_1^2 + (n - 1)y^2 = \gamma. \]
Solving for \( \omega_1 \), and bearing in mind that we are computing a minimum, we obtain

\[
\omega_1 = \frac{\beta_{\bar{x}} - \sqrt{(n - 1)(n\gamma_{\bar{x}} - \beta_{\bar{x}}^2)}}{n}.
\]

The only other solutions satisfying the necessary optimality conditions are solutions with one or more \( \omega_j \) \((j \geq 2)\) equal to zero. However, it can be shown that these would produce a larger value of \( \omega_1 \) and therefore do not represent global minima.

For the original variable \( \xi_1 \) this gives

\[
\xi_1 = \bar{x} + \frac{n}{-\beta_{\bar{x}} + \sqrt{(n - 1)(n\gamma_{\bar{x}} - \beta_{\bar{x}}^2)}},
\]

This solution for \( \xi_1 \) represents one step of Laguerre’s method!

From the minimization problem we can immediately conclude that Laguerre’s method should, at least for the computation of the smallest zero, converge faster than either Newton’s or Ostrowski’s, for the feasible set is now smaller and will therefore yield a higher minimum value. All three methods converge monotonically to the smallest zero \( \xi_1 \) when started from any point to the left of \( \xi_1 \).

**Discrete Laguerre method.** Second derivatives may sometimes be difficult to compute, in which case one might suggest a discrete analog of Laguerre’s method, a method in which both constraints are as in Newton’s method, but each is evaluated at a different point. Assuming that these points are at \( x = x_0 \) and \( x = x_1 \), where \( x_0 < x_1 < \xi_1 \), the corresponding optimization problem is

\[
\min \left\{ \xi_1 : \sum_{j=1}^{n} \frac{1}{x_0 - \xi_j} = \beta_{x_0}; \sum_{j=1}^{n} \frac{1}{x_1 - \xi_j} = \beta_{x_1}; \xi_j > x_1, \ j = 1, 2, \ldots, n \right\}.
\]

The solution was obtained in [15] as

\[
\xi_1 = \frac{x_0 + x_1}{2} + \frac{n - 1}{4} \left( \frac{\Delta \beta}{\Delta x} + S \right) \big( \Delta x \big)^2 - \frac{1}{2} (\beta_{x_0} + \beta_{x_1}) + \sqrt{S \left( 1 - n + S \big( \frac{\Delta x}{2} \big) \right)},
\]

where \( \Delta \beta = \beta_{x_1} - \beta_{x_0}, \Delta x = x_1 - x_0, \) and \( S = \beta_{x_0}^2 - \beta_{x_1}^2 + n \frac{\Delta \beta}{\Delta x} \). The iterates converge monotonically to the smallest zero when started from any point to the left of \( \xi_1 \). This method has very similar properties to Laguerre’s method and was studied in [6] as well as in [2], [3], and [15], where it is called the “quasi-Laguerre method.”

**Improved Newton method.** We consider an improved Newton method by looking at the following minimization problem, whose solutions are finite and bounded away from \( \bar{x} \):

\[
\min \left\{ \xi_1 : \sum_{j=1}^{n} (\bar{x} - \xi_j) = \alpha_{\bar{x}}; \sum_{j=1}^{n} \frac{1}{\bar{x} - \xi_j} = \beta_{\bar{x}}; \xi_j > \bar{x}, \ j = 1, 2, \ldots, n \right\}.
\]

If we denote the Lagrange multipliers corresponding to the first and second constraints \( \mu \) and \( \nu \), respectively, the first-order necessary optimality conditions yield

\[
1 - \mu + \nu/(\bar{x} - \xi_1)^2 = 0 \quad \text{and} \quad -\mu + \nu/(\bar{x} - \xi_1)^2 = 0 \quad (j \neq 1).
\]

