

## GEOMETRIC ASPECTS OF PELLET'S AND RELATED THEOREMS

A. MELMAN

**ABSTRACT.** Pellet's theorem determines when the zeros of a polynomial can be separated into two regions, according to their moduli. We refine one of those regions and replace it with the closed interior of a lemniscate that provides more precise information on the location of the zeros. Moreover, Pellet's theorem is considered the generalization of a zero inclusion region due to Cauchy. Using linear algebra tools, we derive a different generalization that leads to a sequence of smaller inclusion regions, which are also the closed interiors of lemniscates.

**1. Introduction.** Pellet's classical theorem ([15], [10, Theorem (28,1)]) derives a criterion for the separation of the zeros of a general polynomial with complex coefficients into two regions of the complex plane: a disk and the complement of a larger disk, both centered at the origin. Our first result is to replace the latter set by the interior of a lemniscate, which provides more precise information on the location of the zeros. It adds a geometric component to a theorem that is formulated in terms of simple bounds on the moduli of the zeros that obscure the details of the zero distribution. Although Pellet's theorem is often viewed as the generalization of an inclusion region by Cauchy ([3], [10, Theorem (27,1)]), we show that a different generalization can be obtained, leading to smaller regions consisting, once again, of the closed interiors of lemniscates.

Both Pellet's theorem and Cauchy's result can be proven with Rouché's theorem. However, to refine and extend these theorems, we used the Gershgorin set to estimate the eigenvalues of a polynomial's companion matrix. This set is a union of disks in the complex plane, centered at the diagonal elements of the matrix. To be able to extract useful results from this often crude method, we consider a similarity

---

2010 AMS *Mathematics subject classification.* Primary 12D10, 15A18, 30C15.

*Keywords and phrases.* Pellet, Cauchy, Gershgorin, zero, root, polynomial, lemniscate.

Received by the editors on June 1, 2012, and in revised form on March 19, 2013.

DOI:10.1216/RMJ-2015-45-2-603

Copyright ©2015 Rocky Mountain Mathematics Consortium

transformation of an appropriate polynomial of the companion matrix, rather than the companion matrix itself. The main advantage of the Gershgorin set, as we apply it here, is to point out results that might otherwise not be apparent, even if their subsequent proof by Rouché's theorem is relatively straightforward.

Although Pellet's theorem was recently generalized to matrix polynomials in [2, 12], our results here cannot easily be similarly extended because of their heavy dependence on the scalar nature of the coefficients. We will discuss other improvements after the derivation of our results.

A good introduction to Gershgorin disks and other eigenvalue inclusion regions can be found in [5, Chapter 6]. For a more in-depth study of the subject, including its interesting history, we refer to [19] and the many references therein. For results concerning polynomial zeros we refer to the encyclopedic works [10, 16].

In Section 2, we state the aforementioned theorems, together with definitions and lemmas that are needed in Section 3, where we derive and illustrate our main results.

**2. Preliminaries.** We start by stating Pellet's theorem.

**Theorem 2.1.** ([15], [10, Theorem (28,1), page 128]). *Given the polynomial  $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$  with complex coefficients,  $a_0a_k \neq 0$ , and  $n \geq 3$ . Let  $1 \leq k \leq n-1$ , and let the polynomial*

$$f_k(x) = x^n + |a_{n-1}|x^{n-1} + \cdots + |a_{k+1}|x^{k+1} - |a_k|x^k \\ + |a_{k-1}|x^{k-1} + \cdots + |a_0|$$

*have two distinct positive roots  $r$  and  $R$ ,  $r < R$ . Then  $p$  has exactly  $k$  zeros in or on the circle  $|z| = r$  and no zeros in the annular ring  $r < |z| < R$ .*

We will not dwell on the numerical solution of  $f_k(x) = 0$ . A systematic method to do so can be found in [13], while a heuristic method was developed in [18]. An implementation using the Newton polygon can be found in [1, 2] for scalar and matrix polynomials, respectively.

Pellet's theorem is considered the generalization of the following result by Cauchy:

**Theorem 2.2.** ([3], [10, Theorem (27,1), page 122]). *All the zeros of the polynomial  $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$  with complex coefficients lie in the circle  $|z| = r$ , where  $r$  is the positive root of the equation*

$$x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0| = 0.$$

Both these theorems are a direct consequence of Rouché's theorem ([17], [6, Theorem 1.6, page 181]). However, they only provide information on the moduli of the zeros. To introduce more interesting geometry into these results, leading to better information on the distribution of the zeros in the complex plane, we will instead use linear algebra tools, namely, Gershgorin's theorem and the polynomial's companion matrix. We state Gershgorin's theorem next.

**Theorem 2.3.** (Gershgorin, [4], [5, Theorem 6.1.1, page 344]). *All the eigenvalues of the  $n \times n$  complex matrix  $A$  with elements  $a_{ij}$  are located in the union of  $n$  disks*

$$\Gamma(A) = \bigcup_{i=1}^n \{z \in \mathcal{C} : |z - a_{ii}| \leq R'_i(A)\}, \quad \text{with } R'_i(A) = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|.$$

*Moreover, if  $\ell$  disks form a connected region that is disjoint from the remaining  $n - \ell$  disks, then this region contains exactly  $\ell$  eigenvalues.*

$R'_i(A)$  is called the  $i$ th deleted row sum of  $A$ . The spectrum of  $A$  and  $A^T$  is the same, so that the Gershgorin set also has a column version, obtained by applying the theorem to  $A^T$ , where the deleted column sums replace the deleted row sums. In addition, any similarity transformation of  $A$ , namely,  $S^{-1}AS$  for a nonsingular matrix  $S$ , has the same eigenvalues as  $A$ , but may have a smaller Gershgorin set. Frequently,  $S$  is chosen to be a diagonal matrix.

Eigenvalue inclusion sets can be used to estimate zeros of a polynomial by applying them to the polynomial's companion matrix, whose eigenvalues are the zeros of the polynomial. A common choice for a

companion matrix of the monic polynomial  $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$  is given by (see, e.g., [5, page 146]):

$$(2.1) \quad C(p) = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix},$$

where blank entries represent zeros, a convention we will follow throughout. In what follows, we set  $a_n = 1$ , where  $a_n$  is the leading coefficient of the aforementioned polynomial  $p$  and denote an open disk centered at  $a$  with radius  $\rho$  by  $O(a; \rho)$ . The closure and the complement of a set  $\Delta$  will be denoted by  $\bar{\Delta}$  and  $\Delta^c$ , respectively. We also define the following.

