Experimenting with Matrix Identities and the Amitsur-Levitzki Theorem

Joseph Koutsoutis

1 Introduction

The Amitsur-Levitzki theorem is a celebrated result stating that for any $n \times n$ matrices A_1, \ldots, A_{2n} over a commutative ring,

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \cdot A_{\sigma(1)} \cdots A_{\sigma(2n)} = 0.$$

This theorem has received many proofs after the original of Amitsur and Levitzki, which was a direct proof [1]. These include a proof of Kostant by relating it to Lie algebra cohomology [2], a proof of Swan via the interpretation of matrices as directed graphs [6] [7], a proof of Razmyslov related to the Cayley-Hamilton theorem [4], a proof of Rosset using Grassman variables [5], and a proof of Procesi showing that the Amitsur-Levitzki theorem is the Cayley-Hamilton identity for the generic Grassman matrix [3].

In this project, we experiment with various classes of matrices to find identities similar to that of the Amitsur-Levitzki theorem, where we still sum over the symmetric group but do not require the constants to correspond to the permutation signs. In particular, we provide a Maple package to investigate symmetric, antisymmetric, tridiagonal, triangular, Toeplitz, Hankel, and circulant matrices. We finish by proving some easy identities motivated by our experiments and describing some potentially interesting behaviors.

2 Preliminaries and Maple Procedures

For completeness, here are the formal definitions of the standard matrix classes we investigated:

- A matrix A is **symmetric** iff $A = A^t$.
- A matrix A is **antisymmetric** iff $A = -A^t$.
- A matrix A is **tridiagonal** iff all nonzero entries are on the main diagonal, the lower diagonal, and the upper diagonal.
- A matrix A is **upper triangular** iff all entries below the main diagonal are 0.
- A matrix A is **Toeplitz** iff each descending diagonal from left to right is constant.

- A matrix A is *Hankel* iff each ascending skew-diagonal from left to right is constant.
- A matrix A is *circulant* iff each row is rotated one element to the right relative to the preceding row.
- A matrix A is **anticirculant** iff each row is rotated one element to the left relative to the preceding row.

Accompanying this report is the Maple package AL.txt. The primary procedure is ALnm, which tries to find a nontrivial identity involving $m \ n \times n$ matrices of a particular class. The main procedures used for testing particular classes are:

- ALnmG: for arbitrary $n \times n$ matrices
- ALnmS: for arbitrary $n \times n$ symmetric matrices
- ALnmAS: for arbitrary $n \times n$ antisymmetric matrices
- ALnmTD: for arbitrary $n \times n$ tridiagonal matrices
- ALnmUT: for arbitrary $n \times n$ upper triangular matrices
- ALnmToeplitz: for arbitrary n × n Toeplitz matrices
- ALnmHankel: for arbitrary $n \times n$ Hankel matrices
- ALnmC: for arbitrary $n \times n$ circulant matrices
- ALnmAC: for arbitrary $n \times n$ anticirculant matrices
- ALnmCG: for arbitrary $n \times n$ circulant matrices with a specified shift

All of these procedures have optional arguments allowing for the restriction of generated matrices to have certain entries set to zero. There is also an optional argument to test for identities involving raising matrices to different powers in the sum.

3 Results

We will begin with some identities found from our experiments that we were able to prove.

3.1 A Circulant Matrix Identity

The following observations and propositions relate to circulant matrices and are all a consequence of lemma 1 and theorem 1, but we include this discussion to motivate the statement of theorem 1. **Observation 1.** Let A_1, A_2, A_3 be $n \times n$ matrices with constant columns. Then for any c_1, c_2, c_3 ,

$$\sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) \cdot c_{\sigma(3)} \cdot A_{\sigma(1)} A_{\sigma(2)} A_{\sigma(3)} = 0.$$

Proof. It is ETS this holds when A_1, A_2, A_3 each have only one nonzero column consisting of 1's. We observe $A_iA_jA_k = A_k$ for $i, j, k \in [3]$. The statement follows.

We note that matrices with constant columns are circulant matrices with shift 0 and that we have removed some constraints on the constants compared to those in the Amitsur-Levitzki theorem. This is the motivation behind our next observation.

Observation 2. Let A_1, A_2, A_3 be $n \times n$ anticirculant matrices. Then for any c_1, c_2, c_3 ,

$$\sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) \cdot c_{\sigma(2)} \cdot A_{\sigma(1)} A_{\sigma(2)} A_{\sigma(3)} = 0.$$

Proof. It is ETS this holds when A_1, A_2, A_3 each have only one nonzero entry per row which is set to 1. Let j_k satisfy $(A_k)_{0,j_k} = 1$ where we take our matrices to be 0-indexed. We observe by taking indices mod n,

$$(A_1A_2A_3)_{i,j} = 1 \iff (A_2A_3)_{j_1-i,j} = 1 \iff (A_3)_{j_2+i-j_1,j} = 1 \iff j = j_3 - j_2 - i + j_1 \mod n$$

and

$$(A_3A_2A_1)_{i,j}=1\iff (A_2A_1)_{j_3-i,j}=1\iff (A_1)_{j_2+i-j_3,j}=1\iff j=j_1-j_2-i+j_3\mod n.$$

Hence by symmetry, $A_iA_iA_k = A_kA_iA_i$ and the statement follows.

It is well known that circulant matrices (with shift 1) commute. At a first glance, circulant matrices with shift $s \notin \{-1,0,1\}$ do not have behavior as nice as these 3 cases. For example, we experimentally found that circulant matrices with shift 2 of shapes 2×2 , 3×3 , 4×4 , 5×5 , 6×6 , 7×7 , and 8×8 required 3, 3, 4, 5, 4, 4, and 5 matrices respectively. However, the following lemma makes it easier to see what is happening in some cases.

Lemma 1. Let A_1, A_2 be $n \times n$ circulant matrices with shifts s_1 and s_2 respectively. Then A_1A_2 is a circulant matrix with shift $s_1 \cdot s_2$.

Proof. By 0-indexing our matrices and taking indices mod *n*, we compute:

$$egin{aligned} &(A_1A_2)_{i,j} = \sum_{k=0}^{n-1} (A_1)_{i,k} \cdot (A_2)_{k,j} \ &= \sum_{k=0}^{n-1} (A_1)_{0,k-is_1} \cdot (A_2)_{0,j-ks_2} \ &= \sum_{k=0}^{n-1} (A_1)_{0,k} \cdot (A_2)_{0,j-(k+is_1)s_2} \ &= \sum_{k=0}^{n-1} (A_1)_{0,k} \cdot (A_2)_{0,j-is_1s_2-ks_2} \ &= \sum_{k=0}^{n-1} (A_1)_{0,k} \cdot (A_2)_{k,j-is_1s_2} \ &= (A_1A_2)_{0,j-is_1s_2}. \end{aligned}$$

Hence A_1A_2 is a circulant matrix with shift $s_1 \cdot s_2$.

With this lemma, we can state more general versions of observation 1.

Proposition 1. Fix $s \ge 1$. Let $n = s^{\ell}$ for some $\ell \ge 1$. Then for any $n \times n$ circulant matrices with shift $s A_1, \ldots, A_{\ell+2}$ and any $(\ell+2)!/2$ constants c_i ,

$$\sum_{\sigma \in S_{\ell+2}} \operatorname{sgn}(\sigma) \cdot c_{\sigma|_{[3,\ell+2]}} \cdot A_{\sigma(1)} \cdots A_{\sigma(\ell+2)} = 0.$$

Proof. By lemma 1, we know that $A_{\sigma(3)} \cdots A_{\sigma(\ell+2)}$ is a circulant matrix with shift $s^{\ell} = 0 \mod n$. Hence

$$A_{\sigma(1)}A_{\sigma(2)}A_{\sigma(3)}\cdots A_{\sigma(\ell+2)} = A_{\sigma(2)}A_{\sigma(1)}A_{\sigma(3)}\cdots A_{\sigma(\ell+2)}.$$

We can similarly state a more general version of observation 2

Proposition 2. Let $n, s \ge 1$ where gcd(n, s) = 1. Then for any $n \times n$ circulant matrices with shift s $A_1, \ldots, A_{ord_n(s)+1}$ and any $(ord_n(s)+1)!/2$ constants c_i ,

$$\sum_{\sigma \in S_{\operatorname{ord}_n(s)+1}} \operatorname{sgn}(\sigma) \cdot c_{\sigma|_{[2,\operatorname{ord}_n(s)]}} \cdot A_{\sigma(1)} \cdots A_{\sigma(\operatorname{ord}_n(s)+1)} = 0.$$

Proof. It is ETS this holds when $A_1, \ldots, A_{\text{ord}_n(s)+1}$ each have only one nonzero entry per row which is set to 1. Let j_k satisfy $(A_k)_{0,j_k} = 1$ where we take our matrices to be 0-indexed. We observe by taking indices mod n,

$$(A_1A_2\cdots A_{\operatorname{ord}_n(s)}A_{\operatorname{ord}_n(s)+1})_{0,j} = 1 \iff (A_2\cdots A_{\operatorname{ord}_n(s)}A_{\operatorname{ord}_n(s)+1})_{j_1,j} = 1 \iff (A_3\cdots A_{\operatorname{ord}_n(s)}A_{\operatorname{ord}_n(s)+1})_{j_2+sj_1,j} = 1 \iff \cdots \iff (A_{\operatorname{ord}_n(s)+1})_{j_{\operatorname{ord}_n(s)}+sj_{\operatorname{ord}_n(s)-1}+\cdots+s^{\operatorname{ord}_n(s)-1}j_{1,j}} \iff j = j_{\operatorname{ord}_n(s)+1} + sj_{\operatorname{ord}_n(s)} + s^2 j_{\operatorname{ord}_n(s)-1} + \cdots + s^{\operatorname{ord}_n(s)}j_1 \mod n \iff j = j_{\operatorname{ord}_n(s)+1} + sj_{\operatorname{ord}_n(s)} + s^2 j_{\operatorname{ord}_n(s)-1} + \cdots + j_1 \mod n$$

and

$$(A_{\operatorname{ord}_n(s)+1}A_2 \cdots A_{\operatorname{ord}_n(s)}A_1)_{0,j} = 1 \iff (A_2 \cdots A_{\operatorname{ord}_n(s)}A_1)_{j_{\operatorname{ord}_n(s)+1,j}} = 1 \\ \iff (A_3 \cdots A_{\operatorname{ord}_n(s)}A_1)_{j_2+sj_{\operatorname{ord}_n(s)+1,j}} = 1 \\ \iff \cdots \\ \iff (A_1)_{j_{\operatorname{ord}_n(s)}+sj_{\operatorname{ord}_n(s)-1}+\dots+s^{\operatorname{ord}_n(s)-1}j_{\operatorname{ord}_n(s)+1,j}} \\ \iff j = j_1 + sj_{\operatorname{ord}_n(s)} + s^2 j_{\operatorname{ord}_n(s)-1} + \dots + s^{\operatorname{ord}_n(s)}j_{\operatorname{ord}_n(s)+1} \mod n \\ \iff j = j_1 + sj_{\operatorname{ord}_n(s)} + s^2 j_{\operatorname{ord}_n(s)-1} + \dots + j_{\operatorname{ord}_n(s)+1} \mod n.$$

Hence by symmetry,

$$A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(\mathrm{ord}_n(s))}A_{\sigma(\mathrm{ord}_n(s)+1)} = A_{\sigma(\mathrm{ord}_n(s)+1)}A_{\sigma(2)}\cdots A_{\sigma(\mathrm{ord}_n(s))}A_{\sigma(1)}.$$

The proof of proposition 2 then gives us a natural way combine our results for circulant matrices with arbitrary fixed *s*.

Theorem 1. Let $n, s \ge 1$. Let $0 \le x_1 < x_2$ be the first integers for which $s^{x_1} = s^{x_2} \mod n$. Then for any $n \times n$ circulant matrices with shift $s A_1, \ldots, A_{x_2+1}$ and any $(x_2 + 1)!/2$ constants c_i ,

$$\sum_{\sigma\in S_{x_2+1}}\operatorname{sgn}(\sigma)\cdot c_{\sigma|_{[2,x_2-x_1]\cup[x_2-x_1+2,x_2+1]}}\cdot A_{\sigma(1)}\cdots A_{\sigma(x_2+1)}=0.$$

Proof. It is ETS this holds when A_1, \ldots, A_{x_2+1} each have only one nonzero entry per row which is set to 1. Let j_k satisfy $(A_k)_{0,j_k} = 1$ where we take our matrices to be 0-indexed. We observe by taking indices mod n and applying the same logic as in the proof of proposition 2,

$$(A_1A_2\cdots A_{x_2-x_1}A_{x_2-x_1+1}A_{x_2-x_1+2}\cdots A_{x_2+1})_{0,j} = 1 \iff j_{x_2+1} + sj_{x_2} + \cdots + s^{x_1-1}j_{x_2-x_1} + s^{x_1}j_{x_2-x_1+1} + s^{x_1+1}j_{x_2-x_1+2} + \cdots + s^{x_2-1}j_2 + s^{x_2}j_1 \mod n \\ \iff j_{x_2+1} + sj_{x_2} + \cdots + s^{x_1-1}j_{x_2-x_1} + s^{x_1}j_{x_2-x_1+1} + s^{x_1+1}j_{x_2-x_1+2} + \cdots + s^{x_2-1}j_2 + s^{x_1}j_1 \mod n$$

and

$$\begin{aligned} (A_{x_2-x_1+1}A_2\cdots A_{x_2-x_1}A_1A_{x_2-x_1+1}\cdots A_{x_2+1})_{0,j} &= 1 \\ \iff j_{x_2+1}+sj_{x_2}+\cdots +s^{x_1-1}j_{x_2-x_1}+s^{x_1}j_1+s^{x_1+1}j_{x_2-x_1+2}+\cdots +s^{x_2-1}j_2+s^{x_2}j_{x_2-x_1+1} \mod n \\ \iff j_{x_2+1}+sj_{x_2}+\cdots +s^{x_1-1}j_{x_2-x_1}+s^{x_1}j_1+s^{x_1+1}j_{x_2-x_1+2}+\cdots +s^{x_2-1}j_2+s^{x_1}j_{x_2-x_1+1} \mod n. \end{aligned}$$

Hence by symmetry,

$$\begin{aligned} A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(x_{2}-x_{1})}A_{\sigma(x_{2}-x_{1}+1)}A_{\sigma(x_{2}-x_{1}+2)}\cdots A_{\sigma(x_{2}+1)} \\ &= A_{\sigma(x_{2}-x_{1}+1)}A_{\sigma(2)}\cdots A_{\sigma(x_{2}-x_{1})}A_{\sigma(1)}A_{\sigma(x_{2}-x_{1}+2)}\cdots A_{\sigma(x_{2}+1)}. \end{aligned}$$

Based on experimental data, we conjecture that theorem 1 is tight.

3.2 A Toeplitz Matrix Identity

Proposition 3. Let A_1, A_2 be $n \times n$ upper triangular Toeplitz matrices. Then $A_1A_2 = A_2A_1$.

Proof. Taking our matrices to be 0-indexed, we compute:

$$(A_1A_2)_{i,j} = \sum_{k=i}^{j} (A_1)_{i,k} (A_2)_{k,j}$$

 $= \sum_{k=i}^{j} (A_1)_{0,k-i} (A_2)_{0,j-k}$
 $= \sum_{k=i}^{j} (A_2)_{0,k-i} (A_1)_{0,j-k}$
 $= (A_2A_1)_{i,j}.$

3.3 Data and Conjectures

Based on our experimental data, we make the following conjecture:

Conjecture 1. Let A_1, \ldots, A_{n+1} be $n \times n$ Hankel matrices with all zeros below the skew-diagonal. Then

$$\sum_{\sigma \in S_{n+1}} \operatorname{sgn}(\sigma) \cdot A_{\sigma(1)} \cdots A_{\sigma(n+1)} = 0.$$

Despite being quite similar to Toeplitz matrices, we observe that this would imply that we require many more matrices compared to the similar assumptions of proposition 3. However, we conjecture that Toeplitz matrices and Hankel matrices behave the same way when restricted in the following way:

Conjecture 2. Let A_1, \ldots, A_4 be $n \times n$ tridiagonal Toeplitz matrices. Let B_1, \ldots, B_4 be $n \times n$ Hankel matrices with all nonzero entries contained in the main skew-diagonal, the lower skew-diagonal, and the upper skew-diagonal. Then

$$\sum_{\sigma \in S_4} \operatorname{sgn}(\sigma) \cdot A_{\sigma(1)} A_{\sigma(2)} A_{\sigma(3)} A_{\sigma(4)} = 0 = \sum_{\sigma \in S_4} \operatorname{sgn}(\sigma) \cdot B_{\sigma(1)} B_{\sigma(2)} B_{\sigma(3)} B_{\sigma(4)}.$$

Dropping the Toeplitz restricted but keeping the upper triangular shape, we conjecture **Conjecture 3.** Let A_1, \ldots, A_n be $n \times n$ upper triangular matrices with zeros on the main diagonal. Then

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \cdot A_{\sigma(1)} \cdots A_{\sigma(n)} = 0.$$

Less precisely, we found interesting behavior for the following classes of matrices:

- tridiagonal matrices with all zeros on the main diagonal
- symmetric matrices with all zeros on the main diagonal
- antisymmetric matrices with all zeros on the main diagonal
- arbitrary matrices with all zeros on the first row

We were able to find likely identities for these classes with $n \times n$ matrices for small fixed n requiring less than 2n matrices. The tridiagonal, symmetric, and antisymmetric were particularly interesting since the outputs of our Maple procedures suggest that these sums might not work by taking the constants to be the signs of the permutations.

4 Conclusion

References

- Shimshon Amitsur and Jakob Levitzki. "Minimal identities for algebras". In: Proceedings of the American Mathematical Society. Vol. 1. 4. 1950, pp. 449–463.
- [2] Bertram Kostant. "A Theorem of Frobenius, a Theorem of Amitsur-Levitski and Cohomology Theory". In: Indiana Univ. Math. J. 7 (2 1958), pp. 237–264. ISSN: 0022-2518.

- [3] Claudio Procesi. On the theorem of Amitsur-Levitzki. 2013. arXiv: 1308.2421 [math.RA]. URL: https://arxiv.org/abs/1308.2421.
- [4] Ju. P. Razmyslov. "Identities with trace in full matrix algebras over a field of characteristic zero". In: *Mathematics of the USSR-Izvestiya* 8.4 (1974), p. 727.
- [5] Shmuel Rosset. "A new proof of the Amitsur-Levitski identity". In: Israel Journal of Mathematics 23 (1976), pp. 187–188. URL: https://api.semanticscholar.org/CorpusID: 121625182.
- [6] Richard G. Swan. "An application of graph theory to algebra". In: *Proceedings of the American Mathematical Society*. Vol. 14. 3. 1963, pp. 367–373.
- [7] Richard G. Swan. "Correction to: "An application of graph theory to algebra"". In: *Proceed-ings of the American Mathematical Society*. Vol. 21. 2. 1969, pp. 379–380.