Because of the posi-
tivity constraint this means that, at the optimal solution, the \( n - 1 \) quantities \( \bar{x} - \zeta_j \) \((j = 2, \ldots, n)\) are all equal. Denoting their common value by \( y \), we have

\[
\begin{cases}
(\bar{x} - \zeta_1) + (n - 1)y = \alpha_\bar{x}, \\
\frac{1}{\bar{x} - \zeta_1} + \frac{n - 1}{y} = \beta_\bar{x}.
\end{cases}
\] (3)

From the second equation in (3) we retrieve

\[ y = \frac{(n - 1)(\bar{x} - \zeta_1)}{\beta_\bar{x}(\bar{x} - \zeta_1) - 1} \]

(note that \( \bar{x} - \zeta_1 \neq 1/\beta_\bar{x} \)). Substituting this value into the first equation yields the following quadratic equation in \( \bar{x} - \zeta_1 \):

\[ \beta_\bar{x}(\bar{x} - \zeta_1)^2 - (\alpha \beta + 1 - (n - 1)^2)(\bar{x} - \zeta_1) + \alpha_\bar{x} = 0. \] (4)

Since we are minimizing \( \zeta_1 \), we are interested in the largest value for \( \bar{x} - \zeta_1 \), so

\[ \bar{x} - \zeta_1 = \frac{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2) - \sqrt{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2)^2 - 4\alpha_\bar{x} \beta_\bar{x}}}{2\beta_\bar{x}}. \] (5)

Before examining these zeros, we note that \( n/\beta_\bar{x} \leq \alpha_\bar{x}/n \) (i.e., \( \alpha_\bar{x} \beta_\bar{x} \geq n^2 \)) because of the harmonic-arithmetic means inequality. After some algebraic manipulation, we can write the discriminant as

\[ (\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2)^2 - 4\alpha_\bar{x} \beta_\bar{x} = (\alpha_\bar{x} \beta_\bar{x} - n^2)(\alpha_\bar{x} \beta_\bar{x} - (n - 2)^2). \]

Both factors on the right-hand side of the last equation are positive, implying that the discriminant is also positive. (The other zero represents the solution to the maximum problem.) No other solutions satisfy the necessary optimality conditions.

We want to examine \( \bar{x} - \zeta_1 \) more closely. Multiplying by the conjugate, we have

\[
\frac{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2) - \sqrt{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2)^2 - 4\alpha_\bar{x} \beta_\bar{x}}}{2\beta_\bar{x}} = \frac{2\alpha_\bar{x} \beta_\bar{x}}{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2) + \sqrt{(\alpha_\bar{x} \beta_\bar{x} + 1 - (n - 1)^2)^2 - 4\alpha_\bar{x} \beta_\bar{x}}} \cdot \frac{1}{\beta_\bar{x}}
\]

\[ = \frac{2\alpha_\bar{x} \beta_\bar{x}}{(\alpha_\bar{x} \beta_\bar{x} - n(n - 2)) + \sqrt{(\alpha_\bar{x} \beta_\bar{x} - n^2)(\alpha_\bar{x} \beta_\bar{x} - (n - 2)^2)}} \cdot \frac{1}{\beta_\bar{x}}. \]

We can therefore write

\[ \zeta_1 = \bar{x} - \frac{2\alpha_\bar{x} \beta_\bar{x}}{(\alpha_\bar{x} \beta_\bar{x} - n(n - 2)) + \sqrt{(\alpha_\bar{x} \beta_\bar{x} - n^2)(\alpha_\bar{x} \beta_\bar{x} - (n - 2)^2)}} \cdot \frac{1}{\beta_\bar{x}}. \] (6)

Setting \( s = \alpha_\bar{x} \beta_\bar{x} \) and defining

\[ \phi(s) = \frac{2s}{(s - n(n - 2)) + \sqrt{(s - n^2)(s - (n - 2)^2)}}, \]

we obtain \( \zeta_1 = \bar{x} - \phi(s)/\beta_\bar{x} \), with \( s \geq n^2 \). The function \( \phi(s) \) exhibits the following properties in its domain:
\[ \phi'(s) < 0, \quad \phi(n^2) = n, \quad \lim_{s \to +\infty} \phi(s) = 1. \]

