Zero Location by Hermitian Forms: The Singular Case

Vlastimil Pták
Institute of Mathematics
Czechoslovak Academy of Sciences
Žitná 25
115 67 Praha 1, Czechoslovakia

and

N. J. Young
Mathematics Department
The University
Glasgow, Scotland

Submitted by Miroslav Fiedler

ABSTRACT

The Schur-Cohn criterion for the number of zeros of a polynomial inside and outside the unit disc fails if the polynomial has a pair of conjugate zeros or a zero on the unit circle: the corresponding quadratic form is singular. Recently the authors have shown [5] that the classical Schur-Cohn criterion may be deduced from a simple algebraic identity; this yields not only a very simple proof but also a substantial generalization. The method produces a whole family of quadratic forms which may be used for testing the zeros. In the present paper the same algebraic identity is used to show that singularity of these quadratic forms is always due to the presence of pairs of conjugate zeros or zeros on the unit circle. There is a method for ascertaining zero distribution in the singular case by differentiation; we give a derivation of this test on the basis of our matrix-theoretic treatment. The second section deals with the same problem in the case of a general circle or half plane.

In 1856 Hermite wrote a letter to M. Borchardt describing a novel application of Hermitian forms to the theory of equations, and an extract from this letter was subsequently printed in Crelle's journal [3]. Consider a polynomial p of degree n over the complex field $\mathbf{C}$, and define the $n\times n$ matrix $A=[a_{ij}]$ by

$$\sum_{i,j=1}^{n} a_{ij}z^{i-1}w^{j-1}=-i\frac{p(w)p^*(z)-p(z)p^*(w)}{w-z},$$

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/020181+16$02.75

where $p^*(z) = p(\bar{z})^-$. The Hermitian form x^*Ax can be converted by a linear change of variable to the form

$$|y_1|^2 + \cdots + |y_k|^2 - |y_{k+1}|^2 - \cdots - |y_r|^2,$$

and Hermite showed that if $r = n$, then p has k zeros in the upper half plane and $n - k$ zeros in the lower half plane. Similar tests for the numbers of zeros of a polynomial lying in other plane regions have since been given [4]; the best known is that of Schur and Cohn for the case of the unit disc [1]. The criteria are given in numerous textbooks (e.g. [2], [4]). A disadvantage of Hermite's test is the proviso "if $r = n$." It can easily happen that A is singular —indeed, if p has real coefficients, $A = 0$ and the test yields no information. Hermite had no comments to make on this difficulty: at that time there was still a tendency to regard this type of singularity as an aberration that was unworthy of attention. Hermite did not even explicitly make the hypothesis that A is nonsingular, no doubt presuming that his correspondent would interpret the result in the light of the proof, which is in fact only valid when p has no repeated zeros nor any pair of complex conjugate zeros. Later proofs (e.g. [4]) have shown that the presence of repeated zeros is no bar to the validity of Hermite's theorem, but it is not hard to see that A will always be singular for polynomials having a pair of conjugate zeros.

Since Hermite's time many authors have studied zero-location problems and paid attention to singular cases. A notable contribution was that of Cohn [1], who, building on work of Schur, showed how to determine the number of zeros of a polynomial in the unit disc. He also showed how differentiation could be used to handle the singular case.

The authors have recently given a simple proof and substantial generalization of the Schur-Cohn criterion for the nonsingular case, using a simple algebraic identity [5]. In the present paper we extend the method to give a simple unified treatment which also encompasses the singular case.

We begin with the version for the unit disc. Suppose we wish to ascertain how many of the zeros of a given polynomial p lie in the open unit disc $U \triangleq \{z \in \mathbf{C} : |z| < 1\}$. The Schur-Cohn test proceeds as follows. For any polynomial p of degree n,

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0, \qquad a_n \neq 0,$$

we define a polynomial p_0 by

$$\begin{aligned} p_0(z) &= z^n p(1/\bar{z})^- \\ &= \bar{a}_0 z^n + \bar{a}_1 z^{n-1} + \cdots + \bar{a}_n. \end{aligned}$$

Let S denote the $n\times n$ matrix

$$S=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

with ones on the superdiagonal and zeros elsewhere, and form the Hermitian $n\times n$ matrix

$$H=p_0(S)^*p_0(S)-p(S)^*p(S), \tag{1}$$

the stars denoting conjugate transpose. As in Hermite's test, transform the Hermitian form x^*Hx to a real linear combination of squares by a linear change of variable; denote the number of positive coefficients in the resulting expression by $\pi(H)$. Then, provided H is nonsingular, p has $\pi(H)$ zeros in U and no zeros on the unit circle.

