On some properties of Toeplitz matrices

Dan Kucerovsky¹, Kaveh Mousavand² and Aydin Sarraf³*

Abstract: In this paper, we investigate some properties of Toeplitz matrices with respect to different matrix products. We also give some results regarding circulant matrices, skew-circulant matrices and approximation by Toeplitz matrices over the field of complex numbers.

Subjects: Advanced Mathematics; Algebra; Analysis - Mathematics; Functional Analysis; Linear & Multilinear Algebra; Mathematics & Statistics; Operator Theory; Science

Keywords: Toeplitz matrices; circulant matrices

1. Introduction

Toeplitz matrices are important both in theory and application. For example, it is known that a large class of matrices are similar to Toeplitz matrices (Heinig, 2001; Mackey, Mackey, & Petrovic, 1999). Moreover, it is shown that every matrix is a product of Toeplitz matrices (Lim & Ye, 2013).

The notation of this paper is as follows. An \( n \times n \) Toeplitz matrix \( T \) is denoted by \( T = \begin{bmatrix} t_{ij} \end{bmatrix} \) for \( 1 \leq i, j \leq n \), which implies all the entries along each of the \( 2n - 1 \) diagonals are the same. A circulant matrix \( C \) is a Toeplitz matrix where each row is obtained by applying the permutation \((1 \ 2 \ \ldots \ n)\) to the previous row. A skew-circulant matrix \( S \) differs from a circulant matrix \( C \) only by a change in the sign in the subdiagonal entries, i.e. \( s_{ij} = c_{ij} \) when \( j \geq i \) and \( s_{ij} = -c_{ij} \) when \( j < i \). A matrix \( T \) is involutory if \( T^2 = I \), and is called nilpotent of degree \( k \) if \( T^k = 0 \), but \( T^{k-1} \neq 0 \).

This paper is organized as follows. In Section 2, we characterize those Toeplitz matrices over a commutative ring (of any characteristic) having the property that their product is a Toeplitz matrix.
We consider also Kronecker products of Toeplitz matrices. Furthermore, we also give some combinatorial results over the finite prime field $\mathbb{Z}_p$.

In Section 3, we count the number of involutary and degree two nilpotent Toeplitz matrices in $M_n(\mathbb{Z}_p)$ for particular values of $n$ or $p$.

In Section 4, Toeplitz matrices over the field of complex numbers are studied. We show that for any matrix, there exists a closest Toeplitz matrix (with respect to the Frobenius norm) that approximates it. We describe the equivalence classes of Toeplitz matrices and give several results regarding circulant and skew-circulant matrices.

2. Product of Toeplitz matrices over a commutative ring

We consider three different binary operations, the usual matrix multiplication, Kronecker product and Schur product, denoted by $AB$, $A \otimes B$ and $A \circ B$, respectively. With the exception of Schur product, the set of Toeplitz matrices is neither closed under the usual matrix multiplication nor under the Kronecker product. We characterize those pairs of Toeplitz matrices whose product with respect to the usual multiplication or the Kronecker product is Toeplitz again. The following theorem gives a characterization in terms of the usual matrix multiplication:

**Theorem 2.1** If $A = \{a_{i,j}\}$ and $B = \{\beta_{i,j}\}$ are two $n \times n$ Toeplitz matrices over a commutative ring, then $AB$ is a Toeplitz matrix if and only if the following system of equations, of $(n-1)^2$ equations with $4(n-1)$ variables, holds:

$$a_i \beta_{j} - a_{i-n} \beta_{n-j} = 0 \text{ where } 1 \leq i, j \leq n-1$$

**Proof** If we set $C = AB$, by the product formula we have:

$$c_{i+j+1} = \sum_{k=1}^{n} a_{i+k} \beta_{k+j+1} = a_i \beta_{j+1} + \sum_{k=2}^{n} a_{i+k-1} \beta_{k-j+1}$$

$$= a_i \beta_{j+1} + \sum_{k=1}^{n-1} a_{i+k} \beta_{k} = a_i \beta_{j} - a_{i-n} \beta_{n-j} + \sum_{k=1}^{n} a_{i-k} \beta_{k} = a_i \beta_{j} - a_{i-n} \beta_{n-j} + c_j$$

By definition, $C$ is a Toeplitz matrix if and only if $c_{i+j+1} = c_j$. Therefore, $C$ is Toeplitz if and only if $a_i \beta_{j} - a_{i-n} \beta_{n-j} = 0 \text{ where } 1 \leq i, j \leq n-1$.

It is clear that the inverse of an arbitrary nonsingular Toeplitz matrix is not necessarily Toeplitz. However, the following proposition shows how the stability under inversion, with respect to being Toeplitz, is related to stability under multiplication. For an arbitrary $n \times n$ matrix $A$ over a commutative ring, we can compute the coefficients of the characteristic polynomial $p_A(x)$ in many different ways, including the eigenvalues of the matrix, or the entries of the matrix (Brooks, 2006). In particular, we wish to use the following formula to compute the coefficients:

$$p_A(x) = \sum_{m=0}^{n} x^{n-m} (-1)^m \text{tr}(\Lambda^m A)$$

where $\Lambda^m A$ is the $m^{th}$ exterior power of $A$. Furthermore, if the characteristic of the commutative ring is zero then it is known (Winitzki, 2010, 3.9) that $\text{tr}(\Lambda^m A) = \frac{1}{m!} \det(C)$ where

$$C = \begin{bmatrix}
\text{tr}(A) & m-1 & 0 & \ldots & \ldots & \ldots & 0 \\
\text{tr}(A^2) & \text{tr}(A) & m-2 & 0 & \ldots & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\
\text{tr}(A^{m-1}) & \text{tr}(A^{m-2}) & \ldots & \ldots & \ddots & \ddots & 1 \\
\text{tr}(A^m) & \text{tr}(A^{m-1}) & \ldots & \ldots & \text{tr}(A^2) & \text{tr}(A) & \text{tr}(A)
\end{bmatrix}$$
Proposition 2.2  For an \( n \times n \) invertible Toeplitz matrix \( A = \{ a_{ij} \} \) over a commutative ring, if every \( A^m \), for \( m = 2, \ldots, n - 1 \), is Toeplitz, then \( A^{-1} \) is a Toeplitz matrix. Furthermore, if the Toeplitz matrix is considered over a ring of characteristic zero, we have \( \text{tr}(A^m) = n(A^m)_0 \) where

\[
(A^m)_0 = \sum_{k_{n-1}=0}^{n-1} a_{k_{n-1}} \sum_{k_{n-2}=0}^{n-1} a_{k_{n-2}-k_{n-1}} \sum_{k_{n-3}=0}^{n-1} a_{k_{n-3}-k_{n-2}} \cdots ,
\]

\( k_0 = 0 \) and the symbol \( \sum_{k_0}^{n-1} \) disappears.