Since \(-1/\beta_i\) corresponds to the Newton step, this means that this method is never worse than Newton’s method and potentially much faster. In fact, in the best case, it yields a step up to \(n\) times the length of the Newton step! As the iterates approach the smallest zero, the linear constraint contributes less and less, whereas the second constraint becomes more and more relevant. Asymptotically, the method becomes Newton’s method. Its main contribution, therefore, lies in obtaining good initial approximations when the iterates are still far from the zero. Of course, adding the linear constraint would similarly improve the other methods we have mentioned.

Although for all the aforementioned methods we have considered the very specific case of the smallest zero, they can, with varying degrees of success, be applied to the inner zeros as well, or even to functions that are not necessarily polynomials. However, while the methods derived in this way all converge monotonically to the smallest zero, their properties for the computation of the other zeros vary extensively from method to method and would require a separate study to explain.

**Numerical comparison.** For purposes of illustration, we ran a few numerical experiments in which we computed the smallest eigenvalue of a symmetric positive semidefinite Toeplitz matrix, a problem sometimes encountered in signal processing. (A Toeplitz matrix is a matrix whose entries are constant along the diagonals.) We compared Newton’s method, its improved version, and the discrete Laguerre method (with the first iterate generated by the improved Newton method), applied to the characteristic polynomial of the matrix. As it is not really relevant here, we omit the details of how all necessary quantities were computed, referring instead to [10].

We have averaged our results over five hundred random matrices of three types:

1. **CVL matrices.** These are matrices defined in [1] (whence their name) as
   \[ T = \mu \sum_{k=1}^{n} \xi_k T_{2\pi \theta_k}, \]
   where \(n\) is the dimension of \(T\) and where \(\theta_k\) and \(\xi_k\) are uniformly distributed random numbers in \((0, 1)\). The parameter \(\mu\) is chosen so that \(T_{kk} = 1\) for \(k = 1, \ldots, n\), and \((T_\theta)_{ij} = \cos(\theta(i - j))\). These matrices are positive semidefinite and were also used in [10]. Even though they could theoretically be singular, none of our matrices were.

2. **KMS matrices.** These are the Kac-Murdock-Szegö matrices (see [8]) defined by
   \[ T_{ij} = \nu^{\left| i-j \right|} \quad (0 < \nu < 1; i, j = 1, \ldots, n), \]
   where \(n\) is the dimension of the matrix. These matrices are positive definite.

3. **DDM matrices.** We define a diagonally dominant symmetric Toeplitz matrix to be a symmetric Toeplitz matrix whose first row is given by a vector with components that are uniformly distributed on \((0, 1)\) and whose \((1, 1)\)-entry is \(2n\), where \(n\) is the dimension of the matrix.

The three graphs in Figure 1 represent, from small to large, the eigenvalues of (from left to right) a typical CVL, KMS, and DDM matrix of dimension \(n = 100\). Note that the largest eigenvalue of the DDM matrix is around 250.

Table 1 contains the average number of iterations necessary to compute the smallest eigenvalue to a relative accuracy of \(10^{-8}\) using the same error estimate as in [10]. As the reader can see, the discrete Laguerre method easily outperforms the other two,
Figure 1. Typical eigenvalue distribution for CVL, KMS, and DDM matrices.

whereas the improved Newton method sometimes improves dramatically over Newton’s method. Comparing this with the results in [10] and [11], we arrive at a clear recommendation to abandon the use of Newton’s method for these kinds of problems in favor of the discrete Laguerre method.

Table 1. Comparison of the algorithms for \(n = 100\).

