Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

ELA
http://math.technion.ac.il/iic/ela

ON THE CONSTRUCTION OF EXPLICIT SOLUTIONS TO THE MATRIX EQUATION $X^2AX = AXA$*

AIHUA LI† AND EDWARD MOSTEIG‡

Abstract. In a previous article by Aihua Li and Duane Randall, the existence of solutions to certain matrix equations is demonstrated via nonconstructive methods. A recurring example appears in that work, namely the matrix equation $AXA = X^2AX$, where A is a fixed, square matrix with real entries and X is an unknown square matrix. In this paper, the solution space is explicitly constructed for all 2×2 complex matrices using Gröbner basis techniques. When A is a 2×2 matrix, the equation $AXA = X^2AX$ is equivalent to a system of four polynomial equations. The solution space then is the variety defined by the polynomials involved. The ideal of the underlying polynomial ring generated by the defining polynomials plays an important role in solving the system. In our procedure for solving these equations, Gröbner bases are used to transform the polynomial system into a simpler one, which makes it possible to classify all the solutions. In addition to classifying all solutions for 2×2 matrices, certain explicit solutions are produced in arbitrary dimensions when A is nonsingular. In higher dimensions, Gröbner bases are extraordinarily computationally demanding, and so a different approach is taken. This technique can be applied to more general matrix equations, and the focus here is placed on solutions coming from a particular class of matrices.

Key words. Matrix equation, Ideal, Gröbner bases.

AMS subject classifications. 39B42, 15A24, 12Y05, 13B25, 13F20.

1. Introduction.

Let A be a fixed square matrix in $M_n(\mathbb{C})$. The objective of this paper is to produce matrices $X \in M_n(\mathbb{C})$ that satisfy the matrix equation

$$AXA = X^2AX. \tag{1.1}$$

The existence of solutions to (1.1) was originally established by Li and Randall in [5] through the use of topological techniques. In that article, the authors investigated the existence of real, $n \times n$ matrices X that satisfy matrix equations of the form

$$F(X, A_1, A_2, \ldots, A_s) = G(X, A_1, A_2, \ldots, A_s), \tag{1.2}$$

where $A_1, A_2, \ldots, A_s$ are fixed $n \times n$ matrices with real entries, and $F(x, z_1, z_2, \ldots, z_s)$ and $G(x, z_1, z_2, \ldots, z_s)$ are monomials in the polynomial ring $\mathbb{R}[x, z_1, \ldots, z_s]$. Specifically, Li and Randall demonstrated the existence of solutions to (1.2) under various conditions by using the Borsak-Ulam Theorem and the Lefschetz Fixed Point Theorem.

*Received by the editors on June 15, 2009. Accepted for publication on July 31, 2010. Handling Editors: Roger A. Horn and Fuzhen Zhang.

†Department of Mathematical Science, Montclair State University, Montclair, NJ 07043, USA (lia@mail.montclair.edu).

‡Department of Mathematics, Loyola Marymount University, Los Angeles, CA 90045, USA (emosteig@lmu.edu).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

Solving the matrix equation (1.1) is equivalent to solving a system of n^2 polynomial equations in which all monomials involved are of degree either 1 or 3. Seeking solutions to such systems is not an easy task, even when n is small. For example, consider the case $n = 2$ and let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \text{and} \qquad X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

One of the four polynomial equations we need to solve is the following:

$$\begin{aligned} &ax_1^3 + bx_1^2x_3 + ax_1x_2x_3 + bx_2x_3^2 + cx_1^2x_2 + cx_1x_2x_4 + dx_1x_2x_3 + dx_2x_3x_4 \\ &-a^2x_1 - abx_3 - acx_2 - bcx_4 = 0. \end{aligned}$$

Li and Randall discussed in [5] the existence of solutions to (1.1) for certain types of matrices A and constructed some solutions using linear algebraic techniques.