Proof  By the Cayley–Hamilton Theorem, \( A^{-1} = \frac{(-1)^{n-1}}{\det(A)} (A^{n-1} + c_{n-1} A^{n-2} + \cdots + c_1 I) \), where each coefficient \( c_{n-m} \) of the characteristic polynomial could be explicitly computed as the sum of all principal minors of \( A \) of size \( m \). Furthermore, if the matrix is considered over a commutative ring of characteristic zero, then we have \( c_{n-m} = \frac{(-1)^{n-m}}{m} \det(C) \). The trace of a Toeplitz matrix is simply \( n \) times its diagonal entry, but the diagonal entry of the \( m \)th power of \( A \), denoted by \( (A^m)_0 \) here, can be calculated recursively.

Remark 2.3  If an upper (lower) triangular Toeplitz matrix is invertible, then its inverse is Toeplitz, because the product of two upper (lower) triangular Toeplitz matrices is again an upper (lower) triangular Toeplitz matrix. Since an upper (lower) triangular Toeplitz matrix is always invertible and its inverse is an upper (lower) unitriangular matrix, the inverse of any upper (lower) unitriangular Toeplitz matrix is also an upper (lower) unitriangular Toeplitz matrix.

Theorem 2.4  Let \( A = \{ a_{ij} \} \) be an \( n \times n \) Toeplitz matrix and \( B = \{ b_{ij} \} \) be an \( m \times m \) Toeplitz matrix over a commutative ring. The \( mn \times mn \) matrix \( A \otimes B \) is Toeplitz if and only if the following system of equations, of \( 2(n-1)(m-1) \) equations with \( 2(n+m) - 3 \) variables, holds:

\[
\begin{align*}
\alpha_i \beta_j &= \begin{cases} 
\alpha_{i-1} \beta_{m+j} & j < 0 \text{ and } i \neq 1 - n \\
\alpha_{i+1} \beta_{j+m} & j > 0 \text{ and } i \neq n - 1 
\end{cases}
\end{align*}
\]

where \( i \in \{ 1 - n, 2 - n, \ldots, n - 2, n - 1 \} \) and \( j \in \{ 1 - m, 2 - m, \ldots, m - 2, m - 1 \} \) – \( \{0\} \).

Proof  With regard to the Kronecker multiplication of two matrices, \( A \otimes B \) consists of \( n^2 \) blocks of Toeplitz matrices, and it is Toeplitz if and only if each of the \( (2nm - 1) \) diagonals has the same entries, i.e., if and only if for each diagonal, which is formed by concatenation of diagonals of the adjacent blocks, the entries agree with each other.

Clearly, for the top right corner block in \( A \otimes B \), there is no concatenation, and therefore, the top \( m - 1 \) diagonals always satisfy the property, and the same fact holds for the bottom \( m - 1 \) diagonals of \( A \otimes B \). Furthermore, the main diagonal of \( A \otimes B \) is just the concatenation of \( n \) copies of the main diagonal of \( B \), multiplied by \( a_0 \), i.e. it is simply \( a_0 \beta_0 \). We also notice that for the diagonals of \( A \otimes B \) that are the result of concatenation of the main diagonals of some internal blocks, we have no extra conditions and each of them always has the same entry since \( A \) and \( B \) are both Toeplitz.

Regarding the remaining diagonals, however, we certainly have concatenation of different diagonals of adjacent blocks, which should agree. A simple computation shows that the above equations are coming from those concatenations. Thus, out of the \( 2nm - 1 \) diagonals of \( A \otimes B \), always \( (2n - 1) + 2(m - 1) \) of them satisfy the property, from which \( 2n - 1 \) comes from the fact that holds for the main diagonals, and \( 2(m - 1) \) from the diagonals, coming from the single blocks on the top-right and bottom-left corners.

Let \( \mathbb{F}_q \) be a finite field with \( q \) elements and let

\[
(\alpha, \beta) = (\alpha_{-n}, \ldots, \alpha_{-1}, \alpha_1, \ldots, \alpha_n, \beta_{-n}, \ldots, \beta_{-1}, \beta_1, \ldots, \beta_{n-1}).
\]
Define the following set

\[ V_q(n, q) = \{(\bar{a}, \bar{\beta}) \in \mathbb{F}_q^{4n-5} : a_i \beta_{i-j} - a_{i-n} \beta_{n-j} = 0, 1 \leq i, j \leq n - 1\} \]

where \(\mathbb{F}_q^{4n-5}\) is the \((4n - 5)\)-dimensional projective space over \(\mathbb{F}_q\). Since the equations of Theorem 2.1 are homogeneous, \(V_q(n, q)\) is a projective variety and an upper bound of its cardinality can be given by the Lang–Weil estimate (Ghorpade & Lachaud, 2002; Lang & Weil, 1954). Similarly, we can define an affine variety

\[ V_A(n, p) = \{(\bar{a}, \bar{\beta}) \in \mathbb{F}_p^{4n-5} : a_i \beta_{i-j} - a_{i-n} \beta_{n-j} = 0, 1 \leq i, j \leq n - 1\} \]

As an example, one can check that \(|V_A(2, p)| = p^3 + p^2 - p\) and \(|V_A(3, p)| = p^5 + 3p^4 - 2p^3 - 2p^2 + p\).