<table>
<thead>
<tr>
<th></th>
<th>CVL</th>
<th>KMS</th>
<th>DDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newton</td>
<td>10.46</td>
<td>125.39</td>
<td>231.13</td>
</tr>
<tr>
<td>Improved Newton</td>
<td>9.98</td>
<td>97.61</td>
<td>30.40</td>
</tr>
<tr>
<td>Discrete Laguerre</td>
<td>7.20</td>
<td>25.19</td>
<td>14.91</td>
</tr>
</tbody>
</table>

In the following section we prove a series of theorems on how far an increased step, taken with any of the aforementioned methods, can overshoot the smallest zero when the initial point lies to the left of that zero.

4. OVERSHOOTING THEOREMS. A slightly different optimization problem allows us to prove the following theorem, which provides a bound on how far a magnified step for the methods in section 3 can overshoot the smallest zero of a polynomial. For Newton’s method there already exists a theorem on double steps (see [14, Theorem 5.5.9]). However, we are unaware of the existence of any such results for the other methods.

**Theorem 1 (Newton’s Method).** If an \(n\)th degree polynomial \(p(x)\) with only real zeros has \(s\) distinct zeros \(\xi_j\), where \(\xi_1 < \xi_2 < \xi_3 < \cdots < \xi_s\), and if \(\xi_j\) has multiplicity \(q_j\), then the following holds whenever \(\tilde{x} < \xi_1\) and \(0 < \theta \leq 1/2\):

\[
\tilde{x} - \left( (1 - \theta)q_1 + \theta q_2 + 2\sqrt{q_1q_2\theta(1 - \theta)} \right) \frac{1}{\beta_{\tilde{x}}} \leq (1 - \theta)\xi_1 + \theta \xi_2.
\]

**Proof.** Fixing \(\theta\) in \((0, 1/2]\), we consider the following optimization problem:

\[
\min \left\{ (1 - \theta)\xi_1 + \theta \xi_2 : \sum_{j=1}^{n} \frac{1}{\tilde{x} - \xi_j} = \beta_{\tilde{x}}; \, \xi_j > \tilde{x}, \, j = 1, 2, \ldots, n \right\}.
\]

We note that the constraint can be written as

\[
\frac{q_1}{\tilde{x} - \xi_1} + \frac{q_2}{\tilde{x} - \xi_2} + \sum_{j=3}^{s} \frac{q_j}{\tilde{x} - \xi_j} = \beta_{\tilde{x}}.
\]

(7)
It should be mentioned that this optimization problem finds a minimum for the case where any two of the variables appear with multiplicities $q_1$ and $q_2$. The optimality conditions are

$$\begin{align*}
(1 - \theta) + \frac{\mu q_1}{(x - \xi_1)^2} &= 0, \\
\theta + \frac{\mu q_2}{(x - \xi_2)^2} &= 0, \\
\frac{\mu q_j}{(x - \xi_j)^2} &= 0 \quad (j \neq 1, 2).
\end{align*}$$

From these conditions we infer that

$$\tilde{x} - \xi_2 = \left(\frac{q_2(1 - \theta)}{q_1 \theta}\right)^{1/2} (\tilde{x} - \xi_1), \quad \xi_j = +\infty \quad (j \neq 1, 2).$$

Substituting this into (7) gives

$$\begin{align*}
\tilde{x} - \xi_1 &= \left(q_1 + \left(\frac{q_1 q_2 \theta}{1 - \theta}\right)^{1/2}\right) \frac{1}{\beta_\tilde{x}}, \\
\tilde{x} - \xi_2 &= \left(q_2 + \left(\frac{q_1 q_2 (1 - \theta)}{\theta}\right)^{1/2}\right) \frac{1}{\beta_\tilde{x}},
\end{align*}$$

which must yield the minimum value, since solutions are bounded away from $\tilde{x}$ and no other solutions satisfy the necessary optimality conditions. A straightforward computation then gives for the minimum value:

$$(1 - \theta)\xi_1 + \theta \xi_2 = \tilde{x} - \left((1 - \theta)q_1 + \theta q_2 + 2\sqrt{q_1 q_2 (1 - \theta)}\right) \frac{1}{\beta_\tilde{x}},$$

or

$$(1 - \theta)\xi_1 + \theta \xi_2 = \tilde{x} - \left(\sqrt{(1 - \theta)q_1} + \sqrt{\theta q_2}\right) \frac{1}{\beta_\tilde{x}}.$$ 