**Definition 2.4.** The associated polynomials  $\{p_k\}_{k=1}^{n-1}$  of the polynomial

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$$

are defined by  $p_1(z) = z$  and the recursion

$$p_{k+1}(z) = z(p_k(z) + a_{n-k}).$$

Therefore,  $p_k$ ,  $1 \leq k \leq n-1$ , is given by

$$p_k(z) = z^k + a_{n-1}z^{k-1} + \cdots + a_{n-k+1}z.$$

**Definition 2.5.** The polynomial  $P_k$  is obtained from  $p_k$  by replacing its coefficients with their moduli, i.e.,

$$P_k(z) = z^k + |a_{n-1}|z^{k-1} + \cdots + |a_{n-k+1}|z.$$

**Definition 2.6.** The complex  $n \times n$  matrix  $M_k(p)$ ,  $1 \leq k \leq n-1$ , is defined as

$$M_k(p) = p_k(C(p)),$$

where  $C(p)$  is the companion matrix of  $p$ .

The eigenvalues of  $M_k(p)$  are  $\{p_k(z_i)\}_{i=1}^n$ , where  $\{z_i\}_{i=1}^n$  are the eigenvalues of  $C(p)$ , which are also the zeros of  $p$ . Its structure is derived in the following lemma.

**Lemma 2.7.** *The matrix  $M_k(p)$ ,  $1 \leq k \leq n-1$ , is given by*

$$M_k(p) = \begin{pmatrix} 0 & & & & -a_0 & & & & \\ a_{n-k+1} & \ddots & & & -a_1 & \ddots & & & \\ a_{n-k+2} & \ddots & \ddots & & \vdots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & 0 & -a_{n-k-1} & \ddots & \ddots & -a_0 \\ a_{n-1} & \ddots & \ddots & \ddots & a_{n-k+1} & -a_{n-k} & \ddots & \ddots & -a_1 & -a_0 \\ 1 & \ddots & \ddots & \ddots & a_{n-k+2} & \ddots & \ddots & \ddots & \vdots & -a_1 \\ & \ddots & \ddots & \ddots & \vdots & & \ddots & -a_{n-k-1} & \vdots & \\ & & \ddots & \ddots & a_{n-1} & & & -a_{n-k} & -a_{n-k-1} & \\ & & & 1 & & & & & -a_{n-k} & \end{pmatrix},$$

where  $k$  diagonal elements are equal to  $a_{n-k}$ , while the remaining ones are zero.

*Proof.* The proof is by induction. Since  $M_1(p) = C(p)$ , the lemma is obviously true for  $k = 1$ . Now assume that is true for  $M_j(p)$ ,  $1 \leq j \leq n-2$ . A straightforward calculation then shows that  $C(p)M_j(p)$  is given by

$$\begin{pmatrix} 0 & & & & -a_0 & & & & a_{n-j}a_0 \\ 0 & \ddots & & & -a_1 & \ddots & & & a_{n-j}a_1 \\ a_{n-j+1} & \ddots & \ddots & & -a_2 & \ddots & \ddots & & a_{n-j}a_2 \\ a_{n-j+2} & \ddots & \ddots & \ddots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \ddots & -a_0 & a_{n-j}a_{n-j-3} \\ a_{n-1} & \ddots & \ddots & \ddots & 0 & -a_{n-j-2} & \ddots & \ddots & -a_1 & a_{n-j}a_{n-j-2} - a_0 \\ 1 & \ddots & \ddots & \ddots & 0 & -a_{n-j-1} & \ddots & \ddots & -a_2 & a_{n-j}a_{n-j-1} - a_1 \\ & \ddots & \ddots & \ddots & a_{n-j+1} & -a_{n-j} & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & \ddots & a_{n-j+2} & \ddots & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \vdots & & \ddots & -a_{n-j-2} & a_{n-j}a_{n-3} - a_{n-j-3} \\ & & & & a_{n-1} & & & -a_{n-j-1} & a_{n-j}a_{n-2} - a_{n-j-2} \\ & & & & 1 & & & -a_{n-j} & a_{n-j}a_{n-1} - a_{n-j-1} \end{pmatrix},$$

from which one easily deduces that  $M_{j+1}(p) = C(p)M_j(p) + a_{n-j}C(p)$  is of the same form as  $M_j(p)$ .  $\square$

The form of  $M_k(p)$  makes it convenient to apply the column version of Gershgorin's theorem since the deleted column sums are the same for identical diagonal elements. To add flexibility to the Gershgorin set, we will use a diagonal similarity transformation. The next lemma shows its effect on  $M_k(p)$ .

**Lemma 2.8.** *Let  $D_x$  be a diagonal matrix with diagonal  $(x^n, x^{n-1}, \dots, x)$  for  $x > 0$ . Then the matrix  $D_x^{-1}M_k(p)D_x$ ,  $1 \leq k \leq n-1$ , is given by*

$$\begin{pmatrix} 0 & & & & -a_0/x^{n-k} & & & & \\ a_{n-k+1}x & \ddots & & & -a_1/x^{n-k-1} & \ddots & & & \\ a_{n-k+2}x^2 & \ddots & \ddots & & \vdots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & & 0 & -a_{n-k-1}/x & \ddots & \ddots & -a_0/x^{n-k} \\ a_{n-1}x^{k-1} & \ddots & \ddots & a_{n-k+1}x & -a_{n-k} & \ddots & \ddots & -a_1/x^{n-k-1} & -a_0/x^{n-k} \\ x^k & \ddots & \ddots & a_{n-k+2}x^2 & & \ddots & \ddots & \vdots & -a_1/x^{n-k-1} \\ & \ddots & \ddots & \vdots & & & \ddots & -a_{n-k-1}/x & \vdots \\ & & \ddots & a_{n-1}x^{k-1} & & & -a_{n-k} & -a_{n-k-1}/x & \\ & & & x^k & & & & -a_{n-k} & \end{pmatrix}.$$

