

involve

a journal of mathematics

Editorial Board

Kenneth S. Berenhaut, *Managing Editor*

Colin Adams	David Larson
John V. Baxley	Suzanne Lenhart
Arthur T. Benjamin	Chi-Kwong Li
Martin Bohner	Robert B. Lund
Nigel Boston	Gaven J. Martin
Amarjit S. Budhiraja	Mary Meyer
Pietro Cerone	Emil Minchev
Scott Chapman	Frank Morgan
Jem N. Corcoran	Mohammad Sal Moslehian
Toka Diagana	Zuhair Nashed
Michael Dorff	Ken Ono
Sever S. Dragomir	Timothy E. O'Brien
Behrouz Emamizadeh	Joseph O'Rourke
Joel Foisy	Yuval Peres
Errin W. Fulp	Y.-F. S. Pétermann
Joseph Gallian	Robert J. Plemmons
Stephan R. Garcia	Carl B. Pomerance
Anant Godbole	Bjorn Poonen
Ron Gould	Józeph H. Przytycki
Andrew Granville	Richard Rebarber
Jerrold Griggs	Robert W. Robinson
Sat Gupta	Filip Saidak
Jim Haglund	James A. Sellers
Johnny Henderson	Andrew J. Sterge
Jim Hoste	Ann Trenk
Natalia Hritonenko	Ravi Vakil
Glenn H. Hurlbert	Antonia Vecchio
Charles R. Johnson	Ram U. Verma
K. B. Kulasekera	John C. Wierman
Gerry Ladas	Michael E. Zieve



INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

MANAGING EDITOR

Kenneth S. Berenhaut Wake Forest University, USA

BOARD OF EDITORS

Colin Adams	Williams College, USA	Suzanne Lenhart	University of Tennessee, USA
John V. Baxley	Wake Forest University, NC, USA	Chi-Kwong Li	College of William and Mary, USA
Arthur T. Benjamin	Harvey Mudd College, USA	Robert B. Lund	Clemson University, USA
Martin Bohner	Missouri U of Science and Technology, USA	Gaven J. Martin	Massey University, New Zealand
Nigel Boston	University of Wisconsin, USA	Mary Meyer	Colorado State University, USA
Amarjit S. Budhiraja	U of North Carolina, Chapel Hill, USA	Emil Minchev	Ruse, Bulgaria
Pietro Cerone	La Trobe University, Australia	Frank Morgan	Williams College, USA
Scott Chapman	Sam Houston State University, USA	Mohammad Sal Moslehian	Ferdowsi University of Mashhad, Iran
Joshua N. Cooper	University of South Carolina, USA	Zuhair Nashed	University of Central Florida, USA
Jem N. Corcoran	University of Colorado, USA	Ken Ono	Emory University, USA
Toka Diagana	Howard University, USA	Timothy E. O'Brien	Loyola University Chicago, USA
Michael Dorff	Brigham Young University, USA	Joseph O'Rourke	Smith College, USA
Sever S. Dragomir	Victoria University, Australia	Yuval Peres	Microsoft Research, USA
Behrouz Emamizadeh	The Petroleum Institute, UAE	Y.-F. S. Pétermann	Université de Genève, Switzerland
Joel Foisy	SUNY Potsdam, USA	Robert J. Plemmons	Wake Forest University, USA
Erin W. Fulp	Wake Forest University, USA	Carl B. Pomerance	Dartmouth College, USA
Joseph Gallian	University of Minnesota Duluth, USA	Vadim Ponomarenko	San Diego State University, USA
Stephan R. Garcia	Pomona College, USA	Bjorn Poonen	UC Berkeley, USA
Anant Godbole	East Tennessee State University, USA	James Propp	U Mass Lowell, USA
Ron Gould	Emory University, USA	József H. Przytycki	George Washington University, USA
Andrew Granville	Université Montréal, Canada	Richard Rebarber	University of Nebraska, USA
Jerrold Griggs	University of South Carolina, USA	Robert W. Robinson	University of Georgia, USA
Sat Gupta	U of North Carolina, Greensboro, USA	Filip Saidak	U of North Carolina, Greensboro, USA
Jim Haglund	University of Pennsylvania, USA	James A. Sellers	Penn State University, USA
Johnny Henderson	Baylor University, USA	Andrew J. Sterge	Honorary Editor
Jim Hoste	Pitzer College, USA	Ann Trenk	Wellesley College, USA
Natalia Hritonenko	Prairie View A&M University, USA	Ravi Vakil	Stanford University, USA
Glenn H. Hurlbert	Arizona State University, USA	Antonia Vecchio	Consiglio Nazionale delle Ricerche, Italy
Charles R. Johnson	College of William and Mary, USA	Ram U. Verma	University of Toledo, USA
K. B. Kulasekera	Clemson University, USA	John C. Wierman	Johns Hopkins University, USA
Gerry Ladas	University of Rhode Island, USA	Michael E. Zieve	University of Michigan, USA

PRODUCTION

Silvio Levy, Scientific Editor

Cover: Alex Scorpan

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2016 is US \$160/year for the electronic version, and \$215/year (+\$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW[®] from Mathematical Sciences Publishers.

PUBLISHED BY



mathematical sciences publishers

nonprofit scientific publishing

<http://msp.org/>

© 2016 Mathematical Sciences Publishers

Affine hyperbolic toral automorphisms

Colin Thomson and Donna K. Molinek

(Communicated by Michael E. Zieve)

A hyperbolic transformation of the torus is an example of a function that is Devaney chaotic; that is, it is topologically transitive and has dense periodic points. An irrational rotation of the torus, on the other hand, is not chaotic because it has no periodic points. We show that a hyperbolic transformation of the torus followed by a translation (an affine hyperbolic toral automorphism) has dense periodic points and maintains transitivity. As a consequence, affine toral automorphisms are chaotic, even when the translation is an irrational rotation.

1. Introduction

Değirmenci and Koçak [2010] showed that the cross-product of the double-angle map and an irrational rotation, which is a function on the torus, is transitive and has sensitive dependence to initial conditions, but no periodic points, and therefore is not chaotic. Linear hyperbolic toral automorphisms are known to be chaotic, so a natural question in light of [Değirmenci and Koçak 2010] (and the generalizations in [Li and Zhou 2013]) is whether a linear hyperbolic toral automorphism plus a translation is still chaotic. We will refer to such functions as affine hyperbolic toral automorphisms to indicate the translation. Our main goal will be to determine whether such an affine map has periodic points, even in the event that the rotation is irrational.

We find that affine hyperbolic toral automorphisms are chaotic; in fact, we can find the precise locations of periodic points in relation to the periodic points of the corresponding linear map. In this respect, we generalize statements about the transitivity and periodic points of linear hyperbolic toral automorphisms to affine hyperbolic toral automorphisms.

2. Definitions

Throughout this paper, $f : X \rightarrow X$ will be a continuous function on a complete metric space (X, d) . We will examine the iterates of f using the notation f^n to represent the n -th iterate of f ; that is, $f^1 = f$ and $f^{n+1} \equiv f \circ f^n$. The composition

MSC2010: primary 54H20; secondary 37B40.

Keywords: topological dynamics, chaos, toral automorphism.

of f is still a continuous function from X to X . For a specific point $x \in X$, we may refer to the n -th iterate of x under f by x^n , which means $x^0 = x$ is the initial point. In this paper, all points in the space will be specified as vectors, and as such the superscript notation will unambiguously denote an iterate, not raising to a power. In addition, subscripts on points in the space will refer to the corresponding coordinate value, with the basis specified in the case that it is unclear.

A function is *transitive* if for every pair of nonempty open sets $U, V \subseteq X$, there exists a positive integer n such that $f^n(U) \cap V \neq \emptyset$. An example of a transitive function is the irrational rotation on the circle. An irrational rotation is actually *totally transitive*, by which we mean that f^m is transitive for every positive integer m . A property of the irrational rotation that makes it useful for counterexamples is that it is transitive, but has no periodic points.

A *periodic point* $p \in X$ is one for which $f^n(p) = p$ for some n , a positive integer. The least such n is called the *period* of p , and if $n = 1$, we say that p is a *fixed point*. We can locate points with a given period m by finding fixed points of f^m , provided that there is no $k < m$ such that f^k also fixes that point.

A function is *Devaney chaotic* (henceforth, *chaotic*) if it is transitive, has *dense* periodic points, and has *sensitive dependence to initial conditions*. “Dense” refers to the presence of at least one periodic point in every nonempty open set. Sensitivity to initial conditions means that there exists an $\epsilon > 0$ so that for all $\delta > 0$ and $x \in X$, there exists a $y \in X$ with $d(x, y) < \delta$ and an $n \in \mathbb{N}$ such that $d(f^n(x), f^n(y)) > \epsilon$. Banks et al. [1992] proved that the first two hypotheses are sufficient for the third, making transitivity and dense periodic points all that is necessary for chaos. As Crannell [1995] pointed out and by Banks et al. [1992], the elimination of the sensitivity hypothesis makes chaos an entirely topological concept: sensitive dependence on initial conditions is the only hypothesis of the three that relied on the metric.

In general, no other combination of two hypotheses implies the third, but on the unit interval, transitivity guarantees dense periodic points, and is therefore sufficient for chaos [Vellekoop and Berglund 1994]. Contrast this with the irrational rotation on the circle, which is transitive but has no periodic points and is not sensitive to initial conditions.

A torus of d dimensions \mathbb{T}^d is the cartesian product of d copies of the circle, $S^1 \times S^1 \times \dots \times S^1$. Since $S^1 = \mathbb{R}/\mathbb{Z}$, coordinates in \mathbb{T}^d are real numbers from 0, inclusive, to 1, exclusive. A linear automorphism of \mathbb{T}^d is matrix multiplication of the coordinates in $[0, 1) \times [0, 1) \times \dots \times [0, 1)$, taken modulo 1. Since the corners of the unit d -cube are all identified on \mathbb{T}^d , their images under matrix multiplication must all have integer entries to ensure they are each mapped to the origin, modulo 1. Thus the matrix representing the linear transformation must have integer entries. In addition, this matrix must have determinant ± 1 so that the map is a bijection.

This paper is concerned with *hyperbolic* toral automorphisms. If A is the matrix representing the toral automorphism, the product of the d (not necessarily distinct) eigenvalues of A is the determinant, which we require to be ± 1 . A toral automorphism is hyperbolic when none of the eigenvalues are equal in magnitude to 1.

3. Preliminary results

Lemma 3.1 [Katok and Hasselblatt 1995]. *Any hyperbolic toral automorphism with a largest eigenvalue whose eigenvector has rationally independent entries is transitive.*

Proof. Let $U, V \subset \mathbb{T}^d$ be nonempty open sets. The set U must contain a line segment parallel to the eigenvector associated with the largest eigenvalue. Since this eigenvalue is greater than 1, under iteration the line segment grows without bound while remaining parallel to the eigenvector. Since the line “wraps around” the torus whenever the value of a coordinate exceeds 1, the distances between points where the line intersects the i -axis take on values that are multiples of the i -th entry in the eigenvector. As with the irrational rotation of the circle, as the number of iterates tends towards infinity, these intersection points are dense on the i -axis. Since the line stays parallel to the eigenvector, and the entries are rationally independent, the orbit of the line is dense in \mathbb{T}^d . This guarantees that the line intersects V after a finite number of iterations, and therefore U and V have nontrivial intersection for some number of iterations of f . \square

Lemma 3.2 [Katok and Hasselblatt 1995]. *The rational points on the torus are periodic for any hyperbolic toral automorphism.*

Proof. Let

$$p = \left(\frac{p_1}{q}, \dots, \frac{p_d}{q} \right),$$

with $p_1, \dots, p_d, q \in \mathbb{N}$, be a point in \mathbb{T}^d with rational coordinates (not necessarily in lowest terms). Since the entries of the matrix corresponding to the hyperbolic toral automorphism are all integers, the image of p is also a rational point with common denominator q . Since there are precisely q^d rational points in the unit square with denominator q (again, not necessarily in lowest terms), every such point can take on only finitely many values under iterates of the automorphism. Thus, each rational point is either periodic, or preperiodic (in the sense that p is mapped into a periodic orbit, but that orbit does not contain p). Since the automorphism is invertible, no points are preperiodic and therefore must be periodic, with maximum period q^d . \square

In fact, only the rational coordinates are periodic. To see this, consider that periodic points of period n are in the kernel of $A^n - I_d$, where I_d is the identity matrix of dimension d . Since $A^n - I_d$ has integer entries, its kernel is composed only of vectors with rational entries.

Example 3.3 [Elaydi 2008; Katok and Hasselblatt 1995]. The canonical example of a hyperbolic toral automorphism is the Arnold “cat” map

$$L_A : \mathbb{T}^2 \rightarrow \mathbb{T}^2, \quad x \mapsto \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} x \pmod{1}. \quad (1)$$

The eigenvalues of the matrix are $\frac{1}{2}(3 + \sqrt{5})$ and $\frac{1}{2}(3 - \sqrt{5})$ with respective eigenvectors $[\frac{1}{2}(1 + \sqrt{5}), 1]^\top$ and $[\frac{1}{2}(1 - \sqrt{5}), 1]^\top$. You can see that one of the eigenvalues is larger than 1 and the other less, while both eigenvectors have irrational slope.

4. Main results

With the previous two lemmas, we have enough machinery to prove the main theorem pertaining to affine hyperbolic toral automorphisms. As in the introduction, an affine hyperbolic toral automorphism is a hyperbolic toral automorphism followed by a translation. We give two proofs of the result. The first gives the precise location of periodic points. The second relies on the fact that chaos is entirely topological and uses topological conjugacy.

Theorem 4.1. *Any affine hyperbolic toral automorphism is chaotic.*

Proof. Let v_1, v_2, \dots, v_d be the eigenvectors of A associated with $\lambda_1, \lambda_2, \dots, \lambda_d$, respectively. The eigenvectors form a basis for \mathbb{R}^d , so for any translation $b \in \mathbb{R}^d$, b can be written as $b = b_1 v_1 + b_2 v_2 + \dots + b_d v_d$ and any point on $x \in \mathbb{T}^d$ as $\sum_{i=1}^d x_i v_i$. So instead of $x^{n+1} = Ax^n + b$, we may write

$$x^{n+1} = A \sum_{i=1}^d x_i^n v_i + \sum_{i=1}^d b_i v_i.$$

We wish to find a closed form of x^m . For any point $x^0 \in \mathbb{T}^d$,

$$x^1 = Ax^0 + b = A \sum_{i=1}^d x_i^0 v_i + \sum_{i=1}^d b_i v_i = \sum_{i=1}^d \lambda_i x_i^0 v_i + \sum_{i=1}^d b_i v_i,$$

$$\begin{aligned} x^2 = Ax^1 + b &= A \left(\sum_{i=1}^d \lambda_i x_i^0 v_i + \sum_{i=1}^d b_i v_i \right) + \sum_{i=1}^d b_i v_i \\ &= \sum_{i=1}^d \lambda_i^2 x_i^0 v_i + \sum_{i=1}^d \lambda_i b_i v_i + \sum_{i=1}^d b_i v_i, \end{aligned}$$

$$\begin{aligned} x^3 = Ax^2 + b &= A \left(\sum_{i=1}^d \lambda_i^2 x_i^0 v_i + \sum_{i=1}^d \lambda_i b_i v_i + \sum_{i=1}^d b_i v_i \right) + \sum_{i=1}^d b_i v_i \\ &= \sum_{i=1}^d \lambda_i^3 x_i^0 v_i + \sum_{i=1}^d \lambda_i^2 b_i v_i + \sum_{i=1}^d \lambda_i b_i v_i + \sum_{i=1}^d b_i v_i. \end{aligned}$$

The first three iterations suggest that

$$x^n = \sum_{i=1}^d \lambda_i^n x_i^0 v_i + \sum_{j=1}^n \sum_{i=1}^d \lambda_i^{j-1} b_i v_i. \tag{2}$$

Assume (2) as an induction hypothesis. Then we see that it also holds for $n + 1$:

$$\begin{aligned} x^{n+1} &= Ax^n + b = A \left(\sum_{i=1}^d \lambda_i^n x_i^0 v_i + \sum_{j=1}^n \sum_{i=1}^d \lambda_i^{j-1} b_i v_i \right) + \sum_{i=1}^d b_i v_i \\ &= \sum_{i=1}^d \lambda_i^{n+1} x_i^0 v_i + \sum_{j=1}^n \sum_{i=1}^d \lambda_i^j b_i v_i + \sum_{i=1}^d b_i v_i \\ &= \sum_{i=1}^d \lambda_i^{n+1} x_i^0 v_i + \sum_{j=1}^{n+1} \sum_{i=1}^d \lambda_i^{j-1} b_i v_i. \end{aligned}$$

The last expression in (2) is not particularly revealing until we rewrite the double sum as

$$\sum_{j=1}^n \sum_{i=1}^d \lambda_i^{j-1} b_i v_i = \sum_{i=1}^d b_i v_i \sum_{j=1}^n \lambda_i^{j-1} = \sum_{i=1}^d b_i \frac{1 - \lambda_i^n}{1 - \lambda_i} v_i$$

and remember that we are looking for periodic points such that $x^0 = x^m \pmod 1$. We are looking for $x = \sum_{i=1}^n x_i v_i \pmod 1$ such that

$$\sum_{i=1}^d x_i v_i = \sum_{i=1}^d \lambda_i^n x_i v_i + \sum_{i=1}^d b_i \frac{1 - \lambda_i^n}{1 - \lambda_i} v_i \pmod 1,$$

which leads to

$$\begin{aligned} 0 &= \sum_{i=1}^d \lambda_i^n x_i v_i - \sum_{i=1}^d x_i v_i + \sum_{i=1}^d b_i \frac{1 - \lambda_i^n}{1 - \lambda_i} v_i \pmod 1 \\ &= \sum_{i=1}^d \left(\lambda_i^n x_i v_i - x_i v_i + b_i \frac{1 - \lambda_i^n}{1 - \lambda_i} v_i \right) \pmod 1 \\ &= \sum_{i=1}^d \left((\lambda_i^n - 1) x_i v_i + b_i \frac{1 - \lambda_i^n}{1 - \lambda_i} v_i \right) \pmod 1 \\ &= \sum_{i=1}^d (\lambda_i^n - 1) \left(x_i v_i + \frac{b_i}{\lambda_i - 1} v_i \right) \pmod 1 \\ &= \sum_{i=1}^d (\lambda_i^n - 1) \left(x_i + \frac{b_i}{\lambda_i - 1} \right) v_i \pmod 1, \end{aligned}$$

from which we can conclude that the periodic points of the affine map are precisely those of the linear map translated by $\sum_{i=1}^d (b_i / (\lambda_i - 1)) v_i$. Since the periodic points of the linear map are dense, so too are the periodic points of the affine map. In addition, if U, V are open in \mathbb{T}^d , then there exists an n such that the n -th iterate of the linear map of U intersects V . Thus, the affine map is chaotic. \square

There is another proof of the main result that uses far less calculation, but does not give the new locations of periodic points. We use the fact that the linear and affine hyperbolic toral automorphisms $f(x)$ and $g(x) = f(x) + b$ are topologically conjugate, so that the following diagram commutes:

$$\begin{array}{ccc} \mathbb{T}^d & \xrightarrow{f} & \mathbb{T}^d \\ \downarrow h & & \downarrow h \\ \mathbb{T}^d & \xrightarrow{g} & \mathbb{T}^d \end{array}$$

Alternate proof of Theorem 4.1. If we suppose f is defined by multiplication by a matrix A and g is multiplication by A followed by translation by b , we must define the homeomorphism h so that $f = h^{-1} \circ g \circ h$. This homeomorphism h is simply translation by some \tilde{b} :

$$\begin{aligned} Ax = A(x + \tilde{b}) + b - \tilde{b} &= Ax + A\tilde{b} + b - \tilde{b} = Ax + (A - I_d)\tilde{b} + b \\ \implies (A - I_d)\tilde{b} = -b &\implies \tilde{b} = -(A - I_d)^{-1}b. \end{aligned}$$

We know $A - I_d$ is invertible because A has no eigenvalues equal to 1.

Since f and g are topologically conjugate, they have the same properties regarding transitivity and periodic points. The transitivity and dense periodic points of linear f are known, so they hold also for the affine g . \square

Corollary 4.2. *Any affine hyperbolic transformation of \mathbb{T}^2 is chaotic.*

Proof. Suppose the linear part of the transformation is multiplication by a matrix A . Then the roots of the characteristic polynomial are

$$\begin{aligned} \det(A - \lambda I_2) &= \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc) \\ \implies \lambda &= \frac{(a + d) \pm \sqrt{(a + d)^2 - 4(ad - bc)}}{2}. \end{aligned}$$

The discriminant $(a + d)^2 - 4(ad - bc) = (a + d)^2 - 4$ cannot be a perfect square because 2 is not part of any Pythagorean triple. Thus the eigenvalues are irrational, which in turn implies that each of the eigenvectors has irrational slope. This is sufficient since all affine hyperbolic functions on \mathbb{T}^2 meet the criteria of the main result, and are thus chaotic. \square

Example 4.3. We now give a specific example of the previous corollary based on [Example 3.3](#). Let $L_{A,b} : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ be defined as

$$L_{A,b} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{3}/3 \end{bmatrix}. \quad (3)$$

Recall that the eigenvectors of the matrix A are $v_1 = [\frac{1}{2}(1 + \sqrt{5}), 1]^\top$ and $v_2 = [\frac{1}{2}(1 - \sqrt{5}), 1]^\top$, with eigenvalues $\lambda_1 = \frac{1}{2}(3 + \sqrt{5})$ and $\lambda_2 = \frac{1}{2}(3 - \sqrt{5})$, so that $\lambda_1 > 1 > \lambda_2 > 0$. The points

$$p_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad p_3 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \quad \text{and} \quad p_{10} = \begin{bmatrix} 1/5 \\ 3/5 \end{bmatrix}$$

are fixed and periodic of periods 3 and 10 respectively under [Example 3.3](#). To find the corresponding periodic points q_1, q_3, q_{10} under $L_{A,b}$, first calculate b_1, b_2 , which are the projections of b against v_1, v_2 .¹ Then add the translation prescribed by [Theorem 4.1](#):

$$\frac{b_1}{1-\lambda_1} v_1 + \frac{b_2}{1-\lambda_2} v_2 = \frac{\langle b, v_1 \rangle}{1-\lambda_1} v_1 + \frac{\langle b, v_2 \rangle}{1-\lambda_2} v_2 \approx \begin{bmatrix} .4227 \\ .8702 \end{bmatrix},$$

$$q_1 \approx p_1 + \begin{bmatrix} .4227 \\ .8702 \end{bmatrix} = \begin{bmatrix} .4227 \\ .8702 \end{bmatrix},$$

$$q_3 \approx p_3 + \begin{bmatrix} .4227 \\ .8702 \end{bmatrix} = \begin{bmatrix} .9227 \\ .3702 \end{bmatrix},$$

$$q_{10} \approx p_{10} + \begin{bmatrix} .4227 \\ .8702 \end{bmatrix} = \begin{bmatrix} .6227 \\ .4702 \end{bmatrix}.$$

One can check numerically that indeed

$$L_{A,b}(q_1) = q_1, \quad L_{A,b}^3(q_3) = q_3, \quad L_{A,b}^{10}(q_{10}) = q_{10},$$

and for each this is the minimum number of iterations required.

Corollary 4.4. Any function $f_{k,\alpha}$ on the circle given by $f_{k,\alpha} : \theta \mapsto k\theta + \alpha$ with $k \in \mathbb{Z} \setminus \{-1, 0, 1\}$ and $\alpha \in S^1$ is chaotic.

Proof. The slope of an eigenvector degenerates for $S^1 = \mathbb{T}^1$, and in any case the function $f_{k,\alpha}$ is not an automorphism. In the sense that $f_{k,\alpha}$ is a function that is known to be chaotic followed by a (possibly irrational) rotation, the main result holds.

First, an alternative explanation for the transitivity of $f_{k,\alpha}$ is in order. Any open set U in S^1 contains an open arc (θ_1, θ_2) with length $\theta_2 - \theta_1$. Define $m = 2\pi/(\theta_2 - \theta_1)$. After $n \geq k^m$ iterations of $f_{k,\alpha}$ on U , we have $f_{k,\alpha}^n(U) = S^1$. (A function with this

¹Remember that all arithmetic is performed modulo 1.

condition is said to be *locally eventually onto*.) Since every open set is “eventually onto” the circle, any U will certainly have nontrivial intersection with any nonempty open $V \in S^1$ after a finite number of iterations.

The periodic points of $f_k = f_{k,0}$ are the rational points in $[0, 1)$ with denominator one less than a power of k . Moreover, if this denominator is $q = k^n - 1$, then the point $p/q \in \mathbb{Q} \cap [0, 1)$ will have period n if p/q is fully reduced. To see this, note that

$$f_k^n\left(\frac{p}{q}\right) = k^n \frac{p}{k^n - 1} = (k^n - 1 + 1) \frac{p}{k^n - 1} = (k^n - 1) \frac{p}{k^n - 1} + \frac{p}{k^n - 1} = p + \frac{p}{q} = \frac{p}{q}$$

because $p \in \mathbb{N}$ and all of our arithmetic is modulo 1. Since q can be chosen arbitrarily large and $p = 0, 1, \dots, q - 1$, with p/q evenly spaced about the circle, the periodic points of f_k are dense in S^1 .

We can now use the closed form from the proof of [Theorem 4.1](#) to find the periodic points of $f_{k,\alpha}$. We are searching for points such that $f_{k,\alpha}^n(x) = x \pmod{1}$, or

$$(k^n - 1) \left(x + \frac{\alpha}{k - 1} \right) = x \pmod{1}.$$

This shows that the periodic points of $f_{k,\alpha}$ are rational points of f_k rotated about the circle by $\alpha/(k - 1)$. Since the locally eventually onto property is preserved, and periodic points are still dense, $f_{k,\alpha}$ is chaotic. \square

5. Conclusion

Our main result shows that affine hyperbolic toral automorphisms are chaotic. The added translation by a vector b preserves the transitivity of the map and translates all of the periodic points by $\sum_{i=1}^d (b_i/(\lambda_i - 1))v_i$, where the v_i are eigenvectors, λ_i the corresponding eigenvalues, and b_i the coordinates of the translation vector b in the basis defined by the eigenvectors. Note that in the case that $b = 0$, the periodic points are not translated at all, which coincides with a linear hyperbolic toral automorphism.

Using this translation result, one can construct an automorphism of the torus in which any specified point y has a specified period n : Find an x such that x has period n under a linear hyperbolic toral automorphism. By [Lemma 3.2](#), x will have rational coordinates in the standard basis (but not necessarily in the basis defined by the eigenvectors of the linear automorphism). Then define b such that $b_i = (\lambda_i - 1)(y_i - x_i) \pmod{1}$, where b_i, x_i, y_i are the coordinates in the basis defined by the eigenvectors of the linear toral automorphism. The resulting affine hyperbolic toral automorphism will have y as a periodic point with period n .

More generally, the main result shows that the incorporation of an irrational rotation into a toral automorphism does not necessarily eliminate the possibility of periodic points.

References

- [Banks et al. 1992] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey, “On Devaney’s definition of chaos”, *Amer. Math. Monthly* **99**:4 (1992), 332–334. [MR 1157223](#) [Zbl 0758.58019](#)
- [Crannell 1995] A. Crannell, “The role of transitivity in Devaney’s definition of chaos”, *Amer. Math. Monthly* **102**:9 (1995), 788–797. [MR 1357723](#) [Zbl 0849.58046](#)
- [Değirmenci and Koçak 2010] N. Değirmenci and Ş. Koçak, “Chaos in product maps”, *Turkish J. Math.* **34**:4 (2010), 593–600. [MR 2721970](#) [Zbl 1208.37010](#)
- [Elaydi 2008] S. N. Elaydi, *Discrete chaos: With applications in science and engineering*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL, 2008. [MR 2364977](#) [Zbl 1153.39002](#)
- [Katok and Hasselblatt 1995] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications **54**, Cambridge Univ. Press, 1995. [MR 1326374](#) [Zbl 0878.58020](#)
- [Li and Zhou 2013] R. Li and X. Zhou, “A note on chaos in product maps”, *Turkish J. Math.* **37**:4 (2013), 665–675. [MR 3070943](#) [Zbl 1293.37015](#)
- [Vellekoop and Berglund 1994] M. Vellekoop and R. Berglund, “On intervals, transitivity = chaos”, *Amer. Math. Monthly* **101**:4 (1994), 353–355. [MR 1270961](#) [Zbl 0886.58033](#)

Received: 2014-01-30

Accepted: 2015-08-17

cfcthomson@gmail.com

University of North Carolina at Chapel Hill,
Phillips Hall CB#3250, Chapel Hill, NC 27599, United States

domolinek@davidson.edu

Davidson College, Box 6999, Davidson, NC 28035,
United States

Rings of invariants for the three-dimensional modular representations of elementary abelian p -groups of rank four

Théo Pierron and R. James Shank

(Communicated by Ravi Vakil)

We show that the rings of invariants for the three-dimensional modular representations of an elementary abelian p -group of rank four are complete intersections with embedding dimension at most five. Our results confirm the conjectures of Campbell, Shank and Wehlau (*Transform. Groups* 18 (2013), 1–22) for these representations.

Introduction

We continue the investigation of the rings of invariants of modular representations of elementary abelian p -groups initiated in [Campbell et al. 2013]. We show that the rings of invariants for three-dimensional modular representations of groups of rank four are complete intersections and we confirm the conjectures of [loc. cit., §8] for these representations.

Let V denote an n -dimensional representation of a group G over a field \mathbb{F} of characteristic p for a prime number p . We will usually assume that G is finite and that p divides the order of G , in other words, that V is a *modular representation* of G . We view V as a left module over the group ring $\mathbb{F}G$ and the dual, V^* , as a right $\mathbb{F}G$ -module. Let $\mathbb{F}[V]$ denote the symmetric algebra on V^* . The action of G on V^* extends to an action by degree-preserving algebra automorphisms on $\mathbb{F}[V]$. By choosing a basis $\{x_1, x_2, \dots, x_n\}$ for V^* , we identify $\mathbb{F}[V]$ with the algebra of polynomials $\mathbb{F}[x_1, x_2, \dots, x_n]$. Our convention that $\mathbb{F}[V]$ is a right $\mathbb{F}G$ -module is consistent with the convention used by the invariant theory package in the computer algebra software Magma [Bosma et al. 1997]. The ring of invariants, $\mathbb{F}[V]^G$, is the subring of $\mathbb{F}[V]$ consisting of those polynomials fixed by the action of G . Note that elements of $\mathbb{F}[V]$ represent polynomial functions on V and that elements of $\mathbb{F}[V]^G$ represent polynomial functions on the set of orbits V/G . For G finite and \mathbb{F} algebraically

MSC2010: 13A50.

Keywords: modular invariant theory, elementary abelian p -groups.

closed, $\mathbb{F}[V]^G$ is the ring of regular functions on the categorical quotient $V//G$. For background on the invariant theory of finite groups, see [Benson 1993; Campbell and Wehlau 2011; Derksen and Kemper 2002; Neusel and Smith 2002].

Computing the ring of invariants for a modular representation is typically a difficult problem; the rings are often not Cohen–Macaulay. It is natural to take p -groups as a starting point and recent work of David Wehlau [2013] gives us a good understanding in the case of a cyclic group of order p . The next step is to look at elementary abelian p -groups. The rings of invariants for the two-dimensional modular representations of elementary abelian p -groups were computed in Section 2 of [Campbell et al. 2013] and the three-dimensional modular representations were classified in Section 4 of that paper. The only three-dimensional representations for which computing the ring of invariants is not straightforward are those of type $(1, 1, 1)$, in other words, those representations for which $\dim(V^G) = 1$ and $\dim((V/V^G)^G) = 1$. Our goal here is to compute the rings of invariants for representations of type $(1, 1, 1)$ for groups of rank four. The methods we use are essentially the same as the methods used in [loc. cit.]. As the rank increases, the complexity of the required calculations increases; we believe that it is not feasible to use the methods here for rank greater than four.

We denote by $E = \langle e_1, e_2, e_3, e_4 \rangle \cong (\mathbb{Z}/p)^4$ a rank-four elementary abelian p -group. Note that E only has representations of type $(1, 1, 1)$ if $p > 2$, so we make this assumption throughout the paper. As in Section 4 of [loc. cit.], define $\sigma : \mathbb{F}^2 \rightarrow \mathrm{GL}_3(\mathbb{F})$ by

$$\sigma(c_1, c_2) := \begin{pmatrix} 1 & 2c_1 & c_1^2 + c_2 \\ 0 & 1 & c_1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that σ defines a representation of the group $(\mathbb{F}^2, +)$. For a matrix

$$M := \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \end{pmatrix}$$

with $c_{ij} \in \mathbb{F}$, the assignment $e_j \mapsto \sigma(c_{1j}, c_{2j})$ determines a three-dimensional representation of E , which we denote by V_M . The action of E on $\mathbb{F}[x, y, z]$ is given by right multiplication on $x = [0 \ 0 \ 1]$, $y = [0 \ 1 \ 0]$ and $z = [1 \ 0 \ 0]$. Thus $x \cdot \sigma(c_1, c_2) = x$, $y \cdot \sigma(c_1, c_2) = y + c_1x$ and $z \cdot \sigma(c_1, c_2) = z + 2c_1y + (c_1^2 + c_2)x$. The representation V_M is of type $(1, 1, 1)$ if at least one c_{1j} is nonzero. Furthermore, by Proposition 4.1 of [loc. cit.], for every representation of type $(1, 1, 1)$, there exists a choice of basis for which the action is given by some matrix M .

In this paper, we compute $\mathbb{F}[V_M]^E$ for all $M \in \mathbb{F}^{2 \times 4}$. We give a stratification of $\mathbb{F}^{2 \times 4}$ and show that within each stratum there is a uniform computation of $\mathbb{F}[V_M]^E$. Note that the automorphism group of E is isomorphic to $\mathrm{GL}_4(\mathbb{F}_p)$, where \mathbb{F}_p denotes the field of p elements. Since $\mathbb{F}_p \subseteq \mathbb{F}$, there is a natural right action of $\mathrm{GL}_4(\mathbb{F}_p)$

on $\mathbb{F}^{2 \times 4}$. If M and M' lie in the same $\mathrm{GL}_4(\mathbb{F}_p)$ -orbit, then $\mathbb{F}[V_M]^E = \mathbb{F}[V_{M'}]^E$. Essentially, we study subrings of $\mathbb{F}[x, y, z]$ parametrised by points in $\mathbb{F}^{2 \times 4} / \mathrm{GL}_4(\mathbb{F}_p)$ and use elements of $\mathbb{F}[\mathbb{F}^{2 \times 4}]^{\mathrm{SL}_4(\mathbb{F}_p)}$ to describe the stratification.

In [Section 2](#), we work over the field $\mathbb{k} := \mathbb{F}_p(x_{ij} \mid i \in \{1, 2\}, j \in \{1, 2, 3, 4\})$ and compute $\mathbb{k}[V_{\mathcal{M}}]^E$ for the generic matrix

$$\mathcal{M} := \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{pmatrix}.$$

We show that $\mathbb{k}[V_{\mathcal{M}}]^E$ is a complete intersection of embedding dimension five with generators in degrees 1, p^2 , $p^2 + 2p$, $p^3 + 2$ and p^4 , and relations in degrees $p^3 + 2p^2$ and $p^4 + 2p$. Consider the 10×4 matrix

$$\Gamma := \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{11}^p & x_{12}^p & x_{13}^p & x_{14}^p \\ x_{21}^p & x_{22}^p & x_{23}^p & x_{24}^p \\ \vdots & \vdots & \vdots & \vdots \\ x_{11}^{p^4} & x_{12}^{p^4} & x_{13}^{p^4} & x_{14}^{p^4} \\ x_{21}^{p^4} & x_{22}^{p^4} & x_{23}^{p^4} & x_{24}^{p^4} \end{pmatrix}$$

and for a subsequence (i, j, k, ℓ) of $(1, 2, \dots, 10)$, let γ_{ijkl} denote the associated 4×4 minor of Γ . Note that $\gamma_{ijkl} \in \mathbb{F}[\mathbb{F}^{2 \times 4}]^{\mathrm{SL}_4(\mathbb{F}_p)}$ and, for $g \in \mathrm{GL}_4(\mathbb{F}_p)$, we have $g(\gamma_{ijkl}) = \det(g)\gamma_{ijkl}$. We use zero-sets of various γ_{ijkl} to define the stratification of $\mathbb{F}^{2 \times 4} / \mathrm{GL}_4(\mathbb{F}_p)$. In [Section 3](#), we show that for $M \in \mathbb{F}^{2 \times 4}$ with $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$, the generic calculation survives evaluation. In [Sections 4 through 10](#), we compute the rings of invariants for the remaining strata.

[Section 4](#): $\gamma_{1357}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$, $\gamma_{1234}(M) = 0$. We show $\mathbb{F}[V_M]^E$ is a complete intersection with generators in degrees 1, $2p$, p^3 , $p^3 + 2$ and p^4 , and relations in degrees $2p^3$ and $p^4 + 2p$.

[Section 5](#): $\gamma_{1357}(M) \neq 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) \neq 0$. If $\gamma_{1245}(M) \neq 0$ then $\mathbb{F}[V_M]^E$ is a complete intersection with generators in degrees 1, p^2 , $p^2 + p$, $p^3 + p + 2$ and p^4 , and relations in degrees $p^3 + p^2$ and $p^4 + p^2 + 2p$. Otherwise, $\mathbb{F}[V_M]^E$ is a hypersurface with generators in degrees 1, p^2 , $p^2 + 2$ and p^4 , with the relation in degree $p^4 + 2p^2$.

[Section 6](#): $\gamma_{1357}(M) = 0$, $\gamma_{1235}(M) \neq 0$, $\gamma_{1234}(M) \neq 0$. We show $\mathbb{F}[V_M]^E$ is a complete intersection with generators in degrees 1, p^2 , $p^2 + 2p$, $p^3 + 1$ and p^4 , and relations in degrees $p^3 + 2p^2$ and $p^4 + p$.

[Section 7](#): $\gamma_{1357}(M) \neq 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) = 0$. We show $\mathbb{F}[V_M]^E$ is a hypersurface. If $\gamma_{1257}(M) = 0$, then the generators are in degrees 1, 2, p^4 and p^4

and the relation is in degree $2p^4$. Otherwise, the generators are in degrees 1, p , $p^3 + p^2 + p + 2$, p^4 and the relation is in degree $p^4 + p^3 + p^2 + 2p$.

Section 8: $\gamma_{1357}(M) = 0$, $\gamma_{1235}(M) \neq 0$, $\gamma_{1234}(M) = 0$. We show $\mathbb{F}[V_M]^E$ is a complete intersection with generators in degrees 1, $2p$, p^3 , $p^3 + 1$ and p^4 , with relations in degrees $2p^2$ and $p^4 + p$.

Section 9: $\gamma_{1357}(M) = 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) \neq 0$. If $\gamma_{1245}(M) \neq 0$, then $\mathbb{F}[V_M]^E$ is a complete intersection with generators in degrees 1, p^2 , $p^2 + p$, $p^3 + 1$ and p^4 , with relations in degrees $p^3 + p^2$ and $p^4 + p$. Otherwise, $\mathbb{F}[V_M]^E$ is a hypersurface with generators in degrees 1, p^2 , $p^2 + 1$ and p^4 , with a relation in degree $p^4 + p^2$.

Section 10: $\gamma_{1357}(M) = 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) = 0$. If $\gamma_{1246}(M) \neq 0$ then $\mathbb{F}[V_M]^E$ is a hypersurface with generators in degrees 1, p , $p^3 + 1$, p^4 and a relation in degree $p^4 + p$. Otherwise, the representation is either not faithful or not of type (1, 1, 1); in either case, the invariants were computed in [Campbell et al. 2013].

1. Preliminaries

We make extensive use of the theory of SAGBI bases to compute rings of invariants. A SAGBI basis is the subalgebra analogue of a Gröbner basis for ideals, and is a particularly nice generating set for the subalgebra. The concept was introduced independently by Robbiano and Sweedler [1990] and Kapur and Madlener [1989]; a useful reference is Chapter 11 of Sturmfels [1996]. We adopt the convention that a monomial is a product of variables and a term is a monomial with a coefficient. We use the graded reverse lexicographic order with $x < y < z$. For a polynomial $f \in \mathbb{F}[x, y, z]$, we denote the lead monomial of f by $\text{LM}(f)$ and the lead term of f by $\text{LT}(f)$. For $\mathcal{B} = \{h_1, \dots, h_\ell\} \subset \mathbb{F}[x, y, z]$ and $I = (i_1, \dots, i_\ell)$, a sequence of nonnegative integers, denote $\prod_{j=1}^\ell h_j^{i_j}$ by h^I . A *tête-à-tête* for \mathcal{B} is a pair (h^I, h^J) with $\text{LM}(h^I) = \text{LM}(h^J)$; we say that a *tête-à-tête* is *nontrivial* if the support of I is disjoint from the support of J . The reduction of an S -polynomial is a fundamental calculation in the theory of Gröbner bases. The analogous calculation for SAGBI bases is the *subduction* of a *tête-à-tête*. For any $f \in \mathbb{F}[x, y, z]$, if there exists a sequence I such that $\text{LM}(f) = \text{LM}(h^I)$, we can choose $c \in \mathbb{F}$ so that $\text{LT}(f) = \text{LT}(ch^I)$. Then $\text{LT}(f - ch^I) < \text{LT}(f)$. If by iterating this process we can write f as a polynomial in the h_i , we say that f *subducts to zero* (using \mathcal{B}). For a *tête-à-tête* (h^I, h^J) , choose c so that $\text{LT}(h^I) = \text{LT}(ch^J)$. We say that the *tête-à-tête* *subducts to zero* if $h^I - ch^J$ subducts to zero. A subset \mathcal{B} of a subalgebra $A \subset \mathbb{F}[x_1, \dots, x_n]$ is a SAGBI basis for A if the lead monomials of the elements of \mathcal{B} generate the lead term algebra of A or, equivalently, every nontrivial *tête-à-tête* for \mathcal{B} subducts to zero. For background material on term orders and Gröbner bases, we recommend [Adams and Loustau 1994].

The following specialisation of Theorem 1.1 of [Campbell et al. 2013] is our primary computational tool. Note that under the hypotheses of the theorem, $\{x, h_1, h_\ell\}$ is a homogeneous system of parameters and, therefore, $\mathbb{F}[V_M]^E$ is an integral extension of A .

Theorem 1.1. *For homogeneous $h_1, \dots, h_\ell \in \mathbb{F}[V_M]^E$ with $\text{LM}(h_1) = y^i$ for some $i > 0$, $\text{LM}(h_\ell) = z^j$ for some $j > 0$ and $\text{LM}(h_k) \in \mathbb{F}[y, z]$ for $k = 2, \dots, \ell - 1$, define $\mathcal{B} := \{x, h_1, \dots, h_\ell\}$ and let A denote the algebra generated by \mathcal{B} . If $A[x^{-1}] = \mathbb{F}[V_M]^E[x^{-1}]$ and \mathcal{B} is a SAGBI basis for A , then $A = \mathbb{F}[V_M]^E$ and \mathcal{B} is a SAGBI basis for $\mathbb{F}[V_M]^E$.*

Note that, if an algebra is generated by a finite SAGBI basis, then for the corresponding presentation, the ideal of relations is generated by elements corresponding to the subdivisions of the nontrivial tête-à-têtes (see Corollary 11.6 of [Sturmfels 1996]). We use the term *complete intersection* to refer to an algebra with a presentation for which the ideal of relations is generated by a regular sequence. Since the Krull dimension of $\mathbb{F}[V_M]^E$ is three, the ring is a complete intersection if the number of generators minus the number of nontrivial tête-à-têtes is three.

We routinely use the *SAGBI/divide-by- x* algorithm introduced in Section 1 of [Campbell et al. 2013]. The traditional SAGBI basis algorithm proceeds by subtracting tête-à-têtes and adding any nonzero subdivisions to the generating set. For *SAGBI/divide-by- x* , if a nonzero subdivision is divisible by x , we divide by the highest possible power of x before adding the polynomial to the generating set. While the SAGBI algorithm extends the generating set for a given subalgebra, *SAGBI/divide-by- x* extends the subalgebra. If we start with a subalgebra A which contains a homogeneous system of parameters and satisfies the condition that $A[x^{-1}] = \mathbb{F}[V_M]^E[x^{-1}]$, then the *SAGBI/divide-by- x* algorithm will produce a generating set for $\mathbb{F}[V_M]^E$ (see Theorem 1.2 of [loc. cit.]).

For $f \in \mathbb{F}[V_M]$, we define the *norm* of f to be the orbit product

$$N_M(f) := \prod \{f \cdot g \mid g \in E\} \in \mathbb{F}[V_M]^E$$

with the action of E determined by M . When applying Theorem 1.1, we often take h_ℓ to be $N_M(z)$.

Remark 1.2. Note that the action of E restricts to an action on $\mathbb{F}[x, y]$ and that $\mathbb{F}[x, y]^E = \mathbb{F}[x, N_M(y)]$ (see Section 2 of [Campbell et al. 2013]). Therefore, if $h \in \mathbb{F}[x, y]^E$ is homogeneous with $\deg(h) = |\{y \cdot g \mid g \in E\}|$ then h is a linear combination of $N_M(y)$ and $x^{\deg(h)}$.

Define $\delta := y^2 - xz$ and observe that

$$\delta \cdot \sigma(c_1, c_2) = (y + c_1x)^2 - x(z + 2c_1y + (c_1^2 + c_2)x) = \delta - c_2x^2.$$

Note that $\mathbb{F}[x, y, z][x^{-1}] = \mathbb{F}[x, y, -\delta/x][x^{-1}]$ and that the $\mathbb{F}[x, y, -\delta/x]^E$ is a polynomial algebra (see Theorem 3.9.2 of [Campbell and Wehlauf 2011]). This “change of basis” can be a useful way to compute the field of fractions of $\mathbb{F}[V_M]^E$. Form the matrix $\tilde{\Gamma}$ by augmenting Γ with the column

$$\left[\frac{y}{x} \left(-\frac{\delta}{x^2} \right) \left(\frac{y}{x} \right)^p \left(-\frac{\delta}{x^2} \right)^p \cdots \left(\frac{y}{x} \right)^{p^4} \left(-\frac{\delta}{x^2} \right)^{p^4} \right]^T.$$

For a subsequence $J = (j_1, \dots, j_5)$ of $(1, 2, \dots, 10)$, let $\tilde{f}_J \in \mathbb{k}[x, y, z][x^{-1}]$ denote the associated 5×5 minor of $\tilde{\Gamma}$. Let f_J denote the element of $\mathbb{k}[x, y, z]$ constructed by minimally clearing the denominator of \tilde{f}_J . Observe that $f_J \in \mathbb{k}[V_M]^E$. Furthermore, the coefficients of f_J lie in $\mathbb{F}_p[x_{ij}]^{\text{SL}_4(\mathbb{F}_p)}$ and, for an arbitrary $M \in \mathbb{F}^{2 \times 4}$, evaluating the coefficients of f_J at M gives an element $\tilde{f}_J \in \mathbb{F}[V_M]^E$. Invariants constructed in this way are a crucial ingredient in our calculations. Define $f_1 := f_{12345}$ and observe that $\text{LT}(f_1) = \gamma_{1234}y^{p^2}$. Note that $\text{LT}(f_{12346}) = -\gamma_{1234}y^{2p^2}$. A straightforward calculation shows that

$$\text{LT}(f_1^2 + \gamma_{1234}f_{12346}) = 2\gamma_{1234}\gamma_{1235}x^{p^2-2p}y^{p^2+2p}.$$

Therefore,

$$f_2 := \frac{f_1^2 + \gamma_{1234}f_{12346}}{2x^{p^2-2p}} \in \mathbb{k}[V_M]^E$$

has lead term $\gamma_{1234}\gamma_{1235}y^{p^2+2p}$.

We make frequent use of the *Plücker relations* for the minors of Γ and $\tilde{\Gamma}$.

Theorem 1.3. *Let N be an $n \times m$ matrix with $n > m$. Denote by p_{i_1, \dots, i_m} the $m \times m$ minor of N determined by the rows i_1, \dots, i_m . For sequences (i_1, \dots, i_{m-1}) and (j_1, \dots, j_{m+1}) , we have the following Plücker relation*

$$\sum_{a=1}^{m+1} (-1)^a p_{i_1, \dots, i_{m-1}, j_a} p_{j_1, \dots, j_{a-1}, j_{a+1}, \dots, j_{m+1}} = 0.$$

For a proof of the above theorem, see, for example, [Lakshmibai and Raghavan 2008, §4.1.3].

Lemma 1.4. *For $2 < i < 7$,*

$$\gamma_{12i7}\gamma_{1234}^p = \gamma_{12i6}\gamma_{1235}^p - \gamma_{12i5}\gamma_{1245}^p + \gamma_{12i4}\gamma_{1345}^p - \gamma_{12i3}\gamma_{2345}^p.$$

Proof. Since taking p -th powers is \mathbb{F}_p -linear, $\gamma_{(i+2)(j+2)(k+2)(\ell+2)} = \gamma_{ijkl}^p$. For example, $\gamma_{3456} = \gamma_{1234}^p$. The desired result follows from this fact, using the $(1, 2, i)(3, 4, 5, 6, 7)$ Plücker relation for the matrix Γ . \square

For $K = (k_1, k_2, \dots, k_6)$ a subsequence of $(1, 2, \dots, 10)$, let K_i denote the subsequence of K formed by omitting i and let $K_{i,j}$ denote the subsequence of K formed by omitting i and j . The following is Lemma 5.3 from [Campbell et al. 2013].

Lemma 1.5. For any subsequence (i_1, i_2, i_3) of K ,

$$(-1)^{\epsilon_1} \gamma_{K_{i_1, i_2}} \tilde{f}_{K_{i_3}} + (-1)^{\epsilon_2} \gamma_{K_{i_2, i_3}} \tilde{f}_{K_{i_1}} + (-1)^{\epsilon_3} \gamma_{K_{i_1, i_3}} \tilde{f}_{K_{i_2}} = 0$$

for some choice of $\epsilon_\ell \in \{0, 1\}$.

Remark 1.6. Note that $\gamma_{1357}(M) = 0$ if and only if $\{c_{11}, c_{12}, c_{13}, c_{14}\}$ is linearly dependent over \mathbb{F}_p . This follows from the usual construction of the Dickson invariants; see, for example, [Wilkerson 1983]. The key observation is that $\gamma_{1357}(M)^{p-1}$ is the product of the nonzero \mathbb{F}_p -linear combinations of $\{c_{11}, c_{12}, c_{13}, c_{14}\}$.

2. The generic case

In this section we compute $\mathbb{k}[V_{\mathcal{M}}]^E$. With f_1 and f_2 defined as in Section 1, using Theorem 5.2 of [Campbell et al. 2013], we see that

$$\mathbb{k}[V_{\mathcal{M}}]^E[x^{-1}] = \mathbb{k}[x, f_1, f_2][x^{-1}].$$

Thus it is sufficient to extend $\{x, f_1, f_2, N_{\mathcal{M}}(z)\}$ to a SAGBI basis. We use the SAGBI/divide-by- x algorithm of [loc. cit., §1] to do this. We will show that the algorithm produces one new invariant, which we denote by f_3 , and that

$$\text{LT}(f_3) = \gamma_{1357} y^{p^3+2}.$$

For $p = 3$ and $p = 5$, this result follows from a Magma calculation. For the rest of this section, we assume $p > 5$.

Expanding the definitions of f_1 , f_{12346} and f_2 gives

$$\begin{aligned} f_1 &= \gamma_{1234} y^{p^2} + \gamma_{1235} \delta^p x^{p^2-2p} + \gamma_{1245} x^{p^2-p} y^p + \gamma_{1345} \delta x^{p^2-2} + \gamma_{2345} x^{p^2-1} y, \\ f_{12346} &= -\gamma_{1234} \delta^{p^2} + \gamma_{1236} \delta^p x^{2p^2-2p} + \gamma_{1246} x^{2p^2-p} y^p + \gamma_{1346} \delta x^{2p^2-2} + \gamma_{2346} x^{2p^2-1} y \end{aligned}$$

and

$$\begin{aligned} f_2 &= \frac{f_1^2 + \gamma_{1234} f_{12346}}{2x^{p^2-2p}} \\ &= \gamma_{1234} \gamma_{1235} y^{p^2} \delta^p + \gamma_{1234} \gamma_{1245} x^p y^{p^2+p} + \gamma_{1234} \gamma_{1345} \delta x^{2p-2} y^{p^2} \\ &\quad + \gamma_{1234} \gamma_{2345} x^{2p-1} y^{p^2+1} + \frac{1}{2} \gamma_{1234}^2 x^{2p} z^{p^2} + \frac{1}{2} \gamma_{1235}^2 \delta^{2p} x^{p^2-2p} \\ &\quad + \gamma_{1235} \gamma_{1245} \delta^p x^{p^2-p} y^p + \gamma_{1235} \gamma_{1345} \delta^{p+1} x^{p^2-2} + \gamma_{1235} \gamma_{2345} \delta^p x^{p^2-1} y \\ &\quad + \frac{1}{2} \gamma_{1234} \gamma_{1236} x^{p^2} \delta^p + \frac{1}{2} \gamma_{1245}^2 x^{p^2} y^{2p} + \gamma_{1245} \gamma_{1345} \delta x^{p^2+p-2} y^p \\ &\quad + \gamma_{1245} \gamma_{2345} x^{p^2+p-1} y^{p+1} + \frac{1}{2} \gamma_{1234} \gamma_{1246} y^p x^{p^2+p} + \frac{1}{2} \gamma_{1345}^2 \delta^2 x^{p^2+2p-4} \\ &\quad + \gamma_{1345} \gamma_{2345} \delta x^{p^2+2p-3} y + \frac{1}{2} \gamma_{2345}^2 x^{p^2+2p-2} y^2 \\ &\quad + \frac{1}{2} \gamma_{1234} \gamma_{1346} \delta x^{p^2+2p-2} + \frac{1}{2} \gamma_{1234} \gamma_{2346} x^{p^2+2p-1} y. \end{aligned}$$

Subducting the tête-à-tête (f_1^{p+2}, f_2^p) gives

$$\begin{aligned} \tilde{f}_3 = & \underbrace{\gamma_{1235}^p f_1^{p+2}}_{T_1} - \underbrace{\gamma_{1234}^2 f_2^p}_{T_2} + \underbrace{\alpha_1 x^{p^2-2p} f_1^p f_2}_{T_3} \\ & + \underbrace{\alpha_2 x^{p^2} f_1^{p+1}}_{T_4} + \underbrace{\alpha_3 x^{2p^2-2p} f_1^{p-1} f_2}_{T_5} + \underbrace{\alpha_4 x^{2p^2-p} f_1^{(p-3)/2} f_2^{(p+1)/2}}_{T_6}, \end{aligned}$$

where

$$\alpha_1 = -2\gamma_{1235}^p, \quad \alpha_2 = \gamma_{1234}\gamma_{1245}^p, \quad \alpha_3 = \frac{\gamma_{1234}^{p+1}\gamma_{1237}}{\gamma_{1235}}, \quad \alpha_4 = \frac{\gamma_{1234}^{p+3}\gamma_{1257}}{\gamma_{1235}^{(p+3)/2}}.$$

Lemma 2.1. *For $p \geq 5$, we have $\text{LT}(\tilde{f}_3) = \alpha x^{2p^2-2} y^{p^3+2}$ with*

$$\alpha = \frac{\gamma_{1234}^{p+1}}{\gamma_{1235}} (\gamma_{1234}\gamma_{1345}^{p+1} + \gamma_{1235}^p \gamma_{1345}\gamma_{1236} - \gamma_{1235}^{p+1}\gamma_{1346}) = -\frac{\gamma_{1357}\gamma_{1234}^{2p+2}}{\gamma_{1235}}.$$

Proof. We work modulo the ideal in $\mathbb{k}[x, y, z]$ generated by x^{2p^2-1} . By the definition of f_2 , we have

$$T_1 - T_2 + T_3 = -\gamma_{1235}^p \gamma_{1234} f_1^p f_{12346} - \gamma_{1234}^2 f_2^p.$$

As $f_1^p \equiv \gamma_{1234}^p y^{p^3}$ and

$$\begin{aligned} f_2^p \equiv & \gamma_{1234}^p \gamma_{1235}^p \delta^{p^2} y^{p^3} + \gamma_{1234}^p \gamma_{1245}^p x^{p^2} y^{p^3+p^2} \\ & + \gamma_{1234}^p \gamma_{1345}^p \delta^p x^{2p^2-2p} y^{p^3} + \gamma_{1234}^p \gamma_{2345}^p x^{2p^2-p} y^{p^3+p}, \end{aligned}$$

we obtain

$$\begin{aligned} T_1 - T_2 + T_3 \equiv & -\gamma_{1234}^{p+2} \gamma_{1245}^p x^{p^2} y^{p^3+p^2} - \gamma_{1234}^{p+1} (\gamma_{1234}\gamma_{1345}^p + \gamma_{1235}^p \gamma_{1236}) \delta^p x^{2p^2-2p} y^{p^3} \\ & - \gamma_{1234}^{p+1} (\gamma_{1234}\gamma_{2345}^p + \gamma_{1235}^p \gamma_{1246}) x^{2p^2-p} y^{p^3+p} \\ & - \gamma_{1234}^{p+1} \gamma_{1235}^p \gamma_{1346} \delta x^{2p^2-2} y^{p^3}. \end{aligned}$$

Since

$$\begin{aligned} x^{p^2} f_1^{p+1} & \equiv \gamma_{1234}^p y^{p^3} x^{p^2} f_1 \\ & \equiv \gamma_{1234}^{p+1} x^{p^2} y^{p^3+p^2} + \gamma_{1234}^p \gamma_{1235}^p \delta^p x^{2p^2-2p} y^{p^3} \\ & \quad + \gamma_{1234}^p \gamma_{1245}^p x^{2p^2-p} y^{p^3+p} + \gamma_{1234}^p \gamma_{1345}^p \delta x^{2p^2-2} y^{p^3}, \end{aligned}$$

we see that

$$\begin{aligned} T_1 - T_2 + T_3 + T_4 \equiv & \gamma_{1234}^{p+1} (\gamma_{1235}\gamma_{1245}^p - \gamma_{1235}^p \gamma_{1236} - \gamma_{1234}\gamma_{1345}^p) x^{2p^2-2p} y^{p^3} \delta^p \\ & + \gamma_{1234}^{p+1} (\gamma_{1245}^{p+1} - \gamma_{1235}^p \gamma_{1246} - \gamma_{1234}\gamma_{2345}^p) x^{2p^2-p} y^{p^3+p} \\ & + \gamma_{1234}^{p+1} (\gamma_{1245}^p \gamma_{1345} - \gamma_{1235}^p \gamma_{1346}) \delta x^{2p^2-2} y^{p^3}. \end{aligned}$$

Using [Lemma 1.4](#) for $i = 3$ and $i = 4$, along with the analogous result coming from the $(1, 3, 4)(3, 4, 5, 6, 7)$ Plücker relation for Γ , gives

$$T_1 - T_2 + T_3 + T_4 \equiv -\gamma_{1234}^{2p+1} \gamma_{1237} x^{2p^2-2p} y^{p^3} \delta^p \\ - \gamma_{1234}^{2p+1} \gamma_{1247} x^{2p^2-p} y^{p^3+p} - \gamma_{1234}^{2p+1} \gamma_{1347} \delta x^{2p^2-2} y^{p^3}.$$

Since $3p^2 - 4p \geq 2p^2 - 1$ for $p \geq 5$, we have $x^{2p^2-2p} f_1^{p-1} \equiv \gamma_{1234}^{p-1} y^{p^3-p^2} x^{2p^2-2p}$. Using the description of f_2 given above,

$$x^{2p^2-2p} f_2 \equiv \gamma_{1234} x^{2p^2-2p} y^{p^2} (\gamma_{1235} \delta^p + \gamma_{1245} x^p y^p + \gamma_{1345} \delta x^{2p-2}).$$

Thus

$$T_5 \equiv \alpha_3 \gamma_{1234}^p y^{p^3} x^{2p^2-2p} (\gamma_{1235} \delta^p + \gamma_{1245} x^p y^p + \gamma_{1345} \delta x^{2p-2}).$$

Using the $(1, 2, 4)(1, 2, 3, 5, 7)$ and $(1, 3, 5)(1, 2, 3, 4, 7)$ Plücker relations gives

$$T_1 - T_2 + T_3 + T_4 + T_5 \equiv -\frac{\gamma_{1234}^{2p+2} \gamma_{1257}}{\gamma_{1235}} x^{2p^2-p} y^{p^3+p} - \frac{\gamma_{1234}^{2p+2} \gamma_{1357}}{\gamma_{1235}} \delta x^{2p^2-2} y^{p^3}.$$

Expanding and reducing modulo $\langle x^{2p^2-1} \rangle$, we get

$$x^{2p^2-p} f_1^{(p-3)/2} \equiv x^{2p^2-p} \gamma_{1234}^{(p-3)/2} y^{(p^3-3p^2)/2}$$

and

$$x^{2p^2-p} f_2^{(p+1)/2} \equiv \gamma_{1234}^{(p+1)/2} \gamma_{1235}^{(p+1)/2} x^{2p^2-p} y^{(p^3+3p^2)/2+p}.$$

Thus

$$\frac{T_6}{\alpha_4} \equiv \gamma_{1234}^{p-1} \gamma_{1235}^{(p+1)/2} x^{2p^2-p} y^{p^3+p}$$

and

$$\tilde{f}_3 = T_1 - T_2 + T_3 + T_4 + T_5 + T_6 \equiv \alpha x^{2p^2-2} y^{p^3+2}.$$

Using the $(1, 2, 3)(1, 3, 4, 5, 6)$ and $(1, 3, 5)(3, 4, 5, 6, 7)$ Plücker relations, we obtain

$$\alpha = \frac{\gamma_{1234}^{p+2}}{\gamma_{1235}} (\gamma_{1345}^{p+1} - \gamma_{1356} \gamma_{1235}^p) = -\frac{\gamma_{1234}^{2p+2} \gamma_{1357}}{\gamma_{1235}}$$

and, since we are using the grevlex term order with $x < y < z$, the result follows. \square

Define

$$f_3 := -\tilde{f}_3 \frac{\gamma_{1235}}{\gamma_{1234}^{2p+2} x^{2p^2-2}}$$

so that $\text{LT}(f_3) = \gamma_{1357} y^{p^3+2}$. Looking at the exponents of y modulo p , it is clear that there is only one new nontrivial tête-à-tête: $(f_3^p, f_2 f_1^{p^2-1})$. In order to prove that $\mathcal{B} := \{x, f_1, f_2, f_3, N_{\mathcal{M}}(z)\}$ is a SAGBI basis for $\mathbb{k}[V_{\mathcal{M}}]^E$, it is sufficient to show that this tête-à-tête subducts to zero. However, $N_{\mathcal{M}}(z)$ is rather complicated

and it is more convenient to take an indirect approach. Subducting the tête-à-tête using only $\{x, f_1, f_2, f_3\}$ gives

$$\begin{aligned} \tilde{f}_4 := & \underbrace{\beta_1 f_3^p}_{T'_1} - \underbrace{\beta_2 f_1^{p^2-1} f_2}_{T'_2} + \underbrace{\beta_3 x^p f_1^{p^2-(p+3)/2} f_2^{(p+1)/2}}_{T'_3} \\ & + \underbrace{\beta_4 x^{2p-2} f_1^{p^2-p} f_3}_{T'_4} + \underbrace{\beta_5 x^{2p-1} f_1^{(p^2-1)/2-p} f_2^{(p-1)/2} f_3^{(p+1)/2}}_{T'_5}, \end{aligned}$$

where

$$\begin{aligned} \beta_1 &:= \gamma_{1235} \gamma_{1234}^{p^2}, & \beta_2 &:= \gamma_{1357}^p, & \beta_3 &:= \frac{\gamma_{1234}(\gamma_{1245} \gamma_{1357}^p - \gamma_{1235} \gamma_{2357}^p)}{\gamma_{1235}^{(p+1)/2}}, \\ \beta_4 &:= \gamma_{1234}^p \gamma_{1345} \gamma_{1357}^{p-1}, & \beta_5 &:= -\gamma_{1234}^{(p^2+p+2)/2} \gamma_{1235}^{(p+3)/2} \gamma_{1357}^{(p-3)/2}. \end{aligned}$$

The lemma below proves that $\{x, f_1, f_2, f_3, \tilde{f}_4/x^{2p}\}$ is a SAGBI basis. We then use this in the proof of [Theorem 2.3](#).

Lemma 2.2. *For $p \geq 5$, we have $\text{LT}(\tilde{f}_4) = \frac{1}{2} \gamma_{1234}^{p^2} \gamma_{1235}^{p+1} x^{2p} z^{p^4}$.*

Proof. We work modulo the ideal in $\mathbb{k}[x, y, z]$ generated by x^{2p+1} and $x^{2p}y$, which we denote by \mathfrak{n} . Since $p \geq 5$, we have $p^2 - 2p \geq 2p + 1$. Therefore, using the expressions for f_1 and f_2 given above, we have $f_1 \equiv_{\mathfrak{n}} \gamma_{1234} y^{p^2}$ and

$$\begin{aligned} f_2 \equiv_{\mathfrak{n}} & \gamma_{1234} \gamma_{1235} y^{p^2} \delta^p + \gamma_{1234} \gamma_{1245} y^{p^2+p} x^p + \gamma_{1234} \gamma_{1345} \delta x^{2p-2} y^{p^2} \\ & + \gamma_{1234} \gamma_{2345} x^{2p-1} y^{p^2+1} + \frac{1}{2} \gamma_{1234}^2 x^{2p} z^{p^2}. \end{aligned}$$

We will need expressions modulo \mathfrak{n} for f_3^p , $x^{2p-2} f_3$ and $x^{2p-1} f_3^{(p+1)/2}$. Let \mathfrak{m} denote the ideal generated by $x^2 y$ and x^3 . Reworking the calculations of the proof of [Lemma 2.1](#) to keep additional terms of f_3 gives

$$f_3 \equiv_{\mathfrak{m}} \gamma_{1357} \delta y^{p^3} + \gamma_{2357} x y^{p^3+1} + \frac{1}{2} \gamma_{1235} x^2 z^{p^3}.$$

Thus

$$\begin{aligned} f_3^p &\equiv_{\mathfrak{n}} \gamma_{1357}^p \delta^p y^{p^4} + \gamma_{2357}^p x^p y^{p^4+p} + \frac{1}{2} \gamma_{1235}^p x^{2p} z^{p^4}, \\ x^{2p-2} f_3 &\equiv_{\mathfrak{n}} \gamma_{1357} \delta x^{2p-2} y^{p^3} + \gamma_{2357} x^{2p-1} y^{p^3+1} + \frac{1}{2} \gamma_{1235} x^{2p} z^{p^3}, \\ x^{2p-1} f_3^{(p+1)/2} &\equiv_{\mathfrak{n}} \gamma_{1357}^{(p+1)/2} x^{2p-1} y^{(p^3+2)(p+1)/2}. \end{aligned}$$

Therefore

$$\begin{aligned} T'_1 - T'_2 &\equiv_{\mathfrak{n}} \gamma_{1234}^{p^2} (\gamma_{1235} \gamma_{2357}^p - \gamma_{1245} \gamma_{1357}^p) x^p y^{p^4+p} - \gamma_{1234}^{p^2} \gamma_{1345} \gamma_{1357}^p \delta x^{2p-2} y^{p^4} \\ &\quad - \gamma_{1234}^2 \gamma_{2345} \gamma_{1357}^p x^{2p-1} y^{p^4+1} + \frac{1}{2} \gamma_{1234}^{p^2} \gamma_{1235}^{p+1} x^{2p} z^{p^4}. \end{aligned}$$

Since $x^p f_2^{(p+1)/2} \equiv_n \gamma_{1234}^{(p+1)/2} \gamma_{1235}^{(p+1)/2} x^p y^{(p^3+3p^2)/2+p}$, we have

$$T'_1 - T'_2 + T'_3 \equiv_n -\gamma_{1234}^p \gamma_{1345}^p \gamma_{1357}^p \delta x^{2p-2} y^{p^4} \\ - \gamma_{1234}^p \gamma_{2345}^p \gamma_{1357}^p x^{2p-1} y^{p^4+1} + \frac{1}{2} \gamma_{1234}^p \gamma_{1235}^{p+1} x^{2p} z^{p^4}.$$

Using the description of $x^{2p-2} f_3$ given above, we see that

$$T'_1 - T'_2 + T'_3 + T'_4 \\ \equiv_n \gamma_{1234}^p \gamma_{1357}^{p-1} (\gamma_{1345} \gamma_{2357} - \gamma_{1357} \gamma_{2345}) x^{2p-1} y^{p^4+1} + \frac{1}{2} \gamma_{1234}^p \gamma_{1235}^{p+1} x^{2p} z^{p^4}.$$

The $(2, 3, 5)(1, 3, 4, 5, 7)$ Plücker relation gives

$$\gamma_{2345} \gamma_{1357} - \gamma_{2357} \gamma_{1345} = -\gamma_{1235} \gamma_{3457}.$$

Thus

$$T'_1 - T'_2 + T'_3 + T'_4 \equiv_n \gamma_{1234}^p \gamma_{1357}^{p-1} \gamma_{1235} \gamma_{3457} x^{2p-1} y^{p^4+1} + \frac{1}{2} \gamma_{1234}^p \gamma_{1235}^{p+1} x^{2p} z^{p^4}.$$

Observe that

$$x^{2p-1} f_1^{(p^2-1)/2-p} \equiv_n \gamma_{1234}^{(p^2-1)/2-p} x^{2p-1} y^{(p^4-p^2)/2-p^3}$$

and

$$x^{2p-1} f_2^{(p-1)/2} \equiv_n \gamma_{1234}^{(p-1)/2} \gamma_{1235}^{(p-1)/2} x^{2p-1} y^{(p^3+p^2)/2-p}.$$

Therefore, using the description of $x^{2p-1} f_3^{(p+1)/2}$ given above, we obtain

$$\tilde{f}_4 := T'_1 - T'_2 + T'_3 + T'_4 + T'_5 \equiv_n \frac{1}{2} \gamma_{1234}^p \gamma_{1235}^{p+1} x^{2p} z^{p^4},$$

and, since we are using the grevlex term order with $x < y < z$, the result follows. \square

Theorem 2.3. *The set $\mathcal{B} := \{x, f_1, f_2, f_3, N_{\mathcal{M}}(z)\}$ is a SAGBI basis, and hence a generating set, for $\mathbb{k}[V_{\mathcal{M}}]^E$. Furthermore, $\mathbb{k}[V_{\mathcal{M}}]^E$ is a complete intersection with generating relations coming from the subduction of the tête-à-têtes (f_2^p, f_1^{p+2}) and $(f_3^p, f_2 f_1^{p^2-1})$.*

Proof. Define $f_4 := \tilde{f}_4/x^{2p}$, $\mathcal{B}' := \{x, f_1, f_2, f_3, f_4\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-têtes for \mathcal{B}' are (f_2^p, f_1^{p+2}) and $(f_3^p, f_2 f_1^{p^2-1})$. From Lemmas 2.1 and 2.2, these tête-à-têtes subduct to zero. Therefore \mathcal{B}' is a SAGBI basis for A . From Theorem 5.2 of [Campbell et al. 2013], $\mathbb{k}[V_{\mathcal{M}}]^E[x^{-1}] = \mathbb{k}[x, f_1, f_2][x^{-1}]$. Thus $A[x^{-1}] = \mathbb{k}[V_{\mathcal{M}}]^E[x^{-1}]$. Note that $\text{LM}(f_4) = z^{p^4}$. Therefore, by Theorem 1.1, $A = \mathbb{k}[V_{\mathcal{M}}]^E$ and \mathcal{B}' is a SAGBI basis for $\mathbb{k}[V_{\mathcal{M}}]^E$. Hence the lead term algebra of $\mathbb{k}[V_{\mathcal{M}}]^E$ is generated by $\{x, y^{p^2}, y^{p^2+2p}, y^{p^3+2}, z^{p^4}\}$. Since the orbit of z has size p^4 , we see that $\text{LM}(N_{\mathcal{M}}(z)) = z^{p^4}$. Thus $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$ and \mathcal{B} is also a SAGBI basis for $\mathbb{k}[V_{\mathcal{M}}]^E$. For any subalgebra with a SAGBI basis, the relations are generated by the nontrivial tête-à-tête. Hence (f_2^p, f_1^{p+2}) and $(f_3^p, f_2 f_1^{p^2-1})$ generate the ideal of relations and $\mathbb{k}[V_{\mathcal{M}}]^E$ is a complete intersection with embedding dimension five. \square

3. The essentially generic case

In this section we consider representations V_M for $M \in \mathbb{F}^{2 \times 4}$ for which $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$. With this restriction on M , we can evaluate the coefficients of the polynomials $\{f_i \mid i = 1, 2, 3, 4\}$, as defined in [Section 2](#), at M to get $\{\bar{f}_i \mid i = 1, 2, 3, 4\} \subset \mathbb{F}[V_M]^E$. Note that $\text{LT}(\bar{f}_1) = \gamma_{1234}(M)y^{p^2}$ so that $\text{LM}(\bar{f}_1) = y^{p^2}$. Similarly $\text{LM}(\bar{f}_2) = y^{p^2+2p}$, $\text{LM}(\bar{f}_3) = y^{p^3+2}$ and $\text{LM}(\bar{f}_4) = z^{p^4}$. Also, note that $\gamma_{1357}(M) = 0$ if and only if $\{c_{11}, c_{12}, c_{13}, c_{14}\}$ is linearly dependent over \mathbb{F}_p . Thus, if $\gamma_{1357}(M) \neq 0$, the orbit of z has size p^4 and $\text{LM}(N_M(z)) = z^{p^4}$.

Theorem 3.1. *If $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$, then the set $\mathcal{B} := \{x, \bar{f}_1, \bar{f}_2, \bar{f}_3, N_M(z)\}$ is a SAGBI basis, and hence a generating set, for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with generating relations coming from the subduction of the tête-à-têtes $(\bar{f}_2^p, \bar{f}_1^{p+2})$ and $(\bar{f}_3^p, \bar{f}_2\bar{f}_1^{p^2-1})$.*

Proof. Define $\mathcal{B}' := \{x, \bar{f}_1, \bar{f}_2, \bar{f}_3, \bar{f}_4\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-têtes for \mathcal{B}' are $(\bar{f}_2^p, \bar{f}_1^{p+2})$ and $(\bar{f}_3^p, \bar{f}_2\bar{f}_1^{p^2-1})$. The calculations in the proofs of [Lemmas 2.1](#) and [2.2](#) survive evaluation at M , proving that these tête-à-têtes subduct to zero and \mathcal{B}' is a SAGBI basis for A . Thus, to use [Theorem 1.1](#) to prove $A = \mathbb{F}[V_M]^E$, we need only show that $A[x^{-1}] = \mathbb{F}[V_M]^E[x^{-1}]$.

Consider

$$f_{12357} = \gamma_{1235}y^{p^3} - \gamma_{1237}y^{p^2}x^{p^3-p^2} + \gamma_{1257}y^p x^{p^3-p} + \gamma_{1357}\delta x^{p^3-2} + \gamma_{2357}yx^{p^3-1}$$

and evaluate the coefficients at M to get $\bar{f}_{12357} \in \mathbb{F}[V_M]^E$ with lead monomial y^{p^3} . Since $\gamma_{1357}(M) \neq 0$, we know that \bar{f}_{12357} has degree one as a polynomial in z . Furthermore, the coefficient of z is $-\gamma_{1357}(M)x^{p^3-1}$. Therefore, using [Theorem 2.4](#) of [\[Campbell and Chuai 2007\]](#), $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12357}][x^{-1}]$. Thus, to prove $A = \mathbb{F}[V_M]^E$, it is sufficient to show that $\{N_M(y), \bar{f}_{12357}\} \subset A[x^{-1}]$.

Using [Lemma 1.5](#) for the subsequence (1, 2, 4) of (1, 2, 3, 4, 5, 7) shows that

$$\gamma_{1235}^p \tilde{f}_{12357} = \gamma_{3457} \tilde{f}_{12357} \in \text{Span}_{\mathbb{F}_p} \{\gamma_{2357} \tilde{f}_{13457}, \gamma_{1357} \tilde{f}_{23457}\}.$$

Thus $\bar{f}_{12357} \in \text{Span}_{\mathbb{F}[x, x^{-1}]} \{\bar{f}_{13457}, \bar{f}_{23457}\}$. Similarly, using the (1, 6, 7) subsequence of (1, 3, 4, 5, 6, 7), we have that $\bar{f}_{13457} \in \text{Span}_{\mathbb{F}[x, x^{-1}]} \{\bar{f}_{13456}, \bar{f}_{12345}^p\}$. Iterating this process gives

$$\bar{f}_{12357} \in \text{Span}_{\mathbb{F}[x, x^{-1}]} \{\bar{f}_{12345}, \bar{f}_{12345}^p, \bar{f}_{12346}\}.$$

Since $\bar{f}_{12345} = \bar{f}_1$ and $\bar{f}_{12346} = 2\bar{f}_2x^{p^2-2p} - \bar{f}_1^2$, we see that $\bar{f}_{13457} \in A[x^{-1}]$. A similar argument shows that

$$\bar{f}_{13579} \in \text{Span}_{\mathbb{F}[x, x^{-1}]} \{\bar{f}_{12345}^{p^i}, \bar{f}_{12346}^{p^j} \mid i, j \in \{0, 1, 2\}\},$$

giving $\bar{f}_{13579} \in A[x^{-1}]$. Since $\bar{f}_{13579} = \gamma_{1357}(M)N_M(y)$ (see [Remark 1.2](#)), we have $N_M(y) \in A[x^{-1}]$. Therefore $A = \mathbb{F}[V_M]^E$. As in the proof of [Theorem 2.3](#), observe that $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

Remark 3.2. Lemmas 2.1 and 2.2 are only valid for $p > 5$. However, for the Magma calculations used to verify Theorem 2.3 for $p = 3$ and $p = 5$, only γ_{1234} and γ_{1235} are inverted. Thus Theorem 3.1 remains valid for $p = 3$ and $p = 5$.

4. The $\gamma_{1234} = 0$, $\gamma_{1235} \neq 0$, $\gamma_{1357} \neq 0$ stratum

In this section we consider representations V_M for $M \in \mathbb{F}^{2 \times 4}$ for which $\gamma_{1234}(M) = 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$. Evaluating coefficients gives

$$\bar{f}_1 = \bar{\gamma}_{1235} \delta^p x^{p^2-2p} + \bar{\gamma}_{1245} y^p x^{p^2-p} + \bar{\gamma}_{1345} \delta x^{p^2-2} + \bar{\gamma}_{2345} y x^{p^2-1}.$$

Define

$$h_1 := \frac{\bar{f}_1}{\bar{\gamma}_{1235} x^{p^2-2p}} \quad \text{and} \quad h_2 := \frac{\bar{f}_{12357}}{\bar{\gamma}_{1235}}$$

so that $\text{LT}(h_1) = y^{2p}$ and $\text{LT}(h_2) = y^{p^3}$. Note that $h_1, h_2 \in \mathbb{F}[V_M]^E$. Furthermore, arguing as in the proof of Theorem 3.1, $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), h_2][x^{-1}]$.

Lemma 4.1.
$$N_M(y) = h_2^p + \left(\frac{\bar{\gamma}_{1237}^p}{\bar{\gamma}_{1235}^p} - \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}} \right) h_2 x^{p^4-p^3} - \frac{\bar{\gamma}_{1357}^p}{\bar{\gamma}_{1235}^p} h_1 x^{p^4-2p}.$$

Proof. Since $\bar{f}_{13579} = \bar{\gamma}_{1357} N_M(y)$ (see Remark 1.2), we have

$$N_M(y) = y^{p^4} - \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}} y^{p^3} x^{p^4-p^3} + \frac{\bar{\gamma}_{1379}}{\bar{\gamma}_{1357}} y^{p^2} x^{p^4-p^2} - \frac{\bar{\gamma}_{1579}}{\bar{\gamma}_{1357}} y^p x^{p^4-p} + \frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}} y x^{p^4-1}.$$

Using the definition gives

$$h_2 := y^{p^3} - \frac{\bar{\gamma}_{1237}}{\bar{\gamma}_{1235}} y^{p^2} x^{p^3-p^2} + \frac{\bar{\gamma}_{1257}}{\bar{\gamma}_{1235}} y^p x^{p^3-p} + \frac{\bar{\gamma}_{1357}}{\bar{\gamma}_{1235}} \delta x^{p^3-2} + \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1235}} y x^{p^3-1}.$$

Thus

$$\begin{aligned} N_M(y) - h_2^p &= \left(\frac{\bar{\gamma}_{1237}^p}{\bar{\gamma}_{1235}^p} - \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}} \right) y^{p^3} x^{p^4-p^3} - \left(\frac{\bar{\gamma}_{1257}^p}{\bar{\gamma}_{1235}^p} - \frac{\bar{\gamma}_{1379}}{\bar{\gamma}_{1357}} \right) y^{p^2} x^{p^4-p^2} \\ &\quad - \left(\frac{\bar{\gamma}_{2357}^p}{\bar{\gamma}_{1235}^p} + \frac{\bar{\gamma}_{1579}}{\bar{\gamma}_{1357}} \right) y^p x^{p^4-p} - \frac{\bar{\gamma}_{1357}^p}{\bar{\gamma}_{1235}^p} \delta^p x^{p^4-2p} + \frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}} y x^{p^4-1}. \end{aligned}$$

Using the (1, 3, 5)(3, 4, 5, 7, 9), (1, 3, 7)(3, 4, 5, 7, 9) and (1, 5, 7)(3, 4, 5, 7, 9) Plücker relations gives

$$\begin{aligned} N_M(y) - h_2^p &= \frac{\bar{\gamma}_{1357}^{p-1}}{\bar{\gamma}_{1235}^p} (\bar{\gamma}_{1345} y^{p^3} x^{p^4-p^3} - \bar{\gamma}_{1347} y^{p^2} x^{p^4-p^2} \\ &\quad - \bar{\gamma}_{1457} y^p x^{p^4-p} - \bar{\gamma}_{1357} \delta^p x^{p^4-2p}) + \frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}} y x^{p^4-1}. \end{aligned}$$

Using the $(1, 2, 3)(1, 3, 4, 5, 7)$ and $(1, 2, 5)(1, 3, 4, 5, 7)$ Plücker relations,

$$\bar{\gamma}_{1347} = \frac{\bar{\gamma}_{1237}\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}} \quad \text{and} \quad \bar{\gamma}_{1457} = \frac{\bar{\gamma}_{1245}\bar{\gamma}_{1357} - \bar{\gamma}_{1257}\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}}.$$

Thus

$$\begin{aligned} N_M(y) = h_2^p + \frac{\bar{\gamma}_{1357}^{p-1}}{\bar{\gamma}_{1235}^p} & \left(\bar{\gamma}_{1345} h_2 x^{p^4-p^3} - \frac{\bar{\gamma}_{1357}\bar{\gamma}_{1245}}{\bar{\gamma}_{1235}} y^p x^{p^4-p} - \frac{\bar{\gamma}_{1357}\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}} \delta x^{p^4-2} \right) \\ & - \frac{\bar{\gamma}_{1357}^p}{\bar{\gamma}_{1235}^p} \delta^p x^{p^4-2p} + \left(\frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}} - \frac{\bar{\gamma}_{1357}^{p-1}\bar{\gamma}_{1345}\bar{\gamma}_{2357}}{\bar{\gamma}_{1235}^{p+1}} \right) y x^{p^4-1}. \end{aligned}$$

Using the $(1, 3, 5)(2, 3, 4, 5, 7)$ Plücker relation, $\bar{\gamma}_{1345}\bar{\gamma}_{2357} = \bar{\gamma}_{1357}\bar{\gamma}_{2345} + \bar{\gamma}_{1235}^{p+1}$, giving

$$\frac{\bar{\gamma}_{1357}^{p-1}\bar{\gamma}_{1345}\bar{\gamma}_{2357}}{\bar{\gamma}_{1235}^{p+1}} = \frac{\bar{\gamma}_{2345}\bar{\gamma}_{1357}^p}{\bar{\gamma}_{1235}^p} + \bar{\gamma}_{1357}^{p-1}.$$

From the definition of h_1 ,

$$N_M(y) = h_2^p + \frac{\bar{\gamma}_{1357}^{p-1}\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}^p} h_2 x^{p^4-p^3} - \frac{\bar{\gamma}_{1357}^p}{\bar{\gamma}_{1235}^p} h_1 x^{p^4-2p} + \left(\frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}} - \bar{\gamma}_{1357}^{p-1} \right) y x^{p^4-1}.$$

The result follows from the fact that $\bar{\gamma}_{3579} = \bar{\gamma}_{1357}^p$. \square

As a consequence of the lemma, $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, h_2][x^{-1}]$. Thus applying the SAGBI/divide-by- x algorithm to $\{x, h_1, h_2, N_M(z)\}$ produces a generating set for $\mathbb{F}[V_M]^E$. Subducting the tête-à-tête $(h_2^2, h_1^{p^2})$ gives

$$\tilde{h}_3 := h_2^2 - h_1^{p^2} + 2\frac{\bar{\gamma}_{1237}}{\bar{\gamma}_{1235}} h_1^{p(p+1)/2} x^{p^3-p^2} - 2\frac{\bar{\gamma}_{1257}}{\bar{\gamma}_{1235}} h_1^{(p^2+1)/2} x^{p^3-p}.$$

Lemma 4.2. $\text{LT}(\tilde{h}_3) = \frac{2\bar{\gamma}_{1357}}{\bar{\gamma}_{1235}} y^{p^3+2} x^{p^3-2}.$

Proof. We work modulo the ideal in $\mathbb{F}[x, y, z]$ generated by x^{p^3-1} . Therefore $h_1^{p^2} \equiv y^{2p^3}$, $h_1 x^{p^3-p} \equiv y^{2p} x^{p^3-p}$ and

$$h_2^2 \equiv y^{2p^3} - 2\frac{\bar{\gamma}_{1237}}{\bar{\gamma}_{1235}} y^{p^3+p^2} x^{p^3-p^2} + 2\frac{\bar{\gamma}_{1257}}{\bar{\gamma}_{1235}} y^{p^3+p} x^{p^3-p} + 2\frac{\bar{\gamma}_{1357}}{\bar{\gamma}_{1235}} \delta y^{p^3} x^{p^3-2}.$$

Since $x^{p^3-p^2} h_1^p \equiv x^{p^3-p^2} y^{2p^2}$, we have $(h_1^p)^{(p+1)/2} x^{p^3-p^2} \equiv x^{p^3-p^2} y^{p^3+p^2}$. Thus

$$h_2^2 \equiv h_1^{p^2} - 2\frac{\bar{\gamma}_{1237}}{\bar{\gamma}_{1235}} h_1^{p(p+1)/2} x^{p^3-p^2} + 2\frac{\bar{\gamma}_{1257}}{\bar{\gamma}_{1235}} h_1^{(p^2+1)/2} x^{p^3-p} + 2\frac{\bar{\gamma}_{1357}}{\bar{\gamma}_{1235}} \delta y^{p^3} x^{p^3-2}.$$

Hence $\tilde{h}_3 \equiv 2(\bar{\gamma}_{1357}/\bar{\gamma}_{1235})\delta y^{p^3} x^{p^3-2}$, and the result follows. \square

Define $h_3 := \bar{\gamma}_{1235}\tilde{h}_3/(2\bar{\gamma}_{1357}x^{p^3-2})$ so that $\text{LT}(h_3) = y^{p^3+2}$. Subducting the tête-à-tête $(h_3^p, h_2^p h_1)$ gives

$$\begin{aligned} \tilde{h}_4 := h_3^p - h_1 h_2^p - \alpha_1 x^p h_1^{(p^3+1)/2} + \alpha_2 x^{2p-2} h_3 h_1^{(p^3-p^2)/2} \\ - \alpha_3 x^{2p-1} h_1^{(p^2-1)/2} h_2^{(p-3)/2} h_3^{(p+1)/2}, \end{aligned}$$

with

$$\alpha_1 := \left(\frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1357}} \right)^p - \frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1235}}, \quad \alpha_2 := \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}}, \quad \alpha_3 := \alpha_2 \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1357}} - \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1235}}.$$

Lemma 4.3. $\text{LT}(\tilde{h}_4) = \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1357}} \right)^p x^{2p} z^p.$

Proof. We work modulo the ideal $\mathfrak{n} := \langle x^{2p+1}, x^{2p}y \rangle$. Using the definition of h_3 and methods analogous to the proof of Lemma 4.2, it is not hard to show that

$$h_3 \equiv_{\langle x^3, x^2y \rangle} \delta y^{p^3} + \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1357}} x y^{p^3+1} + \frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1357}} x^2 z^{p^3}.$$

Thus

$$h_3^p \equiv_{\mathfrak{n}} \delta^p y^{p^4} + \left(\frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1357}} \right)^p x^p y^{p^4+p} + \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1357}} \right)^p x^{2p} z^{p^4}.$$

Since $h_2 \equiv_{\mathfrak{n}} y^{p^3}$, we have

$$h_1 h_2^p \equiv_{\mathfrak{n}} y^{p^4} \left(\delta^p + \frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1235}} y^p x^p + \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}} \delta x^{2p-2} + \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1235}} y x^{2p-1} \right).$$

Furthermore, since $x^p h_1 \equiv_{\mathfrak{n}} x^p \delta^p$, expanding gives $x^p h_1^{(p^3+1)/2} \equiv_{\mathfrak{n}} x^p y^{p^4+p}$. Thus

$$\begin{aligned} h_3^p - h_1 h_2^p - \alpha_1 x^p h_1^{(p^3+1)/2} \\ \equiv_{\mathfrak{n}} - \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1235}} \delta x^{2p-2} y^{p^4} - \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1235}} x^{2p-1} y^{p^4+1} + \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1357}} \right)^p x^{2p} z^{p^4}. \end{aligned}$$

Note that $x^{2p-2} h_1^{(p^3-p^2)/2} \equiv_{\mathfrak{n}} x^{2p-2} y^{p^4-p^3}$. Thus

$$x^{2p-2} h_3 h_1^{(p^3-p^2)/2} \equiv_{\mathfrak{n}} x^{2p-2} \left(y^{p^4} \delta + \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1357}} x y^{p^4+1} \right).$$

Hence

$$\begin{aligned} h_3^p - h_1 h_2^p - \alpha_1 x^p h_1^{(p^3+1)/2} + \alpha_2 x^{2p-2} h_3 h_1^{(p^3-p^2)/2} \\ \equiv_{\mathfrak{n}} \alpha_3 x^{2p-1} y^{p^4+1} + \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1357}} \right)^p x^{2p} z^{p^4}. \end{aligned}$$

Since $x^{2p-1} h_1^{(p^2-1)/2} h_2^{(p-3)/2} h_3^{(p+1)/2} \equiv_{\mathfrak{n}} x^{2p-1} y^{p^4+1}$, the result follows. \square

Define $h_4 := \bar{\gamma}_{1357}\tilde{h}_4/(\bar{\gamma}_{1357}x^{2p})$ so that $\text{LT}(h_4) = z^{p^4}$.

Theorem 4.4. *If $\gamma_{1234}(M) = 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$, then the set $\mathcal{B} := \{x, h_1, h_2, h_3, N_M(z)\}$ is a SAGBI basis, and hence a generating set, for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with generating relations coming from the subduction of the tête-à-têtes $(h_2^2, h_1^{p^2})$ and $(h_3^p, h_1 h_2^p)$.*

Proof. Define $\mathcal{B}' := \{x, h_1, h_2, h_3, h_4\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-têtes for \mathcal{B}' are $(h_2^2, h_1^{p^2})$ and $(h_3^p, h_1 h_2^p)$. Using Lemmas 4.2 and 4.3, these tête-à-têtes subduct to zero, proving that \mathcal{B}' is a SAGBI basis for A . Since $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, h_2][x^{-1}]$, using Theorem 1.1, $A = \mathbb{F}[V_M]^E$. Finally, observe that $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. □

5. The $\gamma_{1234} \neq 0$, $\gamma_{1235} = 0$, $\gamma_{1357} \neq 0$ strata

In this section we consider representations V_M for $M \in \mathbb{F}^{2 \times 4}$ for which $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) \neq 0$ and $\gamma_{1357}(M) \neq 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$.

Lemma 5.1. *If $\bar{\gamma}_{1234} \neq 0$, $\bar{\gamma}_{1235} = 0$ and $\bar{\gamma}_{1357} \neq 0$, then $\bar{\gamma}_{1345} \neq 0$.*

Proof. Let r_i denote row i of the matrix $\Gamma(M)$. Since $\bar{\gamma}_{1234} \neq 0$, the set $\{r_1, r_2, r_3, r_4\}$ is linearly independent. Using this and the hypothesis that $\bar{\gamma}_{1235} = 0$, we conclude that r_5 is a linear combination of $\{r_1, r_2, r_3\}$, say $r_5 = a_1 r_1 + a_2 r_2 + a_3 r_3$. Since r_3 is nonzero and the entries of r_5 are the p -th powers of the entries of r_3 , we see that r_5 is nonzero. Suppose, by way of contradiction, that $\bar{\gamma}_{1345} = 0$. Then r_5 is a nonzero linear combination of $\{r_1, r_3, r_4\}$, say $r_5 = b_1 r_1 + b_3 r_3 + b_4 r_4$. Thus $b_1 r_1 + b_3 r_3 + b_4 r_4 = a_1 r_1 + a_2 r_2 + a_3 r_3$. Since $\{r_1, r_2, r_3, r_4\}$ is linearly independent, $b_4 = a_2 = 0$, $a_1 = b_1$, $a_3 = b_3$ and $r_5 = a_1 r_1 + a_3 r_3$, contradicting the assumption that $\bar{\gamma}_{1357} \neq 0$. □

Take f_1 as defined in Section 2, evaluate coefficients and divide by $\bar{\gamma}_{1234}$ to get

$$\hat{f}_1 := y^{p^2} + \frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1234}} y^p x^{p^2-p} + \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1234}} \delta x^{p^2-2} + \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1234}} y x^{p^2-1}.$$

Note that \hat{f}_1 is of degree one in z with coefficient $x^{p^2-2} \bar{\gamma}_{1345} / \bar{\gamma}_{1234}$ and so, using Theorem 2.4 of [Campbell and Chuai 2007], $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \hat{f}_1][x^{-1}]$. Define

$$\tilde{h}_2 := N_M(y) - \hat{f}_1^{p^2} + \alpha_1 \hat{f}_1^p x^{p^4-p^3} + \alpha_2 \hat{f}_1^2 x^{p^4-2p^2},$$

with

$$\alpha_1 := \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}} + \frac{\bar{\gamma}_{1245}^{p^2}}{\bar{\gamma}_{1234}^{p^2}} \quad \text{and} \quad \alpha_2 := \frac{\bar{\gamma}_{1345}^{p^2}}{\bar{\gamma}_{1234}^{p^2}}.$$

We work modulo the ideal $\mathfrak{n} := \langle x^{p^4-p^2-1} \rangle$. Since $\bar{\gamma}_{1357} N_M(y) = \bar{f}_{13579}$ (see Remark 1.2), we have $N_M(y) \equiv_{\mathfrak{n}} y^{p^4} - (\bar{\gamma}_{1359} / \bar{\gamma}_{1357}) y^{p^3} x^{p^4-p^3}$. Therefore

$$N_M(y) - \hat{f}_1^{p^2} \equiv_{\mathfrak{n}} \left(\frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}} + \frac{\bar{\gamma}_{1245}^{p^2}}{\bar{\gamma}_{1234}^{p^2}} \right) y^{p^3} x^{p^4-p^3} - \frac{\bar{\gamma}_{1345}^{p^2}}{\bar{\gamma}_{1234}^{p^2}} \delta p^2 x^{p^4-2p^2}.$$

Thus

$$N_M(y) - \hat{f}_1^{p^2} + \alpha_1 \hat{f}_1^p x^{p^4-p^3} \equiv_n - \frac{\bar{\gamma}_{1345}^{p^2}}{\bar{\gamma}_{1234}^{p^2}} \delta^{p^2} x^{p^4-2p^2} \equiv_n - \frac{\bar{\gamma}_{1345}^{p^2}}{\bar{\gamma}_{1234}^{p^2}} y^{2p^2} x^{p^4-2p^2}.$$

Hence

$$\begin{aligned} \tilde{h}_2 &= N_M(y) - \hat{f}_1^{p^2} + \alpha_1 \hat{f}_1^p x^{p^4-p^3} + \alpha_2 \hat{f}_1^2 x^{p^4-2p^2} \\ &\equiv_n \frac{2\alpha_2}{\bar{\gamma}_{1234}} (\bar{\gamma}_{1245} y^{p^2+p} x^{p^4-p^2-p} + \bar{\gamma}_{1345} y^{p^2+2} x^{p^4-p^2-2}). \end{aligned}$$

We first consider the case $\bar{\gamma}_{1245} \neq 0$. Define

$$h_2 := \bar{\gamma}_{1234}^{p^2+1} \tilde{h}_2 / (2x^{p^4-p^2-p} \bar{\gamma}_{1345}^{p^2} \bar{\gamma}_{1245})$$

so that $\text{LT}(h_2) = y^{p^2+p}$. Since $N_M(y) \in \mathbb{F}[x, \hat{f}_1, h_2]$, we have

$$\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, h_2][x^{-1}].$$

Subducting the tête-à-tête (h_2^p, \hat{f}_1^{p+1}) gives

$$\tilde{h}_3 := \hat{f}_1^{p+1} - h_2^p + \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^p \hat{f}_1^{p-2} h_2^2 x^{p^2-2p}.$$

Lemma 5.2. $\text{LT}(\tilde{h}_3) = 2 \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^{p+1} y^{p^3+p+2} x^{p^2-p-2}.$

Proof. We work modulo the ideal $\langle x^{p^2-p-1} \rangle$. Thus $\hat{f}_1 \equiv y^{p^2}$. Reviewing the definition of h_2 , we see that

$$h_2^p \equiv y^{p^3+p^2} + \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^p y^{p^3+2p} x^{p^2-2p}$$

and

$$h_2^2 x^{p^2-2p} \equiv y^{2p^2+2p} x^{p^2-2p} + 2 \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right) y^{2p^2+p+2} x^{p^2-p-2}.$$

Thus

$$\hat{f}_1^{p+1} - h_2^p + \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^p \hat{f}_1^{p-2} h_2^2 x^{p^2-2p} \equiv 2 \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^{p+1} y^{p^3+p+2} x^{p^2-p-2}$$

and the result follows. \square

Define $h_3 := \bar{\gamma}_{1245}^{p+1} \tilde{h}_3 / (2\bar{\gamma}_{1345}^{p+1} x^{p^2-p-2})$ so that $\text{LT}(h_3) = y^{p^3+p+2}$.

Lemma 5.3. *Subducting the tête-à-tête $(h_3^p, \hat{f}_1^{p^2-1} h_2^2)$ gives an invariant with lead term $-\bar{\gamma}_{1245}^p \bar{\gamma}_{1234}^{p^2} z^{p^4} x^{p^2+2p} / (4\bar{\gamma}_{1345}^{p^2+p})$.*

Proof. Modulo the ideal $\langle x^{p^2+2p+1}, x^{p^2+2p}y \rangle$, the expression

$$\begin{aligned} & h_3^p - \hat{f}_1^{p^2-1} h_2^2 + \beta_1 h_3 \hat{f}_1^{p^2-p+1} x^{p-2} + \beta_2 h_2 \hat{f}_1^{p^2} x^p + \beta_3 \hat{f}_1^{p^2+1} x^{2p} \\ & + \beta_4 h_2^4 \hat{f}_1^{p^2-4} x^{p^2-2p} + \beta_5 h_3 h_2^2 \hat{f}_1^{p^2-p-2} x^{p^2-p-2} \\ & + \beta_6 h_3^2 \hat{f}_1^{p^2-2p} x^{p^2-4} + \beta_7 h_2^2 \hat{f}_1^{p^2-2} x^{p^2} + \beta_8 h_3 \hat{f}_1^{p^2-p} x^{p^2+p-2} \\ & + \beta_9 h_2 \hat{f}_1^{p^2-1} x^{p^2+p} + \beta_{10} h_3 h_2^{p-1} \hat{f}_1^{p^2-2p} x^{p^2+2p-2} \\ & + \beta_{11} h_3^{(p+1)/2} h_2^{(p-3)/2} \hat{f}_1^{(p^2+1)/2-p} x^{p^2+2p-1}, \end{aligned}$$

with

$$\begin{aligned} \beta_1 & := 2 \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}}, & \beta_2 & := - \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1234}} \right)^{p+1} \left(\frac{\bar{\gamma}_{1234}}{\bar{\gamma}_{1345}} \right)^p, \\ \beta_3 & := \frac{1}{2} \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^p \left(\left(\frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1345}} \right)^{p^2} - \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^p \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1234}} \right)^p \right), \\ \beta_4 & := - \frac{1}{2} \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^p, & \beta_5 & := 2 \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^{p+1}, & \beta_6 & := -2 \left(\frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1245}} \right)^{p+2}, \\ \beta_7 & := \frac{1}{2} \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^p \left(\frac{\bar{\gamma}_{1234}}{\bar{\gamma}_{1345}} \right)^{p^2-p} \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}}, & \beta_8 & := - \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1234}} \beta_7, \\ \beta_9 & := - \frac{1}{2} \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^p \left(\frac{\bar{\gamma}_{1234}}{\bar{\gamma}_{1345}} \right)^{p^2} \left(\frac{\bar{\gamma}_{1245} \bar{\gamma}_{1379}}{\bar{\gamma}_{1234} \bar{\gamma}_{1357}} + \frac{\bar{\gamma}_{1579}}{\bar{\gamma}_{1357}} \right), \\ \beta_{10} & := \frac{1}{2} \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^{p-1} \left(\frac{\bar{\gamma}_{1234}}{\bar{\gamma}_{1345}} \right)^{p^2} \frac{\bar{\gamma}_{1379}}{\bar{\gamma}_{1357}}, & \beta_{11} & := \frac{1}{2} \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1345}} \right)^p \left(\frac{\bar{\gamma}_{1234}}{\bar{\gamma}_{1345}} \right)^{p^2} \frac{\bar{\gamma}_{3579}}{\bar{\gamma}_{1357}}, \end{aligned}$$

is congruent to $-\bar{\gamma}_{1245}^p \bar{\gamma}_{1234}^{p^2} z^{p^4} x^{p^2+2p} / (4\bar{\gamma}_{1345}^{p^2+p})$. \square

Theorem 5.4. *If $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1357}(M) \neq 0$ and $\gamma_{1245}(M) \neq 0$, then the set $\mathcal{B} := \{x, \hat{f}_1, h_2, h_3, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with generating relations coming from the subduction of the tête-à-têtes (h_2^p, \hat{f}_1^{p+1}) and $(h_3^p, \hat{f}_1^{p^2-1} h_2^2)$.*

Proof. Use the subduction of $(h_3^p, \hat{f}_1^{p^2-1} h_2^2)$ given in Lemma 5.3 to construct an invariant h_4 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, \hat{f}_1, h_2, h_3, h_4\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-têtes for \mathcal{B}' are (h_2^p, \hat{f}_1^{p+1}) and $(h_3^p, \hat{f}_1^{p^2-1} h_2^2)$. Using Lemmas 5.2 and 5.3, these tête-à-têtes subduct to zero, proving that \mathcal{B}' is a SAGBI basis for A . Since $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, h_2][x^{-1}]$, using Theorem 1.1, $A = \mathbb{F}[V_M]^E$. Finally, observe that $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

We now consider the case $\bar{\gamma}_{1245} = 0$. Define $\hat{h}_2 := \bar{\gamma}_{1234}^{p^2+1} \tilde{h}_2 / (2x^{p^4-p^2-2} \bar{\gamma}_{1345}^{p^2+1})$ so that $\text{LT}(\hat{h}_2) = y^{p^2+2}$. Since $N_M(y) \in \mathbb{F}[x, \hat{f}_1, \hat{h}_2]$, we have $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, \hat{h}_2][x^{-1}]$.

Lemma 5.5. *Subducting the tête-à-tête $(\hat{h}_2^{p^2}, \hat{f}_1^{p^2+2})$ gives an invariant with lead term $z^{p^4} (\bar{\gamma}_{1234} x^2 / (2\bar{\gamma}_{1345}))^{p^2}$.*

Proof. Modulo the ideal $\langle x^{p^2+1}, x^{p^2} y \rangle$, the expression

$$\begin{aligned} & \hat{f}_1^{p^2+2} - \hat{h}_2^{p^2} - (\alpha_1 \hat{h}_2 \hat{f}_1^{p^2} x^{p^2-2} + \alpha_2 \hat{f}_1^{p^2+1} x^{p^2} \\ & \quad + \alpha_3 \hat{h}_2^p \hat{f}_1^{p^2-p} x^{2p^2-2p} + \alpha_4 \hat{h}_2^{p(p+1)/2} \hat{f}_1^{(p^2-p-2)/2} x^{2p^2-p} \\ & \quad + \alpha_5 \hat{h}_2 \hat{f}_1^{p^2-1} x^{2p^2-2} + \alpha_6 \hat{h}_2^{(p^2+1)/2} \hat{f}_1^{(p^2-3)/2} x^{2p^2-1}), \end{aligned}$$

with

$$\begin{aligned} \alpha_1 &:= \frac{2\bar{\gamma}_{1345}}{\bar{\gamma}_{1234}}, & \alpha_2 &:= -\frac{\bar{\gamma}_{1379} \bar{\gamma}_{1234}^{p^2}}{\bar{\gamma}_{1357} \bar{\gamma}_{1345}^{p^2}}, & \alpha_3 &:= -\frac{\bar{\gamma}_{1359} \bar{\gamma}_{1234}^{p^2-p}}{\bar{\gamma}_{1357} \bar{\gamma}_{1345}^{p^2-p}}, \\ \alpha_4 &:= \frac{\bar{\gamma}_{1579} \bar{\gamma}_{1234}^{p^2}}{\bar{\gamma}_{1357} \bar{\gamma}_{1345}^{p^2}}, & \alpha_5 &:= \frac{\bar{\gamma}_{1379} \bar{\gamma}_{1234}^{p^2-1}}{\bar{\gamma}_{1357} \bar{\gamma}_{1345}^{p^2-1}}, & \alpha_6 &:= -\frac{\bar{\gamma}_{3579} \bar{\gamma}_{1234}^{p^2}}{\bar{\gamma}_{1357} \bar{\gamma}_{1345}^{p^2}}, \end{aligned}$$

is congruent to $z^{p^4} (\bar{\gamma}_{1234} x^2 / (2\bar{\gamma}_{1345}))^{p^2}$. \square

Theorem 5.6. *If $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1357}(M) \neq 0$ and $\gamma_{1245}(M) = 0$, then the set $\mathcal{B} := \{x, \hat{f}_1, \hat{h}_2, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a hypersurface with the relation coming from the subduction of the tête-à-tête $(\hat{h}_2^{p^2}, \hat{f}_1^{p^2+2})$.*

Proof. Use the subduction of $(\hat{h}_2^{p^2}, \hat{f}_1^{p^2+2})$ given in Lemma 5.5 to construct an invariant \hat{h}_3 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, \hat{f}_1, \hat{h}_2, \hat{h}_3\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-tête for \mathcal{B}' is $(\hat{h}_2^{p^2}, \hat{f}_1^{p^2+2})$, which subducts to zero using Lemma 5.5. Thus \mathcal{B}' is a SAGBI basis for A . Since $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, \hat{h}_2][x^{-1}]$, using Theorem 1.1, $A = \mathbb{F}[V_M]^E$. Finally, observe that $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

6. The $\gamma_{1234} \neq 0$, $\gamma_{1235} \neq 0$, $\gamma_{1357} = 0$ stratum

In this section we consider representations V_M for $M \in \mathbb{F}^{2 \times 4}$ for which $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) = 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$. Evaluating the coefficients of f_1 and dividing by $\bar{\gamma}_{1234}$ gives \hat{f}_1 with lead term y^{p^2} . Since $\bar{\gamma}_{1357} = 0$ and $\bar{\gamma}_{1235} \neq 0$, the orbit of y has size p^3 and $N_M(y) = \bar{f}_{12357} / \bar{\gamma}_{1235}$ (see Remark 1.2). For convenience, write

$$N_M(y) = y^{p^3} + \alpha_2 y^{p^2} x^{p^3-p^2} + \alpha_1 y^p x^{p^3-p} + \alpha_0 y x^{p^3-1}$$

and

$$\hat{f}_1 = y^{p^2} + \beta_3 \delta^p x^{p^2-2p} + \beta_2 y^p x^{p^2-p} + \beta_1 \delta x^{p^2-2} + \beta_0 y x^{p^2-1},$$

with

$$\alpha_2 = -\frac{\bar{\gamma}_{1237}}{\bar{\gamma}_{1235}}, \quad \alpha_1 = \frac{\bar{\gamma}_{1257}}{\bar{\gamma}_{1235}}, \quad \alpha_0 = \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1235}},$$

$$\beta_3 = \frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1234}}, \quad \beta_2 = \frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1234}}, \quad \beta_1 = \frac{\bar{\gamma}_{1345}}{\bar{\gamma}_{1234}}, \quad \beta_0 = \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1234}}.$$

Subducting $N_M(y)$ gives

$$\tilde{h}_2 := N_M(y) - \hat{f}_1^p + \beta_3^p x^{p^3-2p^2} \hat{f}_1^2.$$

Lemma 6.1.
$$\text{LT}(\tilde{h}_2) = 2 \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1234}} \right)^{p+1} y^{p^2+2p} x^{p^3-p^2-2p}.$$

Proof. We work modulo the ideal $\langle x^{p^3-p^2-p} \rangle$. Using the definitions of f_{12357} and f_{12345} , we have $N_M(y) \equiv y^{p^3}$ and $\hat{f}_1^p \equiv y^{p^3} + (\bar{\gamma}_{1235}/\bar{\gamma}_{1234})^p y^{2p^2} x^{p^3-2p^2}$. The result follows from the observation that

$$\hat{f}_1 x^{p^3-2p^2} \equiv y^{p^2} x^{p^3-2p^2} + \left(\frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1234}} \right) y^{2p} x^{p^3-p^2-2p}. \quad \square$$

Define $h_2 := \tilde{h}_2 \bar{\gamma}_{1234}^{p+1} / (2\bar{\gamma}_{1235}^{p+1} x^{p^3-p^2-2p})$ so that $\text{LT}(h_2) = y^{p^2+2p}$ and

$$h_2 \equiv_{(x^{2p})} y^{p^2} \left(\delta^p + \frac{\beta_2}{\beta_3} y^p x^p + \frac{\beta_1}{\beta_3} \delta x^{2p-2} + \frac{\beta_0}{\beta_3} y x^{2p-1} \right). \quad (1)$$

Lemma 6.2.
$$\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, h_2][x^{-1}].$$

Proof. Since $\bar{\gamma}_{1357} = 0$ and the first row of M is nonzero, we can use a change of coordinates, see [Campbell et al. 2013, §4], and the $\text{GL}_4(\mathbb{F}_p)$ -action to write

$$M = \begin{pmatrix} 1 & c_{12} & c_{13} & 0 \\ 0 & c_{22} & c_{23} & c_{24} \end{pmatrix}.$$

Since $\bar{\gamma}_{1235} \neq 0$, we have $c_{24} \neq 0$. With this choice of generators for E , let H denote the subgroup generated by e_1 and e_4 . Using the calculation of $\mathbb{F}[x, y, z]^H$ from Theorem 6.4 of [loc. cit.], we see that $\mathbb{F}[V_M]^H[x^{-1}] = \mathbb{F}[x, N_H(y), N_H(\delta)][x^{-1}]$ with $N_H(y) := y^p - yx^{p-1}$ and $N_H(\delta) = \delta^p - \delta(c_{24}x^2)^{p-1}$. Thus, to compute $\mathbb{F}[V_M]^G[x^{-1}] = (\mathbb{F}[V_M]^H[x^{-1}])^{G/H}$, it is sufficient to compute

$$(\mathbb{F}[x, N_H(y), N_H(\delta)][x^{-1}])^{G/H} = \mathbb{F}[x, N_H(y)/x^{p-1}, N_H(\delta)/x^{2p-1}]^{G/H}[x^{-1}].$$

Note that $\deg(N_H(y)/x^{p-1}) = \deg(N_H(\delta)/x^{2p-1}) = 1$. Furthermore

$$\mathbb{F}[x, N_H(y)/x^{p-1}]^{G/H} = \mathbb{F}[x, N_{G/H}(N_H(y)/x^{p-1})]$$

and $N_{G/H}(N_H(y)/x^{p-1}) = N_M(y)/x^{p^3-p^2}$. Using the form of M given above, we see that $\bar{\gamma}_{1345} = -c_{24}^{p-1} \bar{\gamma}_{1235}$. If we evaluate $\tilde{\Gamma}$ at M and set $x = 1$, $y = 1$

and $z = 1$, then first and last columns of the resulting matrix are equal. Thus $\tilde{f}_{12345}(1, 1, 1) = \bar{\gamma}_{1234} + \bar{\gamma}_{1245} + \bar{\gamma}_{2345} = 0$. Using these two relations, we can write

$$\hat{f}_1 = N_H(y)^p - \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1234}} N_H(y)x^{p^2-p} + \frac{\bar{\gamma}_{1235}}{\bar{\gamma}_{1234}} N_H(\delta)x^{p^2-2p}.$$

Thus we have $\hat{f}_1/x^{p^2-p} \in \mathbb{F}[x, N_H(y)/x^{p-1}, N_H(\delta)/x^{2p-1}]^{G/H}$ is of degree one in $N_H(\delta)/x^{2p-1}$ with coefficient $x^{p-1}\bar{\gamma}_{1235}/\bar{\gamma}_{1234}$. Thus by Theorem 2.4 of [Campbell and Chuai 2007], we have

$$\mathbb{F}[x, N_H(y)/x^{p-1}, N_H(\delta)/x^{2p-1}]^{G/H}[x^{-1}] = \mathbb{F}[x, N_M(y)/x^{p^3-p^2}, \hat{f}_1/x^{p^2-p}][x^{-1}].$$

Therefore $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \hat{f}_1][x^{-1}]$. The result then follows from the fact that $N_M(y) \in \mathbb{F}[x, \hat{f}_1, h_2]$. \square

Subducting the tête-à-tête (h_2^p, \hat{f}_1^{p+2}) gives

$$\begin{aligned} \tilde{h}_3 := & h_2^p - \hat{f}_1^{p+2} + 2\beta_3 \hat{f}_1^p h_2 x^{p^2-2p} \\ & - \beta_3^{-p} (\alpha_2 \hat{f}_1^{p+1} x^{p^2} - \alpha_2 \beta_3 \hat{f}_1^{p-1} h_2 x^{2p^2-2p} + \alpha_1 \hat{f}_1^{(p-3)/2} h_2^{(p+1)/2} x^{2p^2-p}) \end{aligned}$$

for $p \geq 5$ and

$$\begin{aligned} \tilde{h}_3 := & h_2^3 - \hat{f}_1^5 + 2\beta_3 \hat{f}_1^3 h_2 x^3 \\ & - (\alpha_2 \beta_3^{-3} + \beta_3^3) (\hat{f}_1^4 x^9 - \beta_3 \hat{f}_1^2 h_2 x^{12}) - (\alpha_1 \beta_3^{-3} + \alpha_2 \beta_3^{-1} + \beta_3^5) h_2^2 x^{15} \end{aligned}$$

for $p = 3$.

Lemma 6.3. $\text{LT}(\tilde{h}_3) = \alpha_0 \beta_3^{-p} y^{p^3+1} x^{2p^2-1}$.

Proof. For $p = 3$, this is a Magma calculation. Suppose $p \geq 5$. We work modulo the ideal $\langle x^{2p^2} \rangle$. Since $p^3 - 2p^2 > 2p^2$, we have $\hat{f}_1^p \equiv y^{p^3}$. Furthermore, $3p^2 - 4p > 2p^2$, giving $\hat{f}_1 x^{2p^2-2p} \equiv y^{p^2} x^{2p^2-2}$. Using congruence (1) given above, we have

$$h_2 x^{2p^2-2p} \equiv x^{2p^2-2p} y^{p^2} \left(\delta^p + \frac{\beta_2}{\beta_3} y^p x^p + \frac{\beta_1}{\beta_3} \delta x^{2p-2} + \frac{\beta_0}{\beta_3} y x^{2p-1} \right)$$

and

$$h_2^p \equiv y^{p^3} \left(\delta^p + \frac{\beta_2}{\beta_3} y^p x^p + \frac{\beta_1}{\beta_3} \delta x^{2p-2} + \frac{\beta_0}{\beta_3} y x^{2p-1} \right)^p.$$

Using the definition of h_2 , we get

$$\begin{aligned} \hat{f}_1^2 - 2\beta_3 h_2 x^{p^2-2p} &= \beta_3^{-p} x^{2p^2-p^3} (\hat{f}_1^p - N_M(y)) \\ &= \delta^{p^2} + \beta_3^{-p} ((\beta_2^p - \alpha_2) y^{p^2} x^{p^2} + \beta_1^p \delta^p x^{2p^2-2p} \\ &\quad + (\beta_0^p - \alpha_1) y^p x^{2p^2-p} - \alpha_0 y x^{2p^2-1}). \end{aligned}$$

Thus

$$h_2^p - \hat{f}_1^p (\hat{f}_1^2 - 2\beta_3 h_2 x^{p^2-2p}) \equiv \frac{y^{p^3}}{\beta_3^p} (\alpha_2 y^{p^2} x^{p^2} + \alpha_1 y^p x^{2p^2-p} + \alpha_0 y x^{2p^2-1}).$$

Furthermore, using the above expressions,

$$\hat{f}_1^{p+1} x^{p^2} - \beta_3 \hat{f}_1^{p-1} h_2 x^{2p^2-2p} \equiv y^{p^3-p^2} x^{p^2} (y^{p^2} \hat{f}_1 - \beta_3 h_2 x^{p^2-2p}) \equiv x^{p^2} y^{p^3+p^2}.$$

Therefore

$$\begin{aligned} h_2^p - \hat{f}_1^p (\hat{f}_1^2 - 2\beta_3 h_2 x^{p^2-2p}) - \frac{\alpha_2}{\beta_3^p} (\hat{f}_1^{p+1} x^{p^2} - \beta_3 \hat{f}_1^{p-1} h_2 x^{2p^2-2p}) \\ \equiv \frac{y^{p^3}}{\beta_3^p} (\alpha_1 y^p x^{2p^2-p} + \alpha_0 y x^{2p^2-1}). \end{aligned}$$

Note that $h_2 x^{2p^2-p} \equiv y^{p^2+2p} x^{2p^2-p}$ and $\hat{f}_1 x^{2p^2-p} \equiv y^{p^2} x^{2p^2-p}$. Hence

$$\hat{f}_1^{(p-3)/2} h_2^{(p+1)/2} x^{2p^2-p} \equiv y^{p^3+p} x^{2p^2-p},$$

giving $\tilde{h}_3 \equiv \alpha_0 y^{p^3+1} x^{2p^2-1} / \beta_3^p$, as required. \square

Note that $\alpha_0 / \beta_3^p = \bar{\gamma}_{2357} \bar{\gamma}_{1234}^p / \bar{\gamma}_{1235}^{p+1}$. Since $\bar{\gamma}_{1357} = 0$, $\bar{\gamma}_{1235} \neq 0$ and $\bar{\gamma}_{3457} = \bar{\gamma}_{1235}^p \neq 0$, arguing as in the proof of [Lemma 5.1](#), we see that $\bar{\gamma}_{2357} \neq 0$. Define $h_3 := \bar{\gamma}_{1235}^{p+1} \tilde{h}_3 / (x^{2p^2-1} \bar{\gamma}_{2357} \bar{\gamma}_{1234}^p)$ so that $\text{LT}(h_3) = y^{p^3+1}$.

Lemma 6.4. $\text{LM}(h_3^p - h_2^{(p^2+1)/2} \hat{f}_1^{(p^2-2p-1)/2}) = x^p z^{p^4}$.

Proof. Working modulo the ideal $\mathfrak{n} := \langle x^{p+1}, x^p y \rangle$, we see that $\hat{f}_1 \equiv_{\mathfrak{n}} y^{p^2}$ and $h_2 \equiv_{\mathfrak{n}} y^{p^2+2p}$, giving $h_3^p - h_2^{(p^2+1)/2} \hat{f}_1^{(p^2-2p-1)/2} \equiv_{\mathfrak{n}} h_3^p - y^{p^4+p}$. Thus it is sufficient to identify the lead monomial of $h_3 - y^{p^3+1}$. Note that y^{p^3+1} and xz^{p^3} are consecutive monomials in the grevlex term order. Therefore, if xz^{p^3} appears with nonzero coefficient in h_3 , then $\text{LM}(h_3 - y^{p^3+1}) = xz^{p^3}$, and the result follows. Work modulo the ideal $\mathfrak{m} := \langle y \rangle$. Then $\hat{f}_1 \equiv_{\mathfrak{m}} -\beta_3 z^p x^{p^2-p} - \beta_1 z x^{p^2-1}$ and $N_M(y) \equiv_{\mathfrak{m}} 0$. Therefore

$$h_2 \equiv_{\mathfrak{m}} \frac{1}{2\beta_3} \left(z^{p^2} x^{2p} + \frac{\beta_1^p}{\beta_3^p} z^p x^{p^2+p} + x^{p^2} (\beta_3 z^p + \beta_1 z x^{p-1})^2 \right).$$

Hence h_3 has degree p^3 as a polynomial in z , with leading coefficient $x/2\alpha_0$ and the result follows. \square

Theorem 6.5. *If $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) = 0$, then the set $\mathcal{B} := \{x, \hat{f}_1, h_2, h_3, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with generating relations coming from the subduction of the tête-à-têtes (h_2^p, \bar{f}_1^{p+2}) and $(h_3^p, \bar{f}_1^{(p^2-2p-1)/2} h_2^{(p^2+1)/2})$.*

Proof. Use the subduction given in [Lemma 6.4](#) to construct an invariant h_4 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, \hat{f}_1, h_2, h_3, h_4\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-têtes for \mathcal{B}' are

$$(h_2^p, \bar{f}_1^{p+2}) \quad \text{and} \quad (h_3^p, \bar{f}_1^{(p^2-2p-1)/2} h_2^{(p^2+1)/2}).$$

Using [Lemmas 6.3](#) and [6.4](#), these tête-à-têtes subduct to zero, proving that \mathcal{B}' is a SAGBI basis for A . By [Lemma 6.2](#), we have $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, \hat{f}_1, h_2][x^{-1}]$.

Using [Theorem 1.1](#), $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since \mathcal{B}' is a SAGBI basis for $\mathbb{F}[V_E]^E$, this forces $k = 4$, giving $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

7. The $\gamma_{1234} = 0$, $\gamma_{1235} = 0$, $\gamma_{1357} \neq 0$ strata

In this section we consider representations V_M for $M \in \mathbb{F}^{2 \times 4}$ for which $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) = 0$ and $\gamma_{1357}(M) \neq 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$.

We first consider the case $\bar{\gamma}_{1257} = 0$. Let r_i denote row i of the matrix $\Gamma(M)$. Since $\gamma_{1357}(M) \neq 0$, the set $\{r_1, r_3, r_5, r_7\}$ is linearly independent. Thus r_2 is a linear combination of r_1, r_5 and r_7 . Since $\bar{\gamma}_{1235} = 0$, we know that r_2 is a linear combination of r_1, r_3 and r_5 . Using the $(1, 2, 3)(3, 4, 5, 7, 9)$ Plücker relation, $\bar{\gamma}_{1237} = 0$. Thus r_2 is a linear combination of r_1, r_3 and r_7 . Combining these observations, we see that r_2 is a scalar multiple of r_1 . Using a change of coordinates (see Section 4 of [[Campbell et al. 2013](#)]), we may assume that r_2 is zero. If the second row of M is zero, then V_M is a symmetric square representation and the invariants are generated by $x, \delta, N_M(y)$ and $N_M(z)$. Since $\bar{\gamma}_{1357} \neq 0$, we have that $N_M(y)$ and $N_M(z)$ are both of degree p^4 and there is a single relation in degree $2p^4$ which can be constructed by subducting the tête-à-tête $(\delta^{p^4}, N_M(y)^2)$ (see [Theorem 3.3](#) of [[loc. cit.](#)]).

For the rest of this section, we assume $\bar{\gamma}_{1257} \neq 0$. Evaluating coefficients gives the invariant \bar{f}_{12357} . Using the $(1, 2, 3)(3, 4, 5, 7, 9)$ Plücker relation, $\bar{\gamma}_{1237}^{p+1} = 0$. Thus $\bar{\gamma}_{1237} = 0$, and we have $\bar{f}_{12357} = \bar{\gamma}_{1257}y^p x^{p^3-p} + \bar{\gamma}_{1357}\delta x^{p^3-2} + \bar{\gamma}_{2357}yx^{p^3-1}$. Divide by $\bar{\gamma}_{1257}x^{p^3-p}$ to get

$$h_1 := y^p + \frac{\bar{\gamma}_{1357}}{\bar{\gamma}_{1257}}\delta x^{p-2} + \frac{\bar{\gamma}_{2357}}{\bar{\gamma}_{1257}}yx^{p-1}.$$

Observe that $N_M(y) = \bar{f}_{13579}/\bar{\gamma}_{1357}$. Subducting $N_M(y)$ gives

$$\begin{aligned} \tilde{h}_2 = N_M(y) - h_1^{p^3} + \alpha^{p^3} h_1^{2p^2} x^{p^4-2p^3} - 2\alpha^{p^3+p^2} h_1^{p^2+2p} x^{p^4-p^3-2p^2} \\ + 4\alpha^{p^3+p^2+p} h_1^{p^2+p+2} x^{p^4-p^3-p^2-2p}, \end{aligned}$$

with $\alpha := \bar{\gamma}_{1357}/\bar{\gamma}_{1257}$.

Lemma 7.1. $\text{LT}(\tilde{h}_2) = 8\alpha^{p^3+p^2+p+1}y^{p^3+p^2+p+2}x^{p^4-p^3-p^2-p-2}$.

Proof. It will be convenient to work modulo the ideal $\langle x^{p^4-p^3}, x^{p^4-p^3-p^2-p-1}y \rangle$, so that $N_M(y) \equiv y^{p^4}$ and $h_1^{p^3} \equiv y^{p^4} + \alpha^{p^3}\delta^{p^3}x^{p^4-2p^3}$. Thus $N_M(y) - h_1^{p^3} \equiv -\alpha^{p^3}\delta^{p^3}x^{p^4-2p^3}$. Expanding gives

$$x^{p^4-2p^3}(h_1^{p^2})^2 \equiv x^{p^4-2p^3}y^{p^3}(y^{p^3} + 2\alpha^{p^2}\delta^{p^2}x^{p^3-2p^2}).$$

Thus

$$N_M(y) - h_1^{p^3} + \alpha^{p^3}h_1^{2p^2}x^{p^4-2p^3} \equiv 2\alpha^{p^3+p^2}y^{p^3}\delta^{p^2}x^{p^4-p^3-2p^2}.$$

Again expanding gives

$$h_1^{p^2+2p} x^{p^4-p^3-2p^2} \equiv x^{p^4-p^3-2p^2} y^{p^3+p^2} (y^{p^2} + 2\alpha^p \delta^p x^{p^2-2p}).$$

Hence

$$\begin{aligned} N_M(y) - h_1^{p^3} + \alpha^{p^3} h_1^{2p^2} x^{p^4-2p^3} - 2\alpha^{p^3+p^2} h_1^{p^2+2p} x^{p^4-p^3-2p^2} \\ \equiv -4\alpha^{p^3+p^2+p} \delta^p y^{p^3+p^2} x^{p^4-p^3-p^2-2p}. \end{aligned}$$

Since $h_1^{p^2+p+2} x^{p^4-p^3-p^2-2p} \equiv x^{p^4-p^3-p^2-2p} y^{p^3+p^2+p} (y^p + 2\alpha \delta x^{p-2})$, we have

$$\tilde{h}_2 \equiv 8\alpha^{p^3+p^2+p+1} y^{p^3+p^2+p+2} x^{p^4-p^3-p^2-p-2}$$

and the result follows. \square

Define $h_2 := \tilde{h}_2 / (8\alpha^{p^3+p^2+p+1} x^{p^4-p^3-p^2-p-2})$ so that $\text{LT}(h_2) = y^{p^3+p^2+p+2}$.

Lemma 7.2. *Subducting the tête-à-tête $(h_2^p, h_1^{p^3+p^2+p+2})$ gives an invariant with lead term*

$$\left(\frac{\bar{\gamma}_{1257}}{2\bar{\gamma}_{1357}} \right)^{p^3+p^2+p} z^{p^4} x^{p^3+p^2+2p}.$$

Proof. For $p = 3$, this is a Magma calculation. For $p > 3$, the subduction is given by

$$\begin{aligned} h_2^p - h_1^{p^3+p^2+p+2} + 2\alpha h_2 h_1^{p^3} x^{p-2} \\ + \frac{1}{4\alpha^{p^3+p^2+p}} (\beta_1 h_1^{p^3+p^2} x^{p^2+2p} - \beta_1 \alpha^{p^2} h_1^{p^3+2p} x^{p^3-p^2+2p} \\ + 2\beta_1 \alpha^{p^2+p} h_1^{p^3+p+2} x^{p^3} - 4\beta_1 \alpha^{p^2+p+1} h_2 h_1^{p^3-p^2} x^{p^3+p-2} \\ - \beta_2 x^{p^3} (h_1^{p^3+p} x^{2p} - \alpha^p h_1^{p^3+2} x^{p^2} + 2\alpha^{p+1} h_2 h_1^{p^3-p^2-p} x^{p^2+p-2}) \\ + \beta_3 x^{p^3+p^2+p} (h_1^{p^3+1} - \alpha h_2 h_2^{p^3-p^2-p-1} x^{p-2}) \\ - \beta_4 h_2^{(p+1)/2} h_1^{(p^2+p+1)(p-3)/2} x^{p^3+p^2+2p-1}), \end{aligned}$$

with

$$\alpha := \frac{\bar{\gamma}_{1357}}{\bar{\gamma}_{1257}}, \quad \beta_1 := \frac{\bar{\gamma}_{1359}}{\bar{\gamma}_{1357}}, \quad \beta_2 := \frac{\bar{\gamma}_{1379}}{\bar{\gamma}_{1357}}, \quad \beta_3 := \frac{\bar{\gamma}_{1579}}{\bar{\gamma}_{1357}}, \quad \beta_4 := \bar{\gamma}_{1357}^{-1}.$$

To calculate the lead term, work modulo the ideal generated by $x^{p^3+p^2+2p+1}$ and $x^{p^3+p^2+2p} y$. \square

Theorem 7.3. *If $\gamma_{1234}(M) = 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1357}(M) = 0$ and $\gamma_{1257}(M) \neq 0$, then the set $\mathcal{B} := \{x, h_1, h_2, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a hypersurface with the relation coming from the subduction of the tête-à-tête $(h_2^p, h_1^{p^3+p^2+p+2})$.*

Proof. Use the subduction given in [Lemma 7.2](#) to construct an invariant h_3 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, h_1, h_2, h_3\}$ and let A denote the algebra generated by \mathcal{B}' . The only nontrivial tête-à-tête for \mathcal{B}' is $(h_2^p, h_1^{p^3+p^2+p+2})$, which subducts to zero using the definition of h_3 . Thus \mathcal{B}' is a SAGBI basis for A . Since h_1 is of degree one in z with coefficient $-\alpha x^{p-1}$, it follows from [\[Campbell and Chuai 2007\]](#) that $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, N_M(y)][x^{-1}]$. Since $N_M(y) \in \mathbb{F}[x, h_1, h_2]$, we have $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, h_2][x^{-1}]$. Using [Theorem 1.1](#), $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since \mathcal{B}' is a SAGBI basis for $\mathbb{F}[V_E]^E$, this forces $k = 4$, giving $\text{LM}(\mathcal{B}) \subset \text{LM}(\mathcal{B}')$. \square

8. The $\gamma_{1234} = 0$, $\gamma_{1235} \neq 0$, $\gamma_{1357} = 0$ stratum

In this section we consider representations V_M with $\gamma_{1235}(M) \neq 0$, $\gamma_{1234}(M) = 0$ and $\gamma_{1357}(M) = 0$. The results of this section are valid for $p \geq 3$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$. Observe that $N_M(y) = \bar{f}_{12357}/\bar{\gamma}_{1235}$ (see [Remark 1.2](#)). Thus $N_M(y)$ has lead term y^{p^3} . Furthermore, \bar{f}_{12345} has lead term $\bar{\gamma}_{1235}y^{2p}x^{p^2-2p}$. Define $h_1 := \bar{f}_{12345}/(\bar{\gamma}_{1235}x^{p^2-2p})$ so that $\text{LT}(h_1) = y^{2p}$.

Lemma 8.1. $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, N_M(y)][x^{-1}]$.

Proof. We argue as in the proof of [Theorem 4.4](#) of [\[Campbell et al. 2013\]](#). Since $N_M(y)$ and h_1/x^p are algebraically independent elements of $\mathbb{F}[x, y, \delta/x]^E$ with $\deg(N_M(y)) \deg(h_1/x^p) = p^4 = |E|$, applying [Theorem 3.7.5](#) of [\[Derksen and Kemper 2002\]](#) gives $\mathbb{F}[x, y, \delta/x]^E = \mathbb{F}[x, N_M(y), h_1/x^p]$. The result then follows from the observation that

$$\mathbb{F}[x, y, z]^E[x^{-1}] = \mathbb{F}[x, y, \delta/x]^E[x^{-1}]. \quad \square$$

Subducting the tête-à-tête $(N_M(y)^2, h_1^{p^2})$ gives

$$\tilde{h}_2 := N_M(y)^2 - h_1^{p^2} + \frac{2}{\bar{\gamma}_{1235}}(\bar{\gamma}_{1237}x^{p^3-p^2}h_1^{(p^2+p)/2} - \bar{\gamma}_{1257}x^{p^3-p}h_1^{(p^2+1)/2}).$$

Lemma 8.2. $\text{LT}(\tilde{h}_2) = 2\bar{\gamma}_{2357}y^{p^3+1}x^{p^3-1}/\bar{\gamma}_{1235}$.

Proof. We work modulo the ideal $\langle x^{p^3} \rangle$. Expand $N_M(y)^2$ and observe that $h_1^{p^2} \equiv y^{2p^3}$, $h_1^p x^{p^3-p^2} \equiv y^{2p^2}x^{p^3-p^2}$ and $h_1 x^{p^3-p} \equiv y^{2p}x^{p^3-p}$. \square

Using the $(1, 3, 5)(2, 3, 4, 5, 7)$ Plücker relation, we have $\bar{\gamma}_{1345}\bar{\gamma}_{2357} = \bar{\gamma}_{1235}^{p+1}$. Thus $\bar{\gamma}_{2357} \neq 0$. Define $h_2 := \bar{\gamma}_{1235}\tilde{h}_2/(2\bar{\gamma}_{2357}x^{p^3-1})$ so that $\text{LT}(h_2) = y^{p^3+1}$.

Lemma 8.3. $\text{LM}(h_2^p - h_1^{(p^3+1)/2}) = z^{p^4}x^p$.

Proof. A careful calculation shows that

$$\text{LT}(h_2^p - h_1^{(p^3+1)/2}) = \frac{\bar{\gamma}_{1235}^p}{2\bar{\gamma}_{2357}^p}x^p z^{p^4}. \quad \square$$

Theorem 8.4. *If $\gamma_{1234}(M) = 0$, $\gamma_{1235}(M) \neq 0$ and $\gamma_{1357}(M) = 0$, then the set $\mathcal{B} := \{x, h_1, h_2, N_M(y), N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with relations coming from the subduction of the tête-à-têtes $(N_M(y)^2, h_1^{p^2})$ and $(h_2^p, h_1^{(p^3+1)/2})$.*

Proof. Use the subduction from [Lemma 8.3](#) to construct an invariant h_3 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, N_M(y), h_1, h_2, h_3\}$ and let A denote the algebra generated by \mathcal{B}' . The nontrivial tête-à-têtes for \mathcal{B}' subduct to zero using [Lemmas 8.2](#) and [8.3](#). Thus \mathcal{B}' is a SAGBI basis for A . From [Lemma 8.1](#), $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, h_1, N_M(y)][x^{-1}]$. Thus, using [Theorem 1.1](#), $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since \mathcal{B}' is a SAGBI basis for $\mathbb{F}[V_M]^E$, this forces $k = 4$, giving $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. □

9. The $\gamma_{1234} \neq 0$, $\gamma_{1235} = 0$, $\gamma_{1357} = 0$ strata

In this section we consider representations V_M with $\gamma_{1235}(M) = 0$, $\gamma_{1234}(M) \neq 0$ and $\gamma_{1357}(M) = 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$. Using the $(1, 3, 5)(3, 4, 5, 6, 7)$ Plücker relation, $\bar{\gamma}_{1345} = 0$. Thus

$$\bar{f}_1 = \bar{\gamma}_{1234}y^{p^2} + \bar{\gamma}_{1245}y^p x^{p^2-p} + \bar{\gamma}_{2345}yx^{p^2-1} \in \mathbb{F}[x, y].$$

Since $\bar{\gamma}_{1234} \neq 0$, the orbit of y contains at least p^2 elements. Thus $N_M(y) = \bar{f}_1/\bar{\gamma}_{1234}$ (see [Remark 1.2](#)).

Lemma 9.1. $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12346}][x^{-1}]$.

Proof. We argue as in the proof of [Lemma 8.1](#) (and [Theorem 4.4](#) of [\[Campbell et al. 2013\]](#)). Since $N_M(y)$ and \bar{f}_{12346}/x^{p^2} are algebraically independent elements of $\mathbb{F}[x, y, \delta/x]^E$ with $\deg(N_M(y)) \deg(\bar{f}_{12346}/x^{p^2}) = p^4 = |E|$, applying [Theorem 3.7.5](#) of [\[Derksen and Kemper 2002\]](#) gives

$$\mathbb{F}[x, y, \delta/x]^E = \mathbb{F}[x, N_M(y), \bar{f}_{12346}/x^{p^2}].$$

The result then follows from the observation that

$$\mathbb{F}[x, y, z]^E[x^{-1}] = \mathbb{F}[x, y, \delta/x]^E[x^{-1}]. \quad \square$$

We first consider the case $\bar{\gamma}_{1245} \neq 0$. Define $\hat{f}_2 := \bar{f}_2/(\bar{\gamma}_{1234}\bar{\gamma}_{1245}x^p)$ so that $\text{LT}(\hat{f}_2) = y^{p^2+p}$. Subduct the tête-à-tête $(\hat{f}_2^p, N_M(y)^{p+1})$ to get

$$\tilde{h}_3 := N_M(y)^{p+1} - \hat{f}_2^p - \left(\frac{\bar{\gamma}_{1245}}{\bar{\gamma}_{1234}} - \frac{\bar{\gamma}_{2345}^p}{\bar{\gamma}_{1245}^p} \right) \hat{f}_2 N_M(y)^{p-1} x^{p^2-p}.$$

Lemma 9.2. $\text{LT}(\tilde{h}_3) = \left(\frac{\bar{\gamma}_{2345}^{p+1}}{\bar{\gamma}_{1245}^{p+1}} \right) x^{p^2-1} y^{p^3+1}$.

Proof. Expand and reduce modulo the ideal $\langle x^{p^2} \rangle$. □

Define

$$h_3 := \frac{\bar{\gamma}_{1245}^{p+1}}{x^{p^2-1}\bar{\gamma}_{2345}^{p+1}}\tilde{h}_3$$

so that $\text{LT}(h_3) = y^{p^3+1}$.

Lemma 9.3. *Subducting the tête-à-tête $(h_3^p, N_M(y)^{p^2-1}\hat{f}_2)$ gives an invariant with lead monomial $x^p z^{p^4}$.*

Proof. Work modulo the ideal $\langle x^{p+1}, x^p y \rangle$ and expand to get

$$h_3^p - \hat{f}_2 N_M(y)^{p^2-1} + \frac{\bar{\gamma}_{2345}}{\bar{\gamma}_{1234}} x^{p-1} h_3 N_M(y)^{p^2-p} \equiv \left(\frac{\bar{\gamma}_{1234}^{p^2} \bar{\gamma}_{1245}^p}{\bar{\gamma}_{2345}^{p^2+p}} \right) z^{p^4} x^p. \quad \square$$

Theorem 9.4. *If $\gamma_{1234}(M) \neq 0$, $\gamma_{1235}(M) = \gamma_{1357}(M) = 0$ and $\gamma_{1245}(M) \neq 0$, then the set $\mathcal{B} := \{x, N_M(y), \hat{f}_2, h_3, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a complete intersection with relations coming from the subduction of the tête-à-têtes $(\hat{f}_2^p, N_M(y)^{p+1})$ and $(h_3^p, N_M(y)^{p^2-1}\hat{f}_2)$.*

Proof. Use the subduction given in Lemma 9.3 to construct an invariant h_4 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, N_M(y), \hat{f}_2, h_3, h_4\}$ and let A denote the algebra generated by \mathcal{B}' . The nontrivial tête-à-têtes for \mathcal{B}' subduct to zero using Lemmas 9.2 and 9.3. Thus \mathcal{B}' is a SAGBI basis for A . From Lemma 9.1, $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12346}][x^{-1}]$. However, since $f_2 = (f_1^2 + \gamma_{1234} f_{12346}) / (2x^{p^2-2p})$, we see that

$$\mathbb{F}[x, N_M(y), \bar{f}_{12346}][x^{-1}] = \mathbb{F}[x, N_M(y), \hat{f}_2][x^{-1}].$$

Thus, using Theorem 1.1, $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since \mathcal{B}' is a SAGBI basis for $\mathbb{F}[V_M]^E$, this forces $k = 4$, giving $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

Suppose $\bar{\gamma}_{1245} = 0$ and let r_i denote row i of the matrix $\Gamma(M)$. Since $\bar{\gamma}_{1234} \neq 0$, we see that $\{r_1, r_2, r_3, r_4\}$ is linearly independent. Using the assumptions that $\bar{\gamma}_{1235} = \bar{\gamma}_{1245} = 0$, we see that $r_5 \in \text{Span}(r_1, r_2, r_3) \cap \text{Span}(r_1, r_2, r_4)$. Therefore $r_5 \in \text{Span}(r_1, r_2)$. However, since $\bar{\gamma}_{1357} = 0$, using a change of coordinates (see [Campbell et al. 2013, §4]) and the $\text{GL}_4(\mathbb{F}_p)$ -action, we may assume

$$M := \begin{pmatrix} 1 & c_{12} & c_{13} & 0 \\ 0 & c_{22} & c_{23} & c_{24} \end{pmatrix}$$

with $c_{24} \neq 0$. Since $r_5 = r_1^{p^2}$, we conclude that $r_5 = r_1$. Thus $\bar{\gamma}_{2345} = -\bar{\gamma}_{1234}$. Hence $N_M(y) = \bar{f}_1 / \bar{\gamma}_{1234} = y^{p^2} - yx^{p^2-1}$. Define $\hat{h}_2 := -\bar{f}_2 / (\bar{\gamma}_{1234}^2 x^{2p-1})$ so that $\text{LT}(\hat{h}_2) = y^{p^2+1}$.

Theorem 9.5. *If $\gamma_{1234}(M) \neq 0$ and $\gamma_{1235}(M) = \gamma_{1357}(M) = \gamma_{1245}(M) = 0$, then the set $\mathcal{B} := \{x, N_M(y), \hat{h}_2, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a hypersurface with the relation coming from the subduction of the tête-à-tête $(\hat{h}_2^{p^2}, N_M(y)^{p^2+1})$.*

Proof. Using the definition of \hat{h}_2 and the description given above of $N_M(y)$, we see that

$$\text{LT}(\hat{h}_2^{p^2} - N_M(y)^{p^2+1} - \hat{h}_2(xN_M(y))^{p^2-1}) = -\frac{1}{2}z^{p^4}x^{p^2}.$$

Thus we can use the subduction of the tête-à-tête $(\hat{h}_2^{p^2}, N_M(y)^{p^2+1})$ to construct an invariant h_4 with lead term z^{p^4} . Define $B' := \{x, N_M(y), \hat{h}_2, h_4\}$ and let A denote the algebra generated by B' . The only nontrivial tête-à-tête subducts to zero. Therefore B' is a SAGBI basis for A . From [Lemma 9.1](#), $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12346}][x^{-1}]$. However, it follows from the definition of \hat{h}_2 that $\mathbb{F}[x, N_M(y), \bar{f}_{12346}][x^{-1}] = \mathbb{F}[x, N_M(y), \hat{h}_2][x^{-1}]$. Thus, using [Theorem 1.1](#), $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since B' is a SAGBI basis for $\mathbb{F}[V_E]^E$, this forces $k = 4$, giving $\text{LM}(B) = \text{LM}(B')$. \square

10. The $\gamma_{1234} = 0, \gamma_{1235} = 0, \gamma_{1357} = 0$ strata

In this section we consider representations V_M with $\gamma_{1235}(M) = 0, \gamma_{1234}(M) = 0$ and $\gamma_{1357}(M) = 0$. For convenience, we write $\bar{\gamma}_{ijkl}$ for $\gamma_{ijkl}(M)$. We assume that the first row of M is nonzero; otherwise, the representation is of type $(2, 1)$ and the calculation of $\mathbb{F}[V_M]^E$ can be found in Section 4 of [\[Campbell et al. 2013\]](#). Using a change of coordinates, see Proposition 4.3 of [\[loc. cit.\]](#), the $\text{GL}_4(\mathbb{F}_p)$ -action, and the hypothesis that $\bar{\gamma}_{1357} = 0$, we may take

$$M = \begin{pmatrix} 1 & c_{12} & c_{13} & 0 \\ 0 & c_{22} & c_{23} & c_{24} \end{pmatrix}.$$

Since $\bar{\gamma}_{1235} = 0$, either $c_{24} = 0$ or $\{1, c_{12}, c_{13}\}$ is linearly dependent over \mathbb{F}_p . We assume $c_{24} \neq 0$; otherwise the representation is not faithful and we can view V_M as a representation of a group of rank three. Using the $\text{GL}_4(\mathbb{F}_p)$ -action, we replace the third column by a linear combination of the first two columns to get

$$\begin{pmatrix} 1 & c_{12} & 0 & 0 \\ 0 & c_{22} & c_{23} & c_{24} \end{pmatrix}.$$

Expanding gives

$$\bar{\gamma}_{1234} = (c_{12} - c_{12}^p) \det \begin{pmatrix} c_{23} & c_{24} \\ c_{23}^p & c_{24}^p \end{pmatrix}.$$

Since $\bar{\gamma}_{1234} = 0$, either $c_{12} \in \mathbb{F}_p$ or $\{c_{23}, c_{24}\}$ is linearly dependent over \mathbb{F}_p . However, if $\{c_{23}, c_{24}\}$ is linearly dependent over \mathbb{F}_p , then the representation is not faithful. So we may assume $c_{12} \in \mathbb{F}_p$. Using the $\text{GL}_4(\mathbb{F}_p)$ -action to replace the second column with a linear combination of the first two columns gives

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_{22} & c_{23} & c_{24} \end{pmatrix}.$$

If $\bar{\gamma}_{1246} = 0$, then $\{c_{22}, c_{23}, c_{24}\}$ is linearly dependent over \mathbb{F}_p , and again the representation is not faithful. Thus we may assume that $\bar{\gamma}_{1246} \neq 0$. Using the above form for M , it is clear that $\bar{\gamma}_{1236} = 0$, $\bar{\gamma}_{1346} = 0$ and $\bar{\gamma}_{1246} = -\bar{\gamma}_{2346}$. Thus

$$\bar{f}_{12346} = \bar{\gamma}_{1246}(y^p x^{2p^2-p} - yx^{2p^2-1}) \in \mathbb{F}[x, y]^E.$$

Since $\mathbb{F}[x, y]^E = \mathbb{F}[x, N_M(y)]$, we have

$$N_M(y) = \bar{f}_{12346}/(\bar{\gamma}_{1246}x^{2p^2-p}) = y^p - yx^{p-1}.$$

Lemma 10.1. $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12468}][x^{-1}]$.

Proof. The proof is similar to the proof of Theorem 4.4 of [Campbell et al. 2013] (and Lemmas 8.1 and 9.1). Since $N_M(y)$ and \bar{f}_{12468}/x^{p^3} are algebraically independent elements of $\mathbb{F}[x, y, \delta/x]^E$ with $\deg(N_M(y)) \deg(\bar{f}_{12468}/x^{p^3}) = p^4 = |E|$, applying Theorem 3.7.5 of [Derksen and Kemper 2002] gives

$$\mathbb{F}[x, y, \delta/x]^E = \mathbb{F}[x, N_M(y), \bar{f}_{12468}/x^{p^3}].$$

The result then follows from the observation that

$$\mathbb{F}[x, y, z]^E[x^{-1}] = \mathbb{F}[x, y, \delta/x]^E[x^{-1}]. \quad \square$$

Subducting \bar{f}_{12468} gives

$$\tilde{h}_1 := \bar{f}_{12468} + \bar{\gamma}_{1246}(N_M(y)^{2p^2} + 2N_M(y)^{p^2+p}x^{p^3-p^2} + 2N_M(y)^{p^2+1}x^{p^3-p}).$$

Lemma 10.2. $\text{LT}(\tilde{h}_1) = -2\bar{\gamma}_{1246}x^{p^3-1}y^{p^3+1}$.

Proof. We work modulo the ideal $\langle x^{p^3} \rangle$. Using the definition, $\bar{f}_{12468} \equiv -\bar{\gamma}_{1246}y^{2p^3}$. Since $N_M(y) = y^p - yx^{p-1}$, we have

$$N_M(y)^{2p^2} = y^{2p^3} - 2y^{p^3+p^2}x^{p^3-p^2} + y^{2p^2}x^{2p^3-2p^2} \equiv y^{2p^3} - 2y^{p^3+p^2}x^{p^3-p^2}.$$

Expanding and simplifying gives

$$N_M(y)^{p^2+p}x^{p^3-p^2} + N_M(y)^{p^2+1}x^{p^3-p} \equiv y^{p^3+p^2}x^{p^3-p^2} - y^{p^3+1}x^{p^3-1}.$$

Thus

$$\begin{aligned} \tilde{h}_1 &= \bar{f}_{12468} + \bar{\gamma}_{1246}(N_M(y)^{2p^2} + 2N_M(y)^{p^2+p}x^{p^3-p^2} + 2N_M(y)^{p^2+1}x^{p^3-p}) \\ &\equiv -2\bar{\gamma}_{1246}x^{p^3-1}y^{p^3+1}. \end{aligned} \quad \square$$

Define $h_1 := -\tilde{h}_1/(2\bar{\gamma}_{1246}x^{p^3-1})$ so that $\text{LT}(h_1) = y^{p^3+1}$. Note that

$$\mathbb{F}[x, N_M(y), h_1][x^{-1}] = \mathbb{F}[x, N_M(y), \bar{f}_{12468}][x^{-1}].$$

Lemma 10.3. *Subducting the tête-à-tête $(h_1^p, N_M(y)^{p^3+1})$ gives an invariant with lead monomial $x^p z^{p^4}$.*

Proof. Refining the calculation in the proof of the previous lemma gives

$$\tilde{h}_1 \equiv_{\langle x^{p^3+1}, x^{p^3}y \rangle} \bar{\gamma}_{1246}(-2y^{p^3+1}x^{p^3-1} + x^{p^3}z^{p^3}).$$

Thus

$$h_1 \equiv_{\langle x^2, xy \rangle} y^{p^3+1} - \frac{1}{2}z^{p^3}x \quad \text{and} \quad h_1^p \equiv_{\langle x^{p+1}, x^p y \rangle} y^{p^4+p} - \frac{1}{2}z^{p^4}x^p.$$

Furthermore

$$N_M(y)^{p^3+1} \equiv_{\langle x^{p+1}, x^p y \rangle} y^{p^4+p} - y^{p^4+1}x^{p-1}$$

and

$$h_1 N_M(y)^{p^3-p^2} x^{p-1} \equiv_{\langle x^{p+1}, x^p y \rangle} y^{p^4+1} x^{p-1}.$$

Thus $\text{LT}(h_1^p - N_M^{p^3+1} - h_1 N_M(y)^{p^3-p^2}) = -\frac{1}{2}x^p z^{p^4}$. \square

Theorem 10.4. *If $\gamma_{1234}(M) = 0$, $\gamma_{1235}(M) = 0$, $\gamma_{1357}(M) = 0$ and $\gamma_{1246}(M) \neq 0$, then the set $\mathcal{B} := \{x, N_M(y), h_1, N_M(z)\}$ is a SAGBI basis for $\mathbb{F}[V_M]^E$. Furthermore, $\mathbb{F}[V_M]^E$ is a hypersurface with the relation coming from the subduction of the tête-à-tête $(h_1^p, N_M(y)^{p^3+1})$.*

Proof. Use the subduction given in [Lemma 10.3](#) to construct an invariant h_2 with lead term z^{p^4} . Define $\mathcal{B}' := \{x, N_M(y), h_1, h_2\}$ and let A denote the algebra generated by \mathcal{B}' . The single nontrivial tête-à-tête for \mathcal{B}' subducts to zero using [Lemma 10.3](#). Thus \mathcal{B}' is a SAGBI basis for A . From [Lemma 10.1](#), $\mathbb{F}[V_M]^E[x^{-1}] = \mathbb{F}[x, N_M(y), h_1][x^{-1}]$. Thus, using [Theorem 1.1](#), $A = \mathbb{F}[V_M]^E$. Clearly $\text{LT}(N_M(z)) = z^{p^k}$ for $k \leq 4$. Since \mathcal{B}' is a SAGBI basis for $\mathbb{F}[V_E]^E$, this forces $k = 4$, giving $\text{LM}(\mathcal{B}) = \text{LM}(\mathcal{B}')$. \square

References

- [Adams and Loustaunau 1994] W. W. Adams and P. Loustaunau, *An introduction to Gröbner bases*, Graduate Studies in Mathematics **3**, American Mathematical Society, Providence, RI, 1994. [MR 1287608](#) [Zbl 0803.13015](#)
- [Benson 1993] D. J. Benson, *Polynomial invariants of finite groups*, London Mathematical Society Lecture Note Series **190**, Cambridge University Press, 1993. [MR 1249931](#) [Zbl 0864.13001](#)
- [Bosma et al. 1997] W. Bosma, J. J. Cannon, and C. Playoust, “The Magma algebra system, I: The user language”, *J. Symbolic Comput.* **24**:3–4 (1997), 235–265. [MR 1484478](#) [Zbl 0898.68039](#)
- [Campbell and Chuai 2007] H. E. A. E. Campbell and J. Chuai, “Invariant fields and localized invariant rings of p -groups”, *Q. J. Math.* **58**:2 (2007), 151–157. [MR 2334859](#) [Zbl 1134.13002](#)
- [Campbell and Wehlau 2011] H. E. A. E. Campbell and D. L. Wehlau, *Modular invariant theory*, Encyclopaedia of Mathematical Sciences **139**, Springer, Berlin, 2011. [MR 2759466](#) [Zbl 1216.14001](#)
- [Campbell et al. 2013] H. E. A. E. Campbell, R. J. Shank, and D. L. Wehlau, “Rings of invariants for modular representations of elementary abelian p -groups”, *Transform. Groups* **18**:1 (2013), 1–22. [MR 3022756](#) [Zbl 1264.13009](#)
- [Derksen and Kemper 2002] H. Derksen and G. Kemper, *Computational invariant theory*, Encyclopaedia of Mathematical Sciences **130**, Springer, Berlin, 2002. [MR 1918599](#) [Zbl 1011.13003](#)

- [Kapur and Madlener 1989] D. Kapur and K. Madlener, “A completion procedure for computing a canonical basis for a k -subalgebra”, pp. 1–11 in *Computers and mathematics* (Cambridge, MA, 1989), edited by E. Kaltofen and S. M. Watt, Springer, New York, NY, 1989. [MR 1005954](#) [Zbl 0692.13001](#)
- [Lakshmibai and Raghavan 2008] V. Lakshmibai and K. N. Raghavan, *Standard monomial theory: invariant theoretic approach*, Encyclopaedia of Mathematical Sciences **137**, Springer, Berlin, 2008. [MR 2388163](#) [Zbl 1137.14036](#)
- [Neusel and Smith 2002] M. D. Neusel and L. Smith, *Invariant theory of finite groups*, Mathematical Surveys and Monographs **94**, American Mathematical Society, Providence, RI, 2002. [MR 1869812](#) [Zbl 0999.13002](#)
- [Robbiano and Sweedler 1990] L. Robbiano and M. Sweedler, “Subalgebra bases”, pp. 61–87 in *Commutative algebra* (Salvador, 1988), edited by W. Bruns and A. Simis, Lecture Notes in Math. **1430**, Springer, Berlin, 1990. [MR 1068324](#) [Zbl 0725.13013](#)
- [Sturmfels 1996] B. Sturmfels, *Gröbner bases and convex polytopes*, University Lecture Series **8**, American Mathematical Society, Providence, RI, 1996. [MR 1363949](#) [Zbl 0856.13020](#)
- [Wehlau 2013] D. L. Wehlau, “Invariants for the modular cyclic group of prime order via classical invariant theory”, *J. Eur. Math. Soc.* **15**:3 (2013), 775–803. [MR 3085091](#) [Zbl 1297.13011](#)
- [Wilkerson 1983] C. W. Wilkerson, “A primer on the Dickson invariants”, pp. 421–434 in *Proceedings of the Northwestern Homotopy Theory Conference* (Evanston, IL, 1982), edited by H. R. Miller and S. B. Priddy, Contemporary Mathematics **19**, American Mathematical Society, Providence, RI, 1983. [MR 711066](#) [Zbl 0525.55013](#)

Received: 2014-10-22

Revised: 2015-06-30

Accepted: 2015-08-17

theo.pierron@ens-rennes.fr

Département Mathématiques, ENS Rennes, 35170 BRUZ, France

r.j.shank@kent.ac.uk

School of Mathematics, Statistics & Actuarial Science, University of Kent, Canterbury, CT2 7NF, United Kingdom

Bootstrap techniques for measures of center for three-dimensional rotation data

L. Katie Will and Melissa A. Bingham

(Communicated by Mary C. Meyer)

Bootstrapping is a nonparametric statistical technique that can be used to estimate the sampling distribution of a statistic of interest. This paper focuses on implementation of bootstrapping in a new setting, where the data of interest are 3-dimensional rotations. Two measures of center, the mean rotation and spatial average, are considered, and bootstrap confidence regions for these measures are proposed. The developed techniques are then used in a materials science application, where precision is explored for measurements of crystal orientations obtained via electron backscatter diffraction.

1. Introduction

Three-dimensional rotation data is common in the field of materials science, where electron backscatter diffraction (EBSD) can be used to study the microtexture of metals, including the orientation of crystals within the metal. Using EBSD, a fixed beam of electrons is diffracted off of a metal sample, creating an image on a focal plane of sensors. These images reveal information about crystal structure and orientation in the metal [Randle 2003]. One area of interest in regards to EBSD measurements is precision. As Bingham, Nordman, and Vardeman [Bingham et al. 2009a] point out, methods used for quantifying EBSD precision in the materials science literature are not standard, with ad hoc precision estimates often reported (see, for example, [Demirel et al. 2000; Wilson and Spanos 2001]). This led Bingham et al. [2009a] to investigate the precision of measurements obtained via EBSD by developing new statistical distributions for 3-dimensional rotations. While the distributions developed by Bingham et al. [2009a] do allow for some flexibility in modeling, our intent here is development of nonparametric techniques, namely bootstrap confidence regions, that can be used without the need for any distributional assumptions. While bootstrapping techniques are commonly used in

MSC2010: 62G09, 62P30.

Keywords: bootstrap, 3-D rotations, mean matrix, spatial average.

This research was supported by NSF grant DMS-1104409.

one-dimensional nonparametric statistics, these techniques have not been applied to 3-dimensional rotation data.

Suppose that $O_1, \dots, O_n \in \text{SO}(3)$ represent orientations at n scanned locations on a metal specimen as measured by EBSD, where $\text{SO}(3)$ denotes the set of all 3×3 orthogonal rotation matrices. When adjacent locations produce similar EBSD crystal orientations, those locations are considered to be part of the same grain. We are interested in estimating the central rotation of a set of orientations from within the same grain, since the true central orientation would represent the actual grain orientation, with random scatter in O_1, \dots, O_n due to measurement error in the EBSD process. We will investigate two different measures of center for 3-dimensional rotation data.

The mean rotation, M , is a commonly used measure of center [León et al. 2006; Bingham et al. 2009a; Khatri and Mardia 1977] that is defined to be the rotation that maximizes $\text{trace}(M^T \bar{O})$, where $\bar{O} = \frac{1}{n} \sum_{i=1}^n O_i$ for $O_1, \dots, O_n \in \text{SO}(3)$. The mean rotation M can be found by using $M = VW$, where $\bar{O} = V\Sigma W$ is the singular value decomposition of \bar{O} . It is necessary to use these components from the singular value decomposition since \bar{O} may not be an element of $\text{SO}(3)$, but M is.

The second measure of center considered is the spatial average of Ball and Greiner [2012]. The spatial average of a set of rotations $O_1, \dots, O_n \in \text{SO}(3)$ is obtained through an iterative procedure that uses what is referred to as the axis-angle representation of a matrix. For a given matrix, the axis and angle are such that if you rotate the 3×3 identity matrix about the axis by the angle, you will arrive at the specified matrix. The steps to find the spatial average are outlined below, where the end result is the matrix S found in step (4) in the final iteration. The procedure begins by looking at just the first two matrices, O_1 and O_2 . Starting at O_1 , we rotate half of the way towards O_2 , resulting in a matrix S . Then we consider the third matrix in the data set O_3 and rotate S one-third of the way towards this matrix, giving an updated matrix S . We then rotate S one-fourth of the way towards O_4 , again updating S . This process continues until we have been through all n matrices in the data set. Note that for large samples both the mean rotation and spatial average converge to the population central matrix.

- (1) Let $S = O_1$ and let $i = 2$.
- (2) Compute $G = S^T O_i$.
- (3) Let u be the axis of G , let θ be the angle of G , and let $p = \theta/i$.
- (4) Compute $S = SP$, where P is the matrix form of (u, p) , and let $i = i + 1$.
- (5) If $i \leq n$, return to step (2).

In Section 2, development of bootstrap confidence regions for these measures of center for 3-dimensional rotations will be discussed. Accuracy of the bootstrap techniques will be explored through a simulation study in Section 3. Finally, the

bootstrap procedure will be applied to data from a nickel specimen to evaluate EBSD precision in [Section 4](#).

2. Development of bootstrapping technique

Bootstrapping is a nonparametric statistical technique that uses resampling with replacement. It can be used to estimate the sampling distribution of almost any statistic, e.g., mean, proportion, variance. It is commonly used to find confidence regions for population parameters. To find a 95% confidence interval in one-dimension, a large number (say 1000) of samples of size n are drawn from the original sample of size n with replacement and the statistic of interest is computed for each bootstrap sample. This creates a sampling distribution for the statistic of interest. Under the bootstrap percentile method, a 95% confidence interval is then obtained by using the 2.5th and 97.5th percentiles as confidence bounds.

Although bootstrapping has been used to create confidence regions in a wide variety of settings, including analyzing directional data such as p -dimensional unit vectors [[Fisher and Hall 1989](#)], we introduce the concept of bootstrapping for measures of center in the 3-dimensional setting, where the data can be represented by 3×3 orthogonal rotation matrices. To estimate measures of center for 3-dimensional rotation data by bootstrapping, we sample with replacement from the original sample of n matrices 1000 times. Each sample is a bootstrap sample, for which we compute a measure of center (mean rotation or spatial average). We will refer to these 1000 matrices as bootstrap central matrices. To provide an estimate of center for the 3-dimensional rotation data, the mean rotation of the 1000 bootstrap central matrices can be computed. Since this matrix is analogous to what we would consider a “point estimate” when considering 1-dimensional data, we also refer to it as a point estimate here.

After obtaining the point estimate for our central rotation, we want to find a confidence region around this matrix. Because rotation matrices do not have a natural ordering, the percentile bootstrap method of using the 2.5th and 97.5th percentiles as confidence bounds does not translate directly to 3-dimensional rotation data. Instead, we will use a set of three cones centered at the point estimate to give a confidence region for the true central rotation in a similar fashion to [Bingham et al. \[2009a\]](#). [Figure 1](#) illustrates this concept for two different-sized sets of cones. To determine the size of the cones needed to give a 95% confidence region, we first think of each matrix as a set of three axes (x , y , and z) and consider the angles between the axes of the point estimate and the bootstrap central matrices. For each bootstrap central matrix, we find three angles. Each one is the angle between an axis and the corresponding axis of the point estimate. We then take the maximum of these three angles, so that moving a distance of that angle away from all three axes

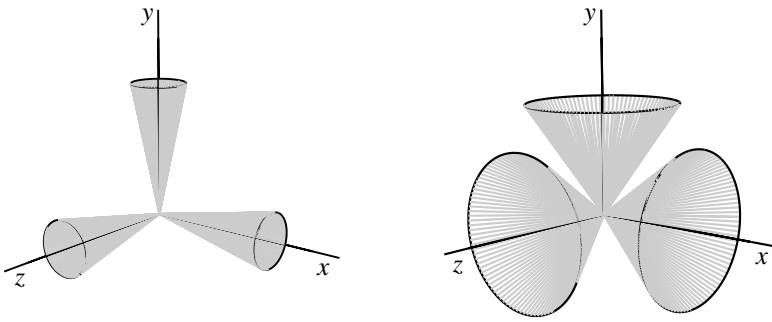


Figure 1. Plot of confidence cones around a point estimate (represented as three axes) with an angle of 0.2 radians (left) and 0.6 radians (right).

of the point estimate would contain all three axes of the bootstrap central matrix. Once these maximum angles are computed for all 1000 bootstrap central matrices, we take the 95th percentile and use this as the cone size. Since this set of three cones centered at the point estimate will capture 95% of the bootstrap central matrices, we think of it as a 95% bootstrap percentile region for the true central matrix.

3. A simulation study

To examine the accuracy of the bootstrap technique developed in Section 2, a simulation study was performed. Data sets were simulated from both the von Mises version of the uniform axis random spin (vM-UARS) distributions [Bingham et al. 2009a] and the matrix Fisher distribution [Khatri and Mardia 1977]. A vM-UARS or matrix Fisher distribution is characterized by a central rotation $S \in \text{SO}(3)$ and a spread parameter $\kappa \in (0, \infty)$. The spread parameter κ is best described as a concentration parameter since larger values of κ yield rotations with less variability.

For this simulation study we used κ values of 1, 5, 20, and 500 and sample size n of 10, 30, and 100. For each combination of κ and n both the mean rotation and spatial average were considered and the bootstrapping procedure was replicated 1000 times (i.e., 1000 different samples were drawn from each of the vM-UARS and matrix Fisher distributions) with 1000 bootstrap samples taken from the original sample each time. For each of the 1000 replications a 95% confidence region as a set of three cones was found. The coverage rates of the confidence cones were then found as the proportion of times out of 1000 that the true central rotation S was captured. Note that our choice of S for simulation purposes was arbitrary, as results are the same regardless of what true central rotation is used. Tables 1 and 2 show the coverage rates along with the median cone size, in radians, for each case.

The coverage rates fluctuate closely around 95%, which validates that the bootstrapping procedure is behaving as desired for the two distributions considered here.

(κ, n)	Mean rotation		Spatial average	
	Coverage rate	Median cone size	Coverage rate	Median cone size
(1, 10)	0.966	0.63441	0.944	0.72272
(1, 30)	0.962	0.32435	0.980	0.51034
(1, 100)	0.948	0.16934	0.979	0.26583
(5, 10)	0.944	0.22523	0.924	0.22695
(5, 30)	0.932	0.13152	0.947	0.13726
(5, 100)	0.954	0.07147	0.965	0.07526
(20, 10)	0.944	0.10876	0.943	0.10816
(20, 30)	0.946	0.06436	0.945	0.06500
(20, 100)	0.956	0.03512	0.959	0.03535
(500, 10)	0.944	0.02157	0.927	0.02176
(500, 30)	0.963	0.01269	0.945	0.01271
(500, 100)	0.949	0.00697	0.961	0.00702

Table 1. Coverage rates and median cone sizes (in radians) for estimating the center of the vM-UARS distribution using the mean rotation and the spatial average.

(κ, n)	Mean rotation		Spatial average	
	Coverage rate	Median cone size	Coverage rate	Median cone size
(1, 10)	0.957	1.05138	0.951	1.35648
(1, 30)	0.950	0.55232	0.995	0.88534
(1, 100)	0.942	0.29247	0.991	0.43596
(5, 10)	0.922	0.27921	0.909	0.27613
(5, 30)	0.945	0.16369	0.939	0.16325
(5, 100)	0.954	0.09025	0.947	0.09023
(20, 10)	0.919	0.13289	0.920	0.13365
(20, 30)	0.932	0.07896	0.952	0.07891
(20, 100)	0.950	0.04319	0.939	0.04318
(500, 10)	0.918	0.02629	0.925	0.02633
(500, 30)	0.940	0.01556	0.948	0.01544
(500, 100)	0.947	0.00851	0.948	0.00851

Table 2. Coverage rates and median cone sizes (in radians) for estimating the center of the matrix Fisher distribution using the mean rotation and the spatial average.

We also see that the values of κ and n impact the median confidence region sizes as expected, with larger κ (less spread) and larger n resulting in smaller regions. It is also important to compare the nonparametric bootstrap techniques developed here to existing parametric methods for the vM-UARS and matrix Fisher distributions. Bingham, Vardeman, and Nordman [Bingham et al. 2009b, Table 5, page 618] provide median cone sizes for the central rotation of the vM-UARS distributions obtained by maximum quasi-likelihood estimation using the same κ and n values considered here. Bingham, Nordman, and Vardeman [Bingham et al. 2010b, Table 5, page 1325] use maximum likelihood estimation to provide similar results for the matrix Fisher distribution. To compare the cone sizes of these works to the results given in Tables 1 and 2 presented here, we calculated the relative difference between the sizes as $d(a_p, a_b) = |a_p - a_b|/a_p$, where a_p is the angle from the parametric approach and a_b is the angle from the bootstrap approach. For the vM-UARS distribution, the largest relative difference was 0.1360 (for $\kappa = 1$ and $n = 100$). For the matrix Fisher distribution, the largest relative difference was 0.1441 (for $\kappa = 1$ and $n = 10$). Both of these differences are small, indicating that the bootstrap techniques developed here produce results that are equivalent to existing parametric approaches.

4. Application to electron backscatter diffraction data

Now that the bootstrapping technique developed in Section 2 has been shown to keep coverage rates as expected and perform similarly to existing parametric methods, we use it to investigate precision of measurements obtained through EBSD. A high-iron-concentration nickel specimen of size $40\ \mu\text{m} \times 40\ \mu\text{m}$ was scanned using EBSD. The scanning was done over a regularly spaced grid with $0.2\ \mu\text{m}$ step size across the top of the specimen, resulting in 4121 crystal orientations. See [Bingham et al. 2010a] for more details regarding the machinery used and data collection process.

For two orientations P and Q , the misorientation angle between them is the smallest angle of rotation needed to get from P to Q when rotating about some axis. When using EBSD, orientations close in proximity are classified as composing a grain when the misorientation angle between them is small, so that a grain is thought of as a homogeneous piece of material that produces observations which generally share a common orientation. Figure 2 gives the grain map that resulted from using EBSD on the nickel specimen [Bingham et al. 2010a]. The grain map can be viewed as if one was looking down on the piece of nickel, so that the axes give the x - and y -locations on a rectangular grid. Each dot in the figure corresponds to a single measured orientation from the total 4121 orientations in this scan. Similar orientations are classified into grains, with each colored block on the map indicating a different grain. The similarity of color within grains makes some of the 4121 dots indistinguishable from others. Dots that clearly stand out represent locations on

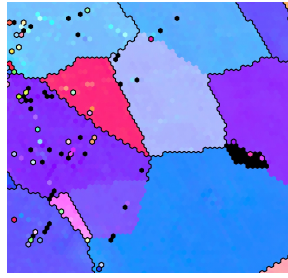


Figure 2. EBSD grain map for the nickel specimen, with grains represented by blocks of similar color.

Grain	Sample size (n)	Original cone size	Reported EBSD precision
1	49	0.0647°	0.4531°
2	31	0.1742°	0.9699°
3	21	0.0876°	0.4016°
4	44	0.0823°	0.5461°
5	22	0.1054°	0.4942°

Table 3. Size of 95% confidence regions for central rotations and reported precision of EBSD measurements (in degrees).

the scan with deformities. Although there are over ten grains present, we will use subsets of data from five of these grains in the analysis here.

For each of the five grains considered, we applied the bootstrapping technique to the 3×3 matrices representing crystal orientations. Using the mean rotation as our measure of center, 95% confidence regions for the central rotation were found. The sizes of the confidence cones are provided in Table 3, in degrees. Because confidence region sizes decrease at a rate of $1/\sqrt{n}$ (which can be verified by examining the cone sizes presented in Tables 1 and 2), before reporting the degree of precision we multiply each of the five cone sizes by \sqrt{n} . The reported precision estimates are also provided in Table 3. We find EBSD precision estimates comparable to the 1° reported by Demirel, El-Dasher, Adams, and Rollett [Demirel et al. 2000] and the 0.5° reported by Wilson and Spanos [2001] by using methods that are much more statistically sound than the methods employed in these works.

5. Conclusion

The study of precision for EBSD measurements considered here is just one of many applications that could benefit from the bootstrapping techniques developed. These bootstrapping techniques, while simple to implement, have been shown to perform as well as existing parametric approaches. Given the complexity of

existing parametric methods, they are likely not easily accessible to practitioners (such as materials scientists) who often collect 3-dimensional rotation data. Further, statistical methods that do not rely on distributional assumptions are important in the area of 3-dimensional rotation data since there are relatively few developed distributions for which parametric methods are even available [Bingham et al. 2009a; Khatri and Mardia 1977; León et al. 2006]. Therefore, the bootstrapping techniques presented here could play an important role in the field of statistics, as well as in areas of study where 3-dimensional rotations are commonly found.

References

- [Ball and Greiner 2012] K. A. Ball and T. M. Greiner, “A procedure to refine joint kinematic assessments: functional alignment”, *Comput. Meth. Biomech. Biomed. Engin.* **15**:5 (2012), 487–500.
- [Bingham et al. 2009a] M. A. Bingham, D. J. Nordman, and S. B. Vardeman, “Modeling and inference for measured crystal orientations and a tractable class of symmetric distributions for rotations in three dimensions”, *J. Amer. Statist. Assoc.* **104**:488 (2009), 1385–1397. MR 2596996 Zbl 1205.62215
- [Bingham et al. 2009b] M. A. Bingham, S. B. Vardeman, and D. J. Nordman, “Bayes one-sample and one-way random effects analyses for 3-D orientations with application to materials science”, *Bayesian Anal.* **4**:3 (2009), 607–629. MR 2551048
- [Bingham et al. 2010a] M. A. Bingham, B. K. Lograsso, and F. C. Laabs, “A Bayes statistical analysis of the variation in crystal orientations obtained through electron backscatter diffraction”, *Ultramicroscopy* **110**:10 (2010), 1312–1319.
- [Bingham et al. 2010b] M. A. Bingham, D. J. Nordman, and S. B. Vardeman, “Finite-sample investigation of likelihood and Bayes inference for the symmetric von Mises–Fisher distribution”, *Comput. Statist. Data Anal.* **54**:5 (2010), 1317–1327. MR 2600834
- [Demirel et al. 2000] M. C. Demirel, B. S. El-Dasher, B. L. Adams, and A. D. Rollett, “Studies on the accuracy of electron backscatter diffraction measurements”, pp. 65–74 in *Electron backscatter diffraction in materials science*, edited by A. J. Schwartz et al., Springer, New York, 2000.
- [Fisher and Hall 1989] N. I. Fisher and P. Hall, “Bootstrap confidence regions for directional data”, *J. Amer. Statist. Assoc.* **84**:408 (1989), 996–1002. MR 1134489
- [Khatri and Mardia 1977] C. G. Khatri and K. V. Mardia, “The von Mises–Fisher matrix distribution in orientation statistics”, *J. Roy. Statist. Soc. Ser. B* **39**:1 (1977), 95–106. MR 0494687 Zbl 0356.62044
- [León et al. 2006] C. A. León, J.-C. Massé, and L.-P. Rivest, “A statistical model for random rotations”, *J. Multivariate Anal.* **97**:2 (2006), 412–430. MR 2234030 Zbl 1085.62066
- [Randle 2003] V. Randle, *Microtexture determination and its applications*, 2nd ed., Maney for the Institute of Materials, Minerals and Mining, London, 2003.
- [Wilson and Spanos 2001] A. W. Wilson and G. Spanos, “Application of orientation imaging microscopy to study phase transformations in steels”, *Mater. Charact.* **46**:5 (2001), 407–418.

Received: 2014-12-31

Revised: 2015-05-26

Accepted: 2015-07-31

kwill@iastate.edu

Department of Statistics, Iowa State University,
Ames, IA 50011, United States

mbingham@uwlax.edu

Department of Mathematics and Statistics,
University of Wisconsin-La Crosse, 1725 State Street,
La Crosse, WI 54601, United States

Graphs on 21 edges that are not 2-apex

Jamison Barsotti and Thomas W. Mattman

(Communicated by Joel Foisy)

We show that the 20-graph Heawood family, obtained by a combination of ∇Y and $Y\nabla$ moves on K_7 , is precisely the set of graphs of at most 21 edges that are minor-minimal with respect to the property “not 2-apex”. As a corollary, this gives a new proof that the 14 graphs obtained by ∇Y moves on K_7 are the minor-minimal intrinsically knotted graphs of 21 or fewer edges. Similarly, we argue that the seven-graph Petersen family, obtained from K_6 , is the set of graphs of at most 17 edges that are minor-minimal with respect to the property “not apex”.

1. Introduction

A graph is n -apex if there is a set of n or fewer vertices whose deletion results in a planar graph. As this property is closed under taking minors, it follows from Robertson and Seymour’s graph minor theorem [2004] that, for each n , the n -apex graphs are characterized by a finite set of forbidden minors. For example, 0-apex is equivalent to planarity, which Wagner [1937] showed is characterized by K_5 and $K_{3,3}$. For the property 1-apex, which we simply call *apex*, there are several hundred forbidden minors (see [Ding and Dziobak 2016], which refers to work of a team led by Kézdy). Since there are likely even more forbidden minors for the 2-apex property, we divide the problem into more manageable pieces by graph size. In an earlier paper [Mattman 2011], the second author showed that every graph on 20 or fewer edges is 2-apex. This means there are no forbidden minors with 20 or fewer edges. In the current paper, we show that there are exactly 20 obstruction graphs for 2-apex of size at most 21.

Following [Hanaki et al. 2011], the *Heawood family* will denote the set of 20 graphs obtained from K_7 by a sequence of zero or more ∇Y or $Y\nabla$ moves. Recall that a ∇Y move consists of deleting the edges of a 3-cycle abc of graph G and adding a new degree-3 vertex adjacent to the vertices a , b , and c . The reverse, deleting a degree-3 vertex and making its neighbors adjacent, is a $Y\nabla$ move. The

MSC2010: primary 05C10; secondary 57M15, 57M25.

Keywords: spatial graphs, intrinsic knotting, apex graphs, forbidden minors.

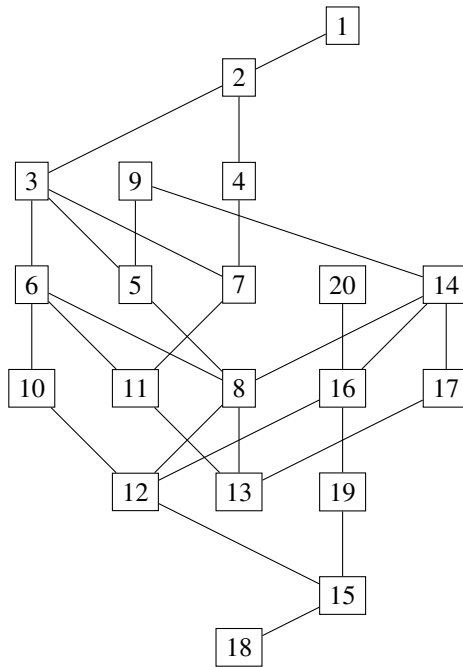


Figure 1. The Heawood family (figure taken from [Goldberg et al. 2014]). Edges represent ∇Y moves.

Heawood family is illustrated schematically in Figure 1, where K_7 is graph 1 at the top of the figure and the $(14, 21)$ Heawood graph is graph 18 at the bottom.

Our main theorem is that the Heawood family is precisely the obstruction set for the property 2-apex among graphs of size at most 21. We will state this in terms of minor-minimality. We say H is a *minor* of graph G if H is obtained by contracting edges in a subgraph of G . The graph G is *minor-minimal* with respect to a graph property \mathcal{P} if G has \mathcal{P} , but no proper minor of G does. We call obstruction graphs for the 2-apex property *minor-minimal not 2-apex* or *MMN2A*.

Theorem 1.1. *The 20 Heawood family graphs are the only MMN2A graphs on 21 or fewer edges.*

As there are no MMN2A graphs of size 20 or less [Mattman 2011] and one easily verifies that the Heawood family graphs are MMN2A, the argument comes down to showing no other 21-edge graph enjoys this property. We give a more complete outline of our proof at the end of this introduction.

Our interest in 2-apex stems from the close connection with intrinsic knotting. A graph is *intrinsically knotted* or *IK* if every tame embedding of the graph in \mathbb{R}^3 contains a nontrivially knotted cycle. Then, a *minor-minimal IK*, or *MMIK*, graph is one that is IK, but such that no proper minor has this property. Again, Robertson and

Seymour’s graph minor theorem [2004] implies a finite list of MMIK graphs, but determining this list or even bounding its size has proved very difficult. Restricting by order, it follows from Conway and Gordon’s seminal paper [1983] that K_7 is the only MMIK graph on seven or fewer vertices; two groups, Campbell et al. [2008] and Blain et al. [2007], independently determined the MMIK graphs of order 8; and we have announced (see [Morris 2008; Goldberg et al. 2014]) a classification of nine-vertex graphs, based on a computer search. In terms of edges, we know ([Johnson et al. 2010] and, independently, [Mattman 2011]) that a graph of size 20 or less is not IK. Using the following lemma (due, independently, to two research teams), this follows from the lack of MMN2A graphs of that size.

Lemma 1.2 [Blain et al. 2007; Ozawa and Tsutsumi 2007]. *If G is IK, then G is not 2-apex.*

The current authors [Barsotti and Mattman 2013] and, independently, Lee et al. [2015] classified the 21-edge MMIK graphs. These are the 14 KS graphs obtained by ∇Y moves on K_7 , first described by Kohara and Suzuki [1992]. In other words, these are the Heawood family graphs except those labeled 9, 14, 16, 17, 19, and 20 in Figure 1. In light of Lemma 1.2, we have a new proof as a corollary to our main theorem.

Corollary 1.3. *The 14 KS graphs are the only MMIK graphs on 21 or fewer edges.*

Proof. Kohara and Suzuki [1992] showed that the KS graphs are MMIK. Suppose G is MMIK of at most 21 edges. Then G is connected. By Lemma 1.2, G has an MMN2A minor and, by Theorem 1.1, this means a Heawood family graph minor. As G has at most 21 edges and is connected, G is a Heawood family graph. Finally, Goldberg et al. [2014] and Hanaki, Nikkuni, Taniyama, and Yamazaki [2011], independently, showed that in the Heawood family only the KS graphs are IK. Therefore, G is a KS graph. \square

We remark that there is considerable overlap in the current paper and our preprint [Barsotti and Mattman 2013]. We have opted for a self-contained presentation here as we will not be publishing the above preprint elsewhere.

The proof of our main theorem relies on our classification of MMNA graphs (i.e., obstructions to the 1-apex, or apex, property) of small size, a result that may be of independent interest. Recall that, in analogy with the Heawood family, the Petersen family is the set of the seven graphs obtained from the Petersen graph by a sequence of ∇Y or $Y\nabla$ moves.

Theorem 1.4. *The seven Petersen family graphs are the only MMNA graphs on 16 or fewer edges.*

Famously, the Petersen family is precisely the obstruction set to intrinsic linking [Robertson et al. 1995]. It would be nice to have a similar description of the

Heawood family. [Theorem 1.1](#) is one such characterization. As a second corollary to our main theorem, we give a characterization of similar flavor. Hanaki et al. [\[2011\]](#) showed that the Heawood family graphs are minor-minimal with respect to the property “intrinsically knotted or completely 3-linked”; that is, Heawood family graphs are MMI(K or C3L).

Corollary 1.5. *The 20 Heawood family graphs are the only MMI(K or C3L) graphs on 21 or fewer edges.*

Proof. Hanaki et al. [\[2011\]](#) proved these graphs are MMI(K or C3L). Let G be MMI(K or C3L) on 21 or fewer edges. Then G is connected. By [\[Hanaki et al. 2011, Remark 4.5\]](#), I(K or C3L) implies N2A, so G must have an MMN2A minor. By [Theorem 1.1](#), this means a Heawood minor. It follows that G has 21 edges and is a Heawood family graph, as required. \square

This gives two characterizations of the Heawood family. However, like our [Theorem 1.4](#), they are less than ideal due to the hypothesis on graph size. Is there a “natural” description of the Heawood family analogous to the way the Petersen family is precisely the obstruction set for intrinsic linking?

Note that the condition on graph size in these three results is necessary. Indeed, for [Theorem 1.4](#), the disjoint union $K_{3,3} \sqcup K_{3,3}$ is an 18-edge MMNA graph outside the Petersen family. On the other hand, a computer search [\[Pierce 2014\]](#) shows that [Theorem 1.4](#) could be extended to 17 edges: there are no MMNA graphs of size 17. Since IK implies both N2A ([Lemma 1.2](#)) and I(K or C3L) (see [\[Hanaki et al. 2011\]](#)) there are many examples of MMN2A and MMI(K or C3L) graphs on 22 edges, including $K_{3,3,1,1}$. Foisy [\[2002\]](#) showed this graph is MMIK, which means it is also N2A and I(K or C3L). As any proper minor of $K_{3,3,1,1}$ would have at most 21 edges, and no Heawood family graph is a minor, it follows from [Theorem 1.1](#) and [Corollary 1.5](#), that $K_{3,3,1,1}$ is both MMN2A and MMI(K or C3L). So, the hypothesis on size is necessary for both the theorem and its corollary.

Thus, $K_{3,3,1,1}$ and the 14 KS graphs are examples of graphs that enjoy all three properties: MMN2A, MMIK, and MMI(K or C3L). On the other hand, the remaining six Heawood graphs show that a graph can be MMN2A and not MMIK. This includes the graph that we have called E_9 [\[Mattman 2011\]](#) and that Hanaki et al. [\[2011\]](#) label N_9 . In [\[Goldberg et al. 2014\]](#) we showed that adding an edge to this graph makes it MMIK. In other words, $E_9 + e$ is MMIK and not MMN2A (as it has the N2A graph E_9 as a subgraph). On the other hand, since IK implies I(K or C3L), every MMIK graph has a minor that is MMI(K or C3L); although E_9 , for example, shows that the set of I(K or C3L) graphs is a strictly larger class than IK. Similarly, I(K or C3L) implies N2A [\[Hanaki et al. 2011\]](#), which means every MMI(K or C3L) has an MMN2A minor, while the disjoint union of three $K_{3,3}$ graphs is an example of a graph that is N2A but not I(K or C3L).

All six of the Heawood graphs that are not MMIK are MMI(K or C3L) and we can ask if a graph that is MMN2A and not MMIK need be I(K or C3L). However, the disjoint union $G = K_6 \sqcup K_5$ is a counterexample. Since K_6 is MMNA and K_5 is nonplanar, G is N2A and, since any proper minor is 2-apex, it is in fact MMN2A. On the other hand, G is neither IK nor I(K or C3L) as each component has fewer than 21 edges.

We conclude this overview of connections between apex graphs and intrinsic knotting with a question. In [Goldberg et al. 2014] we describe the known 263 examples of MMIK graphs. By Lemma 1.2, none of these graphs are 2-apex. However, it is straightforward to verify that each is 3-apex. Does this hold more generally?

Question 1.6. *Is every MMIK graph 3-apex?*

The remainder of our paper is a proof of Theorem 1.1. Let G be an MMN2A graph of size 21. We must show that G is a Heawood family graph. We can assume $\delta(G)$, the *minimum degree*, is at least 3. Indeed, in an N2A graph, deleting a degree-0 vertex or contracting an edge of a vertex of degree 1 or 2 will result in an N2A minor. We can also bound the number of vertices. As G has 21 edges and minimum degree of at least 3, it has at most 14 vertices. On the other hand, we classified MMN2A graphs on nine or fewer vertices in [Mattman 2011]. So we can assume $10 \leq |V(G)| \leq 14$. After introducing some preliminary lemmas, and proving Theorem 1.4, in the next section, we devote one section each to the five cases where the number of vertices runs from 14 down to ten. We opted for this reverse ordering as it roughly corresponds to the increasing lengths of the proofs.

2. Preliminaries

We denote the order of a graph G by $|G|$ and its size by $\|G\|$ and frequently use the pair $(|G|, \|G\|)$ as a way of describing the graph. For $a_i \in V(G)$, we will use $G - a_1, \dots, a_n$ to denote the induced subgraph on $V(G) \setminus \{a_1, \dots, a_n\}$. We will write $G + a$ to denote a graph with vertices $V(G) \cup \{a\}$ that includes G as the induced subgraph on $V(G)$. In the case where $V(G)$ and $\{a\}$ are included in the vertex set of some larger graph, $G + a$ will mean the induced subgraph on $V(G) \cup \{a\}$. We use $N(a)$ to denote the *neighborhood of vertex a* , the set of vertices adjacent to a . We will write NA, MMNA, N2A, and MMN2A for “not apex” (equivalently, “not 1-apex”), “minor-minimal not apex”, “not 2-apex”, and “minor-minimal not 2-apex” respectively.

Vertices of degree less than 3 do not participate in determining whether or not a graph is n -apex, so we next describe a systematic way of deleting those vertices. Recall that in a *multigraph* the edge set is a multiset, so that edges may be repeated. In addition, there may be *loops*, edges that are incident to the same vertex twice.

Definition 2.1. The *simplification* G^s of a graph G is the multigraph obtained by the following procedure:

- (1) Delete all degree-0 vertices.
- (2) Delete all degree-1 vertices and their edges.
- (3) If there remain vertices of degree 0 or 1, go to step (1).
- (4) For each degree-2 vertex v with distinct edges va and vb , delete v and those edges and add the edge ab .
- (5) If there remain any vertices of degree 0 or 1, go to step (1).

The procedure allows us to recognize $V(G^s)$ as a subset of $V(G)$. We call these vertices of G the *branch vertices*.

In step (4), the procedure leaves loops on degree-2 vertices unchanged. On the other hand, it may be that $a = b$ so that va is a doubled edge. In this case, step (4) replaces the doubled edge with a loop on vertex a and deletes vertex v . It's straightforward to verify that G^s is unique, up to isomorphism.

Lemma 2.2. *The graph G is n -apex if and only if G^s is.*

Proof. Just as for a graph, we say that a multigraph is n -apex if there are n or fewer vertices whose deletion results in a planar multigraph. The lemma follows as n -apex is preserved by each step in the definition. \square

This means that graphs where G^s is nonplanar will be of particular interest. An important class is that of *split $K_{3,3}$ graphs*: graphs G such that $G^s = K_{3,3}$.

In this section, we will prove [Theorem 1.4](#): the Petersen family graphs are the MMNA graphs with $\|G\| \leq 16$. Recall that the Petersen family is the set of seven graphs obtained by ∇Y and $Y\nabla$ moves on the (10, 15) Petersen graph P_{10} . In addition to P_{10} , the set includes K_6 , $K_{3,3,1}$, $K_{4,4} - e$, and, by definition, is closed under ∇Y and $Y\nabla$ moves. We first observe that each graph in the family is MMNA.

Lemma 2.3. *The seven graphs in the Petersen family are all MMNA.*

Proof. Aside from describing what is to be checked, we omit most of the details. Let G be a graph in the Petersen family. It's enough to verify that for all $v \in V(G)$, $G - v$ is nonplanar and that for all $e \in E(G)$, deletion and contraction of e both result in apex graphs. \square

The proof of [Theorem 1.4](#) depends on the following lemma that characterizes NA graphs using the idea of a vertex near a branch vertex. If G is a graph and $w \in V(G)$ is such that there is a path from w to a branch vertex, a , of G that contains no other branch vertices of G , then we say w is *near* a . Similarly, if w is a vertex in some $G + v$, then w is near a branch vertex a of G if there is a w - a path independent of the other branch vertices.

For the lemma, we assume that either G is a Kuratowski graph or else it is a multigraph, which we will call a $K_{3,3}$ with a fat edge, denoted by $K_{3,3} + \bar{e}$. This means the multigraph is a $K_{3,3}$ graph but for a single edge that is repeated (possibly many times). Figure 13 (left) is an example. We consider the graph $K_{3,3}$ to be a $K_{3,3} + \bar{e}$. Note that we will use $K_{3,3} + e$ to refer to the graph obtained by adding an edge to $K_{3,3}$; see Figure 13 (right).

Lemma 2.4. *Suppose G simplifies to K_5 or a $K_{3,3} + \bar{e}$. Then $G + v$ is NA if and only if v is near every branch vertex of G .*

Proof. As in the definition above, forming G^s , the simplification of G , determines a set of branch vertices.

First, assume that $G + v$ is NA and v is not near a branch vertex a of G . If we remove a branch vertex b near a , then, we claim, $G + v - b$ is planar, which contradicts that $G + v$ is NA. In the case of a $K_{3,3} + \bar{e}$, choose b to be a vertex of the fat edge, so that it is incident to every repeated edge. To verify the claim, note that $(G - b)^s$ is a Kuratowski graph with one vertex deleted, either K_4 or $K_{3,2}$. The only way that $G + v - b$ could be nonplanar would be for v to take the place of b in the Kuratowski graph. This would require independent paths from v to each of the branch vertices near b . As there is no such v - a path, $G + v - b$ is planar.

Now assume that, in $G + v$, the vertex v is near every branch vertex of G . Then $G^* = (G + v)^s$ is of the form $H + v$, where H is a subdivision of G^s and, by abuse of notation, we again refer to the vertices of H of degree 3 or more as branch vertices (of G). In G^* , the neighbors of v are either branch vertices of G or on edges of G^s that were subdivided to form H . In particular, v is near the same branch vertices in $H + v$ as it was in $G + v$. We wish to show that G^* can, through a series of $Y\nabla$ moves, be transformed into an NA graph. If, in G^* , we have that v is adjacent to all the branch vertices of G , we are done, since if $G^s = K_5$, then G^* has a K_6 minor, and if G^s is a $K_{3,3} + \bar{e}$, then G^* has $K_{3,3,1}$ as a minor. As K_6 and $K_{3,3,1}$ are both NA (see the previous lemma), $G + v$ is as well.

Next, choose a branch vertex a from G . Suppose v is not adjacent to a in G^* . However, we've assumed v is near every branch vertex, including a . Hence there is a vertex w of degree 3 that has both a and v as neighbors. Performing a $Y\nabla$ move on w makes a and v neighbors and will not change the nearness of v to any branch vertices. Repeating this process for the rest of the branch vertices results in a graph where v is adjacent to each branch vertex of G . Again, if $G^s = K_5$, then this series of $Y\nabla$ moves on G^* gives a graph that has a K_6 minor. If G^s is a $K_{3,3} + \bar{e}$ then a series of $Y\nabla$ moves on G^* gives us a graph that has $K_{3,3,1}$ as a minor. Since $Y\nabla$ and ∇Y preserve the Petersen family, we conclude that $G + v$ has a minor from the Petersen family and is, therefore, NA. □

The proof shows that, not only is $G + v$ NA, it has a Petersen family graph as a minor. On the other hand, if $G + v$ has a Petersen family graph minor, then it is NA by [Lemma 2.3](#). Also, Petersen family graph minors characterize intrinsic linking [[Robertson et al. 1995](#)]. The following lemma combines these observations.

Lemma 2.5. *Let G be a graph with vertex v such that $(G - v)^s$ is K_5 or a $K_{3,3} + \bar{e}$. Then the following are equivalent:*

- *The vertex v is near every branch vertex of $G - v$.*
- *G is NA.*
- *G has a Petersen family graph minor.*
- *G is intrinsically linked.*

Lemma 2.6. *Suppose G is NA and there is a vertex a such that $(G - a)^s = K_{3,3} + e$. Then G has a minor in the Petersen family.*

Proof. We use the notation provided by [Figure 13](#) (right). If a is not near v_2 or v_3 then $G - w_3$ is planar. On the other hand, if a is not near one of w_1, w_2 , and w_3 , then $G - v_3$ is planar. So a is near v_2, v_3, w_1, w_2 , and w_3 . If $\{v_2, v_3, w_1, w_2, w_3\} \subset N(a)$, then G has the Petersen family graph P_7 (obtained by a single ∇Y on K_6) as a minor, as required.

Suppose one of these vertices is not in $N(a)$, say $v_2 \notin N(a)$. Then, as in the proof of [Lemma 2.4](#), there is some minor of G in which a $Y\nabla$ move produces a graph that has $v_2 \in N(a)$ (where a and v_2 are the induced vertices from a and v_2 after finding such a minor of G and performing the $Y\nabla$ move) and a is still near each vertex in $\{v_3, w_1, w_2, w_3\}$. Repeat this process for each of those remaining vertices and we see that G has a minor that, following a sequence of $Y\nabla$ moves, becomes P_7 . Since the Petersen family is closed under $Y\nabla$ and ∇Y moves, G has a minor in the Petersen family. \square

Lemma 2.7. *Suppose G is NA and there is a vertex a such that $\|(G - a)^s\| \leq 10$. Then G has a minor in the Petersen family.*

Proof. By assumption, $G - a$ is nonplanar, and by [Lemma 2.2](#), $(G - a)^s$ is as well. So, K_5 or $K_{3,3}$ is a minor, $(G - a)^s$ is either K_5 , $K_{3,3} + e$, or a $K_{3,3} + \bar{e}$, and we can apply [Lemma 2.5](#) or [Lemma 2.6](#). \square

Lemma 2.8. *If $G + a$ is formed by adding a degree-3 vertex a to a split $K_{3,3}$ graph G and $G + a$ is NA, then $(G + a)^s$ is the Petersen graph.*

Proof. By [Lemma 2.4](#), there are paths from a to each branch vertex that avoid all other branch vertices. Up to isomorphism, the only way to arrange this is as in the graph of [Figure 2](#), which is the Petersen graph. \square

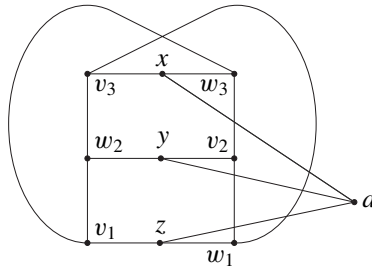


Figure 2. Adding a degree-3 vertex to a split $K_{3,3}$ yields the Petersen graph.

Figure 2 illustrates the idea of a vertex being near an edge. Let G be such that $G^s = K_{3,3}$ or K_5 . As in the proof of Lemma 2.4, if we add a vertex v , then, in general, $(G+v)^s$ will be of the form $H+v$, where H is a subdivision of G^s . We say that v is near the edge xy in G^s , where x and y are branch vertices, if, in $(G+v)^s$, v has a neighbor interior to the (subdivided) edge xy of G^s . In Figure 2, a is near the edges v_iw_i with $i = 1, 2, 3$.

Lemma 2.9. *If $G+a$ is formed by adding a vertex a of degree 4 to a split $K_{3,3}$ graph G and $G+a$ is NA, then $(G+a)^s$ is one of the seven graphs in Figure 3.*

Proof. By Lemma 2.4, there are paths from a to each branch vertex that avoid all other branch vertices. Let $N(a) = \{n_1, n_2, n_3, n_4\}$. As there are six vertices and $d(a) = 4$, there is an n_i , say n_1 , that has an edge, say v_1w_1 , as its nearest part. Since there are four branch vertices left and three neighbors of a , another n_i , say n_2 , must have an edge as its nearest part with vertices disjoint from $\{v_1, w_1\}$; call it v_2w_2 . There are three graphs generated when a has a neighbor whose nearest part is a branch vertex of G and four more when a has no such neighbor. Figure 3 shows the graphs that result from this condition. \square

We conclude this section with a proof of Theorem 1.4. The proof requires one additional lemma. Let $\delta(G)$ and $\Delta(G)$ denote the minimum and maximum degrees of a graph G .

Lemma 2.10. *Suppose G has $\delta(G) = 3$, $\Delta(G) = 4$, and $13 \leq \|G\| \leq 16$. Then either there is a degree-4 vertex with a degree-3 neighbor or else G is the disjoint union $K_5 \sqcup K_4$.*

Proof. For a contradiction, suppose no degree-4 vertex has a degree-3 neighbor. Then G is disconnected with cubic and quartic components. The smallest quartic graph is K_5 with ten edges and the smallest cubic graph is K_4 with six. So, the size of G is at least 16 and $K_5 \sqcup K_4$ is the only way to realize that minimum. \square

Proof of Theorem 1.4. As stated in Lemma 2.3, the Petersen family graphs are all MMNA. What is left is to show that they are the only such graphs on 16 or fewer

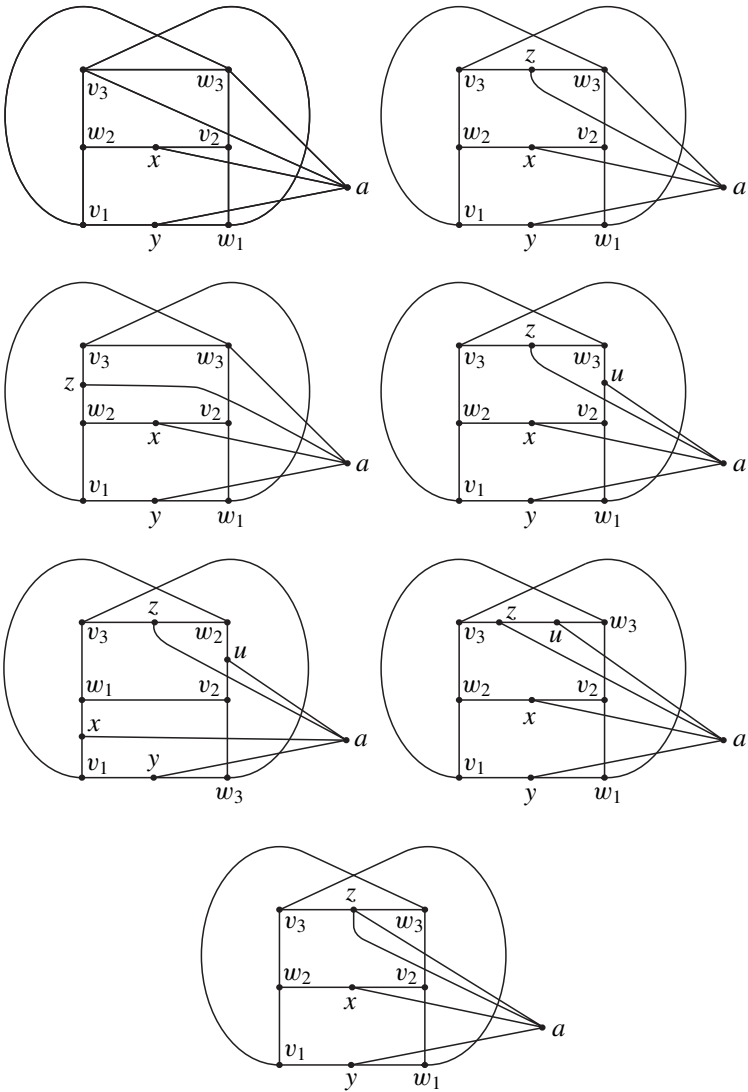


Figure 3. Adding a degree-4 vertex to a split $K_{3,3}$.

edges. Suppose G is an MMNA graph with 16 or fewer edges. Our goal is to show that G is in the Petersen family. If $\delta(G) < 3$, then contracting an edge of a vertex of small degree or deleting an isolated vertex results in a proper minor that is still NA, contradicting minor-minimality. So we assume $\delta(G) \geq 3$.

Further, we can assume that, for every vertex a , we have $\|(G - a)^s\| \geq 11$. Otherwise, by Lemma 2.7, G has a minor in the Petersen family. Since the Petersen family graphs are NA and we're assuming G is MMNA, G must be a Petersen family graph, as required.

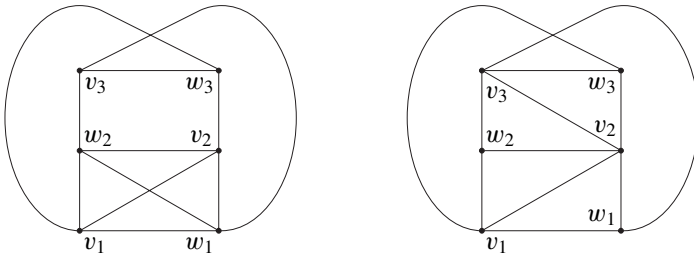


Figure 4. Nonplanar (6, 11) graphs with $\delta(G) \geq 3$.

Combining the assumptions $\delta(G) \geq 3$ and $\|(G - a)^s\| \geq 11$, we see that G has size 14, at least. However, if $\|G\| = 14$, then a minimum degree of 3 and each $G - a$ having size 11 or more imply that G is cubic, which is not possible. In fact, G must have at least 15 edges.

If $\|G\| = 15$, then $\Delta(G) \leq 4$ since we're assuming each $(G - a)^s$ has at least 11 edges. Suppose $\Delta(G) = 4$. Since there are no quartic graphs with 15 edges, by Lemma 2.10, there is a degree-4 vertex a with at least one neighbor of degree 3. Then $\|(G - a)^s\| \leq 10$, contradicting our assumption. So, we can assume G is cubic. In this case, apply Lemma 2.8 to see that G is the Petersen graph.

This leaves the case where $\|G\| = 16$. The assumption that each $(G - a)^s$ has at least 11 edges implies $\Delta(G) \leq 5$. If $\Delta(G) = 5$, let a be a vertex of top degree. We can assume a has no degree-3 neighbor since $\|(G - a)^s\| \geq 11$. Then $G - a$ is a nonplanar simple graph of size 11 and minimum degree 3. The only possibilities are the (6, 11) graphs of Figure 4 or the (7, 11) graph of Figure 16 (top center). As is the case with a , we can assume that no degree-5 vertices have degree-3 neighbors in G . Suppose first that $G - a$ is the (6, 11) graph of Figure 4 (left). Then $N(a)$ must include v_3 and w_3 , the degree-3 vertices of $G - a$, as otherwise there will be a degree-5 vertex with a degree-3 neighbor. Without loss of generality, w_1 is the vertex of $G - a$ missing from $N(a)$. Then $G - v_1$ is planar, a contradiction. Similarly, if $G - a$ is the (6, 11) graph of Figure 4 (right), then, since we assumed $\Delta(G) = 5$, it's v_2 that is missing from $N(a)$, in which case $G - w_2$ is planar. Finally, suppose $G - a$ is the (7, 11) graph of Figure 16 (top center). We see that $v_2 \in N(a)$ as otherwise $G - w_3$ is planar. But then v_2 is a degree-5 vertex in G and can have no degree-3 neighbors. Thus $N(a) = \{u, v_2, w_1, w_2, w_3\}$ and contracting uv_1 gives the Petersen family graph P_7 as a minor. (Recall that P_7 is the result of a ∇Y move on K_6 .)

Next assume $\Delta(G) = 4$. If G is quartic, it is one of the six quartic graphs of order 8 (see [Meringer 1999]). Only two of these are NA. One is $K_{4,4}$, which has the Petersen family graph $K_{4,4} - e$ as a subgraph. The other comes from splitting the degree-6 vertex of the Petersen family graph $K_{3,3,1}$.

Thus we can assume $\delta(G) = 3$ and since each $(G - a)^s$ has at least 11 edges, each degree-4 vertex has at most one degree-3 neighbor. By Lemma 2.10 (note

that $K_5 \sqcup K_4$ is not NA), there is a degree-4 vertex b with a degree-3 neighbor for which $\|(G - b)^s\| = 11$.

Since $\delta(G) = 3$, we have that $|G| \geq 9$. Since $(G - b)^s$ is formed by deleting vertex b and its degree-3 neighbor (which becomes degree-2 and is lost through simplification), it has order 7 at least. Thus, $(G - b)^s$ is either the (7, 11) graph of Figure 16 (top center) or one of the (7, 10) graphs of Figure 15 with a doubled edge, and $G - b$ is formed by a single subdivision.

Suppose $G - b$ is the (7, 11) graph with a single subdivision. Recall that each degree-4 vertex has at most one degree-3 neighbor. So that both $G - w_2$ and $G - w_3$ are nonplanar, the subdivision must be of an edge incident to v_2 . This constitutes a degree-3 neighbor of v_2 and its remaining neighbors must all be adjacent to b . However, this results in a degree-4 vertex with two degree-3 neighbors, a contradiction.

If $(G - b)^s$ is a graph of Figure 15 with a doubled edge, one of those repeated edges is subdivided to form $G - b$. This introduces a new vertex x that must be adjacent to b since $\delta(G) = 3$. If $(G - b)^s$ is the graph of Figure 15 (left), then, since $\delta(G - b) \geq 2$, it must be the edge uv_1 that is doubled. Both u and x are degree-2 in $G - b$. So, both are in $N(b)$ and become degree-3 in G . However, this means the degree-4 vertex b has two degree-3 neighbors in G , which is a contradiction. Similarly, if $(G - b)^s$ is the graph of Figure 15 (right), the doubled edge must be adjacent to u as otherwise $u, x \in N(b)$, which gives b two degree-3 neighbors. So, we can assume it's uv_1 that is doubled. As v_1 is degree-4 in G and x is degree-3, v_1 can have no other degree-3 neighbors. Then $N(b) = \{u, x, w_2, w_3\}$. However, this leaves several degree-4 vertices in G that have two degree-3 neighbors, which is a contradiction.

Having size 16, G is not cubic, so we've completed the argument for graphs of this size, and with it the proof. □

3. 14-vertex graphs

We now show the following (originally proved in [Barsotti and Mattman 2013]):

Proposition 3.1. *If G is a (14, 21) MMN2A graph, then G is in the Heawood family.*

Proof. Let G be a (14, 21) MMN2A graph. We can assume $\delta(G) \geq 3$ as otherwise a vertex deletion or edge contraction on a small-degree vertex will give a proper minor that is also N2A. Then G must have the degree sequence (3^{14}) and for any $a \in V(G)$, we know that $G - a$ has the sequence $(3^{10}, 2^3)$. Now choose another vertex, b , such that $G^* = G - a, b$ has the sequence $(3^6, 2^6)$ (i.e., a and b have no common neighbors). There are enough degree-3 vertices in $G - a$ to assure we can always choose such a b .

Since G is N2A and G^* has the sequence $(3^6, 2^6)$, we have that G^* must be a split $K_{3,3}$. By Lemma 2.8, $(G^* + a)^s$ is the Petersen graph of Figure 2. Then $G' = (G^* + a) - w_3$ is another split $K_{3,3}$.

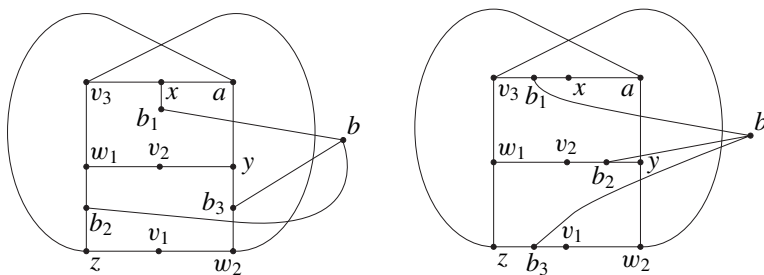


Figure 5. Two possibilities for $G' + b$.

By Lemma 2.4, b must have a path to a that avoids $v_3, w_1, w_2, y,$ and z . Since a and b have no common neighbors, this means b has a neighbor b_1 that is adjacent to x . So, there are two cases: in $G' + b$, either b_1 is of degree 2, or else it has v_3 as a third neighbor. (See Figure 5.)

In either case, b_1 gives paths from b to the branch vertices a and v_3 and there are three ways to split the remaining four branch vertices into two pairs. However, we see that $G - w_2, z$ is planar (and G is 2-apex), unless we make the choices shown in Figure 5. In both cases, adding w_3 back will give us the Heawood graph. Hence the only (14,21) MMN2A graph is the Heawood graph. \square

4. 13-vertex graphs

In this section we prove the following:

Proposition 4.1. *If G is a (13, 21) MMN2A graph, then G is in the Heawood family.*

Proof. Let G be an MMN2A (13, 21) graph. Consider the degree sequences $(3^{12}, 6)$ and $(3^{11}, 4, 5)$. If we remove the vertex of highest degree, the resulting graph simplifies to a graph with fewer than 14 edges, hence (by Theorem 1.4) to an apex graph. So G does not have such a degree sequence.

Then G has the sequence $(3^{10}, 4^3)$. Again, if a is a vertex of degree 4 that has three neighbors of degree 3, then $(G - a)^s$ is apex, so this cannot be the case. We conclude that the degree-4 vertices form a triangle in G and that there is a degree-3 vertex a in G whose neighbors all have degree 3. This means that $G - a$ simplifies to a graph $G^* = (G - a)^s$ with degree sequence $(3^6, 4^3)$. Since G^* must be NA, and has 15 edges, by Theorem 1.4 it is in the Petersen family. There is a unique nine-vertex graph in the family, which we call P_9 ; see Figure 6.

Note that in Figure 6 there is a unique triangle, which we'll denote by xyz and label the corresponding vertices in $G - a$ and G as $x, y,$ and z as well. Notice also that $x, y,$ and z all have degree 4 in G^* so none of them are neighbors of a in G . Moreover, we assumed $x, y,$ and z form a triangle in G , and since the triangle is clearly preserved in G^* , it must also be preserved in $G - a$. In particular, this implies

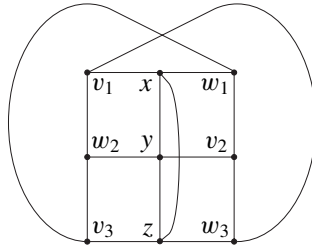


Figure 6. The Petersen family graph P_9 .

that a is not near any of the edges that form this triangle; i.e., none of the degree-2 vertices deleted in simplifying from $G - a$ to G^* are on the edges of the triangle.

Observe that $(G - a, y)^s = K_{3,3}$ and that the induced graph after adding a back must be NA. Hence, by Lemma 2.4, a must have a path to each branch vertex that does not go through any other branch vertex. Since a is not near the edge xz , it must be near either edges xw_1 or xv_1 and zw_3 or zv_3 . Similarly, $(G - a, x)^s$ shows that a must also be near yw_2 or yv_2 .

We claim that a is near $xw_1, yw_2,$ and zw_3 or near $xv_1, yv_2,$ and zv_3 , in which case G is the Heawood family graph C_{13} . (See [Hanaki et al. 2011] for the names, like C_{13} , of the Heawood family graphs. This is the unique order-13 graph in the Heawood family and corresponds to graph 15 in Figure 1). Otherwise, either a is near xv_1 and yw_2 or near xw_1 and yv_2 , in which case $G - v_3, w_3$ is planar, or else a is near zv_3 and yw_2 or near zw_3 and yv_2 in which case $G - v_1, w_1$ is planar. \square

5. 12-vertex graphs

In this section we prove that a $(12, 21)$ MMN2A graph G is in the Heawood family. This means G is one of three graphs that are called $H_{12}, C_{12},$ and N'_{12} by Hanaki et al. [2011] and are represented as graphs 12, 13, and 19, respectively, in Figure 1. We first observe that if G is triangle-free and of the correct degree sequence, it must be H_{12} . This was originally proved in [Barsotti and Mattman 2013].

Lemma 5.1. *Let G be MMN2A of degree sequence $(3^6, 4^6)$ and triangle-free. Then G is H_{12} .*

Proof. Note that if any of the vertices of degree 4 have three or more neighbors of degree 3, removing such a vertex results in an apex graph by Theorem 1.4, so we may assume this doesn't happen. We also notice that we can either single out a degree-3 vertex, all of whose neighbors are degree-3 vertices, or a degree-4 vertex that has two degree-3 neighbors. To see this, suppose it is not the case. Since G has no triangles, the subgraph induced by the degree-4 vertices is $K_{3,3}$ and each of the vertices has a unique neighbor of degree 3. Hence, removing two nonadjacent

vertices of degree 4 results in a graph that simplifies to a graph of size 8, and thus is planar. Hence G would not be 2-apex.

Now assume that we do not have a vertex of degree 4 with two degree-3 neighbors. Say that a is a degree-3 vertex whose neighbors are all of degree 3. Then $(G - a)^s$ has degree sequence $(3^2, 4^6)$. [Theorem 1.4](#) implies that it is $K_{4,4} - e$. Because G has no degree-4 vertex with two degree-3 neighbors, we know that the edge subdivisions from $(G - a)^s$ to $G - a$ are all on edges incident to the degree-3 vertices of $(G - a)^s$. Also, since G is triangle-free, there is at most one subdivision on each edge. Since there are exactly three subdivisions from $(G - a)^s$ to $G - a$, there is one vertex of degree 3 in $(G - a)^s$ that gets at least two subdivisions; call it a_1 . So, a_1 has degree-4 neighbors v_1, v_2 in $(G - a)^s$ so that a_1v_1 and a_1v_2 are subdivided in forming $(G - a)$. Then $G - v_1, v_2$ is planar; indeed $(G - v_1, v_2)^s$ is $K_{4,2}$, and G is 2-apex.

So we may assume that a has degree 4 and there exist $b, c \in N(a)$ such that $d(b) = d(c) = 3$ and $c \neq b$. Then $(G - a)^s$ has degree sequence $(3^6, 4^3)$, which tells us, by [Theorem 1.4](#), that it is P_9 . Furthermore, since G does not have a triangle, we know that one of the subdivisions from $(G - a)^s$ to $G - a$ is on the triangle xyz of [Figure 6](#); say it's xy that is subdivided. Removing either x or y , [Lemma 2.4](#) tells us that the other subdivision from $(G - a)^s$ to $G - a$ must be on an edge incident to z . (Note that $z \notin N(a)$ as it would be a degree-5 vertex.) The subdivision cannot be on the edge yz or xz , otherwise one of x, y , or z would have more than two neighbors of degree 3. Furthermore, we need that either $w_1, w_2 \in N(a)$ or $v_1, v_2 \in N(a)$, since x and y are allowed at most two neighbors of degree 3 and G has no triangles. If $w_1, w_2 \in N(a)$ then considering $(G - a, z)^s$ shows us that a is near w_3 by [Lemma 2.4](#), hence the subdivision is on w_3z . Similarly, if $v_1, v_2 \in N(a)$ the subdivision is on v_3z . Both cases yield H_{12} . □

Proposition 5.2. *If G is a (12, 21) MMN2A graph, then G is in the Heawood family.*

Proof. We assume again that G is MMN2A and that G is a (12, 21) graph. We can assume the maximum degree $\Delta(G)$ is at most 5. A vertex a with $d(a) \geq 6$ in a (12, 21) graph with $\delta(G) \geq 3$ will have at least one neighbor of degree 3. Then $(G - a)^s$ has at most 14 edges and is apex, by [Theorem 1.4](#). This implies $G - a$ is apex and G is 2-apex, a contradiction.

This leaves four possible degree sequences: $(3^9, 5^3)$, $(3^8, 4^2, 5^2)$, $(3^7, 4^4, 5)$, and $(3^6, 4^6)$.

Let G have the degree sequence $(3^9, 5^3)$ or $(3^8, 4^2, 5^2)$. Then any a with $d(a) = 5$ has at least two neighbors of degree 3. This means $(G - a)^s$ simplifies to a graph with fewer than 15 edges and so it is apex ([Theorem 1.4](#)), whence G is 2-apex, a contradiction.

We now focus our attention on the case where G has the degree sequence $(3^7, 4^4, 5)$ and show that the only MMN2A graph with this degree sequence is C_{12} .

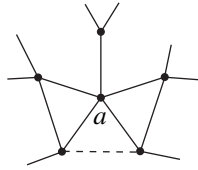


Figure 7. Graph near the degree-5 vertex a . The dotted edge indicates that the degree-4 vertices may form a path.

(See [Hanaki et al. 2011] for the name. This is graph 12 in Figure 1.) Let a denote the vertex of degree 5. Note that a has at most one neighbor of degree 3, as otherwise $\|(G - a)^s\| \leq 14$, meaning $G - a$ is apex (Theorem 1.4) and G is 2-apex. Hence, the neighbors of a are all the vertices of degree 4 and one vertex of degree 3. Moreover, each vertex of degree 4 has at most two neighbors of degree 3. This is illustrated in Figure 7. This implies that $(G - a)^s$ is an NA 3-regular graph with 15 edges, i.e., the Petersen graph (see Figure 2). Since the Petersen graph has no triangles or 4-cycles, we see that $G - a$ has no 4-cycles. This implies that the vertices of degree 4 do not form a triangle or 4-cycle in G . This justifies the specifics of Figure 7.

Then there is a $b \in V(G)$ of degree 4 with exactly two degree-3 neighbors, so that $(G - b)^s$ is a $(9, 15)$ graph with degree sequence $(3^6, 4^3)$. This implies that $(G - b)^s$ is the Petersen family graph P_9 illustrated in Figure 6 (the unique Petersen family graph on nine vertices). In $G - b$, vertex a has degree 4 and without loss of generality is vertex y in the figure. We have deduced that b is adjacent to a as well as to either w_2 or v_2 , say v_2 . At this stage, we see that, in fact, the degree-4 vertices do not form a path. Note that b is not near the edge xz ; otherwise both x and y will have three neighbors of degree 3. In order for $G - a$ to be NA, by Lemma 2.4, b must be near the edges v_1x and v_3z . Adding both a and b back in shows that this graph is C_{12} .

Now let G have the degree sequence $(3^6, 4^6)$. We will show G is either H_{12} or else N'_{12} . (See [Hanaki et al. 2011] for these names. These are graphs 12 and 19 respectively in Figure 1.) By Lemma 5.1, the only triangle-free MMN2A graph with degree sequence $(3^6, 4^6)$ is H_{12} , so we will assume that G has a triangle and show that this implies it is N'_{12} . By Theorem 1.4, each degree-4 vertex in G can have at most two neighbors of degree 3. Notice that in N'_{12} , each degree-4 vertex has exactly one neighbor of degree 3 and vice versa. We argue that G must also share this property in order to be MMN2A.

First, assume there is an $a \in V(G)$ such that a has degree 3 and three degree-3 neighbors. Hence $G^* = (G - a)^s$ has degree sequence $(4^6, 3^2)$ and is an $(8, 15)$ graph. Since G being MMN2A implies that G^* is NA, by Theorem 1.4 it is in the Petersen family. By the degree sequence $(4^6, 3^2)$, we can identify G^* as $K_{4,4} - e$, drawn in Figure 8. Since G^* has no triangles, the triangle of G is formed in reattaching a . Hence there is at least one edge in G^* that is subdivided twice in returning to $G - a$.

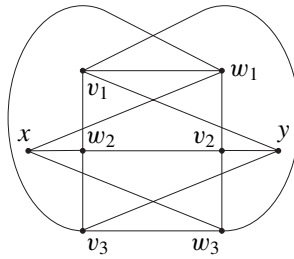


Figure 8. The Petersen family graph $K_{4,4} - e$.

Because of the symmetry of G^* , we may assume without loss of generality that these subdivisions are on the edges v_1w_1 or yv_1 . In the first case, $G - v_1, w_1$ is planar, and the second splits into two cases: either the other subdivision from G^* to $G - a$ occurs on an edge incident to x in G^* or it does not. In the case where it does not, $G - v_i, w_j$ is planar, where v_i and w_j are the vertices in G^* between which the subdivision occurs or v_1 and w_1 if it's on an edge incident to y . In the other case, $G - x, v_1$ is planar since it is essentially the same as the planar graph $G^* - x, v_1$ with an extra path from y to a w_i . So, in an MMN2A graph, every degree-3 vertex has at least one degree-4 neighbor.

Now suppose $a \in V(G)$ is a degree-4 vertex with exactly two neighbors of degree 3. Then $G^* = (G - a)^s$ has degree sequence $(4^3, 3^6)$. Since G^* must be NA, by Theorem 1.4 it is in the Petersen family and hence is the graph P_9 shown in Figure 6. In the following, we use the labeling of that figure.

When we remove $x, y,$ or z separately from G^* , each induced subgraph shows us (by Lemma 2.4) that a must have paths to $x, y,$ and z in G that do not include any of their neighbors in G^* . As these three vertices already have degree 4, the neighborhood of a includes vertices adjacent to x, y, z created by edge subdivisions.

Since there are only two edge subdivisions from G^* to $G - a$, this implies that one has to be on the xyz triangle. By the symmetry of G^* , we can assume without loss of generality that xy is subdivided. The other subdivision is on an edge incident to z in G^* . Since we assume that G contains a triangle, a must be part of that triangle. Observe that $(G^* - y)^s = K_{3,3}$. By Lemma 2.4, a must have paths in $G - y$ to the vertices $v_1, v_3, w_1, w_3, x,$ and z that exclude the others from that list. Now, a is adjacent to exactly two vertices in $G^* - y$ (as the two other neighbors appear only after additional edge subdivisions) and since we have already established that a is near both x and z and possibly v_3 or w_3 , the remaining neighbors of a are either w_2 and v_2, v_1 and $v_2,$ or w_1 and w_2 . Recalling that a is not actually adjacent to x , just simply near it by way of a subdivision of xy in G^* , and since G must have a triangle, none of these cases can be G .

To summarize, we established that if G is MMN2A with degree sequence $(3^6, 4^6)$ and contains a triangle, then each vertex of degree 4 has at most one neighbor of

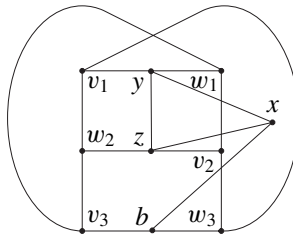


Figure 9. Graph after removing a degree-4 vertex leaving a triangle.

degree 3 and each vertex of degree 3 has at least one neighbor of degree 4. Hence, there is a one-to-one correspondence between the degree-4 vertices and the degree-3 vertices by the relation of being neighbors in G . Note that degree-3 vertices cannot occur on triangles that include degree 4 vertices. Otherwise either the degree-3 vertex is adjacent to two degree-4 vertices, or else there is a degree-4 vertex with two degree-3 neighbors. If the degree-3 vertices form two disjoint triangles, G is 2-apex. Indeed, let a and b be two degree-4 vertices whose neighbors of degree 3 are on distinct triangles. Then $(G - a, b)^s$ is basically a subgraph of the planar graph K_4 . The vertices of the K_4 are the remaining degree-4 vertices of G (besides a and b). In addition to edges between these that were in G , the remnants of the degree-3 vertices contribute two additional paths of length three with the central edge doubled.

Thus, we can assume there is a triangle of vertices of degree 4 in G . Choose some vertex of degree 4 not on this triangle; call it a . Then $G^* = (G - a)^s$ has degree sequence $(3^8, 4^2)$ and contains a triangle. We claim that G^* is the graph illustrated in Figure 9. Note that the two degree-4 vertices in G^* are adjacent. So, if we delete one of them, denote it by y , then $(G^* - y)^s$ has nine edges and must be nonplanar since G^* is NA. Thus $(G^* - y)^s = K_{3,3}$ and, using Lemma 2.9, and the fact that G^* has a triangle and degree sequence $(3^8, 4^2)$, we deduce G^* is as shown in Figure 9.

Now that we have established what G^* looks like (Figure 9), we can determine where a goes. Since both y and z are adjacent to x , we know that x cannot have degree 3 due to the one-to-one correspondence between vertices of degrees 3 and 4. So a is adjacent to x . Then a is adjacent to either v_1 or w_1 since y is adjacent to only one vertex of degree 3, say w_1 . Then, for the same reason x and a were adjacent, a and v_2 are adjacent. Since $G - z$ is NA, by Lemma 2.4, a is near w_2v_3 or v_1w_2 . Similarly, $G - y$ is NA and Lemma 2.4 shows a is near v_1w_2 or v_1w_3 . So a is near v_1w_2 . This graph is N'_{12} . Therefore, the only MMN2A graph with degree sequence $(3^6, 4^6)$ that contains a triangle is N'_{12} . \square

6. 11-vertex graphs

In this section we prove that an $(11, 21)$ MMN2A graph is in the Heawood family. We begin with five lemmas, one for each Heawood family graph of this order: E_{11} ,

C_{11} , H_{11} , N'_{11} , and N_{11} . (See [Hanaki et al. 2011] for the names. These correspond to graphs 8, 10, 11, 16, and 17 respectively in Figure 1.)

Lemma 6.1. *Let G be an $(11, 21)$ MMN2A graph with degree sequence $(3^4, 4^6, 6)$. Then G is C_{11} .*

Proof. Consider $b \in V(G)$ such that $\deg(b) = 6$. Notice that for any $v \in N(b)$ we must have $\deg(v) = 4$; otherwise, by Theorem 1.4, $G - b$ is not NA. This implies that $G - b$ must be the Petersen graph (see Figure 2). Without loss of generality, we can assume that the vertex a in Figure 2 is not a neighbor of b in G . Since $(G - b, x)^s = K_{3,3}$, we have that in $G - x$, by Lemma 2.4, b must be adjacent to z and y . Similarly, if we consider $G - b, z$ we see that b is adjacent to x . Consider again $G - x$. Since b has degree 5 in $G - x$, is adjacent to y and z , and must have paths to v_1, v_2, w_1 , and w_2 that do not go through v_1, v_2, w_1, w_2, x , or y , we see that b is adjacent to either v_3 or w_3 or both. Similarly, considering $G - y$ and $G - z$, we see that b is adjacent to either v_2 or w_2 and v_1 or w_1 . We claim that b is adjacent to v_1, v_2 , and v_3 or w_1, w_2 , and w_3 , in which case we have C_{11} . Otherwise, if $v_2 \in N(b)$ and $w_1 \in N(b)$ then $G - v_3, w_3$ is planar, or if $v_2 \in N(b)$ and $w_3 \in N(b)$ then $G - w_1, v_1$ is planar. Similarly, if $w_2 \in N(b)$ and $v_1 \in N(b)$ then $G - v_3, w_3$ is planar, or if $w_2 \in N(b)$ and $v_3 \in N(b)$ then $G - w_1, v_1$ is planar. Therefore G must be C_{11} . \square

Lemma 6.2. *Let G be an $(11, 21)$ MMN2A graph with degree sequence $(3^5, 4^3, 5^3)$. Then G is E_{11} .*

Proof. We may assume that there exists $a \in V(G)$ such that $\deg(a) = 5$ and there exists $u \in N(a)$ such that $\deg(u) = 3$. If not, then removing any two of the degree-4 vertices results in a K_4 graph with a bridge to a graph of at most seven edges, which is clearly planar. On the other hand, by Theorem 1.4, $G^* = (G - a)^s$ has at least 15 edges, so u is the only degree-3 neighbor. Then G^* has nine vertices.

This means that G^* is the Petersen family graph P_9 shown in Figure 6, the only order-9 graph in the family. By the degree sequence of the original G , we may assume, without loss of generality, that a is adjacent to x and y (referring again to Figure 6), and hence is not adjacent to z . Removing either x or y , Lemma 2.4 shows us that a is near an edge incident to z . If a is near the edge yz or xz , then a is also adjacent to two more vertices in Figure 6. Removing both of these results in a planar graph. Thus a is near the edge v_3z or the edge w_3z . By symmetry, we will assume v_3z .

Applying Lemma 2.4 to $G - y$ shows that a must be adjacent to v_2 and, similarly, considering $G - x$ shows us that a must be adjacent to v_1 . Reassembling G gives E_{11} . \square

Lemma 6.3. *Let G be an $(11, 21)$ MMN2A graph with degree sequence $(3^4, 4^5, 5^2)$. Then G is H_{11} .*

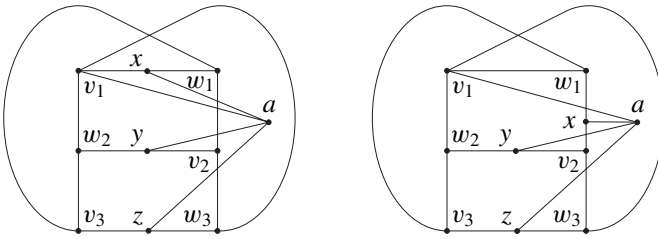


Figure 10. Graphs with degree sequence $(3^8, 4^2)$ by adding a degree-4 vertex a to a split $K_{3,3}$.

Proof. Assume that there exists $a \in V(G)$ such that $\deg(a) = 5$ and there exists $u \in N(a)$ such that $\deg(u) = 3$. Then $G^* = (G - a)^s$ is a $(9, 15)$ NA graph, hence the graph illustrated in Figure 6, with degree sequence $(3^6, 4^3)$. Since G has only two vertices of degree 5, vertex a is adjacent to at most one of x , y , and z in Figure 6. We will assume that it is x and hence $y, z \notin N(a)$. By Lemma 2.4, a must be near edges incident to both y and z (consider $G - z$ and $G - y$, respectively). However, as a has a unique neighbor of degree 2 in $G - a$, it is near only one edge. Therefore, a is near the edge yz . If a is adjacent to v_1, v_2 , and v_3 or w_1, w_2 , and w_3 then G is H_{11} .

We next verify that this must be the case. Note that there are exactly three vertices in $N(a) \cap \{v_1, v_2, v_3, w_1, w_2, w_3\}$. Let us first examine the intersection with $\{v_2, v_3, w_2, w_3\}$. Lemma 2.4 applied to $G - z$ shows that a has at least one neighbor in each of the pairs $\{v_2, w_3\}$, $\{v_3, w_2\}$, and $\{v_3, w_3\}$. The same lemma with $G - x$ shows that $N(a) \cap \{v_2, v_3, w_2, w_3\}$ is not simply $\{v_3, w_3\}$. We conclude that a is adjacent to w_2 and w_3 or v_2 and v_3 , and, by symmetry, we can assume v_2 and v_3 . The last neighbor of a must be v_1 , as otherwise $G - v_3, w_3$ or $G - v_2, w_2$ will be planar.

Let a and b be the degree-5 vertices and suppose neither has a degree-3 neighbor. If a and b are not adjacent, then $(G - a, b)^s$ is a (3^4) multigraph that is clearly planar. Further, a and b can have at most three common neighbors, as otherwise $(G - a, b)^s$ has fewer than nine edges and is therefore planar. On the other hand, since there are only five degree-