A graph-theoretic method for choosing a spanning set for a finite-dimensional vector space, with applications to the Grossman-Larson-Wright module and the Jacobian conjecture

Dan Singer
Department of Mathematics and Statistics
Minnesota State University, Mankato
dan.singer@mnsu.edu

Submitted: Dec 10, 2008; Accepted: Mar 23, 2009; Published: Mar 31, 2009
Mathematics Subject Classifications: 05C99, 05E99, 14R15, 15A03

Abstract

It is well known that a square zero pattern matrix guarantees non-singularity if and only if it is permutationally equivalent to a triangular pattern with nonzero diagonal entries. It is also well known that a nonnegative square pattern matrix with positive main diagonal is sign nonsingular if and only if its associated digraph does not have any directed cycles of even length. Any $m \times n$ matrix containing an $n \times n$ sub-matrix with either of these forms will have full rank. We translate this idea into a graph-theoretic method for finding a spanning set of vectors for a finite-dimensional vector space from among a set of vectors generated combinatorially. This method is particularly useful when there is no convenient ordering of vectors and no upper bound to the dimensions of the vector spaces we are dealing with. We use our method to prove three properties of the Grossman-Larson-Wright module originally described by David Wright: $\overline{\mathcal{M}}(3,\infty)_m = 0$ for $m \geq 3$, $\overline{\mathcal{M}}(4,3)_m = 0$ for $5 \leq m \leq 8$, and $\overline{\mathcal{M}}(4,4)_8 = 0$. The first two properties yield combinatorial proofs of special cases of the homogeneous symmetric reduction of the Jacobian conjecture.

1 Introduction

A classic problem in algebraic combinatorics is to show that the ring of symmetric functions in n variables, $\Lambda_n = \mathbb{Z}[x_1, \ldots, x_n]^{\mathcal{S}_n}$, is generated by the elementary symmetric functions $e_1, \ldots, e_n$, and that the latter are algebraically independent over $\mathbb{Z}$. The proof, as given in [8], is to define $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots$ for each descending partition $\lambda = (\lambda_1, \lambda_2, \ldots)$

with parts of size $\leq n$, then observe that

$$e_{\lambda'} = m_\lambda + \sum_{\mu} a_{\lambda\mu} m_\mu,$$

where λ' is the conjugate partition, m_λ is the monomial symmetric function, the $a_{\lambda\mu}$ are non-negative integers, and the sum is taken over partitions μ which are later than λ in the reverse lexicographic ordering. The crux of the proof is that there is a natural ordering of the m_λ's and the e_λ's in which the corresponding coefficient matrix is unitriangular. Since the monomial symmetric functions form a $\mathbb{Z}$-basis for Λ_n, so do the e_λ.

In this paper we describe a graph-theoretic method for finding a spanning set for a finite-dimensional vector space V from among a set of vectors X generated combinatorially, when it is not readily apparent how to order X or a canonical spanning set of V in a convenient way. The motivation for developing this technique is to make computations in the Grossman-Larson-Wright module which translate into algebraic statements connected with the Jacobian conjecture. In Section 2 we describe the method, which extends existing theorems on square zero and sign pattern matrices which guarantee non-singularity to rectangular zero and sign pattern matrices which guarantee full rank. In Section 3 we provide background information spelling out the connection between the Grossman-Larson-Wright module and the homogeneous symmetric reduction of the Jacobian conjecture. In Section 4 we apply our methods to prove three properties of the Grossman-Larson-Wright module originally described by David Wright: $\overline{\mathcal{M}}(3,\infty)_m = 0$ for $m \geq 3$, $\overline{\mathcal{M}}(4,3)_m = 0$ for $5 \leq m \leq 8$, and $\overline{\mathcal{M}}(4,4)_8 = 0$. The first two properties yield combinatorial proofs of special cases of the Jacobian conjecture.

2 The Graph Method

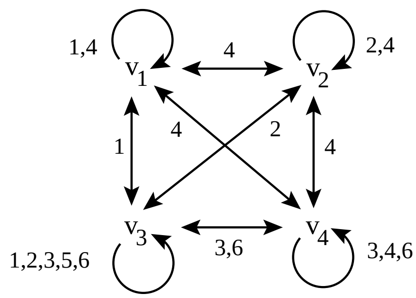
It is well known that a square zero pattern matrix guarantees non-singularity if and only if it is permutationally equivalent to a triangular pattern with nonzero diagonal entries: see ([6], Theorem 4.4). The row and column permutations which bring the matrix into triangular form can be constructed from the edge-labeled digraph G_A and the row selection function r described in Definitions 2.1 and 2.3 below. It is also well known that a nonnegative square pattern matrix with positive main diagonal is sign nonsingular if and only if its associated digraph does not have any directed cycles of even length: see ([4], Corollary 3.2.10, summarizing work of Bassett, Maybee and Quirk [3]). Theorem 2.11 and Corollary 2.12 generalize these results to rectangular zero and sign pattern matrices which guarantee full rank. Corollary 2.13 describes a method for identifying a spanning set in a finite-dimensional vector space based on these results.

Definition 2.1. *Let $A = (a_{ij})$ be a real $m \times n$ matrix. The matrix A gives rise to an edge-labeled digraph $G_A = (V_A, E_A)$, with vertex set $V_A = \{v_1, \ldots, v_n\}$ and for all $(j, i, k) \in [n] \times [m] \times [n]$ a directed edge (v_j, i, v_k) from v_j to v_k labeled i if and only if $a_{ij}a_{ik} \neq 0$.*

Example 2.2.

$$A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 0 & 5 & 6 \\ 7 & 8 & 0 & 9 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 11 & 12 \end{bmatrix}$$

$G_A =$

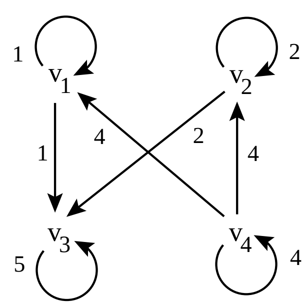


Definition 2.3. *Let $A = (a_{ij})$ be a real $m \times n$ matrix with no zero columns, and let G_A be the associated edge-labeled digraph as in Definition 2.1. For each column $j \leq n$ we define $R_j = \{i \leq m : a_{ij} \neq 0\}$. Since A has no zero columns, every set R_j is non-empty. Given a row selection function $r : V_A \to \{1, \ldots, m\}$ which satisfies $r(v_j) \in R_j$ for all $j \leq n$ we form the row selection subgraph $G_r = (V_A, E_r)$ of G_A with vertex set V_A and edge set*

$$E_r = \{(v, i, v') \in E_A : i = r(v)\}.$$

Example 2.4. *Let A and G_A be as in Example 2.2. Let r be the row selection function defined by $r(v_1) = 1$, $r(v_2) = 2$, $r(v_3) = 5$, $r(v_4) = 4$. Then*

$G_r =$

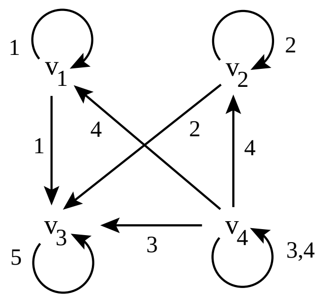


.

Definition 2.5. *Let $A = (a_{ij})$ be a real $m \times n$ matrix and let G_A be the associated edge-labeled digraph as in Definition 2.1. Given a row subset selection function $R : V_A \to 2^{\{1,\ldots,m\}}$ which satisfies $R(v_j) \subseteq R_j$ for all $j \leq n$ we form the row subset selection subgraph $G_R = (V_A, E_R)$ of G_A with vertex set V_A and edge set*

$$E_R = \{(v, i, v') \in E_A : i \in R(v)\}.$$

Example 2.6. *Let A and G_A be as in Example 2.2. Let R be the row subset selection defined by $R(v_1) = \{1\}$, $R(v_2) = \{2\}$, $R(v_3) = \{5\}$, $R(v_4) = \{3, 4\}$. Then*

$G_R =$ 

.