3. Involutory and nilpotent Toeplitz matrices over the finite field \(\mathbb{F}_p\)

Let \(\phi(n, p)\) and \(\psi(n, p)\), respectively, denote the number of involutory Toeplitz matrices and nilpotent Toeplitz matrices of degree two in \(M_n(\mathbb{F}_p)\). In the following theorem, we calculate \(\phi(n, p)\) for specific values of \(p\) or \(n\). Clearly, for a Toeplitz matrix \(A\) in \(M_n(\mathbb{F}_p)\), \(A\) is involutory if and only if any \(\frac{n(l)}{2}\) is idempotent. In this case, since \(\frac{n(l)}{2}\) is the number of idempotent Toeplitz matrices in \(M_n(\mathbb{F}_p)\), for \(p \neq 2\), \(\phi(n, p)\) also gives us the number of idempotent Toeplitz matrices in \(M_n(\mathbb{F}_p)\). Moreover, for \(p = 2\), a Toeplitz matrix \(A\) is a nontrivial involutory matrix if and only if the Toeplitz matrix \((A - I)\) is nilpotent of degree two. Therefore, \(\psi(n, 2) - 1\) is also the number of degree two nilpotent Toeplitz matrices when \(n\) is odd.

**Theorem 3.1**  The following hold:

(i) If \(n\) is odd then \(\phi(n, 2) = 2\phi(n - 2, 2) + 1\) or equivalently \(\phi(n, 2) = 2^{n-1} - 1\)

(ii) \(\phi(n, p) = 2\) if \(p = 3, 5, \ldots\)

(iii) \(\phi(n, p) = 2^{n-2}(p - 1) + 2\) if \(n \in \{2, 4\}\) and \(p \in \{3, 5, 7, \ldots\}\) or if \(n = 3\) and \(p \in \{5, 7, \ldots\}\)

**Proof**  (i) Let \(n = 2k + 1\) where \(k \geq 1\). Let \(T = (t_{i,j})\) be an involutory Toeplitz matrix. The anti-diagonal entries of \(C = T^2\) are of the following form:

\[ c_{n-(i+j)+1} = \sum_{l=1}^{n} t_{n-(i+j)+l} = t_{n-i-1} t_{j-l} + t_{n-i-2} t_{j-l} + \ldots + t_{n-i-n} t_{j-l} + \ldots + t_{n-i-i} t_{j-l} + \ldots \]

where \(i = 0, \ldots, n - 1\). If \(C = I\) then \(c_{n-(i+j)+1} = t_{(n-i-j)+1} = \delta_{(n-i-j)+1}\). Hence, \(c_{n-i-j+1} = t_{l} = 1\) if \(l = \frac{n-1}{2}\) and \(t_{\frac{n-1}{2}} = 0\) if \(l = 0, \ldots, n - 1\) and \(i \neq \frac{n-1}{2}\). Therefore, \(2k + 1\) entries \(t_{l}\) where \(l = -k, \ldots, -2, -1, 0, 1, 2, \ldots, k\) are known, while \(2k\) entries \(t_{l}\) where \(l = -2k, \ldots, -k - 1, k + 1, \ldots, 2k\) are unknown. Therefore, the matrix \(T\) must be of the following form:

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 & t_{k+1} & \cdots & t_{2k} \\
0 & 1 & 0 & \cdots & 0 & t_{k+1} & \cdots & t_{2k+1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
t_{2k-2} & & \cdots & 0 & 1 & 0 & \cdots & 0 \\
t_{2k-1} & & t_{k+1} & \cdots & 0 & 1 & 0 & \cdots \\
t_{2k} & t_{2k-1} & \cdots & t_{k+1} & \cdots & 0 & 1 & \cdots
\end{pmatrix}
\]
It is easy to see that for the matrix $T$ which is illustrated above, the equation $T^2 = I$, in $\mathbb{Z}_p$, gives the following $k^2$ equations in $\mathbb{Z}_p$:

$$t_{i,n}t_{n,j} = 0 \quad 1 \leq i,j \leq k$$

If $t_{i,n} = 0$ for all $i$, then $t_{n,j}$ can be either zero or one for any $j$. Therefore, there are $2^k$ solutions in this case. Similarly, if $t_{n,j} = 0$ for all $j$, then $t_{i,n}$ can be either zero or one for any $i$. Therefore, there are $2^k$ solutions in this case as well. However, we counted the case that $t_{i,n} = t_{n,j} = 0$ for all $1 \leq i,j \leq k$ twice. Hence, there are $2 \times 2^k - 1$ solutions in total. Since $k = \frac{n-1}{2}$, we conclude that there are $2^{\frac{n-1}{2}} - 1$ solutions in total.

(ii) Since trace($T$) = $\rho t_0 = 0$ (mod $p$), the sum of eigenvalues of $T$ must be zero. This is only possible if all eigenvalues of $T$ are $1$ or all are $-1$.

(iii) For $n = 2$, the matrix equation $T^2 = I$, in $\mathbb{Z}_p$, where $p$ is odd, gives the following equations in $\mathbb{Z}_p$:

$$t_0^2 + t_1t_{-1} = 1, \quad t_0t_{-1} = t_0t_1 = 0$$

If $t_0 = 0$ then $t_1t_{-1} = 1$ which has $p - 1$ solutions in $\mathbb{Z}_p$. If $t_0 \neq 0$, then $t_0^2 = 1$ and $t_1t_{-1} = t_0t_{-1} = t_0t_1 = 0$ which has two solutions, $t_0 = 1, t_1 = t_{-1} = 0$ and $t_0 = -1, t_1 = t_{-1} = 0$. Therefore, there are $(p - 1) + 2$ solutions when $n = 2$. For $n = 3$, the matrix equation $T^2 = I$, in $\mathbb{Z}_p$, where $p$ is odd and greater than 3, gives the following equations in $\mathbb{Z}_p$:

$$t_0^2 + 2t_1t_{-1} = 1, \quad t_0t_{-1} + t_1t_{-2} = t_0^2 + 2t_0t_1 = 2t_0t_1 + t_1t_{-2} = t_0^2 + 2t_0t_1 = 0, \quad t_1t_{-1} = t_1t_{-2}$$