*Proof.* For any diagonal matrix  $D$  with diagonal  $(d_1, d_2, \dots, d_n)$ , and matrix  $A$  with elements  $a_{ij}$ ,  $(D^{-1}AD)_{ij} = d_j a_{ij}/d_i$ . The lemma then follows directly by substituting  $D_x$  in  $D_x^{-1}M_k(p)D_x$ .  $\square$

In what follows, we will frequently encounter lemniscates of the form  $|q(z)| = \alpha$ , where  $q$  is a polynomial. The zeros of  $q$  are the foci of the lemniscate, which, depending on the value of  $\alpha$  can consist of at most  $m$  disjoint closed curves, where  $m$  is the order of  $q$ . These curves are simple except for at most  $m-1$  critical values of  $\alpha$ . If a lemniscate has distinct foci  $\{z_j\}_{j=1}^m$ , and if there exists  $\eta > 0$  so that the disks  $\Delta_j = \{z \in \mathcal{C} : |z - z_j| < \eta\}$  are disjoint, then the lemniscate  $\{z \in \mathcal{C} : \prod_{j=1}^m |z - z_j| = \rho^n\}$  is contained in the union  $\cup_{j=1}^m \Delta_j$  for any  $\rho \leq \eta$ . Since a lemniscate must contain all of its foci in its interior, this provides an easily computable sufficient condition for a lemniscate to be composed of  $m$  disjoint simple curves. We refer to [11, Volume I, page 379] for a more detailed discussion of lemniscates.

**3. Main results.** Our first result is Pellet's theorem with a refinement of the region outside the disk with radius  $R$  in Theorem 2.1, where the largest zeros can be found.

**Theorem 3.1.** *Let  $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$  be a polynomial with complex coefficients,  $a_0 \neq 0$ ,  $1 \leq k \leq n-1$ ,  $n \geq 3$ , and with zeros  $\{z_i\}_{i=1}^n$ , labeled so that  $|z_1| \leq |z_2| \leq \dots \leq |z_n|$ . Let  $\{p_j\}_{j=1}^{n-1}$  be the associated polynomials of  $p$ , let  $P_j$  be the polynomial obtained from  $p_j$*

by replacing its coefficients with their moduli, and let

$$\mu(k, x) = \sum_{j=0}^{k-1} |a_j| x^{j-k}.$$

Furthermore, let

$$\begin{aligned} f_k(x) &= x^n + |a_{n-1}|x^{n-1} + \cdots + |a_{k+1}|x^{k+1} - |a_k|x^k \\ &\quad + |a_{k-1}|x^{k-1} + \cdots + |a_0| \end{aligned}$$

have two distinct positive roots  $r$  and  $R$  such that  $0 < r < R$ , and define the disjoint sets

$$\begin{aligned} \Omega_1(k) &= \{z \in \mathcal{D} : |p_{n-k}(z)| \leq P_{n-k}(r) = |a_k| - \mu(k, r)\}, \\ \Omega_2(k) &= \{z \in \mathcal{D} : |p_{n-k}(z) + a_k| \leq \mu(k, R) = |a_k| - P_{n-k}(R)\}. \end{aligned}$$

Then:

- (i) the  $k$  zeros  $\{z_i\}_{i=1}^k$  are contained in the closed disk  $\overline{O}(0; r)$ , whereas the remaining  $n - k$  zeros  $\{z_i\}_{i=k+1}^n$  are contained in  $\Omega_2(k)$ , the closed interior of a lemniscate;
- (ii)  $\overline{O}(0; r) \subseteq \Omega_1(k)$  and  $\Omega_2(k) \subseteq O^c(0; R)$ ;
- (iii) if  $\Omega_2(k)$  consists of disjoint regions whose boundaries are simple closed (Jordan) curves, then each disjoint region contains as many zeros of  $p$  as it contains zeros of  $p_{n-k}(z) + a_k$  (or foci of  $\Omega_2(k)$ ).

*Proof.* We start by observing that  $a_k \neq 0$  since  $f_k$  has positive zeros, and that  $f_k$  can be written as

$$(3.1) \quad f_k(x) = x^k(P_{n-k}(x) - |a_k| + \mu(k, x)).$$

That the zeros  $\{z_i\}_{i=1}^k$  are contained in the closed disk  $\overline{O}(0; r)$  is obtained from Rouché's theorem, exactly as in the proof of Pellet's theorem (see, e.g., [10, Theorem (28,1), page 128]).

The numbers  $\{p_{n-k}(z_i)\}_{i=1}^n$  are the eigenvalues of  $p_{n-k}(C(p)) = M_{n-k}(p)$ , and therefore also of  $D_x^{-1}M_{n-k}(p)D_x$  for any  $x > 0$ , where  $D_x$  is as in Lemma 2.8. From that same lemma, the matrix  $D_x^{-1}M_{n-k}(p)D_x$  is given by

$$\begin{pmatrix} 0 & & & & -a_0/x^k & & & \\ & a_{k+1}x & \ddots & & -a_1/x^{k-1} & \ddots & & \\ & a_{k+2}x^2 & \ddots & \ddots & \vdots & \ddots & \ddots & \\ & \vdots & \ddots & \ddots & & & & \\ & a_{n-1}x^{n-k-1} & \ddots & \ddots & 0 & -a_{k-1}/x & \ddots & -a_0/x^k \\ x^{n-k} & & \ddots & \ddots & a_{k+1}x & -a_k & \ddots & -a_1/x^{k-1} \\ & & \ddots & \ddots & a_{k+2}x^2 & & \ddots & \vdots \\ & & & \ddots & \vdots & & -a_{k-1}/x & \vdots \\ & & & & a_{n-1}x^{n-k-1} & & -a_k & -a_{k-1}/x \\ & & & & x^{n-k} & & & -a_k \end{pmatrix},$$

and its Gershgorin column set is easily seen to be the union of a disk centered at the origin with radius  $P_{n-k}(x)$  and a disk centered at  $-a_k$  with radius  $\mu(k, x)$ . These disks are disjoint if there exists  $\delta > 0$  such that  $P_{n-k}(\delta) + \mu(k, \delta) < |a_k|$ , which, in view of (3.1), is equivalent to  $f_k(\delta) < 0$ . It therefore suffices to choose any  $\delta$  for which  $r < \delta < R$  to obtain disjoint disks. Since the diagonal of  $D_\delta^{-1}M_{n-k}(p)D_\delta$  contains  $k$  zeros, we conclude that exactly  $k$  of the  $n$  numbers  $\{p_{n-k}(z_i)\}_{i=1}^n$  lie in the closed disk  $\overline{O}(0; P_{n-k}(\delta))$ , while the remaining  $n - k$  lie in the disjoint closed disk  $\overline{O}(-a_k; \mu(k, \delta))$ . Since this is true for any  $\delta$  such that  $r < \delta < R$ , the same conclusions hold for the disks  $\overline{O}(0; P_{n-k}(r))$  and  $\overline{O}(-a_k; \mu(k, R))$ , respectively. This concludes the proof of the first part of the theorem.

