The Jordan canonical form for a class of zero-one matrices

David A. Cardon[a], Bradford Tuckfield

[a]*Department of Mathematics, Brigham Young University, Provo, UT 84602*

Abstract

Let $f\colon \mathbb{N} \to \mathbb{N}$ be a function. Let $A_n = (a_{ij})$ be the $n \times n$ matrix defined by $a_{ij} = 1$ if $i = f(j)$ for some i and j and $a_{ij} = 0$ otherwise. We describe the Jordan canonical form of the matrix A_n in terms of the directed graph for which A_n is the adjacency matrix. We discuss several examples including a connection with the Collatz $3n + 1$ conjecture.

Keywords: Jordan canonical form, directed graph, adjacency matrix
2010 MSC: 15A21, 05C20

1. Introduction

Let $f\colon \mathbb{N} \to \mathbb{N}$ be any function. For each $n \in \mathbb{N}$, we define the $n \times n$ matrix $A_n = (a_{ij})$ by

$$a_{ij} = \begin{cases} 1 & \text{if } i = f(j) \text{ for some } i, j \in \{1, \dots, n\}, \\ 0 & \text{otherwise.} \end{cases}$$

The matrix A_n contains partial information about f. We may regard A_n as the adjacency matrix for the directed graph Γ_n with vertices labeled $1, \dots, n$ having a directed edge from vertex j to vertex i if and only if $i = f(j)$. The main purpose of this paper is to describe the Jordan canonical form of A_n in terms of the graph Γ_n. This description is given in Theorem 6.

As a motivating example, let f be the function

$$f(n) = \begin{cases} \frac{3n+1}{2} & \text{if } n \text{ is odd,} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

The Collatz conjecture states that, for each $k \in \mathbb{N}$, the sequence

$$k,\ f(k),\ (f \circ f)(k),\ (f \circ f \circ f)(k), \ldots$$

contains the number 1. In §5, we will discuss this example in more detail where we develop an explicit formula for the number of Jordan blocks for the eigenvalue 0 in the Jordan decomposition of the matrix A_n.

The use of combinatorial and graph theoretic methods for understanding the Jordan canonical form has a long history. In 1837, Jacobi showed that an $n \times n$ matrix is similar to an upper triangular matrix. Many proofs of the Jordan form rely on this result. Brualdi's 1987 expository article [2] explains a graph theoretic interpretation of Turnbull and Aitken's 1932 combinatorial proof of Jacobi's Theorem (see Chap. 6 §4 of [9]). A recent and very accessible treatment of the interplay among matrices, combinatorics, and graphs is given in [3] by Brualdi and Cvetković.

The remainder of this paper is organized as follows: In §2, we describe how to partition the directed graph Γ_n into chains and cycles. These chains and cycles are related to the Jordan form of A_n. In §3, we state and prove our main theorem, Theorem 6, which describes the Jordan block structure of A_n in terms of the cycles and chains of the graph Γ_n. In §5, we apply Theorem 6 to several examples.

2. The directed graph Γ_n associated with A_n

We will form a partition of the directed graph Γ_n, which was defined in §1, into chains and cycles. The Jordan decomposition of the adjacency matrix A_n will be related to the lengths of these chains and cycles. Recall that for the function $f \colon \mathbb{N} \to \mathbb{N}$ and natural number $n \in \mathbb{N}$, we define Γ_n to be the directed graph with vertices $1, \ldots, n$ having a directed edge from j to i if and only if $f(j) = i$.

Definition 1. A *chain* in Γ_n is an ordered list of distinct vertices $C = \{c_1, c_2, \ldots, c_r\}$ such that $f(c_j) = c_{j+1}$ for $1 \leq j < r$ but $f(c_r) \neq c_1$. A *cycle* in Γ_n is an ordered list of distinct vertices $Z = \{z_1, z_2, \ldots, z_r\}$ such that $f(z_j) = z_{j+1}$ for $1 \leq j < r$ and $f(z_r) = z_1$. In either case, we call r the *length* of the chain or cycle and write $r = \operatorname{len} C$ or $r = \operatorname{len} Z$.

Note that in Definition 1 a single vertex $\{i\}$ is a chain or a cycle, but since either $f(i) \neq i$ or $f(i) = i$, it is not both a chain and a cycle. Although an arbitrary directed graph may contain two unequal cycles that share a common vertex, this is not possible for Γ_n. Since Γ_n results from a function $f\colon \mathbb{N} \to \mathbb{N}$, if Z_1 and Z_2 are cycles that share a common vertex, then $Z_1 = Z_2$. Thus, unequal cycles in Γ_n are disjoint.

Definition 2. If $C = \{i_1, \dots, i_s\}$ is a chain of Γ_n, then i_s is called the *terminal point* of the chain. A vertex k of Γ_n such that $f(k) > n$ is a *terminal point* of Γ_n. If k is a vertex of Γ_n such that $f(i) = f(j) = k$ for some i and j with $i \neq j$, then k is a *merge point* of Γ_n.