We now note that for any two pairs $(a, b)$ and $(c, d)$ with $0 \leq a \leq b$ and $0 \leq c \leq d$ we have $ad + bc \leq ac + bd$. To see this, it is sufficient to consider the inequality $(b - a)(d - c) \geq 0$. We have therefore found a lower bound on the quantity

$$\xi_1 \max\{\theta, 1 - \theta\} + \xi_2 \min\{\theta, 1 - \theta\},$$

because $\xi_1$ and $\xi_2$ satisfy more constraints than just the one considered here. Since $0 < \theta \leq 1/2$, this completes the proof. \[\Box\]

In the case of simple zeros ($q_1 = q_2 = 1$) Theorem 1 asserts that at the minimum

$$(1 - \theta)\xi_1 + \theta \xi_2 = \tilde{x} - \left(1 + 2\sqrt{\theta(1 - \theta)}\right) \frac{1}{\beta_\tilde{x}},$$

When $\theta = 1/2$, this means that a double Newton step, when starting from a point to the left of the smallest zero, never overshoots the arithmetic mean of the two smallest zeros. When these smallest zeros are not simple, larger Newton steps can be taken with the same nonovershooting property. Of course, other objective functions could be used for the minimization, such as the harmonic mean of $\xi_1$ and $\xi_2$, etc. In addition, similar results can be obtained for the average of more than two variables.

An analogous result exists for Ostrowski’s method. Since the proof is very similar to the one for Newton’s method, we omit it and simply state the result:

November 2006 | OPTIMIZATION FOR POLYNOMIAL ZEROFINDERS 801
Theorem 2 (Ostrowski’s Method). If an nth degree polynomial \( p(x) \) with only real zeros has \( s \) distinct zeros \( \xi_j \), where \( \xi_1 < \xi_2 < \xi_3 < \cdots < \xi_n \), and \( \xi_j \) has multiplicity \( q_j \), then the following holds whenever \( \bar{x} < \xi_1 \) and \( 0 < \theta \leq 1/2 \):

\[
\bar{x} + \left( (1 - \theta)^2 q_1 + (\theta(1 - \theta)^2 q_1 \sqrt{q_2})^{2/3} + \sqrt{\theta^2 q_2 + ((1 - \theta) \theta^2 q_2 \sqrt{q_1})^{2/3}} \right) \frac{1}{\sqrt[3]{q_1}} 
\leq (1 - \theta) \xi_1 + \theta \xi_2.
\]

In the case of simple zeros \( (q_1 = q_2 = 1) \) Theorem 2 tells us that

\[
(1 - \theta) \xi_1 + \theta \xi_2 = \bar{x} + \left( \sqrt{\theta^2 + (\theta^2(1 - \theta))^2} + \sqrt{(1 - \theta)^2 + (\theta(1 - \theta)^2)^{2/3}} \right) \frac{1}{\sqrt{q_1}}.
\]

When \( \theta = 1/2 \), this means that taking an Ostrowski step multiplied by \( \sqrt{2} \), when starting from the left of the smallest zero, never overshoots the arithmetic mean of the two smallest zeros.