In this paper, we construct all solutions explicitly to (1.1) when $n = 2$, with the aid of Gröbner basis techniques. In this case, the solution space is the variety defined by four polynomials $f_1, f_2, f_3, f_4 \in \mathbb{C}[x_1, x_2, x_3, x_4]$; that is,

$$V(f_1, f_2, f_3, f_4) = \{(x_1, x_2, x_3, x_4) \mid f_i(x_1, x_2, x_3, x_4) = 0 \ \text{ for all } i = 1, 2, 3, 4\}.$$

Let $I = \langle f_1, f_2, f_3, f_4 \rangle$ be the ideal of $\mathbb{C}[x_1, x_2, x_3, x_4]$ generated by the defining polynomials. This ideal plays an important role in solving the system because $V(f_1, f_2, f_3, f_4) = V(I)$, where

$$V(I) = \{(x_1, x_2, x_3, x_4) \mid f(x_1, x_2, x_3, x_4) = 0 \ \text{ for all } f \in I\}.$$

The main result needed for our analysis here is the Elimination Theorem, which is given below (for reference, see [1]).

THEOREM 1.1 (The Elimination Theorem). *Let k be a field, and let I be an ideal in $k[x_1, \ldots, x_n]$. Let G be a Gröbner basis for I with respect to the lexicographic order with $x_1 > x_2 > \cdots > x_n$. Then for any $1 \le t \le n$,*

$$G \cap k[x_t, \ldots, x_n]$$

is a Gröbner basis for the ideal

$$I \cap k[x_t, \ldots, x_n].$$

The key feature in our approach is the identification of appropriate subrings R of $\mathbb{C}[x_1, x_2, x_3, x_4]$ such that $I \cap R$ (called an elimination ideal) consists of "nice" generators g where the equation $g = 0$ can be easily solved. Thus the computation of solutions to the original system can be facilitated by transforming the system into a much simpler one. We use the computer algebra software *Singular* (see [2]) to construct these generators by searching through Gröbner bases of various elimination

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

ELA
http://math.technion.ac.il/iic/ela

ideals. The Elimination Theorem guarantees that a Gröbner basis of $I \cap R$ is simply a subset of a Gröbner basis of I, and hence, is easily computed. By transforming the original system through the computation of Gröbner bases of appropriate elimination ideals, we are able to classify all solutions of (1.1) when $n = 2$.

Similar computations can be made for higher dimensions, though the situation rapidly becomes much more complex. Instead of performing such computations in an attempt to classify all solutions for $n \geq 3$, we use a different technique to produce a small collection of solutions. In Section 3, we consider a special class of matrices, which we call Toeplitz matrices, from which we select solutions to (1.1).

To simplify matters in all dimensions, we consider similarity classes of the matrix A. Let $J \in M_n(\mathbb{C})$ be the Jordan normal form of A and let $P \in M_n(\mathbb{C})$ such that $A = PJP^{-1}$. Every solution $Y \in M_n(\mathbb{C})$ of the matrix equation $JYJ = Y^2JY$ yields a solution to $AXA = X^2AX$ by setting $X = PYP^{-1}$. Therefore, without loss of generality, we assume that A is in Jordan normal form throughout the entire paper.

2. The Two-Dimensional Case. In this section, we find solutions to (1.1) for 2×2 matrices. In particular, we begin by setting

$$X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

We consider two cases, depending on whether or not A is diagonalizable.

2.1. The Matrix A is diagonalizable. Since we are assuming that A is in Jordan normal form, if it is diagonalizable, then in fact, it is diagonal and hence can be written as

$$A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

where $a, d \in \mathbb{C}$. Upon comparing the scalar entries of both sides of (1.1), we find that $f_i = 0$ for $i = 1, 2, 3, 4$, where

$$\begin{aligned} f_1 &= ax_1^3 + ax_1x_2x_3 + dx_1x_2x_3 - a^2x_1 + dx_2x_3x_4, \\ f_2 &= ax_1^2x_2 + dx_1x_2x_4 + ax_2^2x_3 + dx_2x_4^2 - adx_2, \\ f_3 &= ax_1^2x_3 + ax_1x_3x_4 + dx_2x_3^2 + dx_3x_4^2 - adx_3, \\ f_4 &= ax_1x_2x_3 + ax_2x_3x_4 + dx_2x_3x_4 + dx_4^3 - d^2x_4. \end{aligned}$$

Depending on the eigenvalues a and d, the solutions to equation (1.1) can take on different forms. These various possibilities are described in Propositions 2.1, 2.2, 2.3, and 2.4. We begin by considering the case where A is diagonal with two distinct, nonzero eigenvalues.