Definition 3. A *partition* of Γ_n is a collection of disjoint cycles and chains whose union is Γ_n. A *proper partition* of Γ_n is a partition

$$P = \{Z_1, \dots, Z_r, C_1, \dots, C_s\}$$

where $Z_1, \dots, Z_r$ are cycles and $C_1, \dots, C_s$ are chains satisfying the following properties:

1. Each cycle in Γ_n is equal to Z_i for some i.
2. If $\Gamma_n^{(i)}$ is the subgraph of Γ_n obtained by deleting the vertices in the cycles $Z_1, \dots, Z_r$ and in the chains $C_1, \dots, C_i$, then C_{i+1} is a chain of maximal length in $\Gamma_n^{(i)}$.

Lemma 1. *Proper partitions of Γ_n exist.*

Proof. As noted above, the cycles of Γ_n are mutually disjoint. Label them as $Z_1, \dots, Z_r$. Let $\Gamma_n^{(0)}$ be the subgraph of Γ_n obtained by removing all vertices belonging to the cycles $Z_1, \dots, Z_r$. The graph $\Gamma_n^{(0)}$ is an acyclic graph. Then inductively define C_i and $\Gamma_n^{(i)}$ for $i \geq 1$ by choosing C_i to be a chain of maximal length in $\Gamma_n^{(i)}$ and letting $\Gamma_n^{(i+1)}$ be the subgraph of $\Gamma_n^{(i)}$ obtained by deleting the vertices of C_i. □

If

$$P = \{Z_1, \dots, Z_r, C_1, \dots, C_s\} \quad \text{and} \quad P' = \{Z'_1, \dots, Z'_{r'}, C'_{1'}, \dots, C'_{s'}\}$$

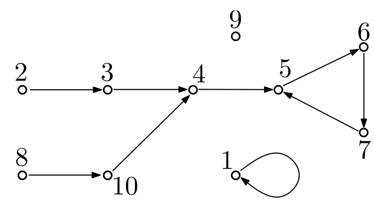
are any two proper partitions of Γ_n, then it is clear that $r = r'$ and the cycles $Z_1, \dots, Z_r$ are the same as the cycles $Z'_1, \dots, Z'_{r'}$ up to reordering. It may be less obvious that $s = s'$ and $\operatorname{len} C_i = \operatorname{len} C'_i$ for $1 \leq i \leq s$. This fact is stated in Corollary 7 below.

Example 1. Suppose $f\colon \mathbb{N} \to \mathbb{N}$ takes the values

$$1 \mapsto 1,\ 2 \mapsto 3,\ 3 \mapsto 4,\ 5 \mapsto 6,\ 6 \mapsto 7,\ 7 \mapsto 5,\ 8 \mapsto 10,\ 9 \mapsto 500,\ 10 \mapsto 4, \ldots$$

The graph Γ_{10} and adjacency matrix A_{10} are:

$$A_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$



The vertices 4 and 5 are merge points of Γ_n. The vertex 9 is a terminal point of Γ_{10}. Two proper partitions of Γ_{10} are:

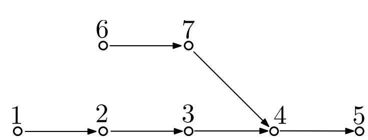
$$\begin{array}{ll} Z_1 = \{1\} & Z_1' = \{5,6,7\} \\ Z_2 = \{5,6,7\} & Z_2' = \{1\} \\ C_1 = \{2,3,4\} & C_1' = \{8,10,4\} \\ C_2 = \{8,10\} & C_2' = \{2,3\} \\ C_3 = \{9\} & C_3' = \{9\} \end{array}$$

Example 2. Suppose $f\colon \mathbb{N} \to \mathbb{N}$ takes the values

$$1 \mapsto 2,\ 2 \mapsto 3,\ 3 \mapsto 4,\ 4 \mapsto 5,\ 5 \mapsto 100,\ 6 \mapsto 7,\ 7 \mapsto 4,\ \ldots$$

The graph Γ_7 and adjacency matrix A_7 are:

$$A_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$



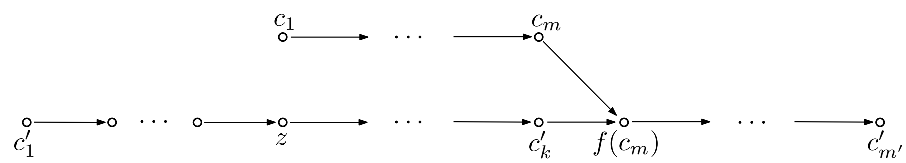


Figure 1: If the chain $C' = \{c'_1, \ldots, c'_{m'}\}$ contains the merge point $f(c_m) = f(c'_k)$, as in Lemma 2, then $k \geq m$. There exists z on the chain C' with $f^m(z) = f^m(c_1) = f(c_m)$.