It follows that the Schur matrix H is singular whenever p has zeros on the unit circle. Under what other circumstances will be the Schur-Cohn test fail? And when it does, how may we determine the zero distribution of p with respect to the unit circle? We shall answer these questions for a class of tests which includes the Schur-Cohn test. We shall show that any polynomial p can be factorized in the form $p=fg$ in a constructive manner and the desired information obtained by testing f and g'.

The first question, as to the singularity of H, is answered by the formula

$$\det H=|a_0|^{2n}\prod_{i,j=1}^{n}(1-\alpha_j\bar{\alpha}_i),$$

where $\alpha_1,\dots,\alpha_n$ are the zeros of p (see [6, Equation (28)]). However, it takes a great deal of calculation to establish this formula, and we shall see that the characterization of the singular case is really quite easy, even for a wider class of tests.

We write U for the open unit disc, and V for the complement of the closed unit disc. We shall call $\alpha\in\mathbf{C}$ a *conjugate zero* of p if $p(\alpha)=0$ and $\alpha\bar{\beta}=1$ for some zero β of p; otherwise a zero α is *nonconjugate*. Every zero on the unit circle is a conjugate zero. We denote the Euclidean norm on $\mathbf{C}^n$ by $\|\cdot\|$. The monic highest common factor of polynomials f,g is written (f,g).

THEOREM 1. *Let p be a polynomial of degree n, let T be an $m\times m$ matrix, $m\geqslant n$, such that $I-T^*T$ is positive semidefinite and has rank one, and let T have no eigenvalue of unit modulus.*

(a) *The Hermitian form*

$$\mathsf{J}(x)=\|p_0(T)x\|^2-\|p(T)x\|^2$$

on $\mathbf{C}^m$ has rank at most n; J has rank n if and only if p has no conjugate zeros. If the canonical form of J is

$$|y_1|^2+|y_2|^2+\cdots+|y_r|^2-|y_{r+1}|^2-\cdots-|y_n|^2, \tag{2}$$

then p has r zeros in U and $n-r$ zeros in V (counting multiplicities).

(b) *If $(p,p_0)(T)$ is nonsingular, the canonical form of J is*

$$|y_1|^2+\cdots+|y_r|^2-|y_{r+1}|^2-\cdots-|y_{r+s}|^2, \tag{3}$$

where r, s are the numbers of nonconjugate zeros of p in U, V respectively.

Statement (a) is a slight sharpening of the main result of [5]; it describes the class of tests we are studying. It includes the Schur-Cohn test, since S has the unique eigenvalue 0 and $I-S^*S=\text{diag}\{1,0,\ldots,0\}$, which is positive semidefinite and of rank one, so that S may be substituted for T to give a Hermitian form J whose matrix is precisely the Schur matrix (1).

Proof. (a): We can assume that p is monic. Let

$$p(z)=(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n),$$

and let

$$B_j=T-\alpha_jI,\quad C_j=I-\bar{\alpha}_jT,\qquad 1\leqslant j\leqslant n.$$

Then

$$p(T)=B_1B_2\cdots B_n,\qquad p_0(T)=C_1C_2\cdots C_n.$$

Since $I-T^*T$ is positive semidefinite and of rank one, it has the form uu^* for some column vector $u \in \mathbf{C}^m \setminus \{0\}$. It is shown in [5, Equation (5)] that if we let

$$\begin{aligned} v_1 &= C_n^* \cdots C_2^* u, \\ v_j &= C_n^* \cdots C_{j+1}^* B_{j-1}^* \cdots B_1^* u, \qquad 2 \leqslant j \leqslant n-1, \\ v_n &= B_{n-1}^* \cdots B_1^* u, \end{aligned} \tag{4}$$

then we can write the matrix J of the Hermitian form J as follows:

$$\begin{aligned} J &= p_0(T)^* p_0(T) - p(T)^* p(T) \\ &= \sum_{j=1}^{n} \left(1 - |\alpha_j|^2\right) v_j v_j^*. \end{aligned} \tag{5}$$

It follows that the rank of J is at most n.

To prove the second half of (a) we need the following technical fact.