We note that $t_0$ cannot be zero because if $t_0 = 0$, then $t_1 = 0$ because of the equation $t_1^2 + 2t_0t_1 = 0$ and the first equation in the above gives a contradiction. Furthermore, it follows from the above equations and invertibility of $t_0$ that if $t_1 = 0$, then $t_{-1} = t_2 = t_{-2} = 0$, $t_0^2 = 1$, and if $t_{-1} = 0$, then $t_1 = t_2 = t_{-2} = 0$, $t_0^2 = 1$. Hence, we have two solutions in this case, namely $t_0 = 1, t_1 = t_{-1} = t_2 = t_{-2} = 0$ and $t_0 = -1$, $t_1 = t_{-1} = t_2 = t_{-2} = 0$. Therefore, the remaining case is when $t_1$ and $t_{-1}$ are both invertible which means that by knowing the values of $t_0$, $t_1$ and $t_{-1}$ we can uniquely determine the values of $t_0$ and $t_{-1}$ from the equations. Thus, it suffices to count the solutions of the equation $t_0^2 + 2t_1t_{-1} = 1$. Since the equation $t_0^2 + 2t_1t_{-1} = 1$ has $(p - 1)$ solutions in $\mathbb{Z}_p$, where $p$ is odd and greater than 3, there are $2(p - 1) + 2$ solutions in total.

For $n = 4$, the matrix equation $T^2 = I$, in $\mathbb{Z}_p$, where $p$ is odd, gives the following equations in $\mathbb{Z}_p$:

$$t_0^2 + 2t_1t_{-1} + t_1t_{-2} = 1, \quad t_0t_{-1} + t_1t_{-2} = t_0t_{-1} + t_1t_{-2} = t_0t_1 + t_1t_{-1} = t_0t_1 + t_2 = 0,
\quad t_2t_3 = t_1t_{-1} + t_2t_{-2} = t_2t_{-3}, \quad t_2t_3 = t_3t_{-2} + t_2t_{-3}, \quad t_2t_3 = t_2t_{-3} + 2t_2t_3 = t_2t_3 + 2t_2t_3 = 0$$

The above equations indicate that if $t_0 = 0$, then either $t_1 = t_1 = t_3 = t_3 = 0$, and $t_2t_3 = 1$ or $t_2 = t_3 = 0$, $t_3 = -t_2^{-1}t_1$, $t_3 = -t_2^{-1}t_1$ and $2t_2t_3 = 1$. Since each of the equations $t_2t_3 = 1$ and $2t_2t_3 = 1$ has $p - 1$ solutions in $\mathbb{Z}_p$, there are $2(p - 1)$ solutions when $t_0 = 0$. If $t_0 \neq 0$, then there are two cases to consider: (1) If either $t_1$ or $t_{-1}$ is zero, then it follows from the above equations that all variables except $t_0$ must be zero. Therefore, this case leads to the equation $t_0^2 = 1$ which has two solutions. (2) If neither $t_1$ nor $t_{-1}$ is zero, then the above equations can be simplified as follows:

$$t_{-2} = -t_1t_0t_{-1}, \quad t_2 = t_1^2t_0t_1t_3^2 = t_1t_3^2t_0t_1t_3^2, \quad t_3 = t_1t_3^2t_0t_1t_3^2, \quad t_3 = t_1t_3^2t_0t_1t_3^2, \quad t_2t_3 = t_2t_3 = 4t_0^2 = 1$$

The equations $4t_0^2 = 1$ and $t_2t_3 = t_0^2$ lead to equations $t_1t_{-1} = \frac{p+1}{2}$ and $t_1t_{-1} = \frac{p-1}{2}$ each of which has $2(p - 1)$ solutions. Therefore, if $t_0 \neq 0$, then there are $2(p - 1) + 2$ solutions, and if $t_0 = 0$, then there are $2(p - 1)$ solutions. Hence, there are $4(p - 1) + 2$ solutions in total.
CONJECTURE 3.2  If n is even then \( \phi(n, 2) = 2\phi(n - 2, 2) + 2 \) or equivalently \( \phi(n, 2) = 3 \times 2^{\frac{n}{2}} - 2 \).

THEOREM 3.3  The following hold:

(i)  
\[
\psi(2, p) = \begin{cases} 
2p - 2 & \text{if } p \neq 2 \\
3 & \text{if } p = 2
\end{cases}
\]

(ii)  
\[
\psi(3, p) = \begin{cases} 
2p - 2 & \text{if } p \neq 3 \\
8 & \text{if } p = 3
\end{cases}
\]

Proof  (i) The matrix equation \( T^2 = 0 \) results in the following equations:
\[
t_0^2 + t_{-1}t_1 = 0, \ 2t_0t_1 = 0, \ 2t_0t_{-1} = 0
\]

If \( p = 2 \), then there is just one equation \( t_0^2 + t_{-1}t_1 = 0 \) which has three solutions. Assume \( p \neq 2 \), if \( t_1 \) and \( t_{-1} \) are both nonzero, then the equation \( t_1t_{-1} = 0 \) is a contradiction. Therefore, either \( t_1 \) and \( t_{-1} \) are both zero which we do not count or one of them is nonzero which gives \( 2(p - 1) \) solutions. Hence, there are \( 2(p - 1) \) solutions in total.

(ii) The matrix equation \( T^2 = 0 \) results in the following equations:
\[
-t_{-1}^2 = -2t_{-1}t_1 = -t_2t_{-2} - t_2t_{-1}, \ t_1^2 = -2t_0t_2, \ t_{-1}^2 = -2t_0t_{-2}, \ t_1t_{-1} = -2t_0t_1, \ t_1t_{-2} = -2t_0t_{-1}
\]

It follows from the above equations that the cases in which \( t_1 = 0 \) and \( t_{-1} \neq 0 \) or \( t_2 = 0 \) and \( t_1 \neq 0 \) result in contradictions. If \( t_1 = t_{-1} = 0 \), then \( t_0 = 0 \) and the equation \( t_2t_{-2} = 0 \) has \( 2p - 2 \) solutions. If \( t_1 \neq 0, t_{-1} \neq 0 \) and \( p \neq 3 \), then the equation \( -9t_1t_{-1} = 0 \) which follows from the above equations is a contradiction. Therefore, the case \( t_1 \neq 0 \) and \( t_{-1} \neq 0 \) is not possible when \( p \neq 3 \) and there are \( 2p - 2 \) solutions in total. If \( t_1 \neq 0, t_{-1} \neq 0 \) and \( p = 3 \), then the equation \( -9t_1t_{-1} = 0 \) is no longer a contradiction and there are four solutions for this case and \( 2 \times 3 = 6 \) solutions for the previous case which amounts to eight solutions when \( p = 3 \).