Vertex 4 is a merge point, and vertex 5 is a terminal point of Γ_7. A proper partition of Γ_7 is

$$C_1 = \{1, 2, 3, 4, 5\} \quad C_2 = \{6, 7\},$$

while an improper partition of Γ_7 is

$$C'_1 = \{6, 7, 4, 5\} \quad C'_2 = \{1, 2, 3\}.$$

Observe that in the proper partition, the chain containing the merge point 4 is longer than the other chain.

The terminal and merge points of Γ_n will play a crucial role in the Jordan decomposition of A_n. The next lemma makes precise the situation in Example 2 as well as the case in which the merge point belongs to a cycle.

Lemma 2. *Let $C = \{c_1, \ldots, c_m\}$ be any chain in a proper partition of Γ_n. Then exactly one of the following occurs:*

1. *The terminal point c_m of the chain is a terminal point of the graph Γ_n.*
2. *The point $f(c_m)$ is a merge point of Γ_n.*

Furthermore, if $f(c_m)$ is a merge point and $f(c_m)$ belongs to another chain $C' = \{c'_1, \ldots, c'_{m'}\}$, then $f(c_m) = f(c'_k)$ where $k \geq m$. (This is illustrated in Figure 1.) Consequently, if $f(c_m)$ is a merge point belonging to either a cycle or a chain, then there is a unique vertex z in the cycle or chain containing $f(c_r)$ such that $f^m(z) = f^m(c_1) = f(c_m)$, where f^m is the composition of f with itself m times.

Proof. Suppose the terminal point c_m of the chain is not a terminal point of the graph Γ_n. If $f(c_m)$ belongs to one of the cycles of Γ_n, then $f(c_m)$ is a merge point since $f(c_m)$ has both c_m and some point of the cycle as preimages. If $f(c_m)$ belongs to another chain C', then either $f(c_m)$ is the initial point of the chain C' or it is not. If $f(c_m)$ is not the initial point of C', then c_m and a point of C' are preimages of $f(c_m)$ making $f(c_m)$ a merge point. If $f(c_m)$ is the initial point of C', then the partition is not proper because C and C' could be joined to form a longer chain, which is a contradiction. This proves that either c_m is a terminal point of Γ_n or $f(c_m)$ is a merge point of Γ_n.

Now, suppose $f(c_m)$ belongs to another chain $C' = \{c'_1, \ldots, c'_{m'}\}$. Then $f(c_m) = f(c'_k)$ for some k. Necessarily $k < m'$. If, by way of contradiction, $k < m$, then the vertices in the two chains C and C' could be repartitioned to belong to the new chains

$$C'' = \{c_1, \ldots, c_m, c'_{k+1}, \ldots, c'_{m'}\} \quad \text{and} \quad C''' = \{c'_1, \ldots, c'_k\}.$$

Since $k < m'$,

$$\operatorname{len} C'' = m' - k + m > m = \operatorname{len} C.$$

Since $k < m$,

$$\operatorname{len} C'' = m' - k + m > m' = \operatorname{len} C'.$$

Thus, the original pair of chains C and C' violate the maximality condition of a proper partition in Definition 3, a contradiction. Therefore $k \geq m$. $\square$

3. The Jordan Structure of A_n

In this section we will state the main result of this paper (Theorem 6) which describes the Jordan canonical form of the adjacency matrix A_n of the graph Γ_n. We will need several standard facts (Propositions 3 and 4 below) about the Jordan canonical form. Good references for this material are [4, Ch. 7], [5, Ch. 3], and [8, Ch. 6].

Definition 4. For a complex number λ and natural number m, $J_m(\lambda)$ will denote the $m \times m$ matrix

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Definition 5. Let A be an $n \times n$ matrix with complex entries. A nonzero vector v is a *generalized eigenvector* of A corresponding to the complex number λ if $(A - \lambda I)^p v = 0$ for some positive integer p.

Definition 6. Let v be a generalized eigenvector of A for the eigenvalue λ and let p be the smallest positive integer such that $(A - \lambda I)^p v = 0$. Then the ordered set

$$\{(A - \lambda I)^{p-1} v, (A - \lambda I)^{p-2} v, \ldots, (A - \lambda I)v, v\} \tag{1}$$

is a *chain of generalized eigenvectors* of A corresponding to λ. Observe that the first elements of the list, $(A - \lambda I)^{p-1} v$, is an ordinary eigenvector.

Note. In the literature, many authors refer to the list of generalized eigenvectors in Definition 6 as a *cycle* of generalized eigenvectors. In the context of this paper, it is better to call it a chain.