Recalling that we defined  $a_n = 1$ , we now observe that  $|z| \leq r$  implies that

$$|p_{n-k}(z)| = \left| \sum_{j=k+1}^n a_j z^{j-k} \right| \leq \sum_{j=k+1}^n |a_j| |z|^{j-k} \leq \sum_{j=k+1}^n |a_j| r^{j-k} = P_{n-k}(r),$$

which means that  $\overline{O}(0; r) \subseteq \Omega_1(k)$ . It also means that it is the  $k$  numbers  $\{p_{n-k}(z_i)\}_{i=1}^k$ , corresponding to the first  $k$  zeros of  $p$ , that lie in  $\overline{O}(0; P_{n-k}(r))$ . To show the second inclusion in the statement of the theorem, assume that  $z \in \Omega_2(k)$ . Since the disk  $\overline{O}(-a_k; \mu(k, R))$  is bounded away from the origin, we have that  $|p_{n-k}(z)| \geq |a_k| - \mu(k, R) = P_{n-k}(R) > 0$ , and therefore that

$$(3.2) \quad P_{n-k}(|z|) \geq |p_{n-k}(z)| \geq P_{n-k}(R) > 0.$$

The polynomial  $P_{n-k}$  is strictly increasing for positive arguments, so that inequality (3.2) implies that  $|z| \geq R$  and therefore that  $z$  must lie



in  $O^c(0; R)$ . This proves the second part of the theorem.

The lemniscate  $\Omega_2(k)$  has  $n - k$  foci, which are the zeros of  $p_{n-k}(z) + a_k$ , and it can consist of, at most,  $n - k$  disjoint regions. Any disjoint region contains one or more foci. Let us now assume that there exists a disjoint region  $\mathcal{H}$  of  $\Omega_2(k)$  with a simple boundary. Define

$$\tilde{\Omega}_2(k) = \{z \in \mathcal{C} : |p_{n-k}(z) + a_k| \leq \mu(k, \delta)\}, \quad r < \delta < R,$$

where  $R - \delta$  is small enough so that  $\tilde{\Omega}_2(k)$  (which contains  $\Omega_2$  since  $\mu(k, \delta) > \mu(k, R)$ ) has the same number of disjoint regions with simple boundary as  $\Omega_2$ , one of which,  $\tilde{\mathcal{H}}$ , contains  $\mathcal{H}$ . Similarly, as before, we have for  $z \in \tilde{\Omega}_2(k)$  that  $|p_{n-k}(z)| \geq |a_k| - \mu(k, \delta) > P_{n-k}(\delta) > 0$ , and therefore that

$$P_{n-k}(|z|) \geq |p_{n-k}(z)| > P_{n-k}(\delta) > 0,$$

so that now  $|z| > \delta$ .

For any  $z$  on  $\partial\tilde{\mathcal{H}}$ , we have that

$$(3.3) \quad |p_{n-k}(z) + a_k| = \mu(k, \delta).$$

Since  $|z| > \delta$ , multiplying both sides of (3.3) by  $|z|^k$  yields

$$\begin{aligned} & |z^n + a_{n-1}z^{n-1} + \cdots + a_kz^k| \\ &= |a_{k-1}| \left| \frac{z}{\delta} \right| |z|^{k-1} + |a_{k-2}| \left| \frac{z}{\delta} \right|^2 |z|^{k-2} + \cdots + |a_0| \left| \frac{z}{\delta} \right|^k \\ &> |a_{k-1}| |z|^{k-1} + |a_{k-2}| |z|^{k-2} + \cdots + |a_0| \\ &\geq |a_{k-1}z^{k-1} + a_{k-2}z^{k-2} + \cdots + a_0|. \end{aligned}$$

Then, by Rouché's theorem, the polynomial  $p$  has as many zeros inside this disjoint region  $\tilde{\mathcal{H}}$  as  $z^k(p_{n-k}(z) + a_k)$ , which is as many zeros as  $p_{n-k}(z) + a_k$  since  $\tilde{\mathcal{H}}$  does not contain 0. Because this remains true as  $\delta \rightarrow R^-$ , the proof follows.  $\square$

It is a direct consequence of part (iii) of this theorem that if  $\Omega_2(k)$  consists of  $n - k$  disjoint regions with simple boundaries, then each region must contain exactly one zero of  $p$ .

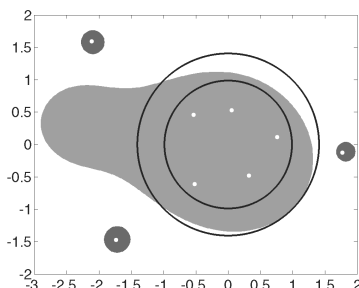


FIGURE 1. The sets  $\Omega_1(5)$  (light gray) and  $\Omega_2(5)$  (dark gray) for  $q_A$ .