LEMMA 1. *Let $\alpha_1, \dots, \alpha_n \in \mathbf{C}$ be such that $\alpha_i \bar{\alpha}_j \neq 1$, $1 \leqslant i, j \leqslant n$, and let polynomials $f_1, \dots, f_n$ be de defined by*

$$\begin{aligned} f_1(z) &= (1 - \alpha_n z)(1 - \alpha_{n-1} z) \cdots (1 - \alpha_2 z), \\ f_j(z) &= (1 - \alpha_n z) \cdots (1 - \alpha_{j+1} z)(z - \bar{\alpha}_{j-1}) \cdots (z - \bar{\alpha}_1), \qquad 2 \leqslant j \leqslant n-1, \\ f_n(z) &= (z - \bar{\alpha}_{n-1}) \cdots (z - \bar{\alpha}_1). \end{aligned}$$

Then $f_1, \dots, f_n$ are linearly independent.

Proof. Suppose not: then $g_1, \dots, g_n$, where $g_j(z) = f_j(z) / \prod_{i=1}^{n} (1 - \alpha_i z)$, are linearly dependent elements of the space E of all rational functions whose poles belong to the set $\{\alpha_1^{-1}, \dots, \alpha_n^{-1}\}$. By our hypothesis, no $\bar{\alpha}_i$ belongs to the latter set, and hence we can define linear operators $P_j \colon E \to E$, $1 \leqslant j \leqslant n$, by

$$P_j f(z) = \frac{(1 - \alpha_j z) f(z) - (1 - \alpha_j \bar{\alpha}_j) f(\bar{\alpha}_j)}{z - \bar{\alpha}_j}.$$

It is clear that

$$P_i \frac{1}{1-\alpha_i z} = 0,$$

$$P_i \left\{ \frac{z-\bar{\alpha}_i}{1-\alpha_i z} g(z) \right\} = g(z) \qquad \text{for all} \quad g \in E.$$

To say that $g_1, \dots, g_n$ are linearly dependent is to assert the existence of $\lambda_1, \dots, \lambda_n$, not all zero, such that

$$\frac{\lambda_1}{1-\alpha_1 z} + \frac{z-\bar{\alpha}_1}{1-\alpha_1 z} \frac{\lambda_2}{1-\alpha_2 z} + \cdots$$

$$+ \frac{z-\bar{\alpha}_1}{1-\alpha_1 z} \frac{z-\bar{\alpha}_2}{1-\alpha_2 z} \cdots \frac{z-\bar{\alpha}_{n-1}}{1-\alpha_{n-1} z} \frac{\lambda_n}{1-\alpha_n z} = 0.$$

On applying $P_{n-1}P_{n-2} \cdots P_1$ to this identity we deduce that $\lambda_n = 0$. We may then apply $P_{n-2} \cdots P_1$ to infer that $\lambda_{n-1} = 0$. Continuing in this way, we conclude that all $\lambda_i = 0$. Hence $f_1, \dots, f_n$ are linearly independent. ■

We return to the proof of Theorem 1(a). Suppose rank $\mathrm{J} < n$: reference to (5) shows that either $v_1, \dots, v_n$ are linearly dependent or some $|\alpha_j| = 1$. In the latter case p has a conjugate zero, while in the former we have, for some scalars $\lambda_1, \dots, \lambda_n$, not all zero,

$$\lambda_1 f_1(T^*)u + \lambda_2 f_2(T^*)u + \cdots + \lambda_n f_n(T^*)u = 0,$$

in the notation of the Lemma. The force of the Lemma is that, if p has no conjugate zero, the polynomial $q = \lambda_1 f_1 + \cdots + \lambda_n f_n$ is not the zero polynomial, and so we have shown that there is a nonzero polynomial q of degree less than n such that $q(T^*)u = 0$. Thus the linear span K of the vectors $\{T^{*j}u : j \geqslant 0\}$ satisfies $\dim K < n$. Hence $K^\perp$ is a nonzero subspace of $\mathbf{C}^m$, and since $T^*K \subseteq K$, $TK^\perp \subseteq K^\perp$. Now T is isometric on $K^\perp$, for $K^\perp \subseteq \{u\}^\perp$ and, for any $x \in \{u\}^\perp$,

$$\|x\|^2 - \|Tx\|^2 = x^*(I - T^*T)x = x^*uu^*x = 0. \qquad (*)$$

We have found a subspace $K^\perp$ of H, invariant under T, of codimension less than n, on which T is isometric. It follows that T has an eigenvalue of unit modulus, contrary to hypothesis. Thus p must have a conjugate zero.