Definition 2.7. *Let V be a vector space with finite spanning set X, let Y be a finite collection of linear combinations of the vectors in X, and for each $x \in X$ let*

$$Y(x) = \{y \in Y : x \text{ appears with non-zero coefficient in } y\}.$$

Then X and Y give rise to an edge-labeled digraph

$$G(X, Y) = (X, E(X, Y))$$

with vertex set X and for all $(x, y, x') \in X \times Y \times X$ a directed edge (x, y, x') from x to x' labeled y if and only if $y \in Y(x) \cap Y(x')$.

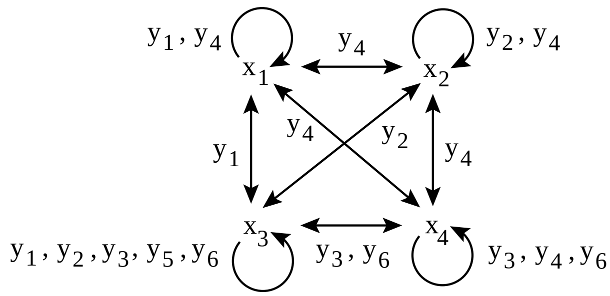
Example 2.8. *Let $V = \mathbb{R}^3$, let $X = \{x_1, x_2, x_3, x_4\}$ where*

$$\begin{aligned} x_1 &= (1, 0, 0), \\ x_2 &= (0, 1, 0), \\ x_3 &= (1, 1, 0), \\ x_4 &= (1, 1, 1), \end{aligned}$$

and let $Y = \{y_1, y_2, y_3, y_4, y_5, y_6\}$ where

$$\begin{aligned} y_1 &= x_1 + 2x_3, \\ y_2 &= 3x_2 + 4x_3, \\ y_3 &= 5x_3 + 6x_4, \\ y_4 &= 7x_1 + 8x_2 + 9x_4, \\ y_5 &= 10x_3, \\ y_6 &= 11x_3 + 12x_4. \end{aligned}$$

Then

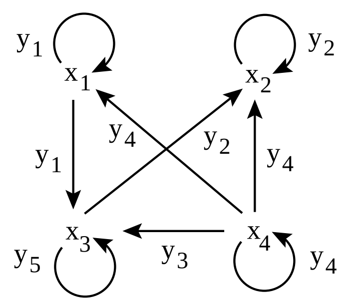
$G(X,Y) =$ 

.

Definition 2.9. *Let V be a vector space with spanning set X, let Y be a finite collection of linear combinations of the vectors in X, and let $G(X,Y)$ be the associated edge-labeled digraph as in Definition 2.7. Given a linear combination subset function $LC : X \to 2^Y$ which satisfies $LC(x) \subseteq Y(x)$ for all $x \in X$ we form the linear combination subgraph $G_{LC}(X,Y) = (X, E_{LC}(X,Y))$ of $G(X,Y)$ with vertex set X and edge set*

$$E_{LC}(X,Y) = \{(x,y,x') \in E(X,Y) : y \in LC(x)\}.$$

Example 2.10. *Let V, X, Y, and $G(X,Y)$ be as in Example 2.8. Let LC be the linear combination subset function defined by $LC(x_1) = \{y_1\}$, $LC(x_2) = \{y_2\}$, $LC(x_3) = \{y_5\}$, $LC(x_4) = \{y_3, y_4\}$. Then*

$G_{LC}(X,Y) =$ 

.

Theorem 2.11. *Let $A = (a_{ij})$ be a $m \times n$ matrix over the reals with no zero columns, let G_A be the associated edge-labeled directed graph described in Definition 2.1, let*

$$r : V_A \to \{1, \ldots, m\}$$

be a row-selection function which satisfies $r(v_j) \in R_j$ for all $j \leq n$, and let G_r be the row selection subgraph of G_A defined by r described in Definition 2.3.

(1) If G_r has no directed cycles of length ≥ 2 then A has n linearly independent rows.

(2) If G_r has no directed cycles of even length, and if A has no negative entries, then A has n linearly independent rows.

In both cases, the rows chosen by the row-selection function r are linearly independent.

Proof. First note that the hypotheses in statements (1) and (2) force r to be injective: suppose $r(v_j) = r(v_k) = i$. Then $a_{ij}a_{ik} \neq 0$, hence the edges (v_j, i, v_k) and (v_k, i, v_j) belong to G_r. Since there are no directed cycles of length 2 in G_r, we must have $v_j = v_k$. Next, observe that permuting the rows of A results in permuting the edge labels of edges in G_A, with no impact on the rank of A or the isomorphism class of G_r. So we can assume without loss of generality that $r(v_j) = j$ for $1 \leq j \leq n$, reordering the rows of A if necessary. This assumption implies that $a_{jj} \neq 0$ for $1 \leq j \leq n$, and allows us to say that $(v_j, j, v_k) \in G_r$ if and only if $a_{jk} \neq 0$ for all $j, k \leq n$. Let B be the matrix which consists of the first n rows of A. Then

$$\det(B) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a(\sigma),$$

where

$$a(\sigma) = a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Given a permutation σ which factors into a product of the disjoint cycles $\tau_1, \ldots, \tau_k$, we have

$$a(\sigma) = a(\tau_1) \cdots a(\tau_k).$$

The non-zero contributions to $\det(B)$ come from permutations $\sigma = \tau_1 \cdots \tau_k$ in which $a(\tau_i) \neq 0$ for each cycle τ_i. Moreover, there is a one-to-one correspondence between cycle permutations τ such that $a(\tau) \neq 0$ and directed cycles in G_r: for a p-cycle τ, we have

$$a(\tau) = a_{j\tau(j)} a_{\tau(j)\tau^2(j)} \cdots a_{\tau^{p-1}(j)j} \neq 0$$

if and only if

$$(v_j, j, v_{\tau(j)}),\ (v_{\tau(j)}, \tau(j), v_{\tau^2(j)}),\ \ldots,\ (v_{\tau^{p-1}(j)}, \tau^{p-1}(j), v_j)$$

are edges in G_r. If G_r has no cycles of length ≥ 2 then the only permutation σ for which $a(\sigma) \neq 0$ is the identity permutation, hence $\det(B) = a_{11} \cdots a_{nn} \neq 0$. If G_r has no directed cycles of even length then the sign of every permutation σ for which $a(\sigma) \neq 0$ is positive, and combined with the hypothesis that A has no negative entries this implies that $\det(B) > 0$. In either case, we conclude that B has linearly independent rows, hence the row selection function r selects n linearly independent rows from A. □

The row selection subgraph G_r can be used to show that an $n \times n$ matrix A is permutationally equivalent to a lower triangular matrix with nonzero diagonal entries when A falls into Case 1. Since G_r has no non-trivial directed cycles, it is possible to relabel the vertices so that $j > k$ whenever there is a directed edge from v_j to a distinct vertex v_k in G_r. Having relabeled the vertices, relabel the edge labels so that $r(v_i) = i$ for each i. The adjacency matrix of the relabeled G_r is lower triangular and permutationally equivalent to A. More generally, the n rows of an $m \times n$ matrix A picked out by the row selection function form a submatrix which is permutationally equivalent to a lower triangular matrix with nonzero diagonal entries when A falls into Case 1 of Theorem 2.11. Of course, a computer can check for the existence of this submatrix in a reasonable amount of time if the matrix is small enough, and by a simple algorithm which has nothing to do with directed graphs, but the graph method may be more suitable for proving full rank if there is no bound to the size of the matrices one is interested in and one has combinatorial information about how the matrices are generated. We will see an example of this in Section 4.