Proposition 2.1. *Suppose $A = diag(a, d)$ where a and d are nonzero and distinct. Then X is a solution to (1.1) if and only if one of the following three conditions hold:*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

(i) The matrix X is of the form

$$\begin{bmatrix} x_1 & 0 \\ 0 & x_4 \end{bmatrix}, \tag{2.1}$$

where $x_1^2 \in \{0, a\}, x_4^2 \in \{0, d\}$.

(ii) The matrix X is of the form

$$\begin{bmatrix} x_1 & x_2 \\ 0 & \left(\frac{a}{a-d}\right) x_1 \end{bmatrix}, \tag{2.2}$$

where $a^3 - a^2d + 2ad^2 - d^3 = 0$, $x_1^2 = a$, and $x_2 \neq 0$.

(iii) The matrix X is of the form

$$\begin{bmatrix} x_1 & 0 \\ x_3 & \left(\frac{a-d}{d}\right) x_1 \end{bmatrix}, \tag{2.3}$$

where $a^3 - 2a^2d + ad^2 - d^3 = 0, x_1^2 = a$, and $x_3 \neq 0$.

Proof. First, it should be noted that all of the forms proposed for X in (2.1), (2.2) and (2.3) above satisfy (1.1), which can be verified through simple algebra.

We now proceed to show that any solution to (1.1) must be one of these forms. To accomplish this, we translate the problem into the language of polynomials. Let $R = \mathbb{C}[x_1, x_2, x_3, x_4]$. There is a one-to-one correspondence between solutions to (1.1) and points in the variety $V(f_1, f_2, f_3, f_4) = V(J)$, where $J = \langle f_1, f_2, f_3, f_4 \rangle$ is the ideal of R generated by f_1, f_2, f_3, f_4. Every point (x_1, x_2, x_3, x_4) satisfying $f_i(x_1, x_2, x_3, x_4) = 0$ for all $i = 1, 2, 3, 4$, also satisfies $g(x_1, x_2, x_3, x_4) = 0$ for all $g \in J$. Our goal is to find some special elements $g \in J$ such that the equation $g = 0$ can be easily solved. We seek such elements from Gröbner bases of various elimination ideals related to J. For our purposes, we are not only interested in the relations induced by J, but also wish to include a few extra conditions. In particular, we wish to require that $a \neq 0, d \neq 0$ and $a \neq d$. To accomplish this, we create an additional variable λ and impose the following condition:

$$\lambda ad(a - d) - 1 = 0.$$

This statement clearly only holds when a, d and $a - d$ are all nonzero. Conversely, whenever a, d and $a-d$ are all nonzero, there exists $\lambda \in \mathbb{C}$ that satisfies this equation. By treating λ, a, b as indeterminates over $\mathbb{C}$ and computing Gröbner bases of the ideal

$$I = \langle f_1, f_2, f_3, f_4, \lambda ad(a - d) - 1 \rangle$$

in $\mathbb{C}[x_1, x_2, x_3, x_4, \lambda, a, d]$ with respect to various monomial orders, we discover via the Elimination Theorem that

$$I \cap \mathbb{C}[x_4, \lambda, a, d] = \langle x_4(x_4^2 - d), \lambda ad(a - d) - 1 \rangle,$$
$$I \cap \mathbb{C}[x_1, \lambda, a, d] = \langle x_1(x_1^2 - a), \lambda ad(a - d) - 1 \rangle,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

and so

$$x_1^2 \in \{0, a\} \qquad \text{and} \qquad x_4^2 \in \{0, d\}. \tag{2.4}$$

Suppose, toward contradiction, that there exists a solution with $x_2 \neq 0$ and $x_3 \neq 0$. We compute the lexicographic Gröbner Basis of $\langle f_1, f_2, f_3, f_4, \lambda ad(a-d)x_2x_3 - 1\rangle$ with respect to $\lambda > x_1 > x_2 > x_3 > x_4 > a > d$, which turns out to be $\{1\}$. The corresponding variety $V(1)$ is empty, which is a contradiction, and so we may assume that at least one of x_2 and x_3 is zero. If $x_2 = x_3 = 0$, then X is of the form given in (2.1). Therefore, we need only consider the case when exactly one of x_2 and x_3 is zero.