If the canonical form of J is (3), then J is congruent to $\operatorname{diag}\{1,1,\ldots,1,-1,\ldots,-1\}$ (r pluses and $n-r$ minuses). On the other hand, (5) can be written

$$J=PDP^*,$$

where $P=[v_1 \ \cdots \ v_n]$ and $D=\operatorname{diag}\{1-|\alpha_1|^2,\ldots,1-|\alpha_n|^2\}$. It follows from Sylvester's law of inertia that $1-|\alpha_j|^2$ is positive for r j's and negative for $n-r$ j's, as required.

(b): We can suppose that the α's are arranged so that

(i) $(\alpha_1,\alpha_2),\ldots,(\alpha_{2k-1},\alpha_{2k})$ are all the pairs of distinct conjugate zeros of p;

(ii) $\alpha_{2k+1},\ldots,\alpha_{2k+t}$ are the zeros of p which lie on the unit circle;

(iii) $\alpha_{2k+t+1},\ldots,\alpha_{2k+t+r}$ are the nonconjugate zeros in U;

(iv) $\alpha_{2k+t+r+1},\ldots,\alpha_{2k+t+r+s}$ are the nonconjugate zeros in V.

Here $2k+t+r+s=n$, but any of k,t,r,s can be zero. Let

$$f(z)=(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_{2k+t}),$$

$$g(z)=(z-\alpha_{2k+t+1})\cdots(z-\alpha_n),$$

so that f,g contain respectively the conjugate and nonconjugate zeros of p, and $p=fg$. Thus $p_0=f_0g_0$.

Now, if $\alpha\bar{\beta}=1$,

$$((z-\alpha)(z-\beta))_0=(1-\bar{\alpha}z)(1-\bar{\beta}z)=(\alpha\beta)^{-1}(z-\alpha)(z-\beta),$$

and if $|\alpha|=1$,

$$(z-\alpha)_0=1-\bar{\alpha}z=-\bar{\alpha}(z-\alpha).$$

Combining such identities, we find that $f_0=\lambda f$ for some $\lambda\in\mathbf{C}$, $|\lambda|=1$. On the other hand the zeros of g_0, being the conjugates with respect to the unit circle of the zeros of g, are distinct from all the α_i—that is, g and g_0 are relatively prime. Thus $p_0=f_0g_0=\lambda fg_0$, $p=fg$ with $(g,g_0)=1$. It follows that $f=(p,p_0)$. By hypothesis, $f(T)$ is nonsingular. We have, further,

$$\begin{aligned}J&=p_0(T)^*p_0(T)-p(T)^*p(T)\\&=\lambda\bar{\lambda}f(T)^*g_0(T)^*g_0(T)^*g_0(T)f(T)-f(T)^*g(T)f(T)\\&=f(T)^*[g_0(T)^*g_0(T)-g(T)^*g(T)]f(T).\end{aligned}$$

Thus J is congruent to $g_0(T)^*g_0(T)-g(T)^*g(T)$, and so has the same numbers of plus and minus signs in its canonical form as the latter matrix. Part (a) applied to g shows that these numbers are r and s respectively. ∎

Theorem 1 tells us that a failure of the Schur-Cohn test (that is, a deficiency in the rank of the Hermitian form x^*Hx) can be caused either by zeros on the unit circle or by pairs of distinct conjugate zeros, but that the test itself cannot distinguish between these two types of singularity. In the worst case, when $p_0=\lambda p$ for some scalar λ of unit modulus, $H=0$ and the test gives no information at all. However, we can still determine the zero distribution of p—by applying the Schur-Cohn test to a derivative.

THEOREM 2. *Let p be a polynomial without repeated zeros, and suppose that $p_0=\lambda p$ where $\lambda\in\mathbf{C}\setminus\{0\}$. The number of zeros of p in U equals the number of zeros of p' in V, and all zeros of p' are nonconjugate.*

Note that necessarily $|\lambda|=1$, as follows from the fact that, whenever $|z|=1$, $\lambda p(z)=p_0(z)=z^np(z)^-$.

We shall deduce Theorem 2 from Theorem 1 and the following.

LEMMA 2. *Let p be as in Theorem 2 and have degree n. Let $h(z)=zp'(z)$, let T be any $n\times n$ matrix such that $I-T^*T$ is positive semidefinite and of rank one, let T have no eigenvalue of unit modulus, and let*

$$K=h_0(T)^*h_0(T)-h(T)^*h(T).$$

Then K is nonsingular, and the number of zeros of p in U is $n-\pi(K)$.