Corollary 2.12. *Let $A = (a_{ij})$ be an $m \times n$ matrix over the reals with no zero columns, let G_A be the associated edge-labeled directed graph as in Definition 2.1, let*

$$R : V_A \to 2^{\{1,\ldots,m\}}$$

be a row subset selection function which satisfies $R(v_j) \subseteq R_j$ and $R(v_j) \neq \emptyset$ for all $j \leq n$, and let G_R be the subgraph of G_A defined by R as in Definition 2.5.

(1) If G_R has no directed cycles of length ≥ 2 then A has n linearly independent rows.

(2) If G_R has no directed cycles of even length, and if A has no negative entries, then A has n linearly independent rows.

Proof. For each vertex v in G_A let $r(v) \in R(v)$ be chosen arbitrarily. This defines a valid row-selection function r for G_A, and G_r is a subgraph of G_R. Therefore G_r falls into Case 1 or Case 2 of Theorem 2.11. Hence A has n linearly independent rows. □

Corollary 2.13. *Let V be a finite-dimensional real vector space with spanning set $X = \{x_1, \ldots, x_n\}$, let $Y = \{y_1, \ldots, y_m\}$ be a collection of linear combinations of the vectors in X, let $G(X, Y)$ be the associated edge-labeled digraph as in Definition 2.7, let $LC : X \to 2^Y$ be a linear combination subset function which satisfies $LC(x) \subseteq Y(x)$ and $LC(x) \neq \emptyset$ for each $x \in X$, and let $G_{LC}(X, Y)$ be the subgraph of $G(X, Y)$ defined by LC as in Definition 2.9.*

(1) If $G_{LC}(X, Y)$ has no directed cycles of length ≥ 2 then Y is a spanning set for V.

(2) If $G_{LC}(X, Y)$ has no directed cycles of even length, and if every linear combination in Y has nonnegative coefficients, then Y is a spanning set for V.

Proof. Let A be an $m \times n$ coefficient matrix which expresses Y in terms of X. Then G_A is isomorphic to $G(X, Y)$, with vertex v_i in G_A corresponding to vertex x_i in $G(X, Y)$ and labeled edge (v_i, k, v_j) in G_A corresponding to labeled edge (x_i, y_k, x_j) in $G(X, Y)$. The linear combination subset function $LC : X \to 2^Y$ gives rise to a valid row subset selection function $R : V_A \to 2^{\{1,\ldots,m\}}$ such that G_R is isomorphic to $G_{LC}(X, Y)$. By construction, $R(v) \neq \emptyset$ for each $v \in V_A$. The subgraph G_R falls into Case 1 or Case 2 of Corollary 2.12, hence A has n linearly independent rows. These rows form an $n \times n$ submatrix of A which is row-equivalent to the identity matrix, which implies that every $x \in X$ can be expressed as a linear combination of the vectors in Y. Hence Y spans V. □

3 A primer on the homogeneous symmetric reduction of the Jacobian conjecture and the Grossman-Larson-Wright module

An algebraic analogue of the inverse function theorem states that if $f_1, \ldots, f_n$ are polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ which satisfy $f_i(0, \ldots, 0) = 0$ for all i and $\det\left(\frac{\partial f_i}{\partial x_j}\right)(0, \ldots, 0) \neq 0$, then there must exist formal power series $g_1, \ldots, g_n$ in $\mathbb{C}[[x_1, \ldots, x_n]]$ which satisfy

$$f_i(g_1, \ldots, g_n) = g_i(f_1, \ldots, f_n) = x_i$$

for all i.

Example 3.1. *Let $n = 1$ and $f_1 = x_1 - x_1^2$. Then $g_1 = \sum_{k=1}^{\infty} \frac{(2k-2)!}{(k-1)!k!} x_1^k$.*

Example 3.2. *Let $n = 2$ and*

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} x_1 - (x_1 + ix_2)^2 \\ x_2 - i(x_1 + ix_2)^2 \end{pmatrix}.$$

Then

$$\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} x_1 + (x_1 + ix_2)^2 \\ x_2 + i(x_1 + ix_2)^2 \end{pmatrix}.$$

The Jacobian conjecture (see [7]) is equivalent to the statement that if $f_i(0, \ldots, 0) = 0$ for all i and if $\det\left(\frac{\partial f_i}{\partial x_j}\right) \in \mathbb{C}^*$ in the set-up above then the expressions $g_1, \ldots, g_n$ are polynomials of finite degree. The polynomial f_1 in Example 3.1 does not meet the hypothesis of the Jacobian conjecture because $\frac{\partial f_1}{\partial x_1} = 1 - 2x_1 \notin \mathbb{C}^*$, but the polynomials f_1 and f_2 in Example 3.2 do because $\det\left(\frac{\partial f_i}{\partial x_j}\right) = 1 \in \mathbb{C}^*$.

There are a number of partial results relating to systems of n polynomials in n variables in which $f_i = x_i - h_i$ for all i, where each h_i is homogeneous of the same total degree $d \geq 2$. Under this scenario, $\det(\partial f) \in \mathbb{C}^*$ implies $(\partial h)^n = 0$. This case is referred to as $J_{n,[d]}$. The Jacobian conjecture is equivalent to $J_{n,[3]}$ [1]. The formal inverse can be expressed in terms of rooted trees. Wright surveyed tree-formula approaches to the Jacobian conjecture in [10]. Singer proposed an alternative approach in terms of Catalan trees [9]. Since the degree of a polynomial inverse can be as large as d^{n-1} in the context of $J_{n,[d]}$, and since the number of trees required grows exponentially with the degree of the inverse, computer runtime and size limitations place severe restrictions on any brute-force search for a solution using these methods.

The most promising approach to the Jacobian conjecture, from a combinatorial point of view, seems to be the homogeneous symmetric reduction due to Michiel de Bondt and Arno van den Essen [2]:

Theorem 3.3. *The Jacobian Conjecture is true if it holds for all polynomial maps F having the form $F = X - H$ with H homogeneous of degree $d \geq 2$ and ∂H is a symmetric matrix. H can be taken to be ∇P, where P is a homogeneous polynomial of degree $d+1$. In fact, it suffices to prove the case $d = 3$.*

Example 3.2 was formed using $P = \frac{1}{3}(x_1 + ix_2)^3$. The formal inverse in the homogeneous symmetric reduction has a combinatorial expression in terms of unrooted trees (Theorem 2.3 in [11]):

Theorem 3.4. *Let $F = X - \nabla P$ be a system of n polynomials in n variables and let G be the inverse system of formal power series. Then $G = X + \nabla Q$ with*

$$Q = \sum_{T \in \mathbb{T}} \frac{1}{|Aut\ T|} \mathcal{Q}_{T,P},$$

where $\mathbb{T}$ is the set of isomorphism classes of unrooted trees,

$$\mathcal{Q}_{T,P} = \sum_{l:E(T) \to \{1,\ldots,n\}} \prod_{v \in V(T)} D_{adj(v)} P,$$

$adj(v)$ is the set $\{e_1, \ldots, e_s\}$ of edges adjacent to v, and

$$D_{adj(v)} = D_{l(e_1)} \cdots D_{l(e_s)}$$

is a product of formal partial differentiation operators.

