Classification of Hereditary Matrices*

Jim Agler, J. William Helton, and Mark Stankus

Department of Mathematics
University of California
San Diego, California

ABSTRACT

A classical approach used to obtain basic facts in the theory of square matrices involves an analysis of the relationship between polynomials \( p \) in one variable and square matrices \( T \) such that \( p(T) = 0 \). We consider matrices and operators which satisfy a different type of polynomial constraint. Let \( \mathcal{H} \) be a complex Hilbert space, \( T \) be a bounded linear transformation of \( \mathcal{H} \), \( T^* \) be the adjoint of \( T \), and \( \mathbb{C}[x, y] \) be the algebra of polynomials in \( x \) and \( y \) with complex coefficients. For a polynomial \( p \in \mathbb{C}[x, y] \) in two variables with complex coefficients, define \( p(T) := \sum_{m, n \geq 0} p^n(m, n)T^*T^n \), where \( p^n(m, n) \) is the coefficient of \( y^nx^m \) in the expansion of \( p \) in a power series about the point \((0,0)\). \( T \) is called a root of \( p \) if and only if \( p(T) = 0 \). Note that if \( p \in \mathbb{C}[x, y] \) is a polynomial in the single variable \( x \), then the definition of \( p(T) \) given here agrees with the classical definition. In this paper, we study the relationships which \( p(T) = 0 \) forces between \( p \) and \( T \) when \( T \) is an algebraic operator (i.e., there exists \( n \geq 1 \) and complex numbers \( a_0, \ldots, a_{n-1} \) such that \( 0 = a_0 + a_1T + \cdots + a_{n-1}T^{n-1} + T^n \)). The classification starts with the following observation: Suppose \( p \in \mathbb{C}[x, y] \) and an algebraic operator \( T \in \mathcal{L}(\mathcal{H}) \) satisfy \( p(T) = 0 \). Then certain subspaces of \( \mathcal{H} \) which are invariant for \( T \) must be orthogonal or certain coefficients of \( p \) must vanish. This leads to the notions of a graph attached to each \( p \in \mathbb{C}[x, y] \) and a graph attached to each square matrix \( T \). For diagonalizable \( T \), a necessary and sufficient graph theoretic condition for solving \( p(T) = 0 \) is given. For nondiagonalizable \( T \), this condition is necessary, but not sufficient. The use of these graphs does, however, reduce the problem to the problem of solving the

*This work was sponsored by the Air Force Office for Scientific Research and by the National Science Foundation.
equation \( p(T) = 0 \) for \( T \) with exactly one or two eigenvalues. For \( T \) with one eigenvalue, we give a necessary and sufficient condition for solving \( p(T) = 0 \). This leaves the case of solving \( p(T) = 0 \) when \( T \) has exactly two eigenvalues. This problem mixes algebra involving polynomials with matrix theory. We show that it is equivalent to the purely algebraic problem of determining if equations of the form

\[
\sum_{(i, j) \in E} c_{i, j} X_{i+r, j+s} = 0
\]

have solutions of finite support with certain nonvanishing properties. We call these equations bi-Hankel equations subordinate to a given subset \( E \) of the lattice of integer pairs \( (i, j) : 0 \leq i \leq m - 1, 0 \leq j \leq n - 1 \). It turns out that there is an algorithm (which uses Gröbner bases) for determining if the type of solution we seek exists and for computing it.

1. ROOTS

Let \( \mathcal{H} \) be a complex Hilbert space, let \( \mathcal{L}(\mathcal{H}) \) be the collection of bounded linear operators on \( \mathcal{H} \), and let \( \mathbb{C}[x, y] \) be the algebra of polynomials in \( x \) and \( y \) with complex coefficients. If \( T \in \mathcal{L}(\mathcal{H}) \) and \( p \in \mathbb{C}[x, y] \), then we define \( p(T) \in \mathcal{L}(\mathcal{H}) \) by

\[
p(T) = \sum_{m, n \geq 0} p^\wedge(m, n) T^m \star T^n,
\]

where \( p^\wedge(m, n) \) is the coefficient of \( y^m x^n \) in the expansion of \( p \) in a power series about the point \((0, 0)\). This functional calculus is termed the hereditary functional calculus in [1], and its properties are described in [1], [6], and [39].

If \( p \in \mathbb{C}[x, y] \) and \( T \in \mathcal{L}(\mathcal{H}) \), then we say \( T \) is a root of \( p \) if \( p(T) = 0 \). The study of roots appears in the literature [1, 3, 6, 9–11, 16, 25, 27, 28, 31, 35–37, 39]. A small collection of polynomials is specified in each of these papers, and roots of each of those polynomials is studied when \( \mathcal{H} \) is an infinite dimensional Hilbert space. The study of such roots has applications to the study of Sobolev space, the Sturm-Liouville equations, differential equations whose coefficients are distributions, disconjugacy theory for both classical differential equations in the real setting and Toeplitz differential operators in the complex setting, and prediction theory for generalized stationary processes. The theory of finding the solution sets \( \{ T : p(T) = 0 \} \) for generic \( p \) (e.g., \( p \) in a dense set of polynomials) and the theory of finding the solution sets \( \{ p : p(T) = 0 \} \) for generic \( T \) (e.g., \( T \) in a dense set of operators) is very limited. This paper centers on the case when \( \mathcal{H} \) is a finite dimensional Hilbert space (and so \( \mathcal{L}(\mathcal{H}) \) is the set of \( n \times n \) matrices, where \( n = \text{dim} \mathcal{H} \), where general results can be obtained. Even though we concentrate on the case when \( \mathcal{H} \) is finite dimensional, many of our results can be generalized to
infinite dimensional Hilbert spaces if one restricts one's attention to algebraic operators. Since every square matrix is an algebraic operator, we state some of the results in this paper in terms of algebraic operators. We will show in a future paper that every root whose spectrum is finite is, in fact, an algebraic operator.

In this paper, we study roots from two different perspectives: first, given an algebraic operator \( T \), find all \( p \) such that \( T \) is a root of \( p \), and secondly, given a polynomial \( p \), find all algebraic operators \( T \) such that \( T \) is a root of \( p \). One goal of this study will be to lay the foundation for studying nonalgebraic roots.

One property which will be used repeatedly is the fact that if \( \mathcal{H} \) is a Hilbert space, \( p \in \mathbb{C}[x, y] \), \( T \in \mathcal{L}(\mathcal{H}) \), and \( \mathcal{M} \) is a subspace of \( \mathcal{H} \) which is invariant for \( T \), then

\[
(2) \quad p(T \mid \mathcal{M}) = P_{\mathcal{M}} p(T) \mid \mathcal{M},
\]

where \( P_{\mathcal{M}} \) is the orthogonal projection from \( \mathcal{H} \) onto \( \mathcal{M} \). Therefore, if \( T \) is a root of \( p \), then \( T \mid \mathcal{M} \) is a root of \( p \).

We now state a key lemma to solving the equation \( p(T) = 0 \) for \( T \) an algebraic operator and \( p \in \mathbb{C}[x, y] \).

**Lemma 3.** Let \( p \in \mathbb{C}[x, y] \), \( \mathcal{H} \) be a Hilbert space, and \( T \in \mathcal{L}(\mathcal{H}) \). If \( I \) is an index set, \( I \) contains at least two elements, and \( \{\mathcal{H}_i\}_{i \in I} \) is a collection of invariant subspaces of \( T \) which span \( \mathcal{H} \), then

\[
(4) \quad p(T) = 0
\]

if and only if

\[
(5) \quad p\left(T \mid \left(\mathcal{H}_i + \mathcal{H}_j\right)^{-}\right) = 0
\]

for every \( i, j \in I \) such that \( i \neq j \).

**Proof.** If \( p(T) = 0 \) and \( i, j \in I \), then \( \mathcal{H}_i + \mathcal{H}_j \) is invariant for \( T \) and, by (2),

\[
p\left(T \mid \left(\mathcal{H}_i + \mathcal{H}_j\right)^{-}\right) = P_{\left(\mathcal{H}_i + \mathcal{H}_j\right)^{-}} p(T) \mid \left(\mathcal{H}_i + \mathcal{H}_j\right)^{-} = 0,
\]

where \( P_{\left(\mathcal{H}_i + \mathcal{H}_j\right)^{-}} \) is the orthogonal projection from \( \mathcal{H} \) onto \( \left(\mathcal{H}_i + \mathcal{H}_j\right)^{-} \). Therefore, (4) implies (5). On the other hand, suppose that (5) holds. For \( i, j \in I \)
such that \( i \neq j \).

\[
p(T)\mathcal{K}_i \subseteq p(T)(\mathcal{K}_i + \mathcal{K}_j) \subseteq (\mathcal{K}_i + \mathcal{K}_j)^\perp \subseteq \mathcal{K}_i^\perp \cap \mathcal{K}_j^\perp,
\]

since (5) holds. Therefore, since \( \{\mathcal{K}_j\}_{j \in J} \) span \( \mathcal{H} \),

\[
p(T)\mathcal{K}_i \subseteq \bigcap_{j \in J, j \neq i} \mathcal{K}_i^\perp \cap \mathcal{K}_j^\perp = \bigcap_{j \in J} \mathcal{K}_j^\perp = \left( \sum_{j \in J} \mathcal{K}_j \right)^\perp = \mathcal{H}^\perp = \{0\}.
\]

Thus, \( p(T) = 0 \) and so (4) holds. This completes the proof of Lemma 3. ■

Since any algebraic operator \( T \in \mathcal{L}(\mathcal{H}) \) has the property that there exists a collection of finite dimensional subspaces \( \{\mathcal{K}_i\} \) such that \( \{\mathcal{K}_i\} \) has dense linear span in \( \mathcal{H} \) and \( T\mathcal{K}_i \subseteq \mathcal{K}_i \) for each \( i \), it is clear that Lemma 3 immediately implies a description of the algebraic roots of \( p \) in terms of a description of the matrix roots of \( p \).

1.1. **Symmetry**

For \( p \in \mathbb{C}[x, y] \), let \( p^\vee \) be the unique element of \( \mathbb{C}[x, y] \) such that

\[
p^\vee(\lambda, \mu) = \overline{p(\bar{\mu}, \bar{\lambda})}
\]

(equivalently, \( (p^\vee)^\wedge(m, n) = \overline{p(n, m)} \) \( m \geq 0 \) and \( n \geq 0 \)).

We now make several key observations involving the \( \vee \) operation. Let \( \mathcal{H} \) be a Hilbert space, \( T \in \mathcal{L}(\mathcal{H}) \) and \( p \in \mathbb{C}[x, y] \). Note that

\[
(6) \hspace{1cm} (p^\vee)^\vee = p,
\]

\[
(7) \hspace{1cm} p(T) = 0 \hspace{0.5cm} \text{if and only if} \hspace{0.5cm} p^\vee(T) = 0,
\]
(8)

\[ p(T) = 0 \quad \text{if and only if} \quad \left( \frac{p + p^\vee}{2} \right)(T) = 0 \quad \text{and} \quad \left( \frac{p - p^\vee}{2i} \right)(T) = 0. \]

Since

(9)

\[ \left( \frac{p + p^\vee}{2} \right)^\vee = \frac{p + p^\vee}{2} \quad \text{and} \quad \left( \frac{p - p^\vee}{2i} \right)^\vee = \frac{p - p^\vee}{2i}. \]

(8) shows that classifying the algebraic roots of \( p \in \mathbb{C}[x, y] \) such that \( p \neq p^\vee \) is equivalent to classifying \( T \in \mathcal{L}(\mathbb{F}) \) which are roots of both \( p_1 \) and \( p_2 \) for some polynomials \( p_1 \in \mathbb{C}[x, y] \) and \( p_2 \in \mathbb{C}[x, y] \) such that \( p_1^\vee = p_1 \) and \( p_2^\vee = p_2 \). For this reason, we shall always restrict our attention to polynomials \( p \in \mathbb{C}[x, y] \) such that \( p = p^\vee \). Note that since, in general, \( p(T)^* = p^\vee(T) \), this assumption implies that \( p(T) \) is self-adjoint for all operators \( T \).