Proof. Let k be the number of zeros of p in U. Let $p_r(z)=p(rz)$, $r\in\mathbf{R}$. The zeros of p_r ($r\neq 0$) are the numbers w/r where w is a zero of p. It follows that there exists $\varepsilon>0$ such that whenever $1-\varepsilon<r<1$, p_r has k zeros in U and no conjugate zeros. Hence $(p_r,(p_r)_0)=1$. Thus, by Theorem 1(b) applied to p_r,

$$k=\lim_{r\to 1-}\pi\big(q_r(T)^*q_r(T)-p_r(T)^*p_r(T)\big),$$

where $q_r=(p_r)_0$. Now, since $p_0=\lambda p$,

$$q_r(z)=z^np_r(1/\bar z)^-=r^n\left(\frac{z}{r}\right)^np(r/\bar z)^-$$

$$=r^np_0(z/r)=\lambda r^np(z/r).$$

Hence, if

$$J(r) \triangleq q_r(T)^* q_r(T) - p_r(T)^* p_r(T),$$

we have

$$J(r) = r^{2n} p(T/r)^* p(T/r) - p(rT)^* p(rT).$$

Now $J(\cdot)$ is a matrix-valued polynomial and so it is differentiable with respect to any norm on the space of $n \times n$ complex matrices, and we have

$$J(r) = J(1) + (r-1)J'(1) + O\big((r-1)^2\big).$$

$J(1) = 0$ and

$$\begin{aligned} J'(r) = {} & 2nr^{2n-1} p(T/r)^* p(T/r) - r^{2n-2} T^* p'(T/r)^* p(T/r) \\ & - r^{2n-2} p(T/r)^* p'(T/r) T - T^* p'(rT)^* p(rT) \\ & - p(rT)^* p'(rT) T. \end{aligned}$$

Hence

$$\begin{aligned} J'(1) &= 2np(T)^* p(T) - 2T^* p'(T)^* p(T) - 2p(T)^* p'(T) T \\ &= \frac{2}{n} \big[(np(T) - p'(T)T)^* (np(T) - p'(T)T) \\ &\quad - T^* p'(T)^* p'(T) T \big]. \end{aligned} \tag{6}$$

Differentiate the relation $p = \bar{\lambda} p_0$, that is

$$p(z) = \bar{\lambda} z^n p(1/\bar{z})^-,$$

to obtain

$$p'(z) = \bar{\lambda} n z^{n-1} p(1/\bar{z})^- - \bar{\lambda} z^{n-2} p'(1/\bar{z})^-.$$

Multiply by z to get

$$h(z) = \bar{\lambda}\{ np_0(z) - (p')_0(z) \},$$

from which one easily deduces that

$$h_0(z)=\lambda\{np(z)-zp'(z)\}.$$

Thus (6) can be written

$$J'(1)=\frac{2}{n}[h_0(T)^*h_0(T)-h(T)^*h(T)]$$
$$=2K/n.$$

Since p has no repeated zeros, p and p' are relatively prime, and hence $zp'(z)$ and $np(z)-zp'(z)$ are relatively prime (the relation $p_0=\lambda p$ implies that z is not a divisor of p). That is, $(h, h_0)=1$, and so h has no conjugate zeros. It follows from Theorem 1(a) that K has rank n and hence is nonsingular.

We now have

$$J(r)=(r-1)\frac{2K}{n}+O((r-1)^2),$$

where K is a nonsingular Hermitian matrix. It is easily seen by diagonalizing K that

$$k=\lim_{r\to 1-}\pi(J(r))=\pi(-K)=n-\pi(K). \qquad \blacksquare$$

Proof of Theorem 2. We have seen in the proof of Lemma 2 that K is nonsingular. Thus, by Theorem 1, h has $\pi(K)$ zeros in U and $n-\pi(K)$ in V, all nonconjugate. Now $h(z)=zp'(z)$, so h and p' have the same number of zeros in V, viz. $n-\pi(K)$, which, by the Lemma, is the number of zeros of p in U.

Remark. p also has $n-\pi(K)$ zeros in V and $n-2[n-\pi(K)]=2\pi(K)-n$ zeros on the unit circle.

Theorems 1 and 2 enable us to determine the zero distribution of any polynomial p with respect to the unit circle. Given p, we can calculate $f=(p, p_0)$ and $g=p/f$; then $(g, g_0)=1$, so that, by theorem 1(b), the Schur-Cohn test will succeed for g. And $f_0=\lambda f$ for some λ, $|\lambda|=1$. f may have repeated zeros, but we can write $f=f_1\cdots f_k$ where each f_i has only simple zeros and $(f_i)_0 = \lambda_i f_i$, $|\lambda_i| = 1$: let $f_1 = f/(f, f')$, $g_1 = (f, f')$, $f_2 = g_1/(g_1, g_1')$, etc. Each f_i has only simple conjugate zeros, and so the zero distribution of each f_i can be found by applying the Schur-Cohn test to f_i'.