We illustrate this theorem at the hand of the polynomial  $q_A$ , defined by

$$q_A(z) = z^8 + 2z^7 - (1-i)z^6 - 11z^5 + \frac{1}{2}z^4 + z^3 + \frac{5}{2}z^2 - z + (1+i),$$

and its sets  $\Omega_1(5)$  and  $\Omega_2(5)$ . We obtain  $r = 0.9872$  and  $R = 1.4065$ , so that

$$\Omega_1(5) = \{z \in \mathcal{D} : |z^3 + 2z^2 - (1-i)z| \leq 4.3065\},$$

$$\Omega_2(5) = \{z \in \mathcal{D} : |z^3 + 2z^2 - (1-i)z - 11| \leq 2.2720\}.$$

Figure 1 shows  $\Omega_1(5)$  in light gray and  $\Omega_2(5)$  in dark gray for  $q_A$ , while its zeros are represented by the white dots. The two circles are the boundaries of the disks  $O(0; r)$  and  $O(0; R)$ . The smaller disk, which contains five zeros, is contained in  $\Omega_1(5)$ , while  $\Omega_2(5)$ , which contains the remaining three zeros, lies outside the larger disk. The disk  $\overline{O}(0; r)$  is clearly preferable to  $\Omega_1(5)$ , whereas  $\Omega_2(5)$  is clearly preferable to  $O^c(0; R)$ , as predicted by the theorem. We remark that  $\Omega_2(k)$  does not necessarily consist of disjoint regions, although it often happens, as in this case. The boundary of  $\Omega_2(5)$  has foci at  $1.8291 - 0.1119i$ ,  $-2.1035 + 1.5937i$  and  $-1.7257 - 1.4818i$ , which can be enclosed in disjoint disks of radius 1.5 centered at these foci. Since  $(1.5)^3 = 3.3750$  and  $2.2720 < 3.3750$ ,  $\Omega_2(5)$  must consist of three disjoint regions with simple boundaries, as explained at the end of Section 2. By part (iii) of Theorem 3.1 each must contain exactly one zero of  $q_A$ .

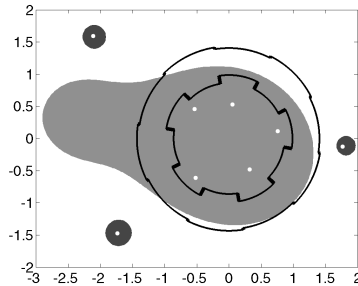


FIGURE 2. The sets  $\Omega_1(5)$  (light gray) and  $\Omega_2(5)$  (dark gray) for  $q_A$ , compared to the refined Pellet region.

In [9] and [10, Theorem (29,1), page 130 and Exercise 2, page 133] a refinement of Pellet's theorem was derived, requiring the solution of an additional equation, related to  $f_k(x) = 0$  in Theorem 3.1, along with a nonzero requirement on one additional coefficient. It was further generalized in [8, Theorem 1]. The refinement leads to a slightly better gear-wheel shaped region (instead of an annulus) that does not contain any zeros of the polynomial. In Figure 2, we have compared  $\Omega_1(5)$  and  $\Omega_2(5)$  for  $q_A$  to the refined region from [10, Theorem (29,1), page 130]. The radii of the circles determining the inner boundary of this region are 0.8612 and 0.9872, whereas for the outer boundary they are given by 1.4065 and 1.4331.

Let us consider two more examples, namely,  $q_B = z^3 + 4z^2 + 2z + 1$  (the example in [8]) for  $k = 2$ , and  $q_C = z^8 + 3iz^5 + z^4 - 8iz^3 + 2iz + 1$  for  $k = 3$ . Similar conclusions as for  $q_A$  can be drawn for these polynomials as can be seen from Figure 3, which shows (using the same conventions as before) the corresponding sets  $\Omega_1(k)$  and  $\Omega_2(k)$  for  $q_B$  (left) and  $q_C$  (right) with  $k = 2$  and  $k = 3$ , respectively, together with the refined Pellet region from [10, Theorem (29,1), page 130]. We leave out the details for brevity.

**Remarks.** (1) By using the reciprocal polynomial, similar lemniscates can be derived for the reciprocals of the zeros of a polynomial, leading to corresponding inclusion regions for the zeros themselves, although these are more complicated. They can be combined with our previous results as is illustrated for  $q_A$  in Figure 4, which also shows the

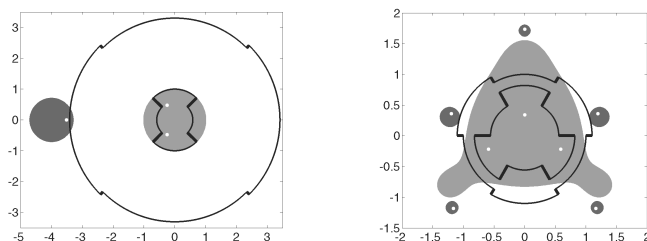


FIGURE 3. The sets  $\Omega_1(k)$  (light gray) and  $\Omega_2(k)$  (dark gray) for  $q_B$  and  $q_C$  with  $k = 2$  and  $k = 3$ , respectively, compared to the refined Pellet region.

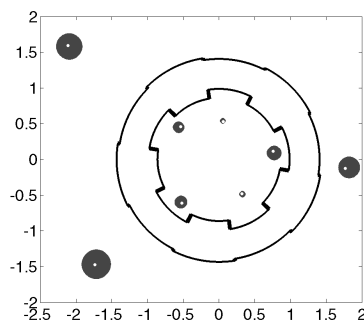


FIGURE 4. Inclusion sets obtained from both the polynomial and its reciprocal for  $q_A$ .

same refined Pellet region as before. Results of a very similar nature are obtained for  $q_B$  and  $q_C$ .

(2) A converse of Pellet's theorem is stated in [20] (see also [10, Theorem (28,3), page 129]), whose proof was later corrected in [14]. It can be formulated as follows. Let  $a_j$  ( $j = 1, 2, \dots, n$ ) be fixed complex coefficients and  $\epsilon_j$  ( $j = 1, 2, \dots, n$ ) arbitrary complex numbers with  $|\epsilon_j| = 1$  for all  $j = 1, 2, \dots, n$ . Let  $\tau$  be any positive number such that it is not a zero of any polynomial  $\sum_{j=1}^n a_j \epsilon_j z^j$ , and let every polynomial of that form have  $k$  zeros ( $0 < k < n$ ) in  $\overline{O}(0; \tau)$ . Then  $f_k$  (with  $f_k$  as defined in Theorem 2.1) has two positive roots  $r$  and  $R$  such that  $r < \rho < R$ .

It may be possible to construct a similar converse of Theorem 3.1,

although its proof would be well beyond the scope of this paper.