Proposition 3 (Linear Independence of Generalized Eigenvectors)**.** *Let λ be an eigenvalue of A and let $\{\gamma_1, \ldots, \gamma_s\}$ be chains of generalized eigenvectors of A corresponding to λ. If the initial vectors of the γ_i's form a linearly independent set, then the γ_i's are disjoint ($\gamma_i \cap \gamma_j = \emptyset$ for $i \neq j$) and the union $\cup_{i=1}^{s} \gamma_i$ is linearly independent.*

Proposition 4 (Jordan Canonical Form)**.** *Let A be an $n \times n$ complex matrix. Then there exists a basis β of $\mathbb{C}^n$ consisting of disjoint chains $\beta_1, \ldots, \beta_r$ of generalized eigenvectors of lengths $n_1, \ldots, n_r$ for the eigenvalues $\lambda_1, \ldots, \lambda_r$ with $n = n_1 + \cdots + n_r$ such that if Q is the matrix whose columns are the members of the basis β then*

$$Q^{-1} A Q = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_r}(\lambda_r).$$

As a preliminary step to determining the Jordan decomposition of the adjacency matrix A_n of the graph Γ_n, we begin with the following simple observation:

Lemma 5. *Every eigenvalue of A_n is either 0 or a root of unity.*

Proof. The jth column of A_n is a zero column if $f(j) > n$. The jth column contains a single 1 if $1 \leq f(j) \leq n$. Thus, for any $k \in \mathbb{N}$, the matrix product A_n^k also consists of either zero columns or columns containing a single 1. Consequently, the infinite sequence

$$I, A_n, A_n^2, A_n^3, \ldots$$

must contain a repetition since there are only finitely many distinct $n \times n$ matrices whose columns are zero columns or contain a single 1. Let $0 \leq i < j$ be exponents such that $A_n^i = A_n^j$. Then A_n satisfies the polynomial $x^j - x^i = x^{j-i}(x^j - 1)$. The eigenvalues of A_n must be a subset of the roots of this polynomial. Hence, all eigenvalues are either zero or roots of unity. □

We are now ready to state the main result of this paper:

Theorem 6. *Let $f\colon \mathbb{N} \to \mathbb{N}$ be a function. Let Γ_n be the directed graph associated with f for the natural number n, and let A_n be its adjacency matrix as defined in §1. Suppose*

$$P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$$

is a proper partition of Γ_n, as in Definition 3, where $Z_1, \ldots, Z_s$ are the cycles and $C_1, \ldots, C_s$ are the chains. Write the lengths of the cycles and chains as

$$\operatorname{len} Z_j = \ell_j \quad (1 \leq j \leq r) \qquad \textit{and} \qquad \operatorname{len} C_j = m_j \quad (1 \leq j \leq s).$$

Let $\omega_j = \exp(2\pi i/\ell_j)$ be a primitive ℓ_jth root of unity. The Jordan decomposition of A_n contains the following 1×1 Jordan blocks for the eigenvalues which are roots of unity:

$$J_1(\omega_j^k) \quad \textit{for } 1 \leq j \leq r \textit{ and } 1 \leq k \leq \ell_j.$$

The Jordan decomposition contains the following blocks associated with the eigenvalue 0:

$$J_{m_1}(0), J_{m_2}(0), \ldots, J_{m_s}(0).$$

Remark. The proof of Theorem 6 (given in §4) will construct an explicit basis (Lemmas 8 and 10) for $\mathbb{C}^n$ consisting of generalized eigenvectors of A_n. Letting Q be the matrix whose columns are these vectors gives $J = Q^{-1}A_nQ$ where J is the Jordan decomposition of A_n.

Corollary 7. *Suppose*

$$P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\} \qquad \textit{and} \quad P' = \{Z'_1, \ldots, Z'_{r'}, C'_{1'}, \ldots, C'_{s'}\}$$

are any two proper partitions of Γ_n. Then $s = s'$ and $\operatorname{len} C_j = \operatorname{len} C'_j$ for $1 \leq j \leq s$.

Proof. This is an immediate consequence of the uniqueness of the Jordan decomposition of A_n. By Theorem 6, the block sizes associated with the eigenvalue 0 are given by the two descending lists of numbers:

$$m_1 \geq m_2 \geq \cdots \geq m_s \quad \text{and} \quad m'_1 \geq m'_2 \geq \cdots \geq m'_{s'}.$$

So, $s = s'$ and $m_j = m'_j$ for $1 \leq j \leq s$. □

4. Proof of Theorem 6

The proof of Theorem 6 will proceed as follows: From Lemma 5, each eigenvalue of A_n is either a root of unity or zero. In Lemma 8 below, we attached an eigenvector of A_n associated with a root of unity to each vertex of each cycle in Γ_n. In Lemma 9, this set of eigenvectors for the roots of unity is shown to be linearly independent. In Lemma 10, we attach chains of generalized eigenvectors of A_n for the eigenvalue 0 to chains of the graph Γ_n. In Lemma 11 we show that these generalized eigenvectors also form a linearly independent set and that the set of all generalized eigenvectors attached to the vertices of Γ_n via Lemmas 8 and 10 is a Jordan basis of $\mathbb{C}^n$ for the matrix A_n.