We now turn to the case of a general half plane or disc. Consider the Hermitian 2×2 matrix

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{10} \\ \gamma_{01} & \gamma_{00} \end{bmatrix}, \tag{7}$$

and write

$$\begin{aligned} \Gamma(z, w) &= [w \quad 1] \Gamma \begin{bmatrix} z \\ 1 \end{bmatrix} \\ &= \gamma_{00} + \gamma_{01} z + \gamma_{10} w + \gamma_{11} z w. \end{aligned}$$

Let

$$\Omega_\Gamma = \{ z \in \mathbf{C} : \Gamma(z, \bar{z}) > 0 \}.$$

This region is only of interest if Γ has rank 2 and signature zero, which we shall therefore assume it to have. Ω_Γ is then a half plane or circle, and all such are of the form Ω_Γ for some Γ. For any $z \in \mathbf{C} \cup \{\infty\}$ the *conjugate of z with respect to* Γ is defined to be the unique number $z^\Gamma \in \mathbf{C} \cup \{\infty\}$ satisfying $\Gamma(z^\Gamma, \bar{z}) = 0$; thus

$$z^\Gamma = -\frac{\gamma_{00} + \gamma_{10} \bar{z}}{\gamma_{01} + \gamma_{11} \bar{z}}.$$

If p is a polynomial of degree n, we define the conjugate of p with respect to Γ, denoted by p^Γ, by

$$p^\Gamma(z) = (-\bar{\gamma}_{01} - \bar{\gamma}_{11} z)^n p(z^\Gamma)^- = (-\gamma_{10} - \gamma_{11} z)^n p(z^\Gamma)^-.$$

Clearly p^Γ is again a polynomial of degree n. We say that α is a *conjugate zero of p with respect to* Γ if $p(\alpha) = 0 = p(\alpha^\Gamma)$.

THEOREM 3. *Let p be a polynomial of degree n, and let Γ be a nonsingular Hermitian 2×2 matrix of zero signature. Let R be an $m \times m$*

matrix, $m \geqslant n$, having no eigenvalue on the boundary of Ω_Γ, such that

$$\gamma_{00} I + \gamma_{01} R + \gamma_{10} R^* + \gamma_{11} R^* R \tag{8}$$

is positive semidefinite and of rank one, where Γ is given by (7).

(a) *The Hermitian form*

$$\mathsf{K}(x) = \| p^\Gamma(R) x \|^2 - |\det \Gamma|^n \| p(R) x \|^2 \tag{9}$$

on $\mathbf{C}^m$ has rank at most n; K has rank n if and only if p has no conjugate zeros with respect to Γ. If the canonical form of K is

$$|y_1|^2 + |y_2|^2 + \cdots + |y_r|^2 - |y_{r+1}|^2 - \cdots - |y_n|^2,$$

then p has r zeros in Ω_Γ and $n - r$ zeros in the complement of the closure of Ω_Γ.

(b) *If $(p, p^\Gamma)(R)$ is nonsingular, the canonical form of K is*

$$|y_1|^2 + \cdots + |y_r|^2 - |y_{r+1}|^2 - \cdots - |y_{r+s}|^2$$

where r, s are the numbers of nonconjugate zeros of p in Ω_Γ and the complement of the closure Ω_Γ respectively.

Proof. Choose a matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad ad - bc \neq 0,$$

such that

$$\Gamma = M^* \operatorname{diag}\{-1, 1\} M. \tag{10}$$

Then $z \in \Omega_\Gamma$ if and only if

$$\left(M \begin{bmatrix} z \\ 1 \end{bmatrix} \right)^* \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} M \begin{bmatrix} z \\ 1 \end{bmatrix} > 0;$$

that is,

$$|cz + d|^2 - |az + b|^2 > 0,$$

or equivalently, if and only if $|\phi(z)|<1$, where $\phi(z)=(az+b)/(cz+d)$. Thus z is a zero of p in Ω_Γ if and only if $w=\phi(z)$ is a zero in U of the polynomial

$$f(w)=(-cw+a)^n p\circ\phi^{-1}(w).$$

Moreover, since the linear fractional transformation ϕ maps conjugates with respect to Γ onto conjugates with respect to $\phi(\Omega_\Gamma)=U$, ϕ maps conjugate zeros of p with respect to Γ onto conjugate zeros of f in the sense of Theorem 1.