(3) A sufficient condition for the  $n-k$  largest zeros to have a modulus strictly larger than  $R$  is for  $\Omega_2(k)$  to be contained in the interior of  $O^c(0; R)$ . A similar condition based on the reciprocal polynomial holds for the  $k$  smallest zeros. In [21], different (but more explicit) conditions were derived for the same situation. The special case  $r = R = \rho$  was considered in [20] (also mentioned in [10, Theorem (28,2), page 129]), where it was shown that the polynomial then has  $\ell$  ( $\ell \geq 0$ ) double roots on the circle  $|z| = \rho$ ,  $k - \ell$  zeros inside and  $n - k - \ell$  zeros outside this circle, respectively. Theorem 3.1 adds to this result a sufficient condition guaranteeing that no double roots lie on the circle  $|z| = \rho$ , as follows: from its proof, we have in this case that the two disks in the Gershgorin set of  $D_x^{-1}M_{n-k}(p)D_x$  are tangent to each other when  $p_{n-k}(\zeta) = -P_{n-k}(\rho)a_k/|a_k|$  for a point  $\zeta \in \Omega_1(k) \cap \Omega_2(k)$ , i.e., when

$$(3.4) \quad p_{n-k}(\zeta) + \frac{P_{n-k}(\rho)}{|a_k|} a_k = 0.$$

This means that  $\Omega_1(k)$  and  $\Omega_2(k)$  touch at exactly  $n-k$  points, namely, the zeros of the polynomial of degree  $n-k$  in (3.4). If those points do not lie on the circle  $|z| = \rho$ , then, because the zeros of a polynomial are continuous functions of their coefficients, none in the group of largest zeros of  $p$  can cross the gap between  $\Omega_2(k)$  and the circle when the coefficients are continuously perturbed from a situation of two very close distinct values for  $r$  and  $R$  to one where  $r = R$ . In view of [20], this means that in such a case no double roots of  $p$  can lie on the circle.

(4) The set  $\Omega_2(k)$  can also be derived without using Gershgorin's theorem in the following way. The equation  $p(z) = 0$  can be written as

$$z^k(z^{n-k} + a_{n-1}z^{n-k-1} + \cdots + a_k) = -a_{k-1}z^{k-1} - \cdots - a_0,$$

from which we have

$$|z^{n-k} + a_{n-1}z^{n-k-1} + \cdots + a_k| \leq \frac{|a_{k-1}|}{|z|} + \cdots + \frac{|a_0|}{|z|^k}.$$

The zeros  $\{z_i\}_{i=k+1}^n$ , which satisfy  $|z_i| \geq R$ , must then also satisfy

$$|z_i^{n-k} + a_{n-1}z_i^{n-k-1} + \cdots + a_k| \leq \frac{|a_{k-1}|}{R} + \cdots + \frac{|a_0|}{R^k},$$

and therefore  $z_i \in \Omega_2(k)$  for  $k+1 \leq i \leq n$ .

On the other hand, Gershgorin's theorem permits a unified and convenient treatment of both the previous and the next theorem, our second result, which represents a generalization of Theorem 2.2.

**Theorem 3.2.** *Let  $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$  be a polynomial with complex coefficients,  $a_0 \neq 0$ ,  $1 \leq k \leq n-1$ ,  $n \geq 3$ , and with zeros  $\{z_i\}_{i=1}^n$ . Let  $\{p_j\}_{j=1}^{n-1}$  be the associated polynomials of  $p$ , let  $P_j$  be the polynomial obtained from  $p_j$  by replacing its coefficients with their moduli, and let*

$$\mu(k, x) = \sum_{j=0}^{k-1} |a_j| x^{j-k}.$$

For  $j = 0, 1, \dots, n-1$ , let  $s_j$  be the unique positive root of

$$h_j(x) = x^n + |a_{n-1}|x^{n-1} + \cdots + |a_{j+1}|x^{j+1} - |a_j|x^j - |a_{j-1}|x^{j-1} - \cdots - |a_0|.$$

Then

- (1) all the zeros of  $p$  are contained in  $\Upsilon_1(k)$  and also in  $\Upsilon_2(k)$ , where

$$\Upsilon_1(k) = \{z \in \mathcal{D} : |p_{n-k}(z)| \leq P_{n-k}(s_k) = \mu(k, s_k) + |a_k|\},$$

$$\Upsilon_2(k) = \{z \in \mathcal{D} : |p_{n-k}(z) + a_k| \leq \mu(k, s_{k-1}) = P_{n-k}(s_{k-1}) + |a_k|\},$$

each of which is the closed interior of a lemniscate;

- (2) if  $\Upsilon_1(k)$  or  $\Upsilon_2(k)$  consists of disjoint regions whose boundaries are simple closed (Jordan) curves and  $\ell$  is the number of foci of the corresponding lemniscate contained in any such region, then that region contains  $\ell$  zeros of  $p$  when that region does not contain the origin, and when it does contain the origin, then it contains  $\ell + k$  zeros of  $p$ .

*Proof.* We begin by observing that  $h_j$  can be written as

$$(3.5) \quad h_j(x) = x^j(P_{n-j}(x) - |a_j| - \mu(j, x)).$$

Let us now consider once more the matrix  $D_x^{-1}M_{n-k}(p)D_x$  for any  $x > 0$ , where  $D_x$  is as in Lemma 2.8, whose eigenvalues are the numbers  $\{p(z_i)\}_{i=1}^k$ . Its Gershgorin column set, which contains these numbers, is the union of the two disks  $\overline{O}(0; P_{n-k}(x))$  and  $\overline{O}(-a_k; \mu(k, x))$ . As  $x$  increases, so does  $P_{n-k}(x)$ , while  $\mu(k, x)$  decreases. The former disk

will then encompass the latter when  $P_{n-k}(x) = |a_k| + \mu(k, x)$ , or, as can be seen from (3.5),  $h_k(x) = 0$ . That is, when  $x = s_k$ . All the zeros of  $p$  are then contained in the set

$$\Upsilon_1(k) = \{z \in \mathcal{D} : |p_{n-k}(z)| \leq P_{n-k}(s_k)\}.$$

Because  $h_k(s_k) = 0$ , the right-hand side of the inequality defining  $\Upsilon_1(k)$  can be replaced by  $\mu(k, s_k) + |a_k|$ .