Suppose $x_2 \neq 0$ and $x_3 = 0$. Using the order $\lambda > x_1 > x_2 > x_3 > x_4 > a > d$ we find that $a^3 - a^2d + 2ad^2 - d^3$, and $x_4(a-d) - x_1a$ are elements of the ideal

$$\langle f_1, f_2, f_3, f_4x_3, \lambda x_2ad(a-d) - 1\rangle,$$

and so $a^3 - a^2d + 2ad^2 - d^3 = 0$ and $x_4(a-d) = x_1a$. Thus, X must be of the form given in (2.2). Similarly, if $x_2 = 0$ and $x_3 \neq 0$, we find that X must be of the form given in (2.3). □

Proposition 2.2. *Suppose $A = diag(a,a)$ where a is nonzero.*

(a) If X is nonsingular, then it is a solution to (1.1) if and only if one of the following two conditions hold:

(i) The matrix X is of the form

$$X = \begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix}, \tag{2.5}$$

where $x_1^2 = a$.

(ii) The matrix X is of the form

$$\begin{bmatrix} x_1 & x_2 \\ x_3 & -x_1 \end{bmatrix}, \tag{2.6}$$

where $x_1^2 + x_2x_3 = a$.

(b) If X is singular, then it is a nontrivial solution (i.e., $X \neq 0$) to (1.1) if and only if it is of the form

$$X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \tag{2.7}$$

where $(x_1 + x_4)^2 = a$.

Proof. It should be noted that all of the forms proposed for X above satisfy (1.1), and so our objective is to prove the converse. Since $A = \text{diag}(a,a)$, it follows from (1.1) that

$$X^3 = aX.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

ELA
http://math.technion.ac.il/iic/ela

(a) If X is nonsingular, then $x_1x_4 - x_2x_3 \neq 0$, and so we consider the ideal

$$I = \langle f_1, f_2, f_3, f_4, \lambda(x_1x_4 - x_2x_3) - 1\rangle.$$

Using the Elimination Theorem with $\lambda > x_1 > x_2 > x_3 > x_4 > a$, we find that I contains the following polynomials:

$$x_2(x_1 + x_4),\ x_3(x_1 + x_4),\ (x_1 - x_4)(x_1 + x_4).$$

The presence of these three polynomials demonstrates that either $x_2 = x_3 = x_1 - x_4 = 0$ or $x_1 + x_4 = 0$. In the former case, if we compute a Gröbner basis for the ideal

$$I \cap \langle x_2, x_3, x_1 - x_4\rangle,$$

we discover that $x_4^2 - a \in I \cap \langle x_2, x_3, x_1 - x_4\rangle$, in which case $x_1^2 = a$, $x_1 = x_4$, $x_2 = x_3 = 0$, and so X is of the form given in (2.5).

In the latter case, if we compute a Gröbner basis for the ideal

$$I \cap \langle x_1 + x_4\rangle,$$

we discover that $x_2x_3 + x_4^2 - a \in I \cap \langle x_1 + x_4\rangle$, in which case $x_1 + x_4 = 0$ and $x_4^2 + x_2x_3 = a$. Since $x_1 = -x_4$, it follows that $x_1^2 + x_2x_3 = a$, and so X must be of the form given in (2.6).

(b) If X is singular, then $x_1x_4 - x_2x_3 = 0$. Using the order $\lambda > a > x_1 > x_2 > x_3 > x_4$, we discover via an application of the Elimination Theorem that the ideal

$$I \cap \langle \lambda a - 1, x_1x_4 - x_2x_3\rangle$$

contains all polynomials of the form $x_i\left((x_1 + x_4)^2 - a\right)$, where $i = 1, 2, 3, 4$. Thus, either $X = 0$ or $(x_1 + x_4)^2 = a$, and so X must be of the form given in (2.7). □

Next, we examine the case where A is a diagonal, singular matrix, which is considered in the following two propositions. The case where A is identically zero is not interesting since it admits all 2×2 matrices as solutions to (1.1).