1.2. Graphs

By a directed graph we shall mean an ordered pair \( G = (V, E) \) where \( E \subseteq V \times V \). The elements of \( V \) will be referred to as the vertices of \( G \), and the elements of \( E \) will be referred to as the edges of \( G \). In this paper all graphs will be undirected i.e., \((v_1, v_2) \in E \) if and only if \((v_2, v_1) \in E \).

To see how Lemma 3 applies to the study of the equation \( p(T) = 0 \), we begin by associating a graph \( G_p \) to every \( p \in \mathbb{C}[x, y] \) and a graph \( G_T \) to every algebraic operator \( T \).

**Definition 10.** For \( p \in \mathbb{C}[x, y] \), let \( G_p \) denote the graph whose set of vertices \( V \) is

\[ \{ \lambda \in \mathbb{C} : p(\lambda, \overline{\lambda}) = 0 \} \]

and such that \((\lambda, \overline{\mu})\) is an edge of \( G \) if \( p(\lambda, \overline{\mu}) = 0 \).

To verify that \( G_p \) is an undirected graph, note that since \( p = p^\vee \), \( p(\lambda, \overline{\mu}) = 0 \) if and only if \( p(\mu, \overline{\lambda}) = 0 \).

In order to define \( G_T \) for an algebraic operator \( T \), we begin by associating a particular invariant subspace of \( T \) to each eigenvalue \( \lambda \) of \( T \). The spectral space for \( T \) at \( \lambda \) is defined to be \( \ker (T - \lambda)^N \), where \( N \) is a sufficiently large
positive integer that for all \( n \geq N \), \( \ker (T - \lambda)^n = \ker (T - \lambda)^N \). The spectral space for \( T \) at \( \lambda \) will be denoted by \( \mathcal{H}_\lambda(t) \). Before continuing, let us note that such an \( N \) exists because \( T \) was assumed to be an algebraic operator. Also note that

\[
(11) \quad \mathcal{H}_\lambda(T) \text{ is an invariant subspace for } T,
\]

\[
(12) \quad T \mid \mathcal{H}_\lambda(T) \text{ has exactly one eigenvalue,}
\]

and

\[
(13) \quad \mathcal{H}_{\lambda_1}(T) + \mathcal{H}_{\lambda_2}(T) + \cdots + \mathcal{H}_{\lambda_m}(T) = \mathcal{H},
\]

where \( \lambda_1, \ldots, \lambda_m \) are the distinct eigenvalues of \( T \).

**Definition 14.** For \( T \in \mathcal{D}(\mathcal{H}) \), let \( G_T \) denote the undirected graph whose set of vertices \( V \) is the set of eigenvalues of \( T \) and set of edges is

\[
\{(\lambda, \mu) \in V \times V : \mathcal{H}_\lambda(T) \text{ is orthogonal to } \mathcal{H}_\mu(T)\}.
\]

1.3. **Complete Characterization of Diagonalizable Roots**

The following theorem classifies diagonalizable roots. For a graph \( G = (V, E) \) and \( V_0 \subseteq V \), let \( G \mid V_0 = (V_0, E \cap (V_0 \times V_0)) \). For graphs \( G_1 = (V_1, E_1) \) and \( G_2 = (V_2, E_2) \), let \( G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2) \). A graph \( G = (V, E) \) is a clique if \( E = V \times V \).

**Theorem 15.** Let \( p \in C[x, y] \), \( \mathcal{H} \) be a Hilbert space, and \( T \in \mathcal{D}(\mathcal{H}) \) such that \( T \) is algebraic and diagonalizable. \( T \) is a root of \( p \) if and only if

\[
(16) \quad \text{the vertex set of } G_T \text{ is a subset of the vertex set of } G_p
\]

and

\[
(17) \quad (G_p \mid \sigma(T)) \cup G_T \text{ is a clique}.
\]

Before proving Theorem 15, we first state and prove two lemmas which hold for (not necessarily diagonalizable) algebraic roots.

The following lemma extends the classical result that the eigenvalues of a matrix \( T \) are zeros of a polynomial.
Lemma 18. Let $\mathcal{H}$ be a Hilbert space, $T \in \mathfrak{L}(\mathcal{H})$, and $p \in \mathbb{C}[x, y]$. If $p(T) = 0$, then every eigenvalue $\lambda$ of $T$ satisfies $p(\lambda, \overline{\lambda}) = 0$.

Proof. Suppose that $p(T) = 0$, $\lambda \in \mathbb{C}$, $h \in \mathcal{H}$, $h \neq 0$, and $Th = \lambda h$. By (1),

$$
\langle p(T)h, h \rangle = \sum_{m, n} p^\wedge(m, n) \langle T^m T^n h, h \rangle
= \sum_{m, n} p^\wedge(m, n) \langle T^m h, T^n h \rangle
= p(\lambda, \overline{\lambda}) \|h\|^2.
$$

Since $p(T) = 0$ and $h \neq 0$, $p(\lambda, \overline{\lambda}) = 0$. This completes the proof of Lemma 18.

The following lemma shows that if $T \in \mathfrak{L}(\mathcal{H})$ is a root of $p \in \mathbb{C}[x, y]$ and both $\lambda$ and $\mu$ are eigenvalues of $T$, then either $p$ has two zeros of a specific type or two subspaces of $\mathcal{H}$ are orthogonal.

Lemma 19. Let $p \in \mathbb{C}[x, y]$, $\mathcal{H}$ be a Hilbert space, and $T$ be an algebraic operator acting on $\mathcal{H}$. If $p(T) = 0$ and $\lambda, \mu \in \sigma(T)$ such that $\lambda \neq \mu$, then either

$$
p(\lambda, \overline{\mu}) = 0
$$

or

$$
\mathcal{H}_{\lambda}(T) \text{ is orthogonal to } \mathcal{H}_{\mu}(T).
$$

Proof. We argue by proof by contradiction. Thus, suppose that there exist $\lambda, \mu \in \sigma(T)$ such that $\lambda \neq \mu$, $\mathcal{H}_{\lambda}(T)$ is not orthogonal to $\mathcal{H}_{\mu}(T)$. $p(\lambda, \overline{\lambda}) = 0$, $p(\mu, \overline{\mu}) = 0$, and either $p(\lambda, \overline{\mu}) \neq 0$ or $p(\mu, \overline{\lambda}) \neq 0$. Without loss of generality, let us assume that $p(\lambda, \overline{\mu}) \neq 0$. Let $T_0 = T \mid (\mathcal{H}_{\lambda}(T) + \mathcal{H}_{\mu}(T))$. Let $c > 0$ and $d > 0$ be such that $(x - \lambda)^c (x - \mu)^d$ is the minimal polynomial of $T_0$. Let $b < d$ be maximal with respect to the condition that $(T - \mu)^b \mathcal{H}_{\mu}(T)$ is not orthogonal to $\mathcal{H}_{\lambda}(T)$. Let $a < c$ be maximal with respect
to the condition that \((T - \lambda)^a \mathcal{H}_\lambda(T)\) is not orthogonal to \((T - \mu)^b \mathcal{H}_\mu(T)\). With this choice of \(a\) and \(b\), \((T - \mu)^b + 1 \mathcal{H}_\mu(T)\) is orthogonal to \(\mathcal{H}_\lambda(T)\), and \((T - \mu)^b \mathcal{H}_\mu(T)\) is orthogonal to \((T - \lambda)^a + 1 \mathcal{H}_\lambda(T)\). Since \(\mathcal{H}_\lambda(T)\) and \(\mathcal{H}_\mu(T)\) are invariant subspaces for \(T\), we find that if \(r > a\) and \(s > b\) but \((r, s) \neq (a, b)\), then \((T - \lambda)^a \mathcal{H}_\lambda(T)\) is orthogonal to \((T - \mu)^b \mathcal{H}_\mu(T)\) and consequently \(P_{\mathcal{H}_\lambda(T)}(T^* - \overline{\mu})^a (T - \lambda)^b P_{\mathcal{H}_\lambda(T)} = 0\). Therefore, by expanding \(p\) as a power series in \(x - \lambda\) and \(y - \overline{\mu}\), it is easy to see that

\[
p(\lambda, \overline{\mu}) P_{\mathcal{H}_\lambda(T)}(T^* - \overline{\mu})^b (T - \lambda)^a P_{\mathcal{H}_\lambda(T)} = 0.
\]

Since \(p(\lambda, \overline{\mu}) \neq 0\), the above equation shows that \((T - \lambda)^a \mathcal{H}_\lambda(T)\) is orthogonal to \((T - \mu)^b \mathcal{H}_\mu(T)\), contrary to the hypothesis. This completes the proof of Lemma 19.

\[
\square
\]

We now prove Theorem 15.

**Proof of Theorem 15.** Recall that \(T\) is assumed to be diagonalizable. In particular, if \(T\) has only one eigenvalue, \(\lambda\), then \(T = \lambda I_T\), where \(I_T\) denotes the identity on \(\mathcal{H}\), and the theorem follows trivially. Accordingly, suppose for the rest of the proof that \(T\) has at least two eigenvalues.

If \(T\) is a root of \(p\), then Lemmas 18 and 19 show that (16) and (17) hold. This completes the “only if” part of the proof.

Now, suppose that (16) and (17) hold. Let \(\lambda\) and \(\mu\) be two distinct eigenvalues of \(T\). Since \(T\) is diagonalizable, \(\mathcal{H}_\lambda(T) = \ker(T - \lambda)\) and \(\mathcal{H}_\mu(T) = \ker(T - \mu)\). If \(\mathcal{H}_\lambda(T)\) and \(\mathcal{H}_\mu(T)\) are orthogonal, then

\[
p(T \mid (\mathcal{H}_\lambda(T) + \mathcal{H}_\mu(T))) = (p(T \mid \mathcal{H}_\lambda(T))) \oplus p(T \mid \mathcal{H}_\mu(T))
\]

\[
= p(\lambda 1_{\mathcal{H}_\lambda(T)}) \oplus p(\mu 1_{\mathcal{H}_\mu(T)})
\]

\[
= 0.
\]

If \(\mathcal{H}(T)\) and \(\mathcal{H}_\mu(T)\) are not orthogonal, then (17) implies that \(p(\lambda, \overline{\mu}) = 0\) and \(p(\mu, \overline{\lambda}) = 0\). Since \(T \mid (\mathcal{H}_\lambda(T) + \mathcal{H}_\mu(T))\) is unitarily equivalent to
\[
\begin{bmatrix}
\lambda & E \\
0 & \mu
\end{bmatrix}
\] for some nonzero operator \(E\) and

\[
p\left(\begin{bmatrix}
\lambda & E \\
0 & \mu
\end{bmatrix}\right) = \begin{bmatrix}
p(\lambda, \lambda) & \frac{p(\lambda, \lambda) - p(\mu, \lambda)}{\lambda - \mu} \\
\frac{p(\lambda, \lambda) - p(\mu, \lambda)}{\lambda - \mu} & c|\lambda|^2 + p(\mu, \mu)
\end{bmatrix},
\]

where \(c = \frac{p(\lambda, \lambda) - p(\mu, \lambda) - p(\mu, \lambda) + p(\mu, \mu)}{|\lambda - \mu|^2}\)

it follows that \(p(T \mid (\mathcal{H}(T) + \mathcal{K}(T))) = 0\). In either case, Lemma 3 implies that \(p(T) = 0\). This completes the proof of Theorem 15.

\[\blacksquare\]

1.4. Examples

In this subsection we will show how results from the literature for \(n\)-symmetric and \(n\)-isometric matrices can be obtained with the results from Section 1.