Remark 3.4  For matrices in \( M_n(\mathbb{Z}_2) \), it is known that the number of invertible, involutory (Hodges, 1958) and nilpotent (Fine & Herstein, 1958) matrices are as follows:

\[
g(n, p) = \prod_{i=0}^{n-1} (p^i - p^i)
\]

\[
I(n, p) = \begin{cases} 
g(n, p) \sum_{i=0}^{\lfloor n/2 \rfloor} \frac{1}{p^{i(p^n - i^2)} + p^{i-1}} & \text{if } p \text{ is odd} \\
g(n, p) \sum_{i=2n+1}^{\lfloor n/2 \rfloor} \frac{1}{p^{i(p^n - i^2)} + p^{i-1}} & \text{if } p \text{ is even}
\end{cases}
\]

\[
l(n, p) = p^{n^2 - n}
\]

Therefore, the probabilities of being invertible, involutory and nilpotent are \( \frac{g(n, p)}{p^n} \), \( \frac{I(n, p)}{p^n} \) and \( \frac{g(n, p)}{p^n} = \frac{1}{p^n} \), respectively. Since the number of invertible Toeplitz matrices is \( (p - 1)p^{2n - 2} \) (Kaltofen & Lobo, 1996, Theorem 4), the probability that a Toeplitz matrix is invertible is \( \frac{1}{p^n} \), which is independent of \( n \). By Theorem 3.1, the probability that a Toeplitz matrix in \( M_n(\mathbb{Z}_2) \), where \( n \) is odd, is involutory is \( 2^{1-2n}(2^{2n} - 1) \leq \frac{1}{2} \).
Theorem 3.5. Let \( G(n, p, m) = \{ T \in \mathbb{M}_n(\mathbb{Z}_p) : T + T^m + \ldots + T = J \text{ and } T \text{ is Toeplitz} \} \) where \( \ast \) is the Schur product and \( J \) is the matrix with all entries equal to 1. The cardinality of the set \( G(n, p, m) \) is \( \text{gcd}(m, p - 1)2^{n-1} \).

Proof. The entries of \( T \) must satisfy the equation \( t_i^m = 1 \) (mod \( p \)). Therefore, we are counting elements of \( \mathbb{Z}_p \) that their order divides \( m \) and their order also divides \( p - 1 \), i.e., their order divides \( \text{gcd}(m, p - 1) \). Hence,

\[
\sum_{d \mid \text{gcd}(m, p - 1)} \phi(d) = \text{gcd}(m, p - 1)
\]

where \( \phi \) is Euler’s totient function. Since there are \( 2n - 1 \) distinct elements, there are \( \text{gcd}(m, p - 1)2^{n-1} \) different matrices \( T \) with the aforementioned properties.

4. Toeplitz matrices over the complex field \( \mathbb{C} \)

In this section, we consider Toeplitz matrices over the complex field \( \mathbb{C} \). We define a metric that measures how far a matrix is from being Toeplitz. We can associate a Toeplitz matrix \( T_A \) to any matrix \( A \) as follows:

\[
t_{ij} = \frac{\sum_{l=0}^{k} a_{i+j+l}}{k+1} 1 \leq j \leq n-1, \quad k = n-j
\]

\[
t_{ii} = \frac{\sum_{l=0}^{k} a_{i+l+l}}{k+1} 1 \leq i \leq n-1, \quad k = n-i, \quad t_{ij} = t_{i+j+l} \quad i + l \leq n, j + l \leq n, \quad t_{in} = a_{in}, \quad t_{nj} = a_{nj}
\]

For example, the associated Toeplitz matrix to \( A = \begin{pmatrix} 3 & -1 \\ 1 & 5 \end{pmatrix} \) is \( T_A = \begin{pmatrix} 4 & -1 \\ 1 & 4 \end{pmatrix} \) and we observe that \( d(A, T_A) = \|A - T_A\| \leq 1 \) where \( \| \| \) is the operator norm. If \( A \) is a Toeplitz matrix, then \( d(A, T_A) \) is zero.

Theorem 4.1. If \( T_A \) is the associated Toeplitz matrix of \( A \) such that \( \|T_A - A\| \leq \epsilon \) where \( \epsilon > 0 \), then

(i) \( \|T_A, A\| \leq 2\epsilon \min(\|T_A\|, \|A\|) \).

(ii) If \( A \) is self-adjoint then \( T_A \) is self-adjoint and furthermore if \( \|A\| \leq 1 \) and \( \|T_A\| \leq 1 \), then \( A \) and \( T_A \) are approximately jointly diagonalizable in the sense of Bronstein and Glashoff (2013).

Proof

(i) We have the following inequalities:

\[
\|T_A - A + T_A^2 + T_A^3\| \leq 2\epsilon \|T_A\| \quad \|T_A - A + A^2\| \leq 2\epsilon \|A\|
\]

Therefore, \( \|T_A, A\| \leq 2\epsilon \min(\|T_A\|, \|A\|) \).

(ii) \( T_A \) is self-adjoint because

\[
t_{ij} = \frac{\sum_{l=0}^{k} a_{i+j+l}}{k+1} = \frac{\sum_{l=0}^{k} \overline{a}_{j+i+l}}{k+1} = \overline{t_{ij}}
\]

and \( t_{in} = a_{in}, t_{nj} = a_{nj} \). It follows from Bronstein and Glashoff (2013, Theorem 4.1) that if \( \|A\| \leq 1 \) and \( \|T_A\| \leq 1 \), then \( A \) and \( T_A \) are approximately jointly diagonalizable.
We now consider involutory Toeplitz matrices over the complex field \( \mathbb{C} \). The following simple Proposition shows that if there exists one such matrix, there exist uncountably many.