In the context of Theorem 3.4, if P is homogeneous of degree $d+1$ then

$$Q = Q^{(1)} + Q^{(2)} + Q^{(3)} + \cdots$$

where

$$Q^{(m)} = \sum_{T \in \mathbb{T}_m} \frac{1}{|\mathrm{Aut}\ T|} \mathcal{Q}_{T,P}$$

and $\mathbb{T}_m$ is the set of isomorphism classes of unrooted trees with m vertices. Each $Q^{(m)}$ is homogenous of degree $m(d-1)+2$. In order to prove that the inverse G is a polynomial system, it suffices to show that $Q^{(m)} = 0$ for all sufficiently large m. In fact, it suffices to prove that

$$Q^{(M+1)} = Q^{(M+2)} = \cdots = Q^{(2M)} = 0$$

for some positive integer M (the Gap Theorem). This is a consequence of Zhao's Formula [13]:

Theorem 3.5. *For $m \geq 1$ let $Q^{(m)}$ be the homogeneous summand of degree $m(d-1)+2$ in the formula for the inverse of $F = X - \nabla P$, where P is homogeneous of degree $d+1$. Then $Q^{(1)} = P$ and for $m \geq 2$,*

$$Q^{(m)} = \frac{1}{2(m-1)} \sum_{\substack{k+l=m \\ k,l \geq 1}} \left(\nabla Q^{(k)} \cdot \nabla Q^{(l)}\right).$$

The hypotheses in the homogeneous symmetric reduction of the Jacobian conjecture supply us with a large source of unrooted trees T for which the expression $\mathcal{Q}_{T,P}$ defined in Theorem 3.4 is equal to zero. Let $P \in \mathbb{C}[X]$ be a polynomial in n variables which is homogeneous of degree ≥ 3, let $H = \nabla P$ and $F = X - H$, and assume $\det(F) \in \mathbb{C}^*$. Then $(\partial H)^n = (\text{Hess } P)^n = 0$. We make the following definitions, adapted from Wright [11]:

Definition 3.6. *Let $e \geq 1$ be given. Then $V(e)$ denotes the set of all tree isomorphism classes which contain at least one vertex of degree $> e$.*

Definition 3.7. *Let $r \geq 2$ be given. A naked r-chain in an unrooted tree T is a path of the form $v_1 - v_2 - \cdots - v_r$ in which $\deg_T(v_1) \leq 2$, $\deg_T(v_r) \leq 2$, and $\deg_T(v_i) = 2$ for $2 \leq i \leq r-1$. $C(r)$ is the set of all unrooted tree isomorphism classes which contain a naked r-chain.*

Definition 3.8. *Let $P \in \mathbb{C}[X]$ be a polynomial in n variables. The function $\rho_P : \mathbb{T} \to \mathbb{C}[X]$ is defined by*

$$\rho_P(T) = \mathcal{Q}_{T,P} = \sum_{l:E(T)\to\{1,\ldots,n\}} \prod_{v\in V(T)} D_{adj(v)} P$$

as in Theorem 3.4.

Wright proved ([11], Proposition 3.6 and Theorem 3.1 respectively)

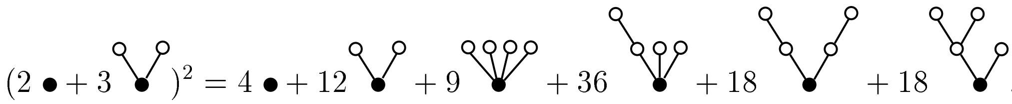
Theorem 3.9. *If $P \in \mathbb{C}[X]$ has degree e then $\rho_P(V(e)) = 0$.*

Theorem 3.10. *Let $P \in \mathbb{C}[X]$ with $(\text{Hess } P)^r = 0$ for some $r \geq 1$. If P is homogeneous of degree ≥ 2 then $\rho_P(C(r)) = 0$.*

The combinatorial program proposed by Wright in [11] is to lift questions related to the homogeneous symmetric reduction of the Jacobian conjecture from the context of differential operators acting on polynomials to that of the Grossman-Larson algebra of rooted trees acting on the module of unrooted trees. The Grossman-Larson algebra $\mathcal{H}$ is a vector space over $\mathbb{Q}$ consisting of all finite linear combinations of trees in $\mathbb{T}_{rt}$, the set of all rooted tree isomorphism classes. Multiplication in $\mathcal{H}$ is defined as follows: Let $S, T \in \mathbb{T}_{rt}$ be given. If S has exactly one vertex, then $S \cdot T = T$. Otherwise, let $S_1, \ldots, S_r$ be the rooted subtrees of S adjacent to the root of S. Then

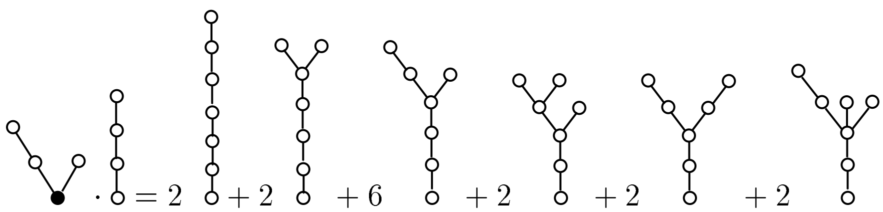
$$S \cdot T = \sum_{(v_1,\ldots,v_r)\in V(T)^r} (S_1, \ldots, S_r) \multimap_{(v_1,\ldots,v_r)} T,$$

where $(S_1, \ldots, S_r) \multimap_{(v_1,\ldots,v_r)} T$ denotes the tree obtained by joining the root of S_i to the vertex v_i in T by a new edge for $1 \le i \le r$. This product is extended by distributivity to all of $\mathcal{H}$. For example,



For more information about the Grossman-Larson algebra, see [5].

The Grossman-Larson-Wright $\mathcal{H}$-module $\mathcal{M}$ is a vector space over $\mathbb{Q}$ consisting of all finite linear combinations of trees in $\mathbb{T}$, the set of all unrooted tree isomorphism classes. The action of $\mathcal{H}$ on $\mathcal{M}$ is defined using the same glueing operation as above, the difference being that the product of a rooted tree with an unrooted tree produces a linear combination of unrooted trees. For example,



(3.1)

All the axioms for a module over an associative $\mathbb{Q}$-algebra are met by $\mathcal{M}$ over $\mathcal{H}$.

The algebra $\mathcal{H}$ is graded: $\mathcal{H} = \bigoplus_{m=0}^{\infty} \mathcal{H}_m$, where $\mathcal{H}_m$ is spanned by rooted trees with m unrooted vertices. The module $\mathcal{M}$ is a graded $\mathcal{H}$-module: $\mathcal{M} = \bigoplus_{m=1}^{\infty} \mathcal{M}_m$, where $\mathcal{M}_m$ is spanned over the rationals by unrooted trees with m vertices. We have $\mathcal{H}_m \mathcal{M}_n \subseteq \mathcal{M}_{m+n}$ for all $m \ge 0$ and $n \ge 1$.

Wright defines the following $\mathcal{H}$-submodules and quotient modules [11]:

Definition 3.11. *Let $e \ge 1$ and $r \ge 2$ be given. Let $\mathcal{V}(e) \subseteq \mathcal{M}$ denote the span of $V(e)$ over the rationals (see Definition 3.6). Let $\mathcal{C}(r) \subseteq \mathcal{M}$ denote the span of all expressions of the form $S \cdot T$ over the rationals, where $S \in \mathbb{T}_{rt}$ and $T \in C(r)$ (see Definition 3.7). Both $\mathcal{V}(e)$ and $\mathcal{C}(r)$ are graded $\mathcal{H}$-submodules of $\mathcal{M}$. Let $\mathcal{N}(r,e) = \mathcal{V}(e) + \mathcal{C}(r)$. Let $\overline{\mathcal{M}}(r,e)$ denote the quotient module $\mathcal{M}/\mathcal{N}(r,e)$. For each $m \ge 1$ let $\overline{\mathcal{M}}(r,e)_m$ denote the image of $\mathcal{M}_m$ in $\overline{\mathcal{M}}(r,e)$.*

The function $\rho_P : \mathbb{T} \to \mathbb{C}[X]$ described in Definition 3.8 can be extended by linearity to a linear transformation $\rho_P : \mathcal{M} \to \mathbb{C}[X]$. When P is homogeneous, ρ_P is a graded $\mathcal{H}$-module homomorphism in the following sense: Let $\mathbb{C}[D_1, \ldots, D_n]$ be the $\mathbb{Q}$-algebra of formal partial differentiation operators acting on the module $\mathbb{C}[x_1, \ldots, x_n]$. Given a polynomial $P \in \mathbb{C}[x_1, \ldots, x_n]$ which is homogenous of degree $d+1$, let $\phi_P : \mathcal{H} \to \mathbb{C}[D_1, \ldots, D_n]$ be the mapping defined by

$$\phi_P(S) = \sum_{l:E(S)\to\{1,\ldots,n\}} \left(\prod_{v \in V(S)-\{\mathrm{root}(S)\}} D_{\mathrm{adj}(v)} P \right) D_{\mathrm{adj}(\mathrm{root}(S))}$$

for all $S \in \mathbb{T}_{rt}$ and extended by linearity to all of $\mathcal{H}$. Then ϕ_P is a $\mathbb{Q}$-algebra homomorphism. Moreover,

$$\rho_P(xy) = \phi_P(x)\rho_P(y) \tag{3.2}$$

for all $(x, y) \in \mathcal{H} \times \mathcal{M}$ and $\deg \rho_P(x) = m(d-1) + 2$ for all $x \in \mathcal{M}_m$.