Although the same kind of theorems can be proved for the new Newton-like method and for the Laguerre and discrete Laguerre methods, the general case proves to be quite complicated. For illustrative purposes, we consider the special case for \( \theta = 1/2 \) and for simple zeros. We then have the following theorem:

Theorem 3 (Improved Newton Method). If an nth degree polynomial \( p(x) \) has all simple and real zeros \( \xi_j \) with \( \xi_1 < \xi_2 < \cdots < \xi_n \), then the following holds whenever \( \bar{x} < \xi_1 \):

\[
\bar{x} - \frac{2\alpha_i \beta_{\bar{x}}}{(\alpha_i \beta_{\bar{x}} - n(n - 4)) + \sqrt{(\alpha_i \beta_{\bar{x}} - n^2)(\alpha_i \beta_{\bar{x}} - (n - 4)^2)}} \cdot \frac{2}{\beta_{\bar{x}}} \leq \frac{1}{2}(\xi_1 + \xi_2).
\]

Proof. We consider the following optimization problem:

\[
\min \left\{ \zeta_1 + \zeta_2 : \sum_{j=1}^{n} (\bar{x} - \zeta_j) = \alpha_{\bar{x}}; \sum_{j=1}^{n} \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \zeta_j > 0, \ j = 1, 2, \ldots, n \right\}.
\]

The solutions are finite and bounded away from \( \bar{x} \). The optimality conditions are

\[
1 - \mu + \frac{\nu}{(\bar{x} - \zeta_j)^2} = 0, \quad 1 - \mu + \frac{\nu}{(\bar{x} - \zeta_j)^2} = 0,
\]

\[
- \mu + \frac{\nu}{(\bar{x} - \zeta_j)^2} = 0 \quad (j \neq 1, 2).
\]

From these conditions we conclude that \( \zeta_1 = \zeta_2 \) and all \( \zeta_j \) are identical when \( j \neq 1, 2 \). Substituting this into the constraints and setting \( y = \bar{x} - \zeta_j \) \((j \neq 1, 2) \) yields

\[
2(\bar{x} - \zeta_1) + (n - 2)y = \alpha_{\bar{x}}, \quad \frac{2}{\bar{x} - \zeta_1} + \frac{n - 2}{y} = \beta_{\bar{x}}.
\]

These are exactly the same equations as in (3), where \( \alpha_{\bar{x}} \) is replaced with \( \alpha_{\bar{x}}/2, \beta_{\bar{x}} \) with \( \beta_{\bar{x}}/2 \), and \( n \) with \( n/2 \). This implies that

\[
\zeta_1 = \bar{x} - \frac{2\alpha_i \beta_{\bar{x}}}{(\alpha_i \beta_{\bar{x}} - n(n - 4)) + \sqrt{(\alpha_i \beta_{\bar{x}} - n^2)(\alpha_i \beta_{\bar{x}} - (n - 4)^2)}} \cdot \frac{2}{\beta_{\bar{x}}}.
\]
No other solutions satisfy the necessary conditions, and we have therefore found the solution that yields the minimum value. Since \( \xi_1 = \xi_2 \), we obtain for the minimum of \( (\xi_1 + \xi_2)/2 \) the same value as for \( \xi_1 \).

Analogous results exist for the Laguerre and the discrete Laguerre method. We once again merely state these results as their proofs are very similar to the one for the improved Newton method.

**Theorem 4 (Laguerre and Discrete Laguerre Method).** If an \( n \)th degree polynomial \( p(x) \) has all simple and real zeros \( \xi_j \), where \( \xi_1 < \xi_2 < \cdots < \xi_n \), then the following holds whenever \( \tilde{x} < \tilde{y} < \xi_1 \):

\[
\bar{x} + \frac{\sqrt{2n}}{-\sqrt{2}\beta_\bar{x} + \sqrt{(n-2)(n\gamma_\bar{x} - \beta_\bar{x}^2)}} \leq \frac{1}{2}(\xi_1 + \xi_2),
\]

\[
\frac{\tilde{x} + \tilde{y}}{2} + \frac{n - \frac{1}{4} \left( \frac{\Delta \beta}{\Delta x} + \frac{\xi}{\Delta x} \right) (\Delta x)^2}{-\frac{1}{4}(\beta_\tilde{x} + \beta_\tilde{y}) + \frac{1}{4} \sqrt{3} S (2 - n + S \frac{(\Delta x)^2}{2})} \leq \frac{1}{2}(\xi_1 + \xi_2),
\]

where \( \Delta \beta = \beta_\tilde{y} - \beta_\tilde{x} \), \( \Delta x = \tilde{y} - \tilde{x} \), and \( S = \beta_\tilde{x} \beta_\tilde{y} + n \frac{\Delta \beta}{\Delta x} \).