On the other hand, we can let  $x$  decrease until the disk centered at  $-a_k$  encompasses the one centered at the origin. This happens when  $\mu(k, x) = |a_k| + P_{n-k}(x)$ . Since, by using (3.5), we can write

$$\begin{aligned} & P_{n-k}(x) + |a_k| - \mu(k, x) \\ &= x^{-1} (x(P_{n-k}(x) + |a_k|) - x\mu(k, x)) \\ &= x^{-1} (P_{n-k+1}(x) - (|a_{k-1}| + \mu(k-1, x))) \\ &= x^{-k} h_{k-1}(x), \end{aligned}$$

we conclude that  $x = s_{k-1}$ . All the zeros of  $p$  are then contained in the set

$$\Upsilon_2(k) = \{z \in \mathcal{D} : |p_{n-k}(z) + a_k| \leq \mu(k, s_{k-1})\}.$$

Because  $h_{k-1}(s_{k-1}) = 0$ , the right-hand side of the inequality defining  $\Upsilon_1(k)$  can be replaced by  $P_{n-k}(s_{k-1}) + |a_k|$ .

If there exists a disjoint region  $\mathcal{H}$  of  $\Upsilon_1(k)$  with a simple boundary, we proceed similarly as in the proof of Theorem 3.1 and first define

$$\tilde{\Upsilon}_1(k) = \{z \in \mathcal{D} : |p_{n-k}(z)| \leq P_{n-k}(s)\}, \quad s > s_k,$$

where  $s - s_k$  is small enough so that  $\tilde{\Upsilon}_1(k)$  (which contains  $\Upsilon_1(k)$  because  $P_{n-k}(s) > P_{n-k}(s_k)$ ) has the same number of disjoint regions with simple boundary as  $\Upsilon_1(k)$ , one of which,  $\tilde{\mathcal{H}}$ , contains  $\mathcal{H}$ . Any  $z \in \partial\tilde{\mathcal{H}}$  satisfies

$$P_{n-k}(|z|) \geq |p_{n-k}(z)| = P_{n-k}(s),$$

implying that  $|z| \geq s$  because  $P_{n-k}$  is increasing for positive arguments. In addition, because  $s > s_k$ , we also have that  $P_{n-k}(s) > \mu(k, s) + |a_k|$ .

Then we can write for any  $z \in \partial\tilde{\mathcal{H}}$ :

$$\begin{aligned} |p_{n-k}(z)| = P_{n-k}(s) &> |a_k| + \frac{|a_{k-1}|}{s} + \cdots + \frac{|a_0|}{s^k} \\ &\geq |a_k| + \frac{|a_{k-1}|}{|z|} + \cdots + \frac{|a_0|}{|z|^k}, \end{aligned}$$

from which it follows that

$$|z|^k |p_{n-k}(z)| > |a_k| |z|^k + |a_{k-1}| |z|^{k-1} + \cdots + |a_0|,$$

and therefore

$$|z^k p_{n-k}(z)| > |a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0|.$$

Since we assumed that  $\partial\tilde{\mathcal{H}}$  is a simple closed (Jordan) curve, we conclude from Rouché's theorem that the polynomial  $p$  has as many zeros in  $\tilde{\mathcal{H}}$  as the polynomial  $z^k p_{n-k}(z)$ . If  $0 \notin \tilde{\mathcal{H}}$ , then that number is the number of foci of the lemniscate forming the boundary of  $\tilde{\Upsilon}_1(k)$  that lie in  $\tilde{\mathcal{H}}$ , i.e., those zeros of  $p_{n-k}$  that are contained in  $\tilde{\mathcal{H}}$ . If  $0 \in \tilde{\mathcal{H}}$ , then  $k$  is added to the number of foci. Because this remains true as  $s \rightarrow s_k^+$ , the proof follows. The proof for  $\Upsilon_2(k)$  is analogous.  $\square$

Following are the four lowest-order lemniscates containing all the zeros of  $p$ :

$$\begin{aligned} \Upsilon_1(n-1) &= \{z \in \mathcal{C} : |z| \leq s_{n-1}\}, \\ \Upsilon_2(n-1) &= \{z \in \mathcal{C} : |z + a_{n-1}| \leq |a_{n-1}| + s_{n-2}\}, \\ \Upsilon_1(n-2) &= \{z \in \mathcal{C} : |z(z + a_{n-1})| \leq s_{n-2}(s_{n-2} + |a_{n-1}|)\}, \\ \Upsilon_2(n-2) &= \{z \in \mathcal{C} : |z(z + a_{n-1}) + a_{n-2}| \\ &\quad \leq s_{n-3}(s_{n-3} + |a_{n-1}|) + |a_{n-2}|\}. \end{aligned}$$

The sets  $\Upsilon_1(n-1)$  and  $\Upsilon_2(n-1)$  are closed disks, whereas the sets  $\Upsilon_1(n-2)$  and  $\Upsilon_2(n-2)$  are the closed interiors of ovals of Cassini ([7, pages 153–155]). The zeros of  $p$  lie in the intersection of all of the aforementioned sets, although, with the exception of the sets that are disks, such an intersection is in general difficult to compute. When all the coefficients of the polynomial are real, then all our inclusion sets are symmetric with respect to the real axis.



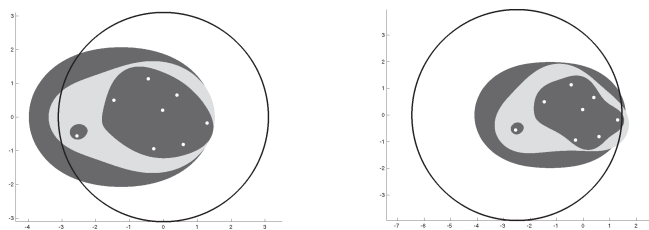


FIGURE 5. The sets  $\Upsilon_1(k)$  (left) and  $\Upsilon_2(k)$  (right) for  $q_D$  with  $k = 6, 5, 3$ .