If P is homogeneous of degree e and $(\text{Hess } P)^r = 0$, then Theorems 3.9 and 3.10 together with Equation 3.2 imply that

$$\mathcal{N}(r, e) \subseteq \ker \rho_P. \tag{3.3}$$

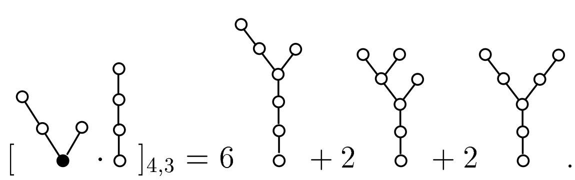
Combining Equation 3.3 with Theorem 3.4 and the Gap Theorem, the link between the homogeneous symmetric reduction of the Jacobian conjecture and the Grossman-Larson-Wright module is summarized as follows:

Theorem 3.12. *Let $P \in \mathbb{C}[x_1, \ldots, x_n]$ be homogeneous of degree $e \geq 3$ and satisfy $(\text{Hess } P)^r = 0$ for some $r \geq 1$. Set $F = X - \nabla P$. If $\overline{\mathcal{M}}(r, e)_m = 0$ for $M + 1 \leq m \leq 2M$ and some positive integer M then the formal inverse of F is a polynomial system.*

4 Applying the graph-theoretic method to 3 examples in the Grossman-Larson-Wright module

Wright states without proof that $\overline{\mathcal{M}}(3, \infty)_m = 0$ for $m \geq 3$ in ([11], Theorem 3.12). Our proof of this in Theorem 4.3 below illustrates the use of Case 1 of Corollary 2.13. This supplies a proof of $J_{n,[d]}$ for all n and $d \geq 2$ when ∂H is symmetric and $(\partial H)^3 = 0$. Wright proves $\overline{\mathcal{M}}(4, 3)_m = 0$ for $5 \leq m \leq 8$ in ([11], Proposition 3.11). Our proof of this in Theorem 4.4 below is different and provides a second example of Case 1 of Corollary 2.13. This supplies a proof of $J_{n,[2]}$ for all n when ∂H is symmetric and $(\partial H)^4 = 0$. Wright announces that $\overline{\mathcal{M}}(4, 4)_m = 0$ for $m = 8, 9, 10, 11, 12, 14$ (but not 13!) in [11] by a computer search, using a program written by Li-Yang Tan [12]. This does not quite supply a proof of $J_{n,[3]}$ for all n when ∂H is symmetric and $(\partial H)^4 = 0$, but Wright finds a way to bridge the gap and complete the proof (see Theorem 3.19 and the paragraph before it in [11]). We have duplicated his results for $\overline{\mathcal{M}}(4, 4)_m = 0$ using Mathematica and can attest to the computational complexity of this problem. We prove $\overline{\mathcal{M}}(4, 4)_8 = 0$ in Theorem 4.5 below using Case 2 of Corollary 2.13.

Definition 4.1. *Let $\mathbb{T}_m(r, e)$ denote the set of unrooted trees with m vertices, no naked r-chains, and all vertex degrees $\leq e$. Let $V_m(r, e) \subseteq \mathcal{M}_m$ be the span of $\mathbb{T}_m(r, e)$ over the rationals. For each $S \in \mathbb{T}_{rt}$ and $T \in \mathbb{T}$ we denote by $[S \cdot T]_{r,e}$ the sum of the terms in $S \cdot T$ which contain no naked r-chains and have all vertex degrees $\leq e$. For example, compare Equation 3.1 with*



We will abbreviate this notation to $[S \cdot T]$ when convenient.

Lemma 4.2. *Let $m \geq 1$, $r \geq 2$, $e \geq 3$ be given. Set $X = \mathbb{T}_m(r, e)$ and*

$$Y = \{[S \cdot T]_{r,e} : (S, T) \in \mathbb{T}_{rt} \times C(r), S \cdot T \in \mathcal{M}_m\}.$$

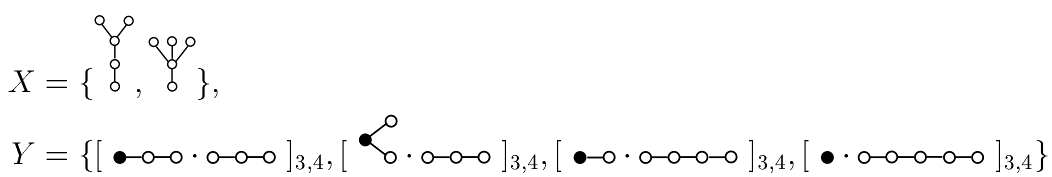
Form the edge-labeled directed graph $G(X, Y)$ as in Definition 2.7. As in Definition 2.9, let $LC : X \to 2^Y$ be a linear combination subset function which satisfies $LC(x) \subseteq Y(x)$ and $LC(x) \neq \emptyset$ for each $x \in X$, and let $G_{LC}(X, Y)$ be the subgraph of $G(X, Y)$ defined by LC. If $G_{LC}(X, Y)$ has no even directed cycles then $\overline{\mathcal{M}}(r, e)_m = 0$.

Proof. Regarded as a collection of vectors in $\mathcal{M}$, the set X spans $V_m(r, e)$. The set Y is a finite collection of vectors in $V_m(r, e)$ with nonnegative coefficients of vectors in X. Hence by Corollary 2.13, Y spans $V_m(r, e)$. Since

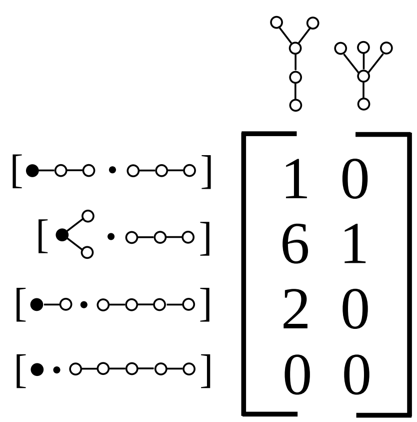
$$[S \cdot T]_{r,e} \equiv S \cdot T \equiv 0 \bmod \mathcal{N}(r, e)$$

for each $[S \cdot T]_{r,e} \in Y$, this implies that $X \subseteq \mathcal{N}(r, e)$. Since the images of X in $\overline{\mathcal{M}}(r, e)$ span $\overline{\mathcal{M}}(r, e)_m$, this in turn implies $\overline{\mathcal{M}}(r, e)_m = 0$. $\square$

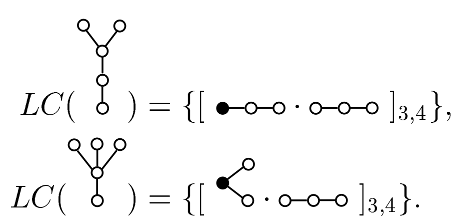
To illustrate the use of Lemma 4.2, here is a proof that $\overline{\mathcal{M}}(3, 4)_5 = 0$: We have



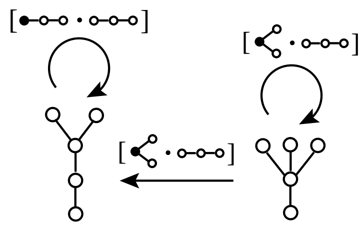
The coefficient matrix for Y has rows indexed by Y, columns indexed by X:



We will choose the linear combination subset function



With this choice we obtain $G_{LC}(X,Y)$ with no non-trivial cycles:

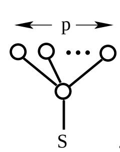


Therefore Y spans $\mathbb{T}_5(3,4)$ and $\overline{\mathcal{M}}(3,4)_5 = 0$.