Throughout the section $P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$ will be a proper partition of Γ_n. Write the lengths of the cycles and chains as

$$\operatorname{len} Z_j = \ell_j \quad (1 \leq j \leq r) \quad \text{and} \quad \operatorname{len} C_j = m_j \quad (1 \leq j \leq s).$$

We have the relationship

$$\ell_1 + \cdots + \ell_s + m_1 + \cdots + m_s = n.$$

The ith standard basis vector of $\mathbb{C}^n$ will be denoted by e_i.

Lemma 8. *Let $Z = \{z_1, \ldots, z_\ell\}$ be any cycle in the partition P, and let $\omega = \exp(2\pi i/\ell)$ be a primitive ℓth root of unity. Then the vector*

$$v_k = \sum_{j=1}^{\ell} \omega^{-kj} e_{z_j} \tag{2}$$

is an eigenvector of A_n for the eigenvalue ω^k. Furthermore,

$$\operatorname{span}\{v_1, \ldots, v_\ell\} = \operatorname{span}\{e_{z_1}, \ldots, e_{z_\ell}\}. \tag{3}$$

We will say that the eigenvector v_k is attached to the vertex z_k.

Proof. Since $A_n e_{z_j} = e_{z_{j+1}}$ for $1 \leq j < \ell$ and $Ae_{z_\ell} = e_{z_1}$ and since $\omega^\ell = 1$, we have

$$\begin{aligned} A_n v_k &= \sum_{j=1}^{\ell} \omega^{-kj} A_n e_j = \sum_{j=1}^{\ell-1} \omega^{-kj} e_{z_{j+1}} + \omega^{-k\ell} e_{z_1} \\ &= e_{z_1} + \sum_{j=2}^{\ell} \omega^{-k(j-1)} e_{z_j} = \omega^k \sum_{j=1}^{\ell} \omega^{-kj} e_{z_j} = \omega^k v_k. \end{aligned}$$

Because the eigenvectors $v_1, \ldots, v_\ell$ all belong to distinct eigenvalues they form a linearly independent set whose span has dimension ℓ. But $\operatorname{span}\{v_1, \ldots, v_\ell\}$ is a subspace of $\operatorname{span}\{e_{z_1}, \ldots, e_{z_\ell}\}$ whose dimension is also ℓ. So, the two subspaces are equal. □

Lemma 9. *In a proper partition $P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$ of Γ_n, the set of all eigenvectors attached to the vertices in the cycles $Z_1, \ldots, Z_r$ is a linearly independent set.*

Proof. This follows immediately from equation (3) in Lemma 8 and the fact that any two cycles in Γ_n are disjoint. □

Next we determine generalized eigenvectors of A_n associated with the eigenvalue 0. Recall from Lemma 2, that if $C = \{c_1, \ldots, c_s\}$ is a chain in a proper partition of Γ_n, then either c_s is a terminal point of Γ_n or $f(c_s)$ is a merge point.

Lemma 10. *Let $C = \{c_1, \ldots, c_m\}$ be any chain in a proper partition of Γ_n.*

1. *If c_m is a terminal point of the graph Γ_n, then*

$$\{e_{c_m}, e_{c_{m-1}}, \ldots, e_{c_2}, e_{c_1}\}$$

   is a chain of generalized eigenvectors of A_n for the eigenvalue 0.

2. *If $f(c_m)$ is a merge point of Γ_n, let z be the vertex in the cycle or chain containing $f(c_m)$ such that $f^m(z) = f(c_m)$. (z exists by Lemma 2.) Then*

$$\{e_{c_m} - e_{f^{m-1}(z)}, \ldots, e_{c_3} - e_{f^2(z)}, e_{c_2} - e_{f(z)}, e_{c_1} - e_z\}$$

   is a chain of generalized eigenvectors of A_n for the eigenvalue 0.

In the first case, we say that the vector e_{c_i} is attached to the vertex c_i. In the second case, we say the vector $e_{c_i} - e_{f^{i-1}(z)}$ is attached to the vertex c_i.

Note. By convention, the *first* element of a chain of generalized eigenvectors, as in Equation (1) is the eigenvector, but the eigenvector corresponds to the *last* element of the chain $\{c_1, \ldots, c_m\}$ in Lemma 10. So, the order of indices in the subscripts is reversed.

Proof. If the first case, c_m is a terminal point of the graph Γ_n. This means that $f(c_m) > n$ which implies that the c_m column of A_n is zero. Then $A_n e_{c_m} = 0$. So, e_{c_m} is an eigenvector of A_n for the eigenvalue 0. Because $C = \{c_1, \ldots, c_m\}$ is a chain in Γ, $A_n e_{c_i} = e_{c_{i+1}}$ for $1 \leq i < m$. Therefore $\{e_{c_m}, \ldots, e_{c_1}\}$ is a chain of generalized eigenvectors of A_n for the eigenvalue 0.