Proposition 2.3. *Suppose $A = diag(a, 0)$ where $a \neq 0$. Then X is a solution to (1.1) if and only if X is of one of the following forms, where $x_1 = 0$ or $x_1^2 = a$:*

$$\begin{bmatrix} 0 & x_2 \\ 0 & x_4 \end{bmatrix},\quad \begin{bmatrix} 0 & 0 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} x_1 & 0 \\ 0 & x_4 \end{bmatrix},\quad \begin{bmatrix} x_1 & 0 \\ x_3 & -x_1 \end{bmatrix}. \tag{2.8}$$

Proof. It is not difficult to show that if X takes on any of the forms above, then it necessarily satisfies (1.1). Towards a justification of sufficiency, we compare the scalar entries of both sides of (1.1). Here we find that $f_i = 0$ for $i = 1, 2, 3, 4$, where

$$\begin{aligned} f_1 &= ax_1^3 + ax_1x_2x_3 - a^2x_1, \\ f_2 &= ax_1^2x_2 + ax_2^2x_3, \\ f_3 &= ax_1^2x_3 + ax_1x_3x_4, \\ f_4 &= ax_1x_2x_3 + ax_2x_3x_4. \end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

ELA
http://math.technion.ac.il/iic/ela

The lexicographic Gröbner Basis of $\langle f_1, f_2, f_3, f_4, \lambda a - 1\rangle$ with respect to $\lambda > x_1 > x_2 > x_3 > x_4 > a$ contains $x_2^2 x_3$, and so either $x_2 = 0$ or $x_3 = 0$. Since $f_1 = 0$, it follows that

$$ax_1(x_1^2 - a) = 0 \tag{2.9}$$

Case 1: If $x_1 = 0$, then since either $x_2 = 0$ or $x_3 = 0$, we know that X must be of one of the first two forms listed in (2.8).

Case 2: We now consider the case $x_1 \neq 0$. Since $a \neq 0$, it follows from (2.9) that $x_1^2 = a$. The lexicographic Gröbner Basis of $\langle f_1, f_2, f_3, f_4, \lambda a x_1 - 1, \rangle$ with respect to $\lambda > x_1 > x_2 > x_3 > x_4 > a$ contains the polynomials x_2 and $x_3(x_1 + x_4)$; thus, $x_2 = 0$, and either $x_3 = 0$ or $x_4 = -x_1$. Therefore, X must be of one of the latter two forms listed in (2.8). $\square$

Similarly, we can prove the following:

Proposition 2.4. *Suppose $A = diag(0, d)$ where $d \neq 0$. Then X is a solution to (1.1) if and only if X is of one of the following forms, where $x_4 = 0$ or $x_4^2 = d$:*

$$\begin{bmatrix} 0 & x_2 \\ 0 & x_4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ x_3 & x_4 \end{bmatrix}, \begin{bmatrix} x_1 & 0 \\ 0 & x_4 \end{bmatrix}, \quad \begin{bmatrix} -x_4 & x_2 \\ 0 & x_4 \end{bmatrix}. \tag{2.10}$$

2.2. The Matrix A is not diagonalizable. Since we are assuming that A is in Jordan normal form, if it is not diagonalizable, then it is a single Jordan block. This leads us to the final piece in terms of classifying all 2×2 solutions to (1.1).

Proposition 2.5. *Suppose A is not diagonalizable, in which case it must be of the form*

$$A = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}.$$

(a) If A is singular, then X is a solution to (1.1) if and only if it is of one of the following forms:

$$\begin{bmatrix} 0 & x_2 \\ 0 & x_4 \end{bmatrix}, \quad \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix}. \tag{2.11}$$

(b) If A is nonsingular, then X is a solution to (1.1) if and only if it is of the form

$$X = \begin{bmatrix} x_1 & \frac{1}{2x_1} \\ 0 & x_1 \end{bmatrix}, \tag{2.12}$$

where $x_1^2 = a$.