**Proposition 4.2** If

\[
\begin{pmatrix}
 a_n & a_{n+1} & \cdots \\
 a_{n-1} & a_n & a_{n+1} \\
 \vdots & \ddots & \ddots \\
 a_{n-1} & a_n & a_{n+1}
\end{pmatrix}
\]

is an involutory Toeplitz matrix over \( \mathbb{C} \), and \( v \) is a nonzero complex scalar, then

\[
\begin{pmatrix}
 a_n & va_{n-1} & \cdots \\
 v^{-1}a_{n-1} & a_n & va_{n-1} \\
 \vdots & \ddots & \ddots \\
 v^{-1}a_{n-1} & a_n & va_{n-1}
\end{pmatrix}
\]

is also an involutory Toeplitz matrix.

**Proof** The two matrices are related by a similarity transformation by a diagonal matrix. See for example, Proposition 3.1 in Kucerovsky and Sarraf (2014).

In view of the above result, it would appear inevitable that we should consider some equivalence relation or some restriction on the class of involutory Toeplitz matrices considered. We give two such results:

**Theorem 4.3** There are \( n+1 \) equivalence classes of involutory Toeplitz matrices in \( M_\mathbb{C} \), where the equivalence relation is similarity. Furthermore, every equivalence class contains at least one (involutory) Hermitian circulant matrix.

**Proof** Let \( T \) be an involutory Toeplitz matrix. Since the polynomial \( p(x) = x^2 - 1 \) annihilates \( T \), it follows that the minimal polynomial of \( T \) divides \( p \). But because \( p \) has distinct roots over \( \mathbb{C} \), it follows that the minimal polynomial of \( T \) has distinct roots over \( \mathbb{C} \), and thus that \( T \) is diagonalizable. Thus, \( T \) is equivalent under similarity to a diagonal matrix \( \Lambda \). Let \( F \) denote the usual Vandermonde matrix representation of the finite Fourier transform (over \( \mathbb{C} \)). Then, \( F^{-1} \Lambda F \) is equivalent to \( T \), is a Hermitian matrix and is a circulant (Driessel, 2011). Clearly, thus, the number of equivalence classes is given by the number of possible involutory diagonal matrices in \( M_\mathbb{C} \), after equivalence, which is \( n+1 \). Two diagonal matrices are in the same class after equivalence if and only if the elements of the diagonals are permutations of one another.

Note: the above proof will work in more general fields: we just need sufficiently many roots of unity and the existence of \( 1 / n \).

We can also look at restricting the class of Toeplitz matrices considered:

**Theorem 4.4** In \( M_n(\mathbb{R}) \), the involutory symmetric Toeplitz matrices are all either symmetric real circulants or are symmetric real skew-circulants. If \( n \) is even and greater than 2, there are a total of \( 3 \cdot 2^{\frac{n}{2}} - 2 \) such matrices. If \( n \) is odd and greater than 1, there is a total of \( 2^{\frac{n+1}{2}} - 2 \) such matrices.

**Proof** We observe that if a symmetric real matrix satisfies \( T^2 = \text{Id} \), it is in fact an orthogonal matrix. Orthogonal and symmetric real Toeplitz matrices are either circulants or skew-circulants (Böttcher, 2008, [section 3]). From Theorem 4.1 and Theorem 4.2 in Böttcher (2008), if \( n \) is odd and greater than
1, there is a total of \(2^{\lceil \frac{n}{2}\rceil} - 2\) such matrices, and if \(n\) is even and greater than 2, there are a total of \(3 \cdot 2^{\lceil \frac{n}{2}\rceil} - 2\) such matrices.

**Theorem 4.5** Consider the Frobenius norm \(\|\cdot\|\) on the complex \(n\)-by-\(n\) matrices, \(M_n(\mathbb{C})\). Given an arbitrary \(A \in M_n(\mathbb{C})\), there exist a unique complex Toeplitz matrix minimizing \(\|A - T\|\) over all Toeplitz matrices \(T\). This unique minimizing matrix is the Toeplitz matrix given in each diagonal by the average of the elements of the corresponding diagonal of \(A\). The same statement holds in the real case.

**Proof** Under this norm, \(M_n(\mathbb{C})\) is a complex Hilbert space, with the complex Toeplitz matrices forming a subspace \(\mathcal{T}\). Let \(e_1, \cdots, e_{2^n-1}\) denote the Toeplitz matrices that have zero in all diagonals except one, and in that \(k\)th diagonal, has elements equal to 1. At the level of Hilbert spaces, this is a finite orthogonal set spanning the subspace \(\mathcal{T}\). Thus, we can explicitly construct the projection from \(M_n(\mathbb{C})\) onto \(\mathcal{T}\) in the form:

\[
P(A) = \sum_{k=1}^{2^n-1} \frac{A e_k}{\|e_k\|} e_k.
\]

It is well known that orthogonal projection onto a closed subspace, in Hilbert space, minimizes the norm. In this case, the subspace \(\mathcal{T}\) is closed as a consequence of its finite-dimensionality. It is clear that \(P(A)\) is given in each diagonal by the average of the elements of the corresponding diagonal of \(A\).

In the real case, exactly the same proof works, but with real Hilbert spaces instead of complex Hilbert spaces.

It seems interesting to consider generalizations of the above Theorem to the C*-norm on \(M_n(\mathbb{C})\). However, projection onto linear subspaces of C*-algebras has poor properties in general. It is not even true that the operation of projection onto a subspace will not increase the norm distance between elements (Silverman, 1969). The situation improves when we consider projection onto subalgebras. Tomiyama (Tomiyama) showed that each projection of norm one of a C*-algebra onto a C*-subalgebra is a conditional expectation, and Takesaki (1972) showed that in a finite von Neumann algebra, there exist faithful conditional expectations onto von Neumann subalgebras. We will make use of this result, and will explicitly construct faithful conditional expectations onto the circulant algebra and the skew-circulant algebra. We will show these expectations are basically unique.