In the second case, $e_{c_m} - e_{f^{m-1}(z)}$ is not the zero vector since $f^{m-1}(z)$ does not belong to the chain C. Then

$$A_n\big(e_{c_m} - e_{f^{m-1}(z)}\big) = e_{f(c_m)} - e_{f^m(z)} = 0.$$

So, $e_{c_m} - e_{f^{m-1}(z)}$ is an eigenvector for the eigenvalue 0. Since C is a chain, $A_n\big(e_{c_i} - e_{f^{i-1}(z)}\big) = e_{c_{i+1}} - e_{f^i(z)}$ for $1 \leq i < m$. Thus,

$$\{e_{c_m} - e_{f^{m-1}(z)}, \ldots, e_{c_3} - e_{f^2(z)}, e_{c_2} - e_{f(z)}, e_{c_1} - e_z\}$$

is a chain of generalized eigenvectors of A_n for the eigenvalue 0. □

With Lemmas 8 and 10 we have attached a generalized eigenvector to each of the n vertices of the graph Γ_n. The final step of the proof of Theorem 6 is to show that this collection of generalized eigenvectors is linearly independent. Then the chains of generalized eigenvectors that these lemmas attach to a proper partition of Γ_n will form a Jordan basis of $\mathbb{C}^n$ for the matrix A_n.

Lemma 11. *Let $P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$ be a proper partition of Γ_n. The set of all generalized vectors attached to vertices of the cycles $Z_1, \ldots, Z_r$ and to the vertices of the chains $C_1, \ldots, C_s$ is a linearly independent set of n vectors. Consequently, this set forms a Jordan basis of $\mathbb{C}^n$ for the matrix A_n.*

Proof. In Lemma 9 it was shown that the set of eigenvectors attached to vertices belonging to cycles is linearly independent. All of these eigenvectors are roots of unity which are, of course, nonzero. If $\text{len}(Z_i) = \ell_i$, then there are

$$\ell_1 + \cdots + \ell_s$$

such eigenvectors.

If $\text{len}(C_i) = m_i$, then there are

$$m_1 + \cdots + m_s$$

generalized vectors attached to the vertices of the chains $C_1, \ldots, C_s$. These generalized eigenvectors all belong to the generalized eigenspace of the eigenvalue 0. If these vectors are linearly independent, then then we will have a total of

$$n = \ell_1 + \cdots + \ell_s + m_1 + \cdots + m_s$$

linearly independent generalized eigenvectors since the union of linearly independent generalized eigenvectors from different generalized eigenspaces is linearly independent.

Thus, it remains to be shown that the generalized eigenvectors attached to the chains $C_1, \ldots, C_s$ form a linearly independent set.

Write $\ell = \ell_1 + \cdots + \ell_r$, and let $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a permutation that maps the numbers $1, \ldots, \ell$ to the vertices belonging to the cycles $Z_1, \ldots, Z_r$ and such that the chains (as ordered lists) are

$$\begin{aligned}
C_1 &= \{\sigma(\ell+1), \ldots, \sigma(\ell+m_1)\} \\
C_2 &= \{\sigma(\ell+m_1+1), \ldots, \sigma(\ell+m_1+m_2)\} \\
&\vdots \\
C_s &= \{\sigma(\ell+m_1+\cdots+m_{s-1}+1), \ldots, \sigma(\ell+m_1+\cdots+m_s)\}.
\end{aligned}$$

By Lemma 10, the eigenvector attached to the *last* element of the chain C_j is one of the following:

$$e_{\sigma(\ell+m_1+\cdots+m_j)} \qquad \text{or} \qquad e_{\sigma(\ell+m_1+\cdots+m_j)} - e_{\sigma(z_j)}, \tag{4}$$

for some appropriate z_j. In the second case, since $f(\sigma(\ell+m_1+\cdots+m_j))$ is a merge point, $\sigma(z_j)$ belongs either to a longer chain than C_j or $\sigma(z_j)$ belongs to a cycle. Either way, from the definition of σ,

$$z_j < \ell + m_1 + \cdots + m_{j-1} + 1.$$

Thus, the $n \times s$ matrix whose jth column is

$$e_{\ell+m_1+\cdots+m_j} \qquad \text{or} \qquad e_{\ell+m_1+\cdots+m_j} - e_{z_j},$$

for $1 \leq j \leq s$, is upper triangular. Since no column is the zero vector, this matrix has linearly independent columns. Permuting the rows of this matrix does not alter the linear independence of the columns. Thus, the set of eigenvectors from (4) for $1 \leq j \leq s$ in a linearly independent set. Since the vectors in (4) were the initial vectors of chains of generalized eigenvectors, Proposition 3 implies that the set of all generalized eigenvectors from those chains is a linearly independent set. Thus, the lemma has been proved. □

The proof of Theorem 6 is now complete.

5. Examples and Applications

We will next illustrate Theorem 6 with several examples.