Let $T=\phi(R)$. We can suppose that this is defined (i.e. $c\lambda+d\neq 0$ for each eigenvalue λ of R), since we can replace M by

$$\left(1-|\alpha|^2\right)^{-1/2}\begin{bmatrix} 1 & \bar{\alpha} \\ \alpha & 1 \end{bmatrix} M$$

for any α, $|\alpha|<1$, without destroying (10). The matrix (8) can be written

$$\begin{aligned}
&\gamma_{00}I+\gamma_{01}R+\gamma_{10}R^*+\gamma_{11}R^*R\\
&\quad=(cR+dI)^*(cR+dI)-(aR+bI)^*(aR+bI)\\
&\quad=(cR+dI)^*(I-T^*T)(cR+dI),
\end{aligned}$$

and hence T satisfies the hypotheses of Theorem 1, which may therefore be applied to f. We consider

$$\mathsf{J}(x)=\|f_0(T)x\|^2-\|f(T)x\|^2$$

where

$$\begin{aligned}
f_0(w)&=w^n f(1/\bar{w})^-\\
&=w^n\left(-\frac{\bar{c}}{w}+\bar{a}\right)^n p\circ\phi^{-1}\left(\frac{1}{\bar{w}}\right)^-\\
&=(-\bar{c}+\bar{a}w)^n p\left(\frac{d-b\bar{w}}{-c+a\bar{w}}\right)^-.
\end{aligned}$$

Now (10) is equivalent to

$$\gamma_{11}=-\bar{a}a+\bar{c}c,\qquad \gamma_{10}=-\bar{a}b+\bar{c}d,$$

$$\gamma_{01}=-\bar{b}a+\bar{d}c,\qquad \gamma_{00}=-\bar{b}b+\bar{d}d.$$

Thus

$$
\begin{aligned}
-\bar{c}+\bar{a}\phi(z) &= \frac{-\bar{c}(cz+d)+\bar{a}(az+b)}{cz+d} \\
&= \frac{(\bar{a}a-\bar{c}c)z+\bar{a}b-\bar{c}d}{cz+d} \\
&= \frac{-\gamma_{10}-\gamma_{11}z}{cz+d}.
\end{aligned}
$$

Note that $\phi(z^\Gamma)$ and $\phi(z)$ are conjugates with respect to U: that is, $z^\Gamma = \phi^{-1}(1/\phi(z)^-)$. Hence

$$
\begin{aligned}
f_0\circ\phi(z) &= [-\bar{c}+\bar{a}\phi(z)]^n p\circ\phi^{-1}(1/\phi(z)^-)^- \\
&= (cz+d)^{-n}(-\gamma_{10}-\gamma_{11}z)^n p(z^\Gamma)^- \\
&= (cz+d)^{-n} p^\Gamma(z) \qquad (11)
\end{aligned}
$$

and

$$
\begin{aligned}
f\circ\phi(z) &= [-c\phi(z)+a]^n p\circ\phi^{-1}\circ\phi(z) \\
&= (cz+d)^{-n}(ad-bc)^n p(z). \qquad (12)
\end{aligned}
$$

Thus

$$
\begin{aligned}
\mathrm{J}(x) &= \|f_0(T)x\|^2 - \|f(T)x\|^2 \\
&= \|f_0\circ\phi(R)x\|^2 - \|f\circ\phi(R)x\|^2 \\
&= \|p^\Gamma(R)(cR+dI)^{-n}x\|^2 \\
&\quad -|ad-bc|^{2n}\|p(R)(cR+dI)^{-n}x\|^2,
\end{aligned}
$$

and since we have, on taking determinants in (10),

$$
\det\Gamma = -|\det M|^2 = -|ad-bc|^2,
$$

comparison with (9) shows that J and K are congruent Hermitian forms. Thus

$$\begin{aligned} \operatorname{rank} \mathsf{K} = n \quad &\Leftrightarrow \quad \operatorname{rank} \mathsf{J} = n \\ &\Leftrightarrow \quad f \text{ has no conjugate zeros} \\ &\Leftrightarrow \quad p \text{ has no conjugate zeros with respect to } \Gamma. \end{aligned}$$

The remainder of the theorem follows similarly. ■

The use of linear fractional transformations also enables us to generalize Theorem 2. Let us denote by Ψ_Γ the complement of the closure of Ω_Γ.