To illustrate Theorem 3.2, we consider the polynomial  $q_D$ , defined by

$$q_D(z) = z^8 + \frac{5}{2}z^7 - \frac{1}{2}z^6 - (2 - 3i)z^5 - \frac{1}{2}z^4 - 4z^3 + 2z^2 - 5z + i.$$

Figure 5 shows the corresponding sets  $\Upsilon_1(k)$  on the left and  $\Upsilon_2(k)$  on the right for  $k = 6, 5, 3$ . To better compare them, the lemniscates have been superimposed in alternating shades of light and dark gray. The circles mark the boundaries of the sets  $\Upsilon_1(n-1) = \Upsilon_1(7)$  (left) and  $\Upsilon_2(n-1) = \Upsilon_2(7)$  (right). The largest proper lemniscates in dark gray, the ones in light gray, and the smallest ones in dark gray correspond to  $k = 6$ ,  $k = 5$ , and  $k = 3$ , respectively. The white dots are the zeros of  $q_D$ .

- Remark 3.3.** (1) The sets  $Y_1(k)$  and  $Y_2(k)$  are not necessarily nested for successive values of  $k$ , although they do tend to become smaller as  $k$  decreases.
- (2) Theorem 3.2 is a generalization of Theorem 2.2, since that theorem is obtained for  $\Upsilon_1(k)$  with  $k = n - 1$  in Theorem 3.2. Lemniscates are more interesting geometric regions than disks, and even low-order lemniscates can already provide significantly improved inclusion regions for the zeros of a polynomial. The computation of the positive root of any  $h_j$  is inexpensive compared to the computation of the (generally complex) zeros of  $p$ .
- (3) The last part of the previous theorem's proof shows that it can be proven with just Rouché's theorem, without using Gershgorin's

theorem. However, the latter provides a natural explanation for the appearance of the functions  $h_j$  that would otherwise be lacking and also creates a larger framework in which both Pellet's and the generalized Cauchy theorems are obtained as special cases for special values of  $x$  in  $D_x^{-1}M_{n-k}(p)D_x$ . This framework generates infinitely many other inclusion regions depending on the values of  $x$ , e.g., the value for which both disks in the proof of Theorem 3.2 have the same radius, leading to a union of the interiors of two lemniscates, to give but one example.

- (4) The  $\Upsilon_1(k)$  and  $\Upsilon_2(k)$  sets can sometimes be simplified if the coefficients of the polynomial exhibit certain patterns. For instance, if the leading coefficients fit the pattern of a power of  $(z - a)$  for some complex number  $a$ , then the corresponding set, defined by a polynomial of the same degree as the power, becomes a simple disk. Consider as an example the set  $\Upsilon_2(3)$  for the polynomial  $z^5 + 2iz^4 - z^3 + z^2 + 3z - 1$ , whose first three coefficients are the same as those of  $(z + i)^2$ . It is given by

$$\Upsilon_2(3) = \{z \in \mathcal{C} : |z + i| = s_2 + 1\}.$$

The inclusion sets for lacunary polynomials with several consecutive leading zero coefficients can be treated in the same way since their leading coefficients fit the pattern of a power of  $(z - 0)$ .

## REFERENCES

1. D.A. Bini, *Numerical computation of polynomial zeros by means of Aberth's method*, Num. Alg. **13** (1996), 179–200.
2. D.A. Bini, V. Noferini and M. Sharify, *Locating the eigenvalues of matrix polynomials*, SIAM J. Matrix Anal. Appl. **34** (2013), 1708–1727.
3. A.L. Cauchy, *Sur la résolution des équations numériques et sur la théorie de l'élimination*, Exer. Math., Quatrième Année, 65–128, de Bure frères, Paris, 1829. Also in: Oeuvres Complètes **9** (1891), 86–161.
4. S. Gerschgorin, *Über die Abgrenzung der Eigenwerte einer Matrix*, Izv. Akad. Nauk SSSR **6** (1931), 749–754.
5. R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1988.
6. S. Lang, *Complex analysis*, Springer-Verlag, New York, 1999.
7. J.D. Lawrence, *A catalog of special plane curves*, Dover Publications, Inc., New York, 1972.

8. B.W. Levinger, *A generalization of Pellet's theorem concerning the zeros of a polynomial*, Proc. Amer. Math. Soc. **18** (1967), 767–774.
9. M. Marden, *A refinement of Pellet's theorem*, Bull. Amer. Math. Soc. **54** (1948), 550–557.
10. ———, *Geometry of polynomials*, Math. Surv. **3**, American Mathematical Society, Providence, R.I., 1966.
11. A.I. Markushevich, *Theory of functions of a complex variable*, Richard A. Silverman, ed., Chelsea Publishing Co., New York, 2005.
12. A. Melman, *Generalization and variations of Pellet's theorem for matrix polynomials*, Linear Alg. Appl. **439** (2013), 1550–1567.
13. ———, *Implementation of Pellet's theorem*, Num. Alg. **65** (2014), 293–304.
14. A. Ostrowski, *On a theorem by J. L. Walsh concerning the moduli of roots of algebraic equations*, Bull. Amer. Math. Soc. **47** (1941), 742–746.
15. M.A. Pellet, *Sur un mode de séparation des racines des équations et la formule de Lagrange*, Bull. Sci. Math. **5** (1881), 393–395.
16. Q.I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, Lond. Math. Soc. Mono. **26**, The Clarendon Press, Oxford University Press, Oxford, 2002.
17. E. Rouché, *Mémoire sur la série de Lagrange*, J. Ecole Polytech. **22** (1862), 217–218.
18. S. Rump, *Ten methods to bound multiple roots of polynomials*, J. Comp. Appl. Math. **156** (2003), 403–432.
19. R.S. Varga, *Geršgorin and his circles*, Springer-Verlag, Berlin, 2004.
20. J.L. Walsh, *On Pellet's theorem concerning the roots of a polynomial*, Ann. Math. **26** (1924), 59–64.
21. Zifang Zhang, Xu Daoyi and Janren Niu, *On the refinement of Cauchy's theorem and Pellet's theorem*, J. Math. Anal. Appl. **291** (2004), 262–269.

DEPARTMENT OF APPLIED MATHEMATICS, SCHOOL OF ENGINEERING, SANTA CLARA UNIVERSITY, SANTA CLARA, CA 95053

Email address: amelman@scu.edu