Example 3. We will apply Theorem 6 to find the Jordan canonical form of the matrix A_{10} from Example 1. In that example, $f\colon \mathbb{N} \to \mathbb{N}$ is defined by

$$1 \mapsto 1,\ 2 \mapsto 3,\ 3 \mapsto 4,\ 5 \mapsto 6,\ 6 \mapsto 7,\ 7 \mapsto 5,\ 8 \mapsto 10,\ 9 \mapsto 500,\ 10 \mapsto 4, \ldots$$

and we found a proper partition of Γ_{10} to be

$$Z_1 = \{1\}, \quad Z_2 = \{5, 6, 7\}, \quad C_1 = \{2, 3, 4\}, \quad C_2 = \{8, 10\}, \quad C_3 = \{9\}.$$

Let $\omega = \exp(2\pi i/3)$. Then ω is a primitive cube root of unity and $\omega^3 = 1$. By the theorem, the Jordan blocks in the Jordan canonical form of A_{10} will be

$$J_1(1),\ J_1(\omega),\ J_1(\omega^2),\ J_1(\omega^3),\ J_3(0),\ J_2(0),\ J_1(0).$$

Using Lemmas 8 and 10 we may find a basis $\beta = \{v_1, \ldots, v_{10}\}$ of generalized eigenvectors (taking care to reverse the order of indices in the chains) as follows:

vertex	generalized eigenvector
1	$v_1 = e_1$
5	$v_2 = \omega^{-1\cdot 1}e_5 + \omega^{-1\cdot 2}e_6 + \omega^{-1\cdot 3}e_7 = \omega^2 e_5 + \omega e_6 + e_7$
6	$v_3 = \omega^{-2\cdot 1}e_5 + \omega^{-2\cdot 2}e_6 + \omega^{-2\cdot 3}e_7 = \omega e_5 + \omega^2 e_6 + e_7$
7	$v_4 = \omega^{-3\cdot 1}e_5 + \omega^{-3\cdot 2}e_6 + \omega^{-3\cdot 3}e_7 = e_5 + e_6 + e_7$
4	$v_5 = e_4 - e_7$ ($f(4) = f(4) = 5$ is a merge point.)
3	$v_6 = e_3 - e_6$ ($f^2(3) = f^2(3) = 5$ is a merge point.)
2	$v_7 = e_2 - e_5$ ($f^3(2) = f^3(5) = 5$ is a merge point.)
10	$v_8 = e_{10} - e_3$ ($f(10) = f(3) = 4$ is a merge point.)
8	$v_9 = e_8 - e_2$ ($f^2(8) = f^2(2) = 4$ is a merge point.)
9	$v_{10} = e_9$ (The chain containing 9 has a terminal point.)

Thus setting

$$Q = (v_1\, v_2\, \cdots\, v_{10}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \omega^2 & \omega & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & \omega & \omega^2 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

gives the Jordan canonical form

$$Q^{-1}A_{10}Q = \begin{pmatrix} \boxed{1} &&&&&&& \\ & \boxed{\omega^2} &&&&&& \\ && \boxed{\omega} &&&&& \\ &&& \boxed{1} &&&& \\ &&&& \boxed{\begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix}} &&& \\ &&&&& \boxed{\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}} && \\ &&&&&& \boxed{0} \end{pmatrix}.$$

Example 4 (The Collatz Problem)**.** Let f be the function

$$f(n) = \begin{cases} \frac{3n+1}{2} & \text{if } n \text{ is odd,} \\ \frac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

The well-known *Collatz conjecture* states that, for each $k \in \mathbb{N}$, the sequence

$$k,\ f(k),\ (f \circ f)(k),\ (f \circ f \circ f)(k), \ldots$$

contains the number 1. For an extensive annotated bibliography of the literature on this problem see Lagarias [6, 7]. For any $n \in \mathbb{N}$, we may consider the $n \times n$ matrix A_n and graph Γ_n associated with the the Collatz function. In this case, we will call A_n the *Collatz matrix*. For example,

$$A_8 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

We may apply Theorem 6 to the study of A_n and the Collatz problem. In [1], Dias et. al. also study the Collatz conjecture from the point of view of finite dimensional matrices, and they establish certain determinantal identities for these matrices.

Working several examples for small values of n (say n up to a few thousand) quickly leads to the following conjecture:

Conjecture 12. *For $n \in \mathbb{N}$, let Γ_n and A_n be the graph and adjacency matrix associated with the Collatz function. Then*

1. *The characteristic polynomial of A_n is $\det(xI_n - A_n) = x^{n-2}(x^2 - 1)$.*
2. *For $n \geq 2$, the only cycle in the graph Γ_n is the two-cycle $\{1, 2\}$.*
3. *For any fixed $k \geq 3$, if n is sufficiently large, then k belongs to the same component of graph as the cycle $\{1, 2\}$.*

We point out that, since this conjecture implies the Collatz conjecture, its proof would likely be quite difficult. The paper of Dias et. al. [1] has some discussion about the characteristic polynomial.

We ask the following questions:

Open Problem 13. *Let $P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$ be a proper partition of the graph Γ_n associated with the Collatz function.*

1. *Is $r = 1$?*
2. *What is the length $\text{len}(C_1) = m_1$ of the longest chain?*
3. *How many connected components does the graph Γ_n have?*

As a consolation prize, we can precisely describe the number s of chains in a proper partition:

Theorem 14. *For $n \geq 2$, let A_n be the $n \times n$ Collatz matrix, let Γ_n be the associated graph, and let $P = \{Z_1, \ldots, Z_r, C_1, \ldots, C_s\}$ be a proper partition of Γ_n. Then the number s of chains in the partition which, by Theorem 6, is also equal to the number of Jordan blocks for the eigenvalue 0 of the matrix A_n is*

$$n - \left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \frac{n-2}{3} \right\rfloor + \left\lfloor \frac{n-4}{6} \right\rfloor = \begin{cases} 2\lfloor n/6 \rfloor & \text{if } n \equiv 0 \pmod 6, \\ 2\lfloor n/6 \rfloor + 1 & \text{if } n \equiv 1 \pmod 6, \\ 2\lfloor n/6 \rfloor & \text{if } n \equiv 2 \pmod 6, \\ 2\lfloor n/6 \rfloor + 1 & \text{if } n \equiv 3 \pmod 6, \\ 2\lfloor n/6 \rfloor + 2 & \text{if } n \equiv 4 \pmod 6, \\ 2\lfloor n/6 \rfloor + 2 & \text{if } n \equiv 5 \pmod 6. \end{cases}$$

Proof. From the Jordan decomposition theorem, the number of Jordan blocks associated with the eigenvalue λ of A_n is $n - \text{rank}(T_n - \lambda I_n)$. In the case $\lambda = 0$, this is $s = n - \text{rank}(A_n)$. So, we need to compute $\text{rank}(A_n)$.

The even numbered columns of A_n consist of the standard basis vectors

$$e_1, e_2, \ldots, e_{\lfloor n/2 \rfloor}.$$

The odd numbered columns of A_n which are also nonzero consist of the standard basis vectors

$$e_2, e_5, e_8, e_{11}, \ldots, e_{3j+2},$$

where $j = \lfloor \frac{n-2}{3} \rfloor$ is the largest integer such that $3j + 2 \leq n$. The elements of the second list in common with the first list are

$$e_2, e_5, \ldots, e_{3k+2}$$

where $k = \lfloor \frac{n/2-2}{3} \rfloor = \lfloor \frac{n-4}{6} \rfloor$ is the largest integer such that $3k+2 \leq n/2$. So, the degree of the column space of A_n is

$$\mathrm{rank}(A_n) = \left\lfloor \frac{n}{2} \right\rfloor + (j-k) = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n-2}{3} \right\rfloor - \left\lfloor \frac{n-4}{6} \right\rfloor,$$

which proves that the number of Jordan blocks for the eigenvalue 0 is

$$s = n - \mathrm{rank}(A_n) = n - \left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \frac{n-2}{3} \right\rfloor + \left\lfloor \frac{n-4}{6} \right\rfloor$$

We obtain the remaining portion of the formula by writing $n = 6\lfloor n/6 \rfloor + \ell$ where $\ell \in \{0,1,2,3,4,5\}$ and considering each case. □

Acknowledgment

The first author is grateful to Professor Donald W. Robinson for teaching him the theory of the Jordan decomposition of a matrix in an exceptionally good course on linear algebra. Also, we thank Richard Brualdi for several suggestions, and we thank the reviewer.

References

[1] J. F. Alves, M. M. Graça, M. E. Sousa Dias, J. Sousa Ramos, A linear algebra approach to the conjecture of Collatz, Linear Algebra Appl. 394 (2005) 277–289.

[2] R. A. Brualdi, The Jordan canonical form: an old proof, Amer. Math. Monthly 94 (3) (1987) 257–267.

[3] R. A. Brualdi, D. Cvetković, A combinatorial approach to matrix theory and its applications, Discrete Mathematics and its Applications (Boca Raton), CRC Press, Boca Raton, FL, 2009.

[4] S. H. Friedberg, A. J. Insel, L. E. Spence, Linear algebra, 3rd Edition, Prentice Hall Inc., Upper Saddle River, NJ, 1997.

[5] R. A. Horn, C. R. Johnson, Matrix analysis, Cambridge University Press, Cambridge, 1990, corrected reprint of the 1985 original.

[6] J. C. Lagarias, The $3x + 1$ problem: An annotated bibliography (1963–1999). URL http://arxiv.org/abs/math/0309224v13

[7] J. C. Lagarias, The $3x+1$ problem: An annotated bibliography, II (2000–2009). URL http://arxiv.org/abs/math/0608208v5

[8] P. Lancaster, M. Tismenetsky, The theory of matrices, 2nd Edition, Computer Science and Applied Mathematics, Academic Press Inc., Orlando, FL, 1985.

[9] H. Turnbull, A. Aitken, An Introduction to the Theory of Canonical Matrices, Blackie and Son, London, 1932.