Proof. Again, one can easily show that if X takes on any of the forms above, then it necessarily satisfies (1.1), and so we only need to show the converse.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

(a) If A is singular, then $a = 0$, and we find that by comparing both sides of (1.1), it turns out that $f_i = 0$ for $i = 1, 2, 3, 4$, where

$$\begin{aligned}
f_1 &= (x_1^2 + x_2x_3)x_3,\\
f_2 &= -x_3 + x_1^2x_4 + x_2x_3x_4,\\
f_3 &= x_3^2(x_1 + x_4),\\
f_4 &= x_3x_4(x_1 + x_4).
\end{aligned}$$

Using the Elimination Theorem, we find that $\langle f_1, f_2, f_3, f_4\rangle$ contains x_3^2, and so $x_3 = 0$. Moreover, since $f_2 = 0$, it follows that $x_1^2x_4 = 0$, and so either $x_1 = 0$ or $x_4 = 0$. Thus, X must be of one of the forms given in (2.11).

(b) If A is nonsingular, then $a \neq 0$, and we find that by comparing both sides of (1.1), it turns out that $f_i = 0$ for $i = 1, 2, 3, 4$, where

$$\begin{aligned}
f_1 &= -a^2x_1 + ax_1^3 - ax_3 + x_1^2x_3 + 2ax_1x_2x_3 + x_2x_3^2 + ax_2x_3x_4,\\
f_2 &= -ax_1 - a^2x_2 + ax_1^2x_2 - x_3 + ax_2^2x_3 - ax_4 + x_1^2x_4 + ax_1x_2x_4 + x_2x_3x_4 + ax_2x_4^2,\\
f_3 &= -x_3(a^2 - ax_1^2 - x_1x_3 - ax_2x_3 - ax_1x_4 - x_3x_4 - ax_4^2),\\
f_4 &= -ax_3 + ax_1x_2x_3 - a^2x_4 + x_1x_3x_4 + 2ax_2x_3x_4 + x_3x_4^2 + ax_4^3.
\end{aligned}$$

Suppose, toward contradiction, that $x_3 \neq 0$. We compute the lexicographic Gröbner Basis of $\langle f_1, f_2, f_3, f_4, \lambda a x_3 - 1\rangle$ with respect to $\lambda > x_1 > x_2 > x_3 > x_4 > a$, which turns out to be $\{1\}$. Thus, the corresponding variety $V(1)$ is empty, and so there are no solutions when $x_3 \neq 0$.

If $x_3 = 0$, then f_1 through f_4 can be rewritten as

$$\begin{aligned}
g_1 &= ax_1(a - x_1^2),\\
g_2 &= ax_1 + a^2x_2 + ax_4 - ax_1^2x_2 - x_1^2x_4 - ax_1x_2x_4 - ax_2x_4^2,\\
g_3 &= 0,\\
g_4 &= ax_4(a - x_4^2).
\end{aligned}$$

Since $g_1 = g_4 = 0$, we conclude $x_1^2 = x_4^2 = a$. The lexicographic Gröbner Basis of $\langle g_1, g_2, g_3, g_4, x_1^2 - a, x_4^2 - a, \lambda a - 1\rangle$ with respect to $\lambda > x_1 > x_2 > x_3 > x_4 > a$ contains the polynomials $x_1 - x_4$ and $2x_2x_4 - 1$. Therefore, $x_1 = x_4$ and $x_4 = \frac{1}{2x_1}$, and so X must be of the form given in (2.12). $\square$

3. Solutions in Higher Dimensions. Although the techniques in this section are used to construct solutions to (1.1), they are sufficiently general that they can be applied to any matrix equation of the form $F(X, A) = G(X, A)$ where F and G are

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

monomials. Since we are assuming that A is in Jordan normal form, we can write

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{bmatrix}$$

where each A_i is a Jordan block. Note that if X_i is a solution to $X_i^2 A_i X_i = A_i X_i A_i$, then

$$X = \begin{bmatrix} X_1 & & & \\ & X_2 & & \\ & & \ddots & \\ & & & X_m \end{bmatrix}$$

is a solution to $X^2AX = AXA$.[1] Thus, we can construct solutions to (1.1) of the block matrix form as shown above. From now on, we assume A is a Jordan block of the form

$$\mathbf{A} = \begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & \ddots & & \\ & & & a & 1 \\ & & & & a \end{bmatrix}$$

where $a \neq 0$ (and hence is nonsingular).