THEOREM 4. *Let p be a polynomial of degree n, let Γ be a nonsingular Hermitian 2×2 matrix of signature zero, and suppose that $p^\Gamma = \lambda p$ for some $\lambda \in \mathbf{C}\setminus\{0\}$.*

(i) *If Ω_Γ is a circular disc, then the number of zeros of p in Ω_Γ is equal to the number of zeros of p' in Ψ_Γ.*

(ii) *If Ω_Γ is a half plane and $\mu \in \Psi_\Gamma$, then the number of zeros of p in Ω_Γ is equal to the number of zeros of $(z-\mu)p'(z) - np(z)$ in Ψ_Γ.*

Proof. We retain the notation of the proof of Theorem 3. It is clear from (11) and (12) that f_0 is a nonzero scalar multiple of f, so that we may apply Theorem 2 to deduce that f, f' have the same numbers of zeros in U,V respectively. It follows that the polynomials $(cz+d)^n f\circ\phi(z)$, $(cz+d)^{n-1}f'\circ\phi(z)$ have the same numbers of zeros in $\Omega_\Gamma, \Psi_\Gamma$ respectively. By (12) the former is $p(z)$, up to a constant multiple, while differentiation of (12) gives

$$\begin{aligned} &(f'\circ\phi(z))\frac{ad-bc}{(cz+d)^2} \\ &\qquad = (ad-bc)^n\left[(cz+d)^{-n}p'(z) - nc(cz+d)^{-n-1}p(z)\right] \end{aligned}$$

and hence

$$(cz+d)^{n-1}f'\circ\phi(z) = (ad-bc)^{n-1}\left[(cz+d)p'(z) - ncp(z)\right].$$

Thus p has the same number of zeros in Ω_Γ as $(cz+d)p'(z) - ncp(z)$ does in Ψ_Γ. Now if Ω_Γ is a circular disc, we can choose M so that $c=0$, $d\neq 0$, which yields statement (i). If Ω_Γ is a half plane we can choose ϕ mapping Ω_Γ onto U to take $\mu\in\Psi_\Gamma$ onto ∞ (or equivalently to take $\mu^\Gamma\in\Omega_\Gamma$ onto 0). That is, $c\mu+d=0$, and clearly $c\neq 0$ [else $\phi(\Omega_\Gamma)$ would be a half plane]. Thus $(cz+d)p'(z)-ncp(z)=c[(z-\mu)p'(z)-np(z)]$, and (ii) is established. ■

We note further that the proof shows (in conjunction with Theorem 2) that all zeros of p' (in the case of a circular disc) or $(z-\mu)p'(z)-np(z)$ (in the case of a half plane) are nonconjugate with respect to Γ.

Finally let us observe that the class of Hermitian forms with the Schur-Cohn property may be further enlarged by allowing the operator T to act on an infinite-dimensional Hilbert space; this requires a slight modification of the condition on the spectrum of T.

Thus the conclusion of Theorem 1 remains valid if T is an operator on an arbitrary (not necessarily finite-dimensional) Hilbert space of dimension at least n provided T satisfies the following three conditions:

(i) $1-T^*T\geqslant 0$,
(ii) $\operatorname{rank}(1-T^*T)=1$,
(iii) T is not isometric on any invariant subspace of finite codimension.

The relevance of condition (iii) is apparent from the sentence following Equation(*) in the proof of Theorem 1(a). Let us remark that condition (iii) is satisfied if the spectrum of T is disjoint from the unit circle.

Similar comments apply to Theorems 3 and 4 with appropriate modifications.

REFERENCES

1 A. Cohn, Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise, *Math. Z.* 14:110–148 (1922).
2 F. R. Gantmacher, *Applications of the Theory of Matrices*, Interscience, New York, 1959.
3 C. Hermite, Sur le nombre des racines d'une équation algébrique comprises entre des limites données, *J. Reine Angew. Math.* 52:39–51 (1856); English transl. by P. C. Parks, *Internat. J. Control* 26:183 (1977).
4 M. Marden, *The Geometry of the Zeros of a Polynomial in a Complex Variable*, A.M.S. Math. Surveys, No. 3, New York, 1949.
5 Vlastimil Pták and N. J. Young, A generalization of the zero location theorem of Schur and Cohn, *IEEE Trans. Automatic Control* AC25:978–980 (1980).
6 N. J. Young, An identity which implies Cohn's theorem on the zeros of a polynomial, *J. Math. Anal. Appl.* 70:240–248 (1979).

Received 1 October 1980; revised 18 June 1981