**Remark.** The results we obtained in these overshooting theorems are the same as would be obtained for regular steps of these respective methods if the first zero were a double zero, as can easily be seen from the optimality conditions.

**Numerical comparison.** To illustrate the effect of taking larger steps, we have run some numerical experiments with the discrete Laguerre method. We used the same type of matrices as before and averaged over five hundred experiments. The results are reported in Tables 2 and 3: they contain the average number of iterations to compute the smallest eigenvalue with a relative accuracy of \( 10^{-6} \) for matrices of dimensions one hundred and five hundred, respectively. As one can see, taking enhanced steps is most useful when the number of iterations is large. When only a small number of iterations are needed for the computation, it may even increase the amount of work because of backtracking after overshooting, which wastes an iterate. Although we are not concerned with it here, it may be possible to avoid this waste by adopting a more sophisticated backtracking procedure.

**Table 2.** Comparison of the regular and enhanced discrete Laguerre methods for \( n = 100 \).

<table>
<thead>
<tr>
<th></th>
<th>CVL</th>
<th>KMS</th>
<th>DDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete Laguerre</td>
<td>7.20</td>
<td>25.19</td>
<td>14.91</td>
</tr>
<tr>
<td>Enhanced Discrete Laguerre</td>
<td>8.41</td>
<td>20.11</td>
<td>13.26</td>
</tr>
</tbody>
</table>

**Table 3.** Comparison of the regular and enhanced discrete Laguerre methods for \( n = 500 \).

<table>
<thead>
<tr>
<th></th>
<th>CVL</th>
<th>KMS</th>
<th>DDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete Laguerre</td>
<td>8.77</td>
<td>60.24</td>
<td>23.21</td>
</tr>
<tr>
<td>Enhanced Discrete Laguerre</td>
<td>9.89</td>
<td>44.46</td>
<td>18.96</td>
</tr>
</tbody>
</table>
5. CONCLUSION. In this paper we have derived polynomial zerofinders by computing a lower bound on the smallest zero of a polynomial with only real zeros, subject to one or more identities that the zeros must satisfy. We have shown that the different bounds correspond to different iterative methods for computing zeros. Some of these are well known, such as Newton’s and Laguerre’s methods, but others are less known, or new. Although we derived all of these methods to compute the smallest zero, they are also applicable to the computation of the other zeros, albeit with different properties. Our framework also allowed us to establish results about the effect of an extended iteration step for the aforementioned methods.

REFERENCES

6. L. V. Foster, Generalizations of Laguerre’s method: Lower order methods, unpublished manuscript.

AARON MELMAN received his M.S. from the Technion–Israel Institute of Technology in 1986 and Ph.D. from Caltech in 1992. He taught at Ben-Gurion University in Beer Sheva, Israel, and is currently teaching at Santa Clara University in California. His research interests are in numerical analysis and linear algebra with, lately, a particular fondness for Toeplitz matrices.

Department of Applied Mathematics, School of Engineering, Santa Clara University, Santa Clara, CA 95053 amelman@scu.edu

BILL GRAGG (M.S. Stanford ’59, Ph.D. UCLA ’64) studied numerical analysis with George Forsythe, Peter Henrici, and Alston Householder. He created the alternating midpoint-trapezoidal rule for solving ordinary initial value problems numerically. He studied the Padé table and its relation to numerical algorithms. This extended naturally to Laurent-Padé tables and is thus related to Toeplitz matrices, polynomials orthogonal on the unit circle, and eigenproblems for unitary Hessenberg matrices. He has contributed to the development of fast algorithms for related highly structured problems.

Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93940 gragg@nps.edu

804 © THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 113