Definition 3.1. *(Refer to [4].) We say that an $n \times n$ matrix $M = \{m_{ij}\}$ is* **Toeplitz** *if for all $1 \leq i, j, k, \ell \leq n$, we have*

$$x_{i,j} = x_{k,\ell}$$

whenever $i - j \equiv k - l \mod n$; that is, the entries along any diagonal are identical. All Toeplitz matrices are of the form

$$\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & a_n \\ a_{-1} & a_0 & a_1 & \cdots & a_{n-3} & a_{n-2} & a_{n-1} \\ a_{-2} & a_{-1} & a_0 & \cdots & a_{n-4} & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{-(n-2)} & a_{-(n-3)} & a_{-(n-4)} & \cdots & a_0 & a_1 & a_2 \\ a_{-(n-1)} & a_{-(n-2)} & a_{-(n-3)} & \cdots & a_{-1} & a_0 & a_1 \\ a_{-n} & a_{-(n-1)} & a_{-(n-2)} & \cdots & a_{-2} & a_{-1} & a_0 \end{bmatrix}.$$

If the matrix above is upper-triangular (i.e., $a_i = 0$ for $i < 0$), we denote it by

$$\mathfrak{t}[a_0, a_1, a_2, \ldots, a_{n-1}].$$

[1]Note that not all solutions to $X^2AX = AXA$ will be of this form. In the 2×2 case, this is demonstrated in Section 2.1 where A is diagonalizable, or equivalently, where A is comprised of two Jordan blocks of dimension one.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

Each solution X to (1.1) that we produce in this section will be an upper-triangular, Toeplitz matrix in $M_n(k)$. To better understand the arithmetic of such matrices, we state the following simple lemma without proof.

Lemma 3.2. *For any field k, let $\mathcal{T}_n(k)$ denote the space of all upper-triangular, Toeplitz $n \times n$ matrices. The function*

$$\begin{aligned} \mathcal{T}_n(k) &\to k[x]/\langle x^n \rangle \\ \mathfrak{t}\,[a_0, a_1, a_2, \ldots, a_{n-1}] &\mapsto \sum_{i=0}^{n-1} a_i x^i \end{aligned}$$

is a ring isomorphism.

As a consequence, upper-triangular Toeplitz matrices commute. In particular, if A is the Jordan block $[a, 1, 0, 0, 0 \ldots, 0]$ and $X = \mathfrak{t}[c_0, c_1, c_2, \ldots, c_{n-1}]$ is an upper-triangular Toeplitz matrix, they must commute. Suppose further that $X^2AX = AXA$ with $a \neq 0$ so A is non-singular. Since A and X commute, we know that $X^2(AX) = A(AX)$, and so $(X^2 - A)(AX) = 0$. However, since A is nonsingular, it follows that

$$(X^2 - A)X = 0. \tag{3.1}$$

Note that the matrix solution X can only have the following two forms:

Case 1: $X^2 = A$ when X is non-singular. This can be seen immediately from (3.1).

Case 2: $X = 0$ when X is singular. Indeed, if X is singular, then $c_0 = 0$, and so $X^2 - A$ is non-singular because the diagonal elements of $X^2 - A$ are $c_0^2 - a = -a \neq 0$. Thus from equation (3.1) again, $X = 0$.

Now, much has been written on the computation of square roots of matrices (for example, see [3] and [6]). For our case in particular, suppose we are given a nonsingular Jordan block whose eigenvalue a is a perfect square in the field of constants (which is $\mathbb{C}$ in our case). It can be shown that this Jordan block has exactly two square roots, which are necessarily Toeplitz. The approach we take to constructing square roots involves computing Taylor series and can be applied more generally to any equation of the form $F(X, A) = G(X, A)$ where F and G are monomials. We begin with a preparatory lemma that allows us to translate the problem to the language of polynomials.

Lemma 3.3. *If*

$$x^n \,\Bigg|\, \left(\left(\sum_{i=0}^{n-1} c_i x^i \right)^2 - (a + x) \right), \tag{3.2}$$

then $X = \mathfrak{t}[c_0, c_1, \ldots, c_{n-1}]$ is a solution to (1.1).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

Proof. Condition (3.2) can be rewritten as $\left(\sum c_i x^i\right)^2 \equiv a+x \mod x^n$. Under the isomorphism in Lemma 3.2, this translates to $X^2 = A$, and so (1.1) directly follows. ∎

The following lemma follows directly from the Factor Theorem (i.e., for any polynomial $P(x)$ with coefficients in a commutative ring, if $P(\alpha) = 0$, then $x - \alpha$ divides $P(x)$).