An operator \(T\) is called \(n\)-symmetric provided \(T\) is a root of

\[
p(x, y) = (x - y)^n.
\]

A theorem of Ball and Helton [16] says that such matrices have the form \(T = S + N\) when

\(S\) is self-adjoint, \(N\) is nilpotent, and \(SN = NS\).

This was proved using Wiener-Hopf factorization of matrix valued functions. As we shall see, classification follows immediately from results in this paper.

Let \(T\) be \(n\)-symmetric and algebraic. It follows from Lemma 18 that all eigenvalues of \(T\) are real. Since \(\lambda \in \mathbb{R}, \mu \in \mathbb{R}\), and \((\lambda - \mu)^n = 0\) implies \(\lambda = \mu\). Lemma 19 implies that the spectral subspaces of \(T\) are mutually orthogonal and so

\[
T \cong \left( T \mid \mathcal{H}_\lambda(T) \right) \oplus \cdots \oplus \left( T \mid \mathcal{H}_\mu(T) \right),
\]

where \(\lambda_1, \ldots, \lambda_m\) are the distinct eigenvalues of \(T\). For \(1 \leq j \leq m\), since \(\sigma(T \mid \mathcal{H}_\lambda(T)) = \{\lambda_j\}\), \(T \mid \mathcal{H}_\lambda(T)\) has the form \(T \mid \mathcal{H}_\lambda(T) = \lambda_j + N_j\) for some
nilpotent $N$. Therefore, by (21), $T$ has the form $S + N$ for some self-adjoint operator $S$ and some nilpotent operator $N$. We will see that Theorem 27 implies that $N$ is nilpotent of order less than or equal to $n - 1$.

Another class of operators, the $n$-isometries, are defined to be roots of $p(x, y) = (xy - 1)^n$. While infinite dimensional examples include the unilateral shift, which has much different structure than we have seen here, the finite dimensional $n$-isometries have a structure like that of $n$-symmetric matrices. Namely, they equal a unitary plus a nilpotent which commutes with it.

2. THE CLASSIFICATION OF NONDIAGONALIZABLE ROOTS

This section reduces consideration of nondiagonalizable algebraic operators to consideration of algebraic operators with one or two eigenvalues.

Theorem 15 completely analyzes the equation $p(T) = 0$ via a graph theoretic characterization when $p \in \mathbb{C}[x, y]$ and $T$ is diagonalizable. If $T$ is not assumed to be diagonalizable and $p(T) = 0$, then, just as in the proof of Theorem 15, Lemmas 18 and 19 imply that

the vertex set of $G_T$ is a subset of the vertex set of $G_p$

and

$$\left( G_p \mid \sigma(T) \right) \cup G_T$$

is a clique.

To see that the above condition is not sufficient to guarantee that $p(T) = 0$ in the nondiagonalizable case, one can consider $p(x, y) = x^2$ for any nonzero nilpotent $T$ of index 2. Nevertheless, the analysis of nondiagonalizable roots can be reduced to the study of roots which have at most two eigenvalues as seen in the following immediate corollary of Lemma 3.

Corollary 22. Let $\mathcal{H}$ be a Hilbert space, $T$ an algebraic operator acting on $\mathcal{H}$, and $p \in \mathbb{C}[x, y]$. If the eigenvalues of $T$ are $\lambda_1, \ldots, \lambda_m$ and $m \geq 2$, then $T$ is a root of $p$ if and only if $T \restriction (\mathcal{H}_i(T) + \mathcal{H}_j(T))$ is a root of $p$ for every $i$ and $j$ such that $1 \leq i \leq m$ and $1 \leq j \leq m$ and $i \neq j$.

The analysis of roots which have exactly two eigenvalues can be further simplified by a second corollary of Lemma 3. We first give an intrinsic
operator theoretic description of the well-known notion of a Jordan cell. We then introduce a generalization of this notion, the concept of a bi-Jordan cell. Bi-Jordan cells will be used throughout the rest of the paper.

**Definition 23.** Let $\mathcal{H}$ be a finite dimensional Hilbert space. For $m \geq 1$ and $\lambda \in \mathbb{C}$, we say that $T \in \mathcal{L}(\mathcal{H})$ is a Jordan cell of type $(\lambda, m)$ if the dimension of $\mathcal{H}$ is $m$ and the minimal polynomial of $T$ is $(x - \lambda)^m$.

**Definition 24.** Let $\mathcal{H}$ be a finite dimensional Hilbert space. For $m \geq 1$, $n \geq 1$, $\lambda \in \mathbb{C}$, and $\mu \in \mathbb{C}$, we say that $T \in \mathcal{L}(\mathcal{H})$ is a bi-Jordan cell of type $(\lambda, \mu, m, n)$ if $\dim \mathcal{H} = m + n$ and there exist subspaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of $\mathcal{H}$ such that $\mathcal{H}_1$ and $\mathcal{H}_2$ are invariant for $T$, $\dim \mathcal{H}_1 = m$, $\dim \mathcal{H}_2 = n$, $\mathcal{H}_1 + \mathcal{H}_2 = \mathcal{H}$, $T | \mathcal{H}_1$ is a Jordan cell of type $(\lambda, m)$, and $T | \mathcal{H}_2$ is a Jordan cell of type $(\mu, n)$.

Note that if $\lambda \neq \mu$, then $T$ is a bi-Jordan cell of type $(\lambda, \mu, m, n)$ if and only if $(x - \lambda)^m(x - \mu)^n$ is the minimal polynomial of $T$ and the dimension of $\mathcal{H}$ is $m + n$. If $\lambda = \mu$, then $T$ is a bi-Jordan cell of type $(\lambda, \mu, m, n)$ if and only if $T$ is similar to a direct sum of a Jordan cell of type $(\lambda, m)$ and an Jordan cell of type $(\mu, n)$.

**Definition 25.** Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{L}(\mathcal{H})$. If $\mathcal{H}_0$ is a Hilbert space and $T_0 \in \mathcal{L}(\mathcal{H}_0)$, then $T_0$ is a bi-Jordan restriction of $T$ if $T_0$ is a bi-Jordan cell, $\mathcal{H}_0 \subseteq \mathcal{H}$, $\mathcal{H}_0$ is invariant for $T$, and $T_0 = T | \mathcal{H}_0$.

**Corollary 26.** Let $\mathcal{H}$ be a Hilbert space, $T$ be an algebraic operator acting on $\mathcal{H}$ which is not a Jordan cell, and $p \in \mathbb{C}[x, y]$. $T$ is a root of $p$ if and only if every bi-Jordan restriction of $T$ is a root of $p$.

**Proof.** If $T$ is a root of $p$, then every bi-Jordan restriction of $T$ is a root of $p$ by (2).

Now suppose every bi-Jordan restriction of $T$ is a root of $p$. Let $\{\mathcal{H}_i\}_{i \in I}$ denote the collection of all maximal elements of

$$\{\mathcal{M} \subseteq \mathcal{H} : \mathcal{M} \text{ is a closed invariant subspace for } T \text{ and } T | \mathcal{M} \text{ is a Jordan cell}\}.$$

Since $T$ is algebraic, $\mathcal{H}$ spans $\mathcal{H}$, and $T$ is assumed not to be a Jordan cell, $\{\mathcal{H}_i\}$ contains at least two elements. Thus, $\{\mathcal{H}_i\}$ satisfies the hypothesis of Lemma 3. Consequently, since $T | (\mathcal{H}_i + \mathcal{H}_j)$ is a bi-Jordan cell whenever
Lemma 3 implies that $T$ is a root of $p$. This completes the proof of Corollary 26.

We now turn to the classification of roots which have only one or two eigenvalues. Since we analyzed diagonalizable roots in the previous section, we concentrate solely on roots which are nondiagonalizable. Moreover, in lieu of the fact that every square matrix, when viewed abstractly, is an algebraic operator, Corollaries 22 and 26 guarantee that general roots can always be simply decomposed into roots with at most two eigenvalues. The following sections focus on such roots.

If $E \subseteq \mathbb{C}[x, y]$, then we let $\langle E \rangle$ denote the smallest ideal of $\mathbb{C}[x, y]$ which contains $E$.

3. ROOTS WITH ONE EIGENVALUE

3.1. A Complete Classification

The following theorem classifies roots $T$ which have only one eigenvalue.

**Theorem 27.** Let $\mathcal{H}$ be a Hilbert space, $T \in \mathcal{L}(\mathcal{H})$, and $p \in \mathbb{C}[x, y]$. If $\lambda \in \mathbb{C}$ and $T - \lambda$ is nilpotent of order $n$, then

\begin{equation}
(28) \quad p(T) = 0
\end{equation}

if and only if

\begin{equation}
(29) \quad \frac{\partial^{i+j}}{\partial^i y \partial^j x} p(\lambda, \lambda) = 0 \quad \text{for all} \quad 0 \leq i \leq n - 1 \text{ and } 0 \leq j \leq n - 1.
\end{equation}

Furthermore, (29) holds if and only if

\begin{equation}
(30) \quad p \in \langle \{(x - \lambda)^n, (y - \lambda)^n\} \rangle.
\end{equation}
In particular, if \( \lambda \in \mathbb{C} \), \( T - \lambda \) is nilpotent, \( S \in \mathcal{L}(\mathcal{H}) \), and \( \lambda \) is invertible, then \( p(T) = 0 \) if and only if \( p(STS^{-1}) = 0 \).

Proof. By expanding \( p \) in a power series in \( x - \lambda \) and \( y - \bar{\lambda} \), we see that

\[
(31) \quad p(x, y) = \sum_{i \geq 0, j \geq 0} \frac{1}{i!j!} \left( \frac{\partial^{i+j}}{\partial^i y \partial^j x} p \right)(\lambda, \bar{\lambda})(x - \lambda)^i(y - \bar{\lambda})^j.
\]

Now, if (29) holds and \((T - \lambda)^n = 0\), then (31) implies that \( p(T) = 0 \). On the other hand, let us suppose that \( p(T) = 0 \), that \((T - \lambda)^n = 0\), and that (29) does not hold. Let \( i_0 \) be the minimal positive integer such that

\[
(32) \quad \left( \frac{\partial^{i_0+j}}{\partial^j y \partial^{i_0} x} p \right)(\lambda, \bar{\lambda}) \neq 0
\]

for some \( j \leq n - 1 \). Let \( j_0 \) be the minimal positive integer such that

\[
(33) \quad \left( \frac{\partial^{i_0+j_0}}{\partial^{j_0} y \partial^{i_0} x} p \right)(\lambda, \bar{\lambda}) \neq 0.
\]

Now (31) yields

\[
(y - \bar{\lambda})^{n-j_0-1} p(x, y)(x - \lambda)^{i_0-1}
\]

\[
= \frac{1}{i_0!j_0!} \left( \frac{\partial^{i_0+j_0}}{\partial^{j_0} y \partial^{i_0} x} p \right)(\lambda, \bar{\lambda})(x - \lambda)^{n-1}(y - \bar{\lambda})^{n-1}
\]

modulo the ideal generated by \((x - \lambda)^n, (y - \bar{\lambda})^n\). Therefore, since \( p(T) = 0 \), \((T - \lambda)^n = 0 \), and \((T^* - \bar{\lambda})^n = 0 \), we have

\[
\frac{1}{i_0!j_0!} \left( \frac{\partial^{i_0+j_0}}{\partial^{j_0} y \partial^{i_0} x} p \right)(\lambda, \bar{\lambda})(T^* - \bar{\lambda})^{n-1}(T - \lambda)^{n-1} = 0.
\]
By (33), we conclude that \((T - \lambda)^{n-1} = 0\). But this contradicts our assumption that \(T - \lambda\) is nilpotent of order \(n\). In view of (31), (29) holds if and only if (30) holds. This completes the proof of Theorem 27. 