In the proof of Theorems 4.3 and 4.4 we refer to the diameter of an unrooted tree and the height of a rooted tree. These are standard terms from graph theory. The distance between two vertices u, v in a graph G is the minimal number of edges in a path connecting u and v in G, and the diameter of G is the greatest distance between any of pair of vertices in G. The height of a rooted tree S is the greatest distance between the root vertex of S and any other vertex in S. The idea of the proof in Theorems 4.3 and 4.4 is to construct $G_{LC}(X,Y)$ in such a way that it has two properties: (1) every directed edge (x, y, x') satisfies diameter$(x) \leq$ diameter(x'), and (2) any walk of sufficient length along non-loop edges from any vertex x must encounter a vertex x' of strictly greater diameter. These two properties guarantee that there are no directed cycles of length ≥ 2 in the graph: if there were a non-trivial directed cycle through vertex x along non-loop edges, then walking around the cycle starting from x we must eventually encounter a vertex x' of strictly larger diameter, and walking from x' to x along the cycle we would find that diameter$(x) <$ diameter$(x') \leq$ diameter(x), a contradiction.

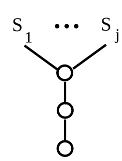
Theorem 4.3. $\overline{\mathcal{M}}(3,\infty)_m = 0$ *for* $m \geq 3$.

Proof. The statement $\overline{\mathcal{M}}(3,\infty)_3 = 0$ is trivially true. Fix $m \geq 4$, $r = 3$, $e = \infty$. We will apply Lemma 4.2. Let $X = \mathbb{T}_m(3,\infty)$, the set of trees with m vertices and no naked 3-chains. Trees in X fall into two disjoint categories. Trees in Category I have a decomposition of the form

$T =$ 

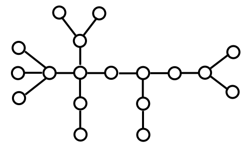
,

where height$(S) =$ diameter$(T) - 2$ and $p > 1$. The remaining trees fall into Category II and have a decomposition of the form

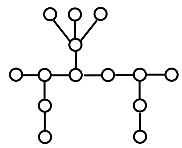
$T =$ 

,

where height(S_1) = diameter$(T) - 3$, $j \geq 2$, and S_1 has a maximal number of vertices. We emphasize that a tree can fall into a category in more than one way. For example, the tree



falls into Category I in three ways, and the tree

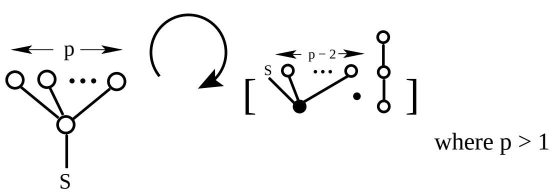


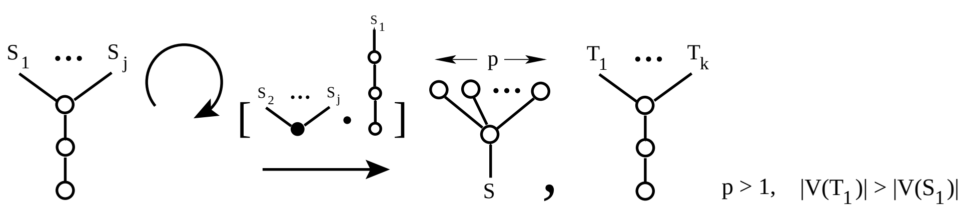
falls into Category II in two ways.

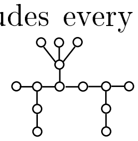
As in Lemma 4.2, let

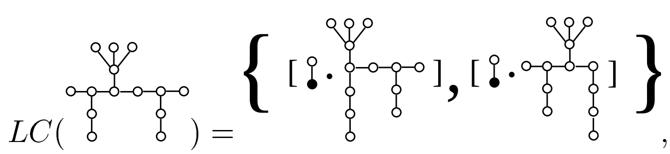
$$Y = \{[S \cdot T]_{3,\infty} : (S, T) \in \mathbb{T}_{rt} \times C(3), S \cdot T \in \mathcal{M}_m\}.$$

We will define a linear combination subset function $LC : X \to 2^Y$ such that $\emptyset \neq LC(X) \subseteq Y(x)$ for each $x \in X$ implicitly by specifying the edges (x, y, x') in $G_{LC}(X, Y)$, organized by the category of x. *The edges to strictly larger diameter trees have been suppressed for simplicity in the following depiction of* $G_{LC}(X, Y)$*:*

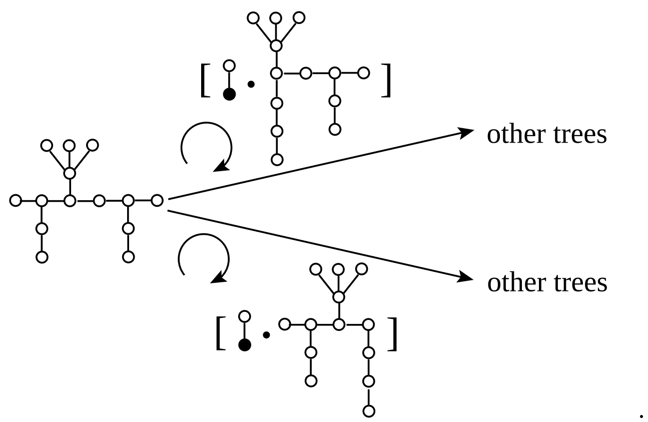




Note that $LC(T)$ includes every bracketed product suggested by the figure above. For example, since the tree  belongs to Category II in two ways, we have



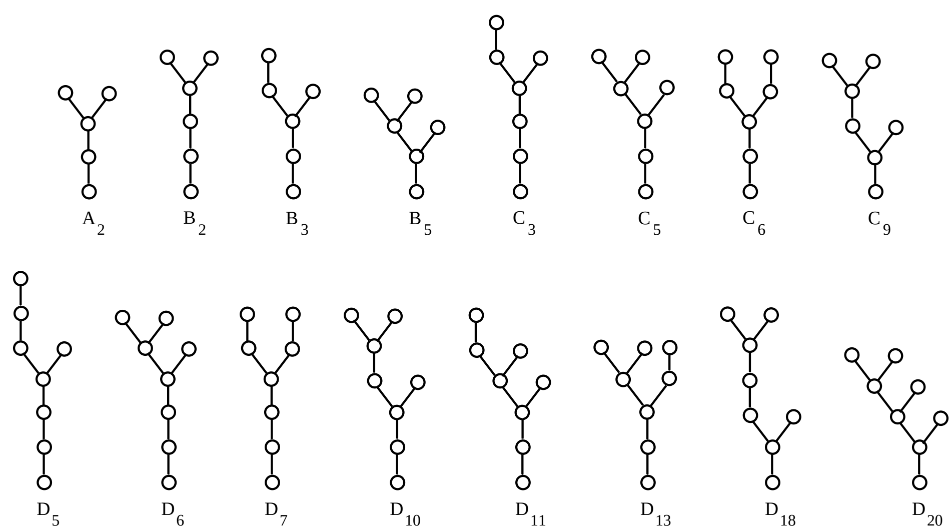
which generates the edges



Every edge (x, y, x') in $G_{LC}(X, Y)$ satisfies diameter$(x) \leq$ diameter(x'). Any walk of length $|X| + 1$ from a vertex x along non-loop edges must encounter a vertex x' of strictly greater diameter. Therefore $G_{LC}(X, Y)$ has no cycles of length ≥ 2. By Lemma 4.2, $\overline{\mathcal{M}}(3, \infty)_m = 0$. $\square$

Theorem 4.4. $\overline{\mathcal{M}}(4, 3)_m = 0$ *for* $5 \leq m \leq 8$.