Lemma 3.4. *Let $P(x) \in R[x]$, where R is a commutative ring. Then for any $A, B \in R$, we have*

$$(A - B) \mid (P(A) - P(B)).$$

This leads us to the following proposition, which allows us to find solutions to (1.1) by examining appropriate power series.

Proposition 3.5. *Let $P(x), F(x) \in \mathbb{C}[x]$. Suppose $\tilde{Q}(x) = \sum_{i=0}^{\infty} \beta_i x^i \in \mathbb{C}[[x]]$ is a formal power series such that $P(\tilde{Q}(x)) = F(x)$. For each positive integer n, define $Q_n(x) \in \mathbb{C}[x]$ to be the finite partial series given by $Q_n(x) = \sum_{i=0}^{n-1} \beta_i x^i$. Then for any positive integer n, we have*

$$x^n \mid P(Q_n(x)) - F(x).$$

Proof. Now, $P(\tilde{Q}(x)) = F(x)$, and so it follows that

$$x^n \mid P(\tilde{Q}(x)) - F(x). \tag{3.3}$$

From the previous lemma, we see that $P(\tilde{Q}(x)) - P(Q_n(x))$ is divisible by $\tilde{Q}(x) - Q_n(x)$. Since x^n clearly divides $\tilde{Q}(x) - Q_n(x) = \sum_{i=n}^{\infty} \beta_i x^i$, we have

$$x^n \mid P(\tilde{Q}(x)) - P(Q_n(x)). \tag{3.4}$$

The conclusion follows from (3.3) and (3.4). ∎

Using Proposition 3.5, we can construct solutions to (1.1).

Theorem 3.6. *The uppper-triangular, Toeplitz matrix $X = \mathfrak{t}[c_0, c_1, c_2, \ldots, c_{n-1}]$ is a solution to (1.1), where*

$$c_i = \frac{(-1)^i (2i)! a^{1/2}}{(1-2i) i!^2 (4a)^i}.$$

Proof. From calculus, the Taylor series expansion of $\sqrt{x+a}$ about $x = 0$ is of the form $\sum_{i=0}^{\infty} c_i x^i$ where

$$c_i = \frac{(-1)^i (2i)! a^{1/2}}{(1-2i) i!^2 (4a)^i}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 21, pp. 142-153, October 2010

The careful reader will note that this is where we really capitalize on the fact that A is nonsingular. Now define $\tilde{Q}(x) = \sum_{i=0}^{\infty} c_i x^i$, $F(x) = a + x$, and $P(x) = x^2$. Since $\tilde{Q}(x)$ is the Taylor expansion of $\sqrt{x+a}$ about $x = 0$, it follows that

$$P(\tilde{Q}(x)) = F(x).$$

Thus, an application of Proposition 3.5 yields

$$x^n \mid P(Q_n(x)) - F(x),$$

which can be rewritten as

$$x^n \left| \left(\left(\sum_{i=0}^{n-1} c_i x^i \right)^2 - (a + x) \right) \right. .$$

Therefore, by Lemma 3.2, it follows that $X = \mathfrak{t}[c_0, c_1, \ldots, c_{n-1}]$ is a solution to equation (1.1) $\square$

REFERENCES

[1] David Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms, Third Edition*, Springer, New York, 2007.

[2] G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3-1-0 — A computer algebra system for polynomial computations. http://www.singular.uni-kl.de (2009).

[3] Mohammed A. Hasan. Power Method for Computing Square Roots of Complex Matrices. *Journal of Mathematical Analysis and Applications*, 213:393–405, 1997.

[4] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, USA, 1990.

[5] Aihua Li and Duane Randall. Non-Trivial Solutions to Certain Matrix Equations. *Electronic Journal of Linear Algebra*, 9:282–289, 2002.

[6] Zhongyun Liu, Yulin Zhang, and Rui Ralha. Computing the square roots of matrices with central symmetry. *Applied Mathematics and Computations*, 186:715–726, 2007.