3.2. Examples

We now show how results from the literature for isosymmetric matrices can be obtained with the results from Sections 1 and 3. See Section 1.4 for examples involving the techniques in Section 1.

An isosymmetry is defined to be a root \(T\) of a polynomial of the form

\[ p(x, y) = (x - y)(xy - 1). \]

Let \(T\) be an isosymmetric matrix. By Lemma 18, each eigenvalue of \(T\) lies on the set \( \mathbb{R} \cup \partial D \), where \( \partial D = \{ z \in \mathbb{C} : |z| = 1 \} \).

If \(\lambda\) and \(\mu\) are vertices of the graph \(G_p\) and \((\lambda, \mu)\) is an edge of \(G_p\), then \(\lambda, \mu \in \mathbb{R} \cup \partial D\) and

\[ (\lambda \mu - 1)(\mu - \lambda) = 0. \]

The above equation holds for \(\lambda, \mu \in \mathbb{R} \cup \partial D\) if and only if either \(\lambda = \mu\) or one of the following two conditions holds:

1. \(\lambda \in \mathbb{R}, \mu \in \mathbb{R}\), and \(\lambda \mu = 1\);
2. \(\lambda \in \partial D, \mu \in \partial D\), and \(\lambda \mu = 1\).

Therefore, a connected component of the graph \(G_T\) consists of at most two points. Therefore, Theorem 28 implies that \(T \upharpoonright \mathcal{H}(T) = \lambda 1_{\mathcal{H}(T)}\) for every eigenvalue \(\lambda\) of \(T\) such that \(\lambda \neq 1\) and \(\lambda \neq -1\). Also, if \(\lambda\) is an eigenvalue of \(T\) and either \(\lambda = 1\) or \(\lambda = -1\), then Theorem 27 implies that \((T - \lambda)(T)^2 = 0\). We have shown the following theorem.

**Theorem 34.** If \(T\) is an isosymmetry and \(T\) is algebraic, then there exist Hilbert spaces \(\mathcal{H}, \mathcal{K}, \mathcal{M}\), and \(\mathcal{N}\) and operators

\[ A, B \in \mathcal{L}(\mathcal{H}), \quad U, C \in \mathcal{L}(\mathcal{K}), \quad A_0 \in \mathcal{L}(\mathcal{M}), \quad \text{and} \quad U_0 \in \mathcal{L}(\mathcal{N}) \]

such that \(A, A_0\) are self-adjoint, \(U, U_0\) are unitary, \(B\) commutes with \(A, C\) commutes with \(U\), and \(T\) is (unitarily equivalent) to

\[
\begin{bmatrix}
A & B \\
0 & A^{-1}
\end{bmatrix} \oplus
\begin{bmatrix}
U & C \\
0 & U^{-1}
\end{bmatrix} \oplus A_0 \oplus U_0.
\]
B and C can be taken to be injective and positive. Any of the summands in the above expression may be absent. (That is, the corresponding Hilbert space may be \{0\}.)

This result, and the more general case for isosymmetries on an infinite dimensional Hilbert space, is contained in [39].

4. ROOTS WITH TWO EIGENVALUES

The classification of roots with two eigenvalues is much more difficult than the classification of roots with exactly one eigenvalue. The next theorem splits this classification problem into two subproblems.

The first subproblem is to relate the orthogonality conditions given in Theorem 36 below to particular properties of polynomials. This leads to an assignment of an \(m \times n\) grid \(G_{p,T}\) of symbols to each pair \((p, T)\), which may be viewed as a refinement of the graphs considered in Section 1. These symbols encode information about the angles between the generalized eigenvectors of \(T\) and the ideal generated by \(p\).

The second subproblem is to present a canonical set of generators for the intersection of the ideals given in (37) and (38) of the next theorem. This is done in Proposition 86.

THEOREM 36. Let \(\lambda\) and \(\mu\) be distinct complex numbers, \(m \geq 1\), \(n \geq 1\), \(\mathcal{H}\) be a Hilbert space, \(T\) be an algebraic operator acting on \(\mathcal{H}\) whose (traditional) minimal polynomial is \((x - \lambda)^m(x - \mu)^n\), and \(p \in \mathbb{C}[x, y]\). The operator \(T\) is a root of \(p\) if and only if the following three conditions hold:

\[ p \in \left\langle \left\{(x - \lambda)^m, (y - \bar{\lambda})^m\right\} \right\rangle, \quad (37) \]

\[ p \in \left\langle \left\{(x - \mu)^n, (y - \bar{\mu})^n\right\} \right\rangle, \quad (38) \]

and

\[ p(T)\mathcal{H}(T) \text{ is orthogonal to } \mathcal{H}_\mu(T) \quad (39) \]

Proof. If \(T\) is a root of \(p\), then, by (2) and Theorem 27, we obtain (37), (38), and (39). Recall our standing assumption that \(p = p^\vee\) so that \(p(T)\) is self-adjoint. Now if (37), (38), and (39) hold, then whenever \(u_1, u_2 \in \mathcal{H}(T)\)
and $v_1, v_2 \in \mathcal{H}_\mu(T)$,

$$\langle p(T)(u_1 + v_1), (u_2 + v_2) \rangle = \langle p(T)u_1, u_2 \rangle + \langle p(T)u_1, v_2 \rangle + \langle p(T)v_1, u_2 \rangle + \langle p(T)v_1, v_2 \rangle = \langle p(T|\mathcal{H}_\lambda(T))u_1, u_2 \rangle + \langle p(T)u_1, v_2 \rangle + \langle p(T)v_1, u_2 \rangle + \langle p(T|\mathcal{H}_\mu(T))v_1, v_2 \rangle = 0.$$ 

Since $\mathcal{H}_\lambda(T) + \mathcal{H}_\mu(T) = \mathcal{H}$, this computation shows that $p(T) = 0$. This completes the proof of Theorem 36. \hfill \blacksquare

4.1. \textit{0-Patterns, }$\perp$\textit{-Patterns, and Grids}

In view of Theorem 36 and Corollary 26, the description of roots has been reduced to the analysis of the condition (39) when $T$ is a bi-Jordan cell. We now give the computations which motivate the upcoming definitions of 0-pattern, $\perp$-pattern, and grid.

As in the hypothesis of Theorem 36, let $\lambda$ and $\mu$ be distinct complex numbers, $m \geq 1$, $n \geq 1$, $\mathcal{H}$ be a Hilbert space, $T$ be an algebraic operator acting on $\mathcal{H}$ whose (traditional) minimal polynomials is $(x - \lambda)^m(x - \mu)^n$, and $p \in \mathbb{C}[x, y]$. Since the minimal polynomial of $T$ is $(x - \lambda)^m(x - \mu)^n$, there exists a nonzero vector $u_0 \in \mathcal{H}_\lambda(T)$ which is a member of ran($((T - \lambda)^m - \mathcal{H}_\lambda(T))^\perp$) and a nonzero vector $v_0 \in \mathcal{H}_\mu(T)$ which is a member of ran($((T - \mu)^n - \mathcal{H}_\mu(T))^\perp$). Let $\sigma_{rs}$ be defined as follows:

\begin{equation}
\sigma_{rs} := \sum_{0 \leq i \leq m-1 \atop 0 \leq j \leq n-1} \frac{1}{i!j!} \left( \frac{\partial^{i+j}}{\partial^j y \partial^i x} p \right)(\lambda, \mu) \langle u_{i+r}, v_{j+s} \rangle,
\end{equation}

where $u_r = (T - \lambda)^r u_0$ and $v_s = (T - \mu)^s v_0$ for $r \geq 0$ and $s \geq 0$.

**Lemma 41.** If $\mathcal{H}$ is a Hilbert space, $T \in \mathcal{L}(\mathcal{H})$, and both $T|\mathcal{H}_\lambda(T)$ and $T|\mathcal{H}_\mu(T)$ are similar to Jordan cells, then $p(T|\mathcal{H}_\lambda(T)$ is orthogonal to $\mathcal{H}_\mu(T)$ if and only if $\sigma_{rs} = 0$ for $0 \leq r \leq m-1$ and $0 \leq s \leq n-1$. 
Proof. With the notation above,

\[
\langle p(T)u_r, v_s \rangle = \langle p(T)(T - \lambda)^ru_0, (T - \mu)^sv_0 \rangle \\
- \langle (T - \mu)^{s} p(T)(T - \lambda)^ru_0, v_0 \rangle \\
= \sum_{0 \leq i \leq m - 1 \atop 0 \leq j \leq n - 1} \left[ \frac{1}{i!j!} \frac{\partial^{i+j}}{\partial^i y \partial^j x} p(\lambda, \mu) \right] \\
\times \langle (T - \lambda)^i(T - \lambda)^ru_0, (T - \mu)^i(T - \mu)^sv_0 \rangle \\
= \sigma_{rs}
\]

For \(0 \leq r \leq m - 1\) and \(0 \leq s \leq n - 1\).

Now suppose that \(p(T)\mathcal{A}_\lambda(T)\) is orthogonal to \(\mathcal{A}_\mu(T)\). Fix \(0 \leq r \leq m - 1\) and \(0 \leq s \leq n - 1\). Since \(u_r \in \mathcal{A}_\lambda(T)\) and \(v_s \in \mathcal{A}_\mu(T)\), \(\sigma_{rs} = 0\).

On the other hand, suppose that \(\sigma_{rs} = 0\) for all \(0 \leq r \leq m - 1\) and \(0 \leq s \leq n - 1\). Since \(T | \mathcal{A}_\lambda(T)\) and \(T | \mathcal{A}_\mu(T)\) are similar to Jordan cells, \(\{u_r\}_{r=0}^{m-1}\) is a spanning set for \(\mathcal{A}_\lambda(T)\) and \(\{v_s\}_{s=0}^{n-1}\) is a spanning set for \(\mathcal{A}_\mu(T)\). Since \(\langle \cdot, \cdot \rangle\) is linear in its first argument and conjugate linear in its second argument, it follows that \(\langle p(T)h, k \rangle = 0\) for all \(h \in \mathcal{A}_\lambda(T)\) and all \(k \in \mathcal{A}_\mu(T)\). This completes the proof of Lemma 4.1. \(\blacksquare\)

Now, (40) defines \(\sigma_{rs}\) to be a sum of products. To analyze the condition \(\sigma_{rs} = 0\), we look at each factor of each term of the sum. In particular, for positive integers \(a\) and \(b\),

1. if \((T - \lambda)^a\mathcal{A}_\lambda(T)\) is orthogonal to \((T - \mu)^b\mathcal{A}_\mu(T)\), then

\[
\langle (T - \lambda)^i u_0, (T - \mu)^j v_0 \rangle = 0
\]

whenever both \(i + r \geq a\) and \(j + s \geq b\), and

2. if \((x - \lambda)^i(y - \mu)^j\) divides \(p(x, y)\), then

\[
\frac{\partial^{i+j}}{\partial^i y \partial^j x} p(\lambda, \mu) = 0
\]

when both \(0 \leq i \leq a\) and \(0 \leq j \leq b\) hold.
The following definition of $\perp$-pattern is motivated by (1) (for the case when both $r = 0$ and $s = 0$), and the following definition of the 0-pattern is motivated by (2). Taking advantage of (1) for the case of $r > 0$ or $s > 0$ will be addressed by shift rules given in Section 5.