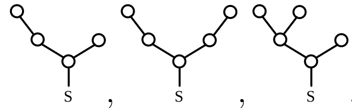
Proof. The trees in $\bigcup_{m=5}^{8} \mathbb{T}_m(4, 3)$ are labeled in [11] as follows:



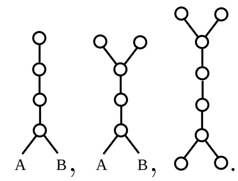
Fix $m \in \{5, 6, 7, 8\}$. Let $X = \mathbb{T}_m(4, 3)$, the set of trees with m vertices, no naked 4-chains, and no vertices of degree > 3. The trees in X can be sorted into two disjoint categories. Trees in Category I have a representation of the form

$$\begin{array}{c} \mathrm{A} \\ | \\ \mathrm{S}, \end{array}$$

where A is a rooted tree of height 2, S is a rooted tree with height equal to diameter $-$ 3, the root of A has degree 2 in A, and S is maximal with respect to number of vertices. The trees in Category I can be sorted into the disjoint subcategories



The remaining trees fall into Category II and can be sorted into the disjoint subcategories

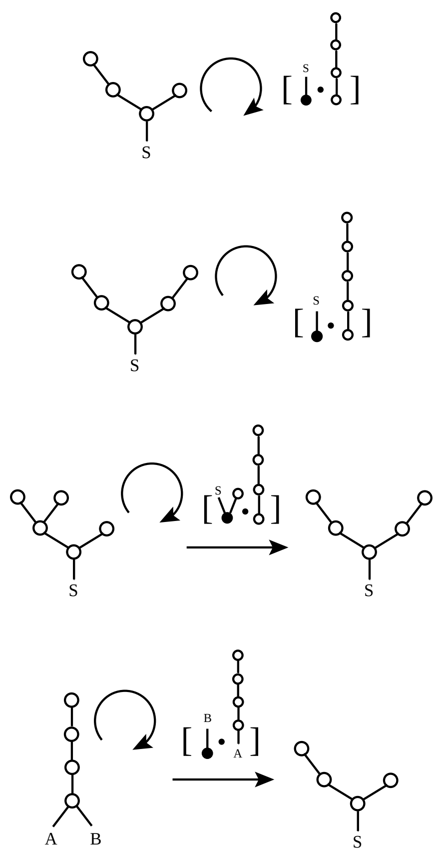


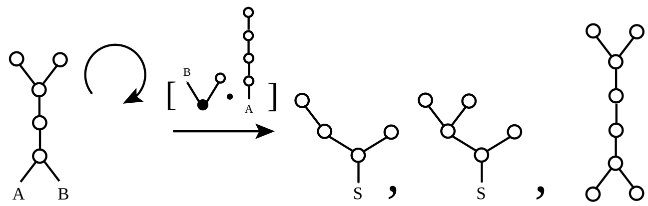
Here A represents a rooted subtree such that height$(A) =$ diameter $-$ 4.

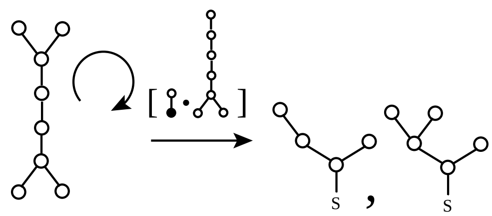
Let

$$Y = \{[S \cdot T]_{4,3} : (S, T) \in \mathbb{T}_{rt} \times C(4), S \cdot T \in \mathcal{M}_m\}.$$

As before, we will define a linear combination subset function $LC : X \to 2^Y$ such that $\emptyset \neq LC(X) \subseteq Y(x)$ for each $x \in X$ implicitly by specifying the edges (x, y, x'), in $G_{LC}(X, Y)$, organized by the category of x. *The edges to strictly larger diameter trees have been suppressed for simplicity in the following depiction of* $G_{LC}(X, Y)$*:*



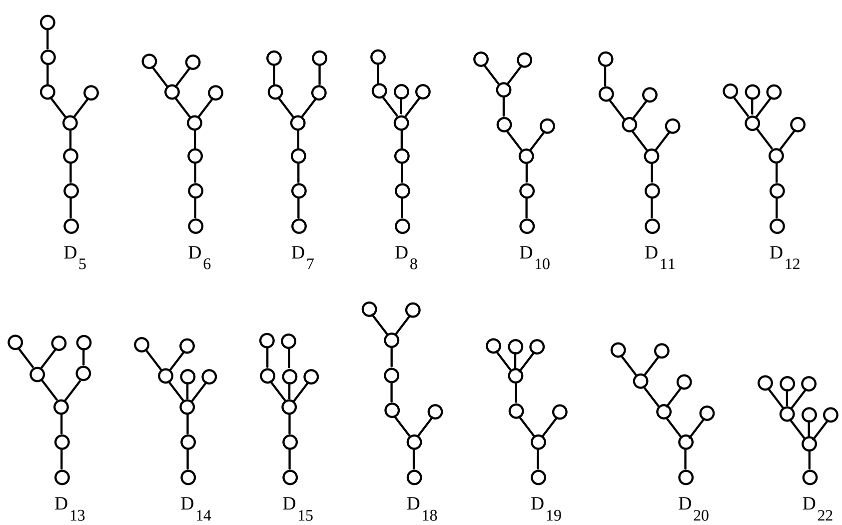




As in the proof of Theorem 4.3, $LC(T)$ includes every bracketed product suggested by the figure above. All edges (x, y, x') in $G_{LC}(X, Y)$ satisfy diameter$(x) \leq$ diameter(x'). Since any walk in $G_{LC}(X, Y)$ of the form $x_0 \to x_1 \to x_2 \to x_3 \to x_4$ along non-loop edges satisfies diameter$(x_0) <$ diameter(x_4), $G_{LC}(X, Y)$ has no directed cycles of length ≥ 2. By Lemma 4.2, $\overline{\mathcal{M}}(4, 3)_m = 0$. $\square$

Theorem 4.5. $\overline{\mathcal{M}}(4, 4)_8 = 0$.