For information on circulants and skew-circulants, see Driessel, 2011. The circulants of a given order form an abelian subalgebra of \(M_n(\mathbb{C})\), as do the skew-circulants. The circulants can be simultaneously diagonalized by (a scalar multiple of) the DFT matrix \(F = [\omega^{(i-1)j-1}]\) where \(\omega\) is a primitive (complex) root of unity. The skew-circulants can be simultaneously diagonalized by \(F \Omega^{-1/2}\) where \(\Omega\) is a suitable diagonal unitary matrix, see Proposition 31 in Driessel (2011). We will write \(W = H' \Lambda H\), where \(W\) is a skew-circulant and \(\Lambda\) is a diagonal matrix, or \(\mathcal{C} = F' \Lambda F\), for a circulant \(\mathcal{C}\).

**Theorem 4.6** A faithful, trace-preserving, idempotent, conditional expectation of operator norm 1 of \(M_n(\mathbb{C})\) onto the circulants is given by

\[
\phi(M) = F' D(FMF'^*) F,
\]

where \(D\) takes all elements that are not on the main diagonal to zero, and preserves the elements of the main diagonal and \(F\) is the discrete Fourier transform matrix.

**Proof** Since the map \(m \mapsto FMF^*\) is a unitary transformation diagonalizing the circulants, we must show that the map \(D\) is a faithful, idempotent, trace-preserving, conditional expectation of norm 1 of \(M_n(\mathbb{C})\) onto the subalgebra of diagonal matrices. To show that \(D\) is a conditional expectation, we verify that the defining property \(D(dm'd^*) = d\phi(m)d^*\) holds, where \(d\) is a diagonal matrix, and \(m\) is in \(M_n(\mathbb{C})\). It is apparent that \(D: M_n(\mathbb{C}) \rightarrow M_n(\mathbb{C})\) preserves the trace. Since the diagonal elements of a positive definite matrix are real and nonnegative, it follows that \(D\) maps a positive definite matrix.
to a positive definite (diagonal) matrix. If $D$ were not faithful, there would be some nonzero positive definite matrix that maps to zero under $D$. But a positive definite matrix with principal diagonal zero is in fact the zero matrix (Horn & Johnson, 1990, p. 398). Conditional expectations are completely positive maps, and a completely positive map of unital algebras has norm one if and only if it takes the unit to the unit. Evidently, the map $D$ is unital, and thus has norm 1. It is clear that $D$ is idempotent.

**Corollary 4.7** There is only one idempotent linear mapping of $M_n(\mathbb{C})$ onto the circulant subalgebra in $M_n(\mathbb{C})$. It is given by the map in the above Theorem.

**Proof** By Theorem 1 in Tomiyama (1957), any given idempotent linear mapping $\mathcal{L}$ is a conditional expectation of a finite-dimensional von Neumann algebra onto a finite-dimensional von Neumann algebra. By finite-dimensionality, the mapping is necessarily normal. By Proposition 1.2 in Tomiyama (1972), the mapping $\mathcal{L}$ is necessarily faithful. By Theorem 6.2.2 (and Theorem 6.2.3) in Arveson (1967), the mapping $\mathcal{L}$ is unique. It therefore follows that any two maps satisfying the hypotheses of the Corollary are in fact the same map, and clearly, the map provided by Theorem 4.6 satisfies the hypotheses.

From the uniqueness, we have the following further Corollary:

**Corollary 4.8** The map of Theorem 4.6 minimizes the error in (operator) norm among all idempotent projections of $M_n(\mathbb{C})$ onto the circulant subalgebra in $M_n(\mathbb{C})$.

Replacing circulants by skew-circulants in the proofs of the last three results, we obtain analogous results for skew-circulants. We summarize as follows:

**Theorem 4.9** A faithful, trace-preserving, idempotent, conditional expectation of operator norm 1 of $M_n(\mathbb{C})$ onto the skew-circulant subalgebra in $M_n(\mathbb{C})$. $H$ is the matrix provided by Driessel (2011, [Proposition 31]).

Since every Toeplitz matrix can be uniquely decomposed as the sum of a circulant and a skew-circulant, we have the following Remark:

**Remark 4.10** For each matrix $m$, the sum of the maps from Theorems 4.6 and 4.9, $m \mapsto \phi(m) + \psi(m)$, gives a Toeplitz matrix.

We next consider how the usual product and Schur product interact with the circulant property. The finite Fourier transform over the complex numbers can be described most briefly as the unitary transformation that is (up to a scalar multiple) represented by the $N$-by-$N$ matrix $[a_{ij}]$ with $a_{ij} = \omega^{i(j-1)}$. The complex number $\omega$ is a principal $N^{th}$ root of unity. This definition can be made in a finite field. If $\omega$ is a principal $N^{th}$ root of unity in a finite field, it follows that $\omega^N = 1$ and $\sum_{j=0}^{N-1} \omega^j = 0$ for $1 \leq i < N$, and these are the main properties used in establishing the well-known properties of the (complex) finite Fourier transform with respect to convolution. In order to be able to invert the generalized finite Fourier transform, it is necessary that $N$ be not divisible by the characteristic of the field. See Chapter 11 of Elliott and Rao (1982) for more information. If these conditions are met, then the generalized finite Fourier transform takes the convolution operation defined by multiplication by a circulant matrix to multiplication by a diagonal matrix, see section 11.2 and 11.3 of Elliott and Rao (1982). Thus, we are able to diagonalize circulant matrices, even over a finite field. The diagonal matrix obtained will be referred to as the eigenvalue matrix of the circulant, and denoted $\Lambda(A)$ where
A is the given circulant. Let us say that a circular convolution of two sequences is the first row of the matrix product of the circulants having the given sequences as their first row. Thus, if $\text{Circ}(a_1, \ldots, a_n)\text{Circ}(b_1, \ldots, b_n) = \text{Circ}(c_1, \ldots, c_n)$, then the sequence $(c_1, \ldots, c_n)$ is the circular convolution of $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$.

**Proposition 4.11** Let $A$ and $B$ be two circulants, over a finite field with characteristic relatively prime to the matrix dimension $N$ of the circulant matrix. The eigenvalues of $AB$ and $A \ast B$ can be determined in terms of the eigenvalues of $A$ and $B$. The eigenvalues of $AB$ are given by $\Lambda(A) \ast \Lambda(B)$, and the eigenvalues of $A \ast B$ are given by $\frac{1}{N}C$ where $C$ is the circulant matrix whose eigenvalue sequence is the circular convolution of the eigenvalue sequences of $A$ and $B$.

**Proof** Since the generalized Fourier transform is a unitary transformation, $W$, it is clear that $WABW^{-1} = WAW^{-1}WBW^{-1}$. Thus, the eigenvalues of $AB$ are the elementwise product of the eigenvalues of $A$ and $B$. The main issue is the behaviour of the Schur product.

To reduce notation, let us consider first the case where both $A$ and $B$ have the property that all the eigenvalues except one are zero, and that eigenvalue is $1$. Then, $A = W^{-1}\Lambda(A)W$ is an outer product of a row of $W$ and a column of $W^{-1}$. Since $W = [\omega^{i-1}j^{-1}]$ and $W^{-1} = [\frac{1}{N}\omega^{-i-1}j^{-1}]$ we thus have that $A = [\frac{1}{N}\omega^{i-1}j^{-1}]$ where $k$ is the index of the nonzero eigenvalue of $A$. Similarly, $B = [\frac{1}{N}\omega^{-i-1}j^{-1}]$ where $\epsilon$ is the index of the nonzero eigenvalue of $B$. Taking the Schur product, we have $A \ast B = [\frac{1}{N}\omega^{i-1}j^{-1}]$ where $C$ is the circulant matrix with all eigenvalues zero except for a 1 in the $(\epsilon + k - 1)^{th}$ place. Since $\omega^N = 1$, the values of the exponent in $\omega^{i-1}j^{-1}$ may be taken as being modulo $N$. Thus, we see that the eigenvalues of $A \ast B$ are given by $\frac{1}{N}C$, where the eigenvalue sequence of the circulant matrix $C$ are given by the circular convolution of the eigenvalue sequence of $A$ and the eigenvalue sequence of $B$. The case of general circulants $A$ and $B$ follows by taking linear combinations.

The above result can be rephrased as follows: if we diagonalize a circulant using the generalized discrete Fourier transform, and then put the eigenvalue sequence into the first row of a circulant matrix, then up to a scalar multiple, Schur products are transformed into matrix products. This transformation is invertible.

**Funding**

This work was partially supported by NSERC [grant number 1-108036-40-01].

**Author details**

Dan Kucerovsky1
E-mail: dukucerov@unb.ca

Kaveh Mousavand2
E-mail: k.mousavand@lacin.ca

Aydin Sarraf3
E-mail: aydin.sarraf@resson.com

1 Department of Mathematics, University of New Brunswick, Fredericton, Canada, NB, E3B 5A3.

2 Département de mathématiques, L’Université du Québec à Montréal, Montréal, Canada, H3C 3P8.

3 Department of Computer Science, University of New Brunswick, Fredericton, Canada, NB, E3B 5A3.

**Citation information**

Cite this article as: On some properties of Toeplitz matrices, Dan Kucerovsky, Kaveh Mousavand & Aydin Sarraf, Cogent Mathematics (2016), 3: 1154705.

**Correction**

This article was originally published with errors, which have now been corrected in the online version. Please see Correction (http://dx.doi.org/10.1080/25742558.2018.1520444).

**References**

Arveson, W. B. (1967). Analyticity in operator algebras. American Journal of Mathematics, 89, 578–642.

Böttcher, A. (2003). Orthogonal symmetric Toeplitz matrices. Complex Analysis and Operator Theory, 2, 285–298.

Brooks, B. P. (2006). The coefficients of the characteristic polynomial in terms of the eigenvalues and the elements of an $n \times n$ matrix. Applied Mathematics Letters, 19, 511–515.

Bronstein, M. M., & Glashoff, K. (2013). Almost-commuting matrices are almost jointly diagonalizable, arXiv:1305.2135.

Driessel, K. R. (2013). A relation between some special centroskew, near-Toeplitz, tridiagonal matrices and circulant matrices. arXiv:1102.1953v2.

Elliott, D. F., & Rao, K. R. (1982). Fast transforms, algorithms, analyses, applications. New York, NY: Academic Press.

Fine, N. J., & Herstein, I. N. (1958). The probability that a matrix be nilpotent. Illinois Journal of Mathematics, 2, 499–504.

Ghorpade, S. R., & Lochaud, G. (2002). Number of solutions of equations over finite fields and a conjecture of Lang and Weil. Number theory and discrete mathematics (pp. 269–291). Basel: Springer.

Heing, G. (2001). Not every matrix is similar to a Toeplitz matrix. Linear Algebra and its Applications, 332, 519–531.

Hodges, J. H. (1958). The matrix equation $X^N = I$ over a finite field. The American Mathematical Monthly, 65, 518–520.

Horn, R. A., & Johnson, C. R. (1990). Matrix analysis. Cambridge: Cambridge University Press.

Kato, E., & Lobo, A. (1996). On rank properties of Toeplitz matrices over finite fields. In Proceedings of the 1996 international symposium on Symbolic and algebraic computation, ACM (pp. 241–249), New York.

Kucerovsky, D., & Sarraf, A. (2014). Schur multipliers and matrix products. Houston Journal of Mathematics, 40, 837–850.
Lang, S., & Weil, A. (1954). Number of points of varieties in finite fields. American Journal of Mathematics, 76, 819–827.
Lim, L. H., & Ye, K. (2013). Every matrix is a product of Toeplitz matrices. arXiv:1307.5132.
Mackey, D. S., Mackey, N., & Petrovic, S. (1999). Is every matrix similar to a Toeplitz matrix? Linear Algebra and its Applications, 297, 87–105.
Silverman, E. (1969). A weak projection of $C$ onto a Euclidean subspace. Transactions of the AMS, 136, 381–390.
Takesaki, M. (1972). Conditional expectations in von Neumann algebras. Journal of Functional Analysis, 9, 306–321.
Tomiyama, J. (1957). On the projection of norm one in $W^*$-algebras. Proceedings of the Japan Academy, 33, 608–612.
Tomiyama, J. (1972). On some types of maximal Abelian subalgebras. Journal of Functional Analysis, 10, 373–386.
Winitzki, S. (2010). Linear algebra via exterior products. San Bernardino: Lulu.