Note that the sum of the right hand side of (40) has fewer nonzero summands as $r$ and $s$ increase, since $(T - \lambda)^mu_0 = 0$ and $(T - \mu)^nv_0 = 0$.

**Definition 43.** Let $\mathcal{H}$ be a finite dimensional Hilbert space, let $T \in \mathcal{L}(\mathcal{H})$ be a bi-Jordan cell with distinct eigenvalues $\lambda$ and $\mu$, and let $p \in \mathbb{C}[x, y]$. The $\perp$-pattern of $T$ is defined to be the set

\[(a) \quad \{(i, j) : 0 \leq i < \dim \mathcal{H}(T) - 1, 0 \leq j < \dim \mathcal{H}(T) - 1, \quad (T - \lambda)^i \mathcal{H}(T) \perp (T - \mu)^j \mathcal{H}(T)\},\]

and the 0-pattern of $p$ is defined to be the set

\[(b) \quad \{(i, j) : 0 \leq i, 0 \leq j, (x - \lambda)^i (y - \mu)^j \text{ divides } p(x, y)\}.\]

**Definition 44.** Let $X \subseteq \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\}$. $X$ is **backwardly invariant** if $(i, j) \in X$ whenever $(i, j) \in X$, $0 \leq i \leq i$ and $0 \leq j \leq j$, and $X$ is **forwardly invariant** if $(i, j) \in X$ whenever $(i, j) \in X$, $i \leq i \leq m - 1$, and $j \leq j \leq n - 1$.

Clearly, the 0-pattern of a polynomial is backwardly invariant, and the $\perp$-pattern of an operator $T$ is forwardly invariant. In fact, the converses of these observations hold.

**Proposition 45.** Let $X \subseteq \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\}$. $X$ is backward invariant if and only if $X$ is the 0-pattern of some polynomial $p$, and $X$ is forward invariant if and only if $X$ is the $\perp$-pattern of some bi-Jordan cell of type $(\lambda, \mu, m, n)$ with $\lambda \neq \mu$.

The first assertion of Proposition 45 is elementary, and the second assertion is shown in Lemma 57.

We will discuss the concepts of forward- and backward-invariant sets further in Section 5.

We shall be interested at an abstract level in which pairs of patterns can arise, so we introduce the next definition.
DEFINITION 46. If \( m \geq 0 \) and \( n \geq 0 \), then we define an \((m, n)\) grid to be an \( m \times n \) matrix \( G = (g_{r,s}) \) such that \( g_{r,s} \in \{0, *, \perp\} \) for each \((r, s) \in \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\}\).

For \( G = (g_{r,s}) \) an \((m, n)\) grid, we define \( \mathcal{P}_0(G)(\mathcal{P}_*(G), \mathcal{P}_\perp(G)) \) as the set of all \((r, s)\) such that \( g_{r,s} = 0 \) \((g_{r,s} = *, g_{r,s} = \perp)\). Note that \( \mathcal{P}_*(G) \) is the complement of \( \mathcal{P}_0(G) \cup \mathcal{P}_\perp(G) \).

DEFINITION 47. \( G = (g_{r,s}) \) is a representable \((m, n)\) grid if there exist distinct \( \lambda \) and \( \mu \), a bi-Jordan cell of type \((\lambda, \mu, m, n)\), and \( p \in \mathbb{C}[x, y] \) such that
\[
48) \quad p(T) = 0,
\]
\[
49) \quad g_{r,s} = 0 \quad \text{if and only if} \quad (r, s) \text{ is a member of the 0-pattern of } p,
\]
and
\[
50) \quad g_{r,s} = \perp \quad \text{if and only if} \quad (r, s) \text{ is a member of the } \perp \text{-pattern of } T.
\]

In the above case, we shall write \( G = G_{p,T} \).

Since the 0-pattern of any polynomial \( p \) is backward invariant and the \( \perp \) -pattern of any bi-Jordan cell is forward invariant, the following necessary condition for the representability of a grid is immediate.

PROPOSITION 51. If a grid \( G \) is representable, then \( \mathcal{P}_0(G) \) is backward invariant and \( \mathcal{P}_\perp(G) \) is forward invariant.

Unfortunately, examples we consider in Section 4.3 show that the above necessary conditions for representability are far from sufficient. Indeed, deciding whether or not a given grid is representable is a hard question.

4.2. Algebraic Characterization of Representable Grids

In this section we convert the problem of whether or not a grid is representable to a problem about the solvability of a certain type of system of linear equations. We now introduce these equations.
Definition 5.2. Let \( E \subseteq \{0, \ldots, m-1\} \times \{0, \ldots, n-1\} \). We refer to a set of equations as bi-Hankel subordinate to \( E \) if the equations have the form

\[
\sum_{(i, j) \in E} c_{i,j} x_{i+r, j+s} = 0
\]

where \( 0 \leq r \leq m-1, 0 \leq s \leq n-1 \). For complex numbers \( d_{ij}, y_{ij} \in \mathbb{C} \), we say that \((\{d_{ij}\}_{i \geq 0, j \geq 0}, \{y_{ij}\}_{i \geq 0, j \geq 0})\) is a solution of (53) if (1) \( d_{ij} = 0 \) and \( y_{ij} = 0 \) whenever \((i, j) \notin E\) and (2) for each \( r \) and \( s \), one obtains 0 when one replaces \( c_{ij} \) with \( d_{ij} \) and replaces \( x_{ij} \) with \( y_{ij} \) for all \( i \) and \( j \) in the left hand side of the equations of (53).

Note that in the above definition, we have considered an infinite number of indeterminates \( c_{ij} \) and \( x_{ij} \) and the solutions are infinite sets of complex numbers \( d_{ij} \) and \( y_{ij} \) where only finite many are nonzero. Only a finite number of indeterminates and a finite number of complex numbers are actually used in the above definition. This reduces the complexity of the definitions.

Our main theorem of this section is the following theorem which connects representability of grids to solutions of bi-Hankel equations.

Theorem 5.4. A grid \( G \) is representable if and only if there exists a solution \((\{d_{ij}\}_{i \geq 0, j \geq 0}, \{y_{ij}\}_{i \geq 0, j \geq 0})\) to the set of bi-Hankel equations subordinate to \( \mathcal{P}_* (G) \) such that

\[
\text{The largest backward-invariant subset of } \mathcal{P}_0 (G) \cup \{(i, j) \in \mathcal{P}_* (G) : d_{ij} = 0\} \text{ is } \mathcal{P}_0 (G)
\]

and

\[
\text{The largest forward-invariant subset of } \mathcal{P}_\perp (G) \cup \{(i, j) \in \mathcal{P}_* (G) : y_{ij} = 0\} \text{ is } \mathcal{P}_\perp (G).
\]

Proof of the “only if” side of Theorem 5.4. Let \( G \) be an \( m \times n \) representable grid. There exist a Hilbert space \( \mathcal{H}, T \in \mathcal{L}(\mathcal{H}) \), and \( p \in \mathbb{C}[x, y] \) such that \( \dim \mathcal{H} = m + n \), the minimal polynomial of \( T \) is \((x - \lambda)^m (x - \mu)^n\), \( p = p^\vee \), \( p(T) = 0 \), and both (49) and (50) hold. Now for \( i \geq 0 \) and \( j \geq 0 \),
\[ d_{i,j} = \frac{1}{i!j!} \left( \frac{\partial^i y}{\partial y^i \partial x^j} p \right) \quad \text{and} \quad y_{i,j} = \left\langle (T - \lambda)^i u_0, (T - \mu)^j v_0 \right\rangle. \]

Since \( p(T) = 0 \), Lemma 41 implies that \((\{d_{i,j} \geq 0, j \geq 0\}, \{y_{i,j} \geq 0, j \geq 0\})\) is a solution of the bi-Hankel equations subordinate to \( \mathcal{P}_x(G) \). Since (49) and (50) hold, (55) and (56) hold. This completes the proof of the "only if" side of Theorem 54.

The proof of the "if" side of Theorem 54 requires two lemmas which we now state. The proof of the first lemma gives a correspondence between the \( m \times n \) matrices with entries in \( \mathbb{C} \) and the bi-Jordan cells of type \((\lambda, \mu, m, n)\) for fixed \( \lambda \neq \mu \).

**Lemma 57.** Fix \( m \geq 1, n \geq 1 \), and \( \lambda, \mu \in \mathbb{C} \) with \( \lambda \neq \mu \). If \( x_{ij} \in \mathbb{C} \) for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \), then there exists a bi-Jordan cell \( T \in \mathcal{L}(\mathcal{H}) \) of type \((\lambda, \mu, m, n)\), \( u_0 \in \mathcal{K}_\lambda(T) \), and \( v_0 \in \mathcal{K}_\mu(T) \) such that

\[ x_{ij} = \left\langle (T - \lambda)^i u_0, (T - \mu)^j v_0 \right\rangle \tag{58} \]

for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \).

**Proof.** Let \( \mathcal{H} \) be the \( m + n \) dimensional Hilbert space \( \mathbb{C}^m \oplus \mathbb{C}^n \), \( \{e_i\}_{i=1}^n \) be the standard basis of \( \mathbb{C}^m \), and \( \{\tilde{e}_j\}_{j=1}^n \) be the standard basis of \( \mathbb{C}^n \). Let \( Y \in \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m) \) be such that

\[ \left\langle Y \tilde{e}_{n-s}, e_{m-r} \right\rangle = -\tilde{\lambda}_{s,r} \]

for \( 0 \leq r \leq m - 1 \) and \( 0 \leq s \leq n - 1 \). Let \( J_\lambda \in \mathcal{L}(\mathbb{C}^n) \) be the \( m \times m \) Jordan cell with eigenvalue \( \lambda \), and let \( J_\mu \in \mathcal{L}(\mathbb{C}^n) \) be the \( n \times n \) Jordan cell with eigenvalue \( \mu \). If we set

\[ u_0 = \begin{bmatrix} e_m \\ 0 \end{bmatrix}, \quad v_0 = \begin{bmatrix} -Y \tilde{e}_n \\ \tilde{e}_n \end{bmatrix}, \quad \text{and} \quad T = \begin{bmatrix} J_\lambda & J_\lambda Y - Y J_\mu \\ 0 & J_\mu \end{bmatrix}, \]
then direct computation shows that

\[ \mathcal{H}(T) = \mathbb{C}^n \oplus \{0\}, \]

\[ \mathcal{H}(T) = \begin{bmatrix} -Yw \\ w \end{bmatrix} : w \in \mathbb{C}^n \}

\[ u_0 \in \mathcal{H}(T), \]

\[ v_0 \in \mathcal{H}(T), \]

\[ (T - \lambda)^i u_0 = \begin{bmatrix} \epsilon_{m-i} \\ 0 \end{bmatrix}, \text{ and} \]

\[ (T - \mu)^j v_0 = \begin{bmatrix} -Y\epsilon_{m-i} \\ \epsilon_{m-i} \end{bmatrix}. \]

Therefore,

\[ \langle (T - \lambda)^i u_0, (T - \mu)^j v_0 \rangle = \langle e_{m-i}, -Y\epsilon_{m-i} \rangle = x_{ij} \]

for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \). Therefore, (58) holds and the proof of Lemma 57 is complete.

**Lemma 59.** If \( p \in \mathbb{C}[x, y] \) and \( p = p \uparrow \), then \( p \) is in the ideal generated by \( (x - \lambda)^m(y - \overline{\mu})^n \) and \( (x - \mu)^n(y - \overline{\lambda})^m \) if and only if there exists \( p_0 \in \mathbb{C}[x, y] \) such that

\[ (60) \quad p(x, y) = (x - \lambda)^m(y - \overline{\mu})^n p_0(x, y) \]

\[ + (x - \mu)^n(y - \overline{\lambda})^m p_0 \uparrow (x, y). \]

Here \( p_0 \) need not equal \( p_0 \uparrow \). Moreover, if \( T \) is a bi-Jordan cell of type \((\lambda, \mu, m, n)\) with \( \lambda \neq \mu \) and \( p \) is given by (60), then the 0-pattern of \( p \) equals the 0-pattern of \( p_0 \) and the following three conditions are equivalent:

\[ (61) \quad p(T) = 0, \]

\[ (62) \quad p(T) \mathcal{H}(T) \text{ is orthogonal to } \mathcal{H}(T), \]
and

\[ p_0(T) \mathcal{Z}_\alpha(T) \text{ is orthogonal to } \mathcal{Z}_\mu(T). \]

**Proof.** Observe that the first assertion of the lemma is an immediate consequence of Lemma 87 with \( \alpha = (x - \lambda)^m(y - \overline{\mu})^n \).

Now let \( p \) and \( p_0 \) satisfy (60). Thus, for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \), (60) implies that \((x - \lambda)^i(y - \overline{\mu})^j \) divides \( p(x, y) \) if and only if \((x - \lambda)^i(y - \overline{\mu})^j \) divides \((x - \mu)^n(y - \lambda)^m p_0(x, y) \). Also, \((x - \lambda)^i(y - \overline{\mu})^j \) divides \((x - \mu)^n(y - \lambda)^m p_0(x, y) \) if and only if \((x - \lambda)^i(y - \overline{\mu})^j \) divides \( p_0(x, y) \), by \( m + n \) applications of Lemma 76.

Now let \( T \) be a bi-Jordan cell of type \((\lambda, \mu, m, n)\). By Theorem 36 and Proposition 86, (61) holds if and only if (62) holds. To see that (62) holds if and only if (63) holds, note first that \((T - \lambda)^m \mathcal{Z}_\alpha(T) = \{0\} \) and \((T - \mu)^n \mathcal{Z}_\mu(T) = \{0\} \). Thus, if \( h \in \mathcal{Z}_\alpha(T) \) and \( k \in \mathcal{Z}_\mu(T) \),

\[ \langle p(T)h, k \rangle = \langle (T^* - \overline{\mu})^n p_0^\vee(T)(T - \lambda)^m h, k \rangle \]

\[ + \langle (T^* - \overline{\lambda})^m p_0(T)(T - \mu)^n h, k \rangle \]

\[ = \langle p_0^\vee(T)(T - \lambda)^m h, (T - \mu)^n k \rangle \]

\[ + \langle p_0(T)(T - \mu)^n h, (T - \lambda)^m k \rangle \]

\[ = \langle p_0(T)(T - \mu)^n h, (T - \lambda)^m k \rangle. \]

Since \((T - \lambda)^m \mathcal{Z}_\mu(T) = \mathcal{Z}_\mu(T) \) and \((T - \mu)^n \mathcal{Z}_\alpha(T) = \mathcal{Z}_\alpha(T) \), (64) implies that (62) holds if and only if (63) holds. This completes the proof of Lemma 59. \( \blacksquare \)

**Proof of the “if” side of Theorem 54.** Suppose that \( G \) is an \( m \times n \) grid and that \((d_{ij})_{i \geq 0, j \geq 0}, \{y_{ij}\}_{i \geq 0, j \geq 0}\) is a solution of the bi-Hankel equations subordinate to \( \mathcal{P}_*(G) \) such that (55) and (56) hold. Let \( p_0(x, y) \in \mathbb{C}[x, y] \) be defined by

\[ p_0(x, y) = \sum_{0 \leq i \leq m} \sum_{0 \leq j \leq n - 1} d_{ij}(x - \lambda)^i(y - \mu)^j, \]

and both \( \mathcal{Y} \) and \( T \in \mathcal{L}(\mathcal{Y}) \) be given as in Lemma 57. Since \((d_{ij})_{i \geq 0, j \geq 0}, \{y_{ij}\}_{i \geq 0, j \geq 0}\) is a solution of the bi-Hankel equations subordinate to \( \mathcal{P}_*(G) \),
Lemma 41 implies that \( p_0(T) \mathcal{X}(T) \) is orthogonal to \( \mathcal{X}(T) \). Now, if we let 
\( p(x, y) \in C[x, y] \) be as in Lemma 59, then Lemma 59 implies that \( p(T) = 0 \)
and the 0-pattern of \( p \) equals the 0-pattern of \( p_0 \). Therefore, \( G = G_{p, T} \) and
\( G \) is a representable grid.

4.3. Examples

One way to construct representable grids is to choose a polynomial \( p \) and
a bi-Jordan cell \( T \) of type \((\lambda, \mu, m, n)\) with \( \lambda \neq \mu \) such that \( p(T) = 0 \) and
with the additional property that the 0-pattern of \( p \) is disjoint from the
\( \perp \)-pattern of \( T \). To any such pair we can associate the unique grid \( G = (g_{r,s}) \)
such that (49) and (50) hold. Furthermore, \( G \) is representable \((G = G_{p, T})\),
and, obviously, any representable grid can be constructed in this way. The
real problem is to determine whether a given grid is representable.

It is easy to use Theorem 54 to show that both

\[
\begin{bmatrix}
0 & * \\
* & \perp
\end{bmatrix}
\quad \text{and} \quad 
\begin{bmatrix}
0 & * & * \\
* & * & \perp \\
* & \perp & \perp
\end{bmatrix}
\]

are representable and that

\[
\begin{bmatrix}
0 & * \\
* & *
\end{bmatrix}
\quad \text{and} \quad 
\begin{bmatrix}
0 & * \\
* & \perp
\end{bmatrix}
\]

are not representable grids.

4.4. Obtaining Nondegenerate Solutions to Sets of Bi-Hankel Equations

Let \( E \) be a subset of \( \{0, \ldots, m-1\} \times \{0, \ldots, n-1\} \), and consider the
set of bi-Hankel equations subordinate to \( E \) given by (53). Obviously, this set
of bi-Hankel equations has a solution regardless of what \( E \) is; namely, if we
set \( d_{ij} = 0 \) and \( y_{ij} = 0 \) for all \( i \) and \( j \). However, our main concern in this
section will be nonvanishing solutions. To be more precise, we now describe
an algebraic method which, for given subsets \( D \) and \( Y \) of \( E \), determines if a
solution \((\{d_{ij}\}, \{y_{ij}\})\) to (53) exists with \( d_{ij} \neq 0 \) for each \((i, j) \in D\)
and \( y_{ij} \neq 0 \) for each \((i, j) \in Y \). The method is based on the next theorem.

Let \( \text{POLY} \) be the set of all polynomials with complex coefficients in the
indeterminates \( c_{ij} \) and \( x_{ij} \) for \( 0 \leq i \leq m-1 \) and \( 0 \leq j \leq n-1 \).

If \( p \in \text{POLY} \) and \( v = (d_{11}, \ldots, d_{1n}, d_{21}, \ldots, d_{mn}, y_{11}, \ldots, y_{1n}, y_{21}, \ldots,
y_{mn}) \in C^{2mn} \), then we say that \( p \) vanishes on \( v \) if \( p \) becomes zero when \( c_{ij} \)
is replaced by \( d_{ij} \) and \( x_{ij} \) is replaced by \( y_{ij} \). If \( V \subset \mathbb{C}^{2mn} \), then we say that \( p \) vanishes on \( V \) if \( p \) vanishes on \( v \) for every \( v \in V \).

If \( I \) is an ideal of a polynomial ring \( R \), the \textit{radical ideal} of \( I \) is the set

\[
\{ f \in R : \text{there exists } N \geq 1 \text{ such that } f^N \in I \}.
\]

It is easy to verify that this set is an ideal.

We view the bi-Hankel equations in (53) as polynomials in the indeterminates \( c_{ij} \) and \( x_{ij} \) for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \). Let \( P \) be the set of polynomials in \( \text{POLY} \) which are the left hand sides of the bi-Hankel equations in (53), and define \( p_{\text{test}} \in \text{POLY} \) by

\[
(67) \quad p_{\text{test}} = \left( \prod_{(i,j) \in D} c_{ij} \right) \left( \prod_{(i,j) \in Y} x_{ij} \right).
\]

\textbf{Theorem 68.} There is a solution \( \{(d_{ij})_{i \neq 0, j \neq 0}, (y_{ij})_{i \neq 0, j \neq 0}\} \) to the bi-Hankel equations (53) subordinate to \( E \) such that

\[
(69) \quad d_{ij} \neq 0 \quad \text{if} \quad (i,j) \in D,
\]

\[
(70) \quad y_{ij} \neq 0 \quad \text{if} \quad (i,j) \in Y
\]

if and only if no power of \( p_{\text{test}} \) is in the ideal generated by the union of \( P \) and

\[
(71) \quad \{c_{ij} : (i,j) \notin E \} \cup \{x_{ij} : (i,j) \notin E \}.
\]

\textit{Proof.} Let \( I \) be the ideal generated by \( P \) and \( \{c_{ij} : (i,j) \notin E \} \cup \{x_{ij} : (i,j) \notin E \} \). Thus, if \( v \in \mathbb{C}^{2mn} \), then \( v \in \mathcal{V}_I \), the variety of \( I \), if and only if \( v \) solves the bi-Hankel equations subordinate to \( E \). Also note that if \( v \in \mathbb{C}^{2mn} \), then \( p_{\text{test}}(v) \neq 0 \) if and only if (69) and (70) hold. Finally, recall that Hilbert's \textit{Nullstellensatz} asserts that \( p \) vanishes on \( \mathcal{V}_I \) if and only if some power of \( p \) is an element of \( I \). This completes the proof of Theorem 68. \( \blacksquare \)

This theorem converts the issue of whether a grid is representable or not into one of determining whether or not a specific polynomial is in a specific ideal. Specifically, let \( G \) be an \((m, n)\) grid. We associate to \( G \) two algebraic
objects, a polynomial \( p_C \) and an ideal \( I_G \). First, let

\[ E = \mathcal{P}_*(G), \]

\[ D = \{(i, j) \in \mathcal{P}_*(G) : \mathcal{P}_0(G) \cup \{(i, j)\} \text{ is backward invariant}\}, \] and

\[ Y = \{(i, j) \in \mathcal{P}_*(G) : P_\perp(G) \cup \{(i, j)\} \text{ is forward invariant}\}. \]

Recall that \( \text{POLY} \) is the set of all polynomials with complex coefficients with the indeterminates \( c_{ij} \) and \( x_{ij} \) for \( 0 \leq i \leq m - 1 \) and \( 0 \leq j \leq n - 1 \). Finally, let \( p_C \) be the polynomial in \( \text{POLY} \) defined by

\[ p_C = \left( \prod_{(i, j) \in D} c_{ij} \right) \left( \prod_{(i, j) \in Y} x_{ij} \right), \]

and let \( I_G \subset \text{POLY} \) be the ideal generated by the polynomials

\[ \sum_{(i, j) \in E} c_{ij} x_{i+r, j+s} \quad \text{for} \quad 0 \leq r \leq m - 1, \quad 0 \leq s \leq n - 1 \]

and the monomials

\[ \{c_{ij} : (i, j) \notin E\} \quad \text{and} \quad \{x_{ij} : (i, j) \notin E\}. \]

With this notation Theorem 54 implies that \( G \) is representable if and only if there exists a solution to the bi-Hankel equations subordinate to \( E \) such that the conditions (69) and (70) of Theorem 68 hold. Thus an immediate consequence of Theorem 68 is the following theorem.

**Theorem 72.** Let \( G \) be a grid. \( G \) is representable if and only if no power of \( p_C \) lies in \( I_G \).

The type of condition that occurs in Theorem 72 is known as the *ideal membership* problem and is exactly the subject which Gröbner basis methods address and solve computationally. If an ideal is specified via its generating set, then one can determine whether or not a particular polynomial lies in this ideal with the assistance of a Gröbner basis. In this particular case, the question is whether or not a particular polynomial is in the radical of an ideal or not. Proposition 8 in Section 2 of [22] gives a condition under which a
polynomial $p$ lies in the radical of an ideal $I$. This condition is as follows:

Let $I$ be the ideal generated by $\{f_1, \ldots, f_n\}$. Then $f$ lies in the radical of the ideal $I$ if and only if the constant polynomial 1 belongs to the ideal of $\text{POLY}[\alpha]$ generated by $\{f_1, \ldots, f_n, 1 - \alpha\}$.

Therefore, if one is given the generating set for an ideal, one can use a Gröbner basis calculation to determine whether or not a particular polynomial lies in the radical of this ideal. A Gröbner basis can be computed via many symbolic manipulators, including Mathematica, Maple, and MacCaulay.

### 4.5. Robust Grids

Another notion is that of a robust grid.

**Definition 73.** A grid $G$ is **polynomially robustly representable** if there exists a subset $\mathcal{P}$ of $C[x, y]$ such that the 0-pattern of $p$ equals $\mathcal{P}_0(G)$ for all $p \in \mathcal{P}$, such that $\mathcal{P}$ is dense in the set of all polynomials whose 0-pattern equals $\mathcal{P}_0(G)$, and such that, for each $p \in \mathcal{P}$, there exists an operator $T$ such that $G = G_{p, T}$. In dual fashion, we say that an $(m, n)$ grid $G$ is **matricially robustly representable** if there exist $\lambda, \mu \in C$, and $m + n$ dimensional complex Hilbert space $\mathcal{H}$, and a subset $\mathcal{P}$ of $\mathcal{L}(\mathcal{H})$ such that $\lambda \neq \mu$, such that $\mathcal{P}$ is dense in the set

$$\{(T \in \mathcal{L}(\mathcal{H}) : T \text{ is a bi-Jordan cell of type } (\lambda, \mu, m, n) \text{ and the } \perp \text{-pattern of } T \text{ equals } \mathcal{P}_\perp(G)\},$$

and such that, for each $T \in \mathcal{P}$, there exists a polynomial $p$ such that $G = G_{p, T}$. Finally, we say $G$ is **robustly representable** if $G$ is both polynomially robustly representable and matricially robustly representable.

Just as the issue of representability for grids reduces to whether a certain structured system of linear equation is solvable, so also can the issue of robust representability be resolved in terms of the generic solvability of a system of structured linear equations. Specifically, the following result follows immediately from Theorem 54.

**Theorem 75.** Let $G$ be a grid, let $D$ denote the set of all tuples $\{d_{ij}\}$ such that (55) holds, let $Y$ denote the set of all tuples such that (56) holds, and let $S$ denote the set of all solutions $((d_i), (y_i))$ to the set of bi-Hankel equations subordinate to $\mathcal{P}_\ast(G)$. Then $G$ is polynomially robustly repre-
sentable (respectively, matrically robustly representable) if and only if the set of first components (respectively, the set of second components) from \( S \cap (D \times Y) \) is dense in \( D \) (respectively, \( Y \)).

The concept of robust representability arises in response to examples such as

\[
G = \begin{bmatrix}
0 & * & * \\
* & * & \bot \\
* & \bot & \bot
\end{bmatrix},
\]

which, while representable, can only be represented as \( G = G_{p,T} \) with \( T \) having a certain degenerate relation on the geometry of its generalized eigenvectors. Such examples suggest that a really clean characterization of when a grid \( G \) is representable entirely in terms of the 0, *, and \( \bot \) patterns of \( G \) would be difficult without the additional requirement of robustness. The situation is similar to algebraic geometry, where often the clean theorems hold only for a generic case. We discuss the above example and related results further in Section 5.

4.6. Assorted Proofs

This section contains statements and proofs of facts that were deferred from previous sections. It is self-contained.

Lemma 76. If \( p \in \mathbb{C}[z,w], q \in \mathbb{C}[w], a \in \mathbb{C}, b \in \mathbb{C}, a \neq b, m \geq 1, \) and \( I = \langle((z-a)^m, q(w))\rangle \) then \((z-b)p(z,w) \in I \) if and only if \( p(z,w) \in I \).

Proof. Clearly, if \( p \in I \), then \((z-b)p \in I \).

If \((z-b) \in I \), then there exist \( r_1(z,w), r_2(z,w) \in \mathbb{C}[z,w] \) such that

\[
(z-b)p(z,w) = (z-a)^m r_1(z,w) + q(w)r_2(z,w).
\]

Since \( r_1, r_2 \in \mathbb{C}[z,w] \), there exist polynomials \( r_3, r_4 \in \mathbb{C}[w] \) and \( r_5, r_6 \in \mathbb{C}[z,w] \) such that

\[
(77) \quad r_1(z,w) = r_3(w) + (z-b)r_4(z,w) \quad \text{and}
\]

\[
(78) \quad r_2(z,w) = r_5(w) + (z-b)r_6(z,w).
\]
By (77), (78), and (79),

\[(80) \quad (z - b)p(z, w) = (z - a)^m [r_3(w) + (z - b)r_4(z, w)]
+ q(w)[r_5(w) + (z - b)r_6(z, w)],\]

and if we substitute \(b\) for \(z\), then

\[(81) \quad 0 = (b - a)^m r_3(w) + q(w)r_5(w).\]

Therefore, \(r_3(w) = -(b - a)^{-m}q(w)r_5(w)\), and so

\[(82) \quad (z - b)p(z, w)
= q(w)r_5(w)\left[1 - \left(\frac{z - a}{b - a}\right)^m\right]
+ (z - b)\left[(z - a)^m r_4(z, w) + q(w)r_6(z, w)\right].\]

Now, \((z - a)^m r_4(z, w) \in I\) and \(q(w)r_5(z, w) \in I\). Also, \(q(w)r_5(z, w) \in I\),
and \(z - b\) divides \(1 - [(z - a)^m (b - a)^m]\). Therefore, \(p \in I\), and the proof of Lemma 76 is complete. 

If \(p \in \mathbb{C}[x, y]\), then let \(\text{deg}_x(p)\) denote the degree of \(p\) with respect to \(x\), and let \(\text{deg}_y(p)\) denote the degree of \(p\) with respect to \(y\). We use the convention that \(\text{deg}_x(p) = -\infty\) and \(\text{deg}_y(p) = -\infty\) if \(p\) is the zero polynomial of \(\mathbb{C}[x, y]\).

**Lemma 83.** Let \(\lambda \in \mathbb{C}\) and \(\mu \in \mathbb{C}\) be such that \(\lambda \neq \mu\). Let \(m\) and \(n\) be positive integers. Let

\[
V_1 = \left\{(x - \lambda)^m (y - \overline{\lambda})^m p_1(x, y) : \right. \\
\left. p_1 \in \mathbb{C}[x, y], \text{deg}_x(p_1) < n, \text{and deg}_y(p_1) < n\right\},
\]

\[
V_2 = \left\{(x - \mu)^n (y - \overline{\lambda})^m p_2(x, y) : \right. \\
\left. p_2 \in \mathbb{C}[x, y], \text{deg}_x(p_2) < m, \text{and deg}_y(p_2) < n\right\},
\]
\[ V_3 = \{(x - \lambda)^m(y - \bar{\mu})^n p_3(x, y) : \]
\[ p_3 \in \mathbb{C}[x, y], \text{deg}_x(p_3) < n, \text{and deg}_y(p_3) < m\}, \]
\[ V_4 = \{(x - \mu)^n(y - \bar{\mu})^n p_4(x, y) : \]
\[ p_4 \in \mathbb{C}[x, y], \text{deg}_x(p_4) < m, \text{and deg}_y(p_4) < m\}, \]
\[ V_5 = \{(x - \lambda)^m(x - \mu)^n(y - \bar{\lambda})^m(y - \bar{\mu})^n p_5(x, y) : p_5 \in \mathbb{C}[x, y]\}. \]

If \( q \in \mathbb{C}[x, y] \), then there exist unique polynomials \( q_j \in V_j \) for \( 1 \leq j \leq 5 \) such that
\[ q = q_1 + q_2 + q_3 + q_4 + q_5. \]

In particular, the ideal generated by \( \{(x - \lambda)^m, (y - \bar{\lambda})^n\} \) (by \( \{(x - \mu)^n, (y - \bar{\mu})^n\} \)) equals \( V_1 + V_2 + V_3 + V_5 \) \((V_2 + V_3 + V_4 + V_5)\).

Proof. Let \( W_1 \) be the vector space spanned by \( \{x^j : 0 \leq j \leq m - 1\} \), and \( W_2 \) be the vector space spanned by \( \{x^j : 0 \leq j \leq n - 1\} \). We first show that for every \( q \in \mathbb{C}[x] \), there exist unique \( p_1 \in W_1, p_2 \in W_2, \) and \( p_3 \in \mathbb{C}[x] \) such that
\[ p = (x - \mu)^n p_1(x) + (x - \lambda)^m p_2(x) \]
\[ + (x - \lambda)^m(x - \mu)^n p_3(x). \]

Let \( V \) be the vector space spanned by \( (x - \lambda)^j(x - \mu)^k \) for \( 0 \leq j \leq m - 1, \)
\( 0 \leq k \leq n - 1, \) and \( (j, k) \neq (m, n) \). Since \( \{(x - \lambda)^m x^j : 0 \leq j \leq n - 1\} \cup \{(x - \mu)^n x^k : 0 \leq k \leq m - 1\} \) has cardinality \( m + n \) and \( \text{dim} V = m + n \), the linear independence of this set (i.e., that for \( p_1 \in W_1 \) and \( p_2 \in W_2, \)
\((x - \lambda)^m p_1(x) + (x - \mu)^n p_2(x) = 0 \) implies \( p_1 = 0 \) and \( p_2 = 0 \)) implies the above claim. Now, if \( (x - \lambda)^m p_1(x) + (x - \mu)^n p_2(x) = 0 \), then \( (x - \lambda)^m p_1(x) \) is in the ideal of \( \mathbb{C}[x] \) generated by \( (x - \mu)^n \). Since \( \lambda \neq \mu \), it follows that \( p_1 \) is in the ideal of \( (x - \mu)^n \). Since the degree of \( p_1 \) is less than \( n \), we have \( p_1 = 0 \). Therefore, \( (x - \mu)^n p_2 = 0 \) and so \( p_2 = 0 \). Therefore, there exist unique \( p_1 \in W_1, p_2 \in W_2, \) and \( p_3 \in W_3 \) such that (85) holds.

The claim in the above paragraph shows that every polynomial of the form \( r_1(x)r_2(y) \) can be represented as in the statement of the theorem.
Since \( \{r_1(x)r_2(y) : r_1 \in \mathbb{C}[x] \text{ and } r_2 \in \mathbb{C}[y] \} \) spans \( \mathbb{C}[x, y] \), we have \( \mathbb{C}[x, y] = V_1 + V_2 + V_3 + V_4 + V_5 \).

To show uniqueness, let \( W_3 = V_1 + V_2 + V_3 + V_4 \). Since the degree of \((x - \lambda)^m(y - \overline{\lambda})^n(x - \mu)^n(y - \overline{\mu})^n\) is \(2(m + n)\), the dimension of the vector space \( \mathbb{C}[x, y]/V_5 \) is \(2(m + n)\). Since \( \mathbb{C}[x, y] = W_3 + V_5 \) and \( W_3 \cap V_5 = \{0\} \), the second isomorphism theorem implies that

\[
\dim(W_3) = \dim(W_3/(W_3 \cap V_5)) = \dim((W_3 + V_5)/V_5) = \dim(\mathbb{C}[x, y]/W_3) = 2(m + n).
\]

Therefore, \( \dim(W_3) = 2(m + n) \), and so if \( q \in \mathbb{C}[x, y] \), the degree of \( q \) in \( x \) is \( \leq m + n \) and the degree of \( q \) in \( y \) is \( \leq m + n \), then there exist unique \( q_j \in V_j \) for \( 1 \leq j \leq 4 \). Since \( W_3 \cap V_5 = \{0\} \), every polynomial \( q \) equals \( q_1 + q_2 + q_3 + q_4 + q_5 \) for unique choices of \( q_j \in V_j \) for \( 1 \leq j \leq 5 \).

We now return to the second assertion of the lemma. Let \( I \) be the ideal generated by \((x - \lambda)^m\) and \((y - \overline{\lambda})^m\). Since \( V_1 + V_2 + V_3 + V_5 \subseteq I \), \( V_1 + V_2 + V_3 + V_4 + V_5 = \mathbb{C}[x, y] \), we have \( I = V_1 + V_2 + V_3 + V_5 \) if and only if \( V_4 \cap I = \{0\} \). Now, if \( V_4 \cap I = \{0\} \), then there exists \( p_4 \in \mathbb{C}[x, y] \) such that \( p_4 \) is not the zero polynomial, \((x - \mu)^n(y - \overline{\mu})^np_4 \in I \), \( \deg_x(p_4) < m \) and \( \deg_y(p_4) < m \). By \( 2n \) applications of Lemma 76, \( p_4 \in I \). Every nonzero element of \( q \) of \( I \), however, has \( \deg_x(q) \geq m \) or \( \deg_y(q) \geq m \), a contradiction. Therefore, \( V_4 \cap I = \{0\} \) and \( I = V_1 + V_2 + V_3 + V_5 \).

A similar argument shows that if \( I \) is the ideal generated by \((x - \mu)^n\) and \((y - \overline{\mu})^n\), then \( I = V_2 + V_3 + V_4 + V_5 \). \( \blacksquare \)

An easy corollary of Lemma 83 is the following proposition.

**Proposition 86.** Let \( \lambda \in \mathbb{C} \) and \( \mu \in \mathbb{C} \) be such that \( \lambda \neq \mu \). Let \( m \) and \( n \) be positive integers. If \( I_1 \) is the ideal generated by \((x - \lambda)^m, (y - \overline{\lambda})^m\), \( I_2 \) is the ideal generated by \((x - \mu)^n, (y - \overline{\mu})^n\), and \( I_3 \) is the ideal generated by \((x - \lambda)^m(y - \overline{\mu})^n, (x - \mu)^n(y - \overline{\lambda})^m\), then \( I_3 = I_1 \cap I_2 \).

**Lemma 87.** If \( \alpha \in \mathbb{C}[x, y], \ p \in \mathbb{C}[x, y] \), and \( p = p^\vee \), then \( p \in \langle \{\alpha, \alpha^\vee\} \rangle \), the ideal generated by \( \alpha \) and \( \alpha^\vee \), if and only if there exists \( p_0 \in \mathbb{C}[x, y] \) such that

\[
(88) \quad p = \alpha p_0 + \alpha^\vee p_0^\vee.
\]
Here, \( p_0 \) need not equal \( p_0^\vee \).

**Proof.** Clearly, if (88) holds, then \( p \in \langle \{\alpha, \alpha^\vee\} \rangle \). On the other hand, if \( p \in \langle \{\alpha, \alpha^\vee\} \rangle \), then there exist \( p_1, p_2 \in C[x, y] \) such that

\[
p = \alpha p_1 + \alpha^\vee p_2.
\]

Since \( p = p^\vee \), we also have

\[
p = \alpha^\vee p_1^\vee + \alpha p_2^\vee.
\]

Combining the above two formulas for \( p \) gives

\[
p = \alpha \frac{p_1 + p_2^\vee}{2} + \alpha^\vee \frac{p_1^\vee + p_2}{2},
\]

and we see that (88) holds with \( p = (p_1 + p_2^\vee)/2 \). \( \blacksquare \)

5. **PROPERTIES OF REPRESENTABLE GRIDS**

This section involves an idea for analyzing the representability and the robust representability of grids using the intrinsic geometry of the grid itself.

**Definition 89.** If \((r, s)\) is an ordered pair of integers, then we can define a mapping \( S_{(r, s)} \) of the subsets \( X \) of \( \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\} \) by the formula

\[
S_{(r, s)}(X) = (X + (r, s)) \cap \{(0, \ldots, m - 1) \times \{0, \ldots, n - 1\}\},
\]

where \( X + (r, s) = \{(i + r, j + s) : (i, j) \in X\} \). We say a mapping \( S \) on subsets of \( \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\} \) is a **shift** if \( S = S_{(r, s)} \) for some integers \( r \) and \( s \). If \( S = S_{(r, s)} \) is a shift, then we say \( S \) is a **left** (respectively **right**) shift if \( r = 0 \) and \( s \leq 0 \) (respectively \( s \geq 0 \)). We say \( S \) is an **upward** (respectively **downward**) shift if \( s = 0 \) and \( r \leq 0 \) (respectively \( r \geq 0 \)). We say \( S \) is a **forward** (respectively **backward**) shift if \( r \geq 0 \) and \( s \geq 0 \) (respectively \( r \leq 0 \) and \( s \leq 0 \)). Finally, if \( X \subseteq \{0, \ldots, m - 1\} \times \{0, \ldots, n - 1\} \), we say \( X \) is **left** (respectively **right**, **upward**, **downward**, **forward**, **backward**) **invariant** if \( S(X) \subseteq X \) whenever \( S \) is a left (respectively right, upward, downward, forward, backward) shift.
One necessary condition for representability in terms of shifts is given by the following result.

**Theorem 90 (Shift rules).** If G is a representable \((m, n)\) grid, then no forward shift of \(\mathcal{P}_0(G)^c\) intersects \(\mathcal{P}_\perp(G)^c\) in exactly one point.

Suppose that \(G\) is representable. By Theorem 54, there exists a solution \(\{(d_{ij}, y_{ij})\}\) to the set of bi-Hankel equation subordinate to \(\mathcal{P}_*(G)\) such that (55) and (56) hold. We argue by contradiction. Accordingly, assume that \(S_{(i, j)}\) is a forward shift and \(S_{(r, s)}(\mathcal{P}_0(G)^c)\) meets \(\mathcal{P}_\perp(G)^c\) in exactly one point, say \((i_0 + r, j_0 + s)\). Thus,

\[
S_{(r, s)}(\mathcal{P}_*(G)) \cap \mathcal{P}_*(G) = \{(i_0 + r, j_0 + s)\}. \tag{91}
\]

Now, since a solution of the bi-Hankel equations must satisfy condition (1) from Definition 52, we see that (53) can be written after substitution of \(\{d_{ij}\}\) and \(\{y_{ij}\}\) as

\[
\sum_{(i, j) \in \mathcal{P}_*(G)} d_{i, j}y_{i+r, j+s} = 0.
\]

Thus, (91) implies that

\[
d_{i_0, j_0} y_{i_0+r, j_0+s} = 0.
\]

But (91) also implies that \(\mathcal{P}_0(G) \cup \{(i_0, j_0)\}\) is backward invariant and \(\mathcal{P}_\perp(G) \cup \{(i_0 + r, j_0 + s)\}\) is forward invariant, so that by (55) and (56) \(d_{i_0, j_0} \neq 0\) and \(y_{i_0+r, j_0+s} \neq 0\). This contradiction establishes Theorem 90. \(\blacksquare\)

Note that Theorem 90 allows one to immediately see the examples given in (66) are indeed not representable. The authors know of no other geometric necessary condition for representability. This would be a highly interesting area for further research.

Perhaps a more promising approach would be the attempt to find geometric conditions for robust representability. Here the idea would be to find more general shift obstructions than appear in Theorem 90. Theorem 90 involved a single shift that isolated a single unsolvable bi-Hankel equation. More generally, one could consider \(k\) shifts that single out a system of \(k\) unsolvable bi-Hankel equations. Here, the system would in general only be
unsolvable generally, and thus the conditions would only be approving in the context of robust representability. Specifically, we propose the following conditions as worthy of analysis in the context of matricially robustly representable grids $G$.

**Conditions 92** (General shift conditions). There exists $X \subseteq \mathcal{P}_*(G)$ such that $X$ has $k$ points and there are $k$ different forward shifts $S_1, \ldots, S_k$ which satisfy

\[(93) \ S_p(X) \cap \mathcal{P}_*(G) \neq \emptyset \quad \text{and} \quad S_p(\mathcal{P}_*(G)/X) \subset \mathcal{P}_\perp(G).\]

To see why one might believe that the above shift conditions are incompatible with robust matricial representability, consider the grid $G$ of (65). Here, if $X = \{(1, 0), (0, 1)\}$, $S_1 = S_{(1, 0)}$, and $S_2 = S_{(0, 1)}$, then it is clear that Conditions 92 hold with $k = 2$. Observe that corresponding to the two shifts are the two bi-Hankel equations

\[c_{01}x_{02} + c_{10}x_{11} = 0\]

\[c_{01}x_{11} + c_{10}x_{20} = 0.\]

Since the above equations have a nonzero solution if and only if $x_{11}^2 - x_{02}x_{20} = 0$, it is clear that $G$ is representable (simply choose $x_{11}^2 = x_{02}x_{20}$). On the other hand, since for generic $T$ one has $x_{11}^2 - x_{02}x_{20} \neq 0$, $G$ cannot be matricially robustly representable.

For a number of small sized grids similar reasoning proves Conditions 92 both necessary and sufficient for nonmatricially robust representability. Dual shift conditions could be introduced that would be connected with polynomially robust representability.

**References**

8 J. Agler, Family theory, Preprint.