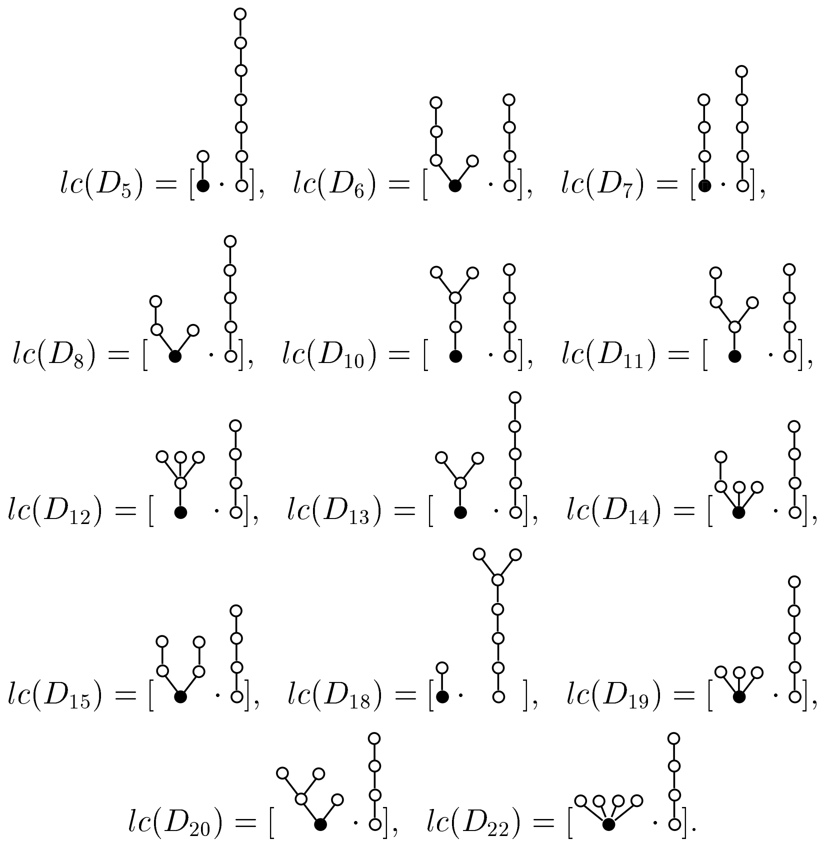
Proof. The trees in $\mathbb{T}_8(4, 4)$ are labeled in [11] as follows:



Let $X = \mathbb{T}_8(4, 4)$ and

$$Y = \{[S \cdot T]_{4,4} : (S, T) \in \mathbb{T}_{rt} \times C(4), S \cdot T \in \mathcal{M}_8\}.$$

Computer calculations show that if A is any coefficient matrix representing Y in terms of X, then A does not have a 14×14 submatrix which is permutationally equivalent to a triangular matrix with non-zero diagonal entries. We will associate with each $x \in X$ a unique $y = lc(x) \in Y$ such that x appears in the support of y, and in each case set $LC(x) = \{lc(x)\}$:



Let Y_0 be the set of the vectors described above. Let A_0 be the zero-one matrix with columns indexed by

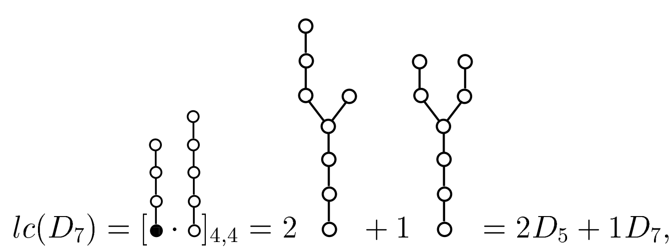
$$\{D_5, D_6, D_7, D_8, D_{10}, D_{11}, D_{12}, D_{13}, D_{14}, D_{15}, D_{18}, D_{19}, D_{20}, D_{22}\},$$

rows indexed by

$$\{lc(D_5), lc(D_6), lc(D_7), lc(D_8), lc(D_{10}), lc(D_{11}), lc(D_{12}),$$

$$lc(D_{13}), lc(D_{14}), lc(D_{15}), lc(D_{18}), lc(D_{19}), lc(D_{20}), lc(D_{22}\},$$

and a 1 in row $lc(D_i)$, column D_j if and only if D_j appears in the support of $lc(D_i)$. For example,



therefore the third row of A_0, corresponding to $lc(D_7)$, contains 1s in columns 1 and 3, corresponding to D_5 and D_7, and 0s elsewhere. These 1s can also be regarded as representing the directed edges $(D_7, lc(D_7), D_5)$ and $(D_7, lc(D_7), D_7)$ in $G_{LC}(X, Y)$. The matrix A_0 represents both the sign pattern of the coefficient matrix which represents Y_0 in terms of X and the adjacency matrix of $G_{LC}(X, Y)$. We have

$$A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

There is exactly one non-trivial directed cycle in $G_{LC}(X, Y)$, and it has odd length: the sequence of labeled edges

$$(D_6, lc(D_6), D_8), \ (D_8, lc(D_8), D_{13}), \ (D_{13}, lc(D_{13}), D_6).$$

Hence by Lemma 4.2, $\overline{\mathcal{M}}(4, 4)_8 = 0$. □

These examples raise several questions:

1. Is there a systematic way to categorize trees as we have done in Theorems 4.3 and 4.4 to prove that $\overline{\mathcal{M}}(4, 4)_m = 0$ for other values of m using Corollary 2.13?

2. Does a sufficiently large value of m guarantee that we can find a spanning set $Y \subseteq \mathcal{N}(4, 4)$ for $V_m(4, 4)$ with a corresponding $G_{LC}(X, Y)$ digraph that contains no non-trivial directed cycles?

3. For which other values of r, e, and m can we apply these methods?

4. In the proof of Lemma 4.2 we have used a basis $X = \mathbb{T}_m(r, e)$ for $V_m(r, e)$ and have found a spanning set Y for $V_m(r, e)$, so we have not used the full force of Corollary 2.13, which allows X to be a spanning set. Is there a way to use a spanning set $X \subseteq \mathcal{N}(r, e)$ for $V_m(r, e)$ to generate a spanning set $Y \subseteq \mathcal{N}(r + 1, e)$ for $V_{m'}(r + 1, e)$?

5. Are there other combinatorial problems that are solvable using these methods?

Acknowledgements

The author would like to thank David Wright for taking the time to read and comment on this paper. Many thanks are also due the reviewer who suggested ways to improve the exposition.

References

[1] H. Bass, E. H. Connell and D. Wright, *The Jacobian conjecture: reduction of degree and formal expansion of the inverse,* Bull. Amer. Math. Soc. 7 (2) (1982), 287–330.

[2] Michiel de Bondt and Arno van den Essen, *A reduction of the Jacobian Conjecture to the symmetric case,* Proc. Amer. Math. Soc. 133 (2005), no. 8, 2201-2205 (electronic).

[3] L. Bassett, J. Maybee, and J. Quirk, *Qualitative economics and the scope of the correspondence principle,* Econometrica 36 (1968), 544-563.

[4] Richard A. Brualdi and Bryan L. Shader, *Matrices of Sign-Solvable Linear Systems,* Cambridge Tracts in Mathematics 116, Cambridge University Press, Cambridge (1995).

[5] Robert Grossman and Richard G. Larson, *Hopf-algebraic structure of families of trees,* J. Algebra 126 (1989), no. 1, 184-210.

[6] D. Hershkowitz and H. Schneider, *Ranks of zero patterns and sign patterns,* Linear and Multilinear Algebra 34 (1993), 3–19.

[7] O.H. Keller, *Ganze Cremona-Transformationen,* Monatsh. Math. Phys., 47(1939), 229-306.

[8] I. G. Macdonald, *Symmetric Functions and Hall Polynomials,* Oxford Univ. Press, Oxford (1979).

[9] D. Singer, *On Catalan trees and the Jacobian conjecture,* The Electronic Journal of Combinatorics 8(1) (2001), #R2.

[10] D. Wright, *Reversion, trees, and the Jacobian conjecture,* Combinatorial and computational algebra (Hong Kong, 1999), 249–267, Contemp. Math. 264, Amer. Math. Soc., Providence, RI, 2000.

[11] D. Wright, *The Jacobian conjecture as a problem in combinatorics,* arXiv:math.CO/0511214v2 22 Mar 2006.

[12] Li-Yang Tan, *Combinatorial calculations in the tree quotient modules of the Grossman-Larson algebra,* `http://www.cs.wustle.edu/~lt1/jc.html`.

[13] W. Zhao, *Inversion problem, Legendre transform and inviscid Burgers' equations,* J. Pure Appl. Algebra 199 (2005), no. 1-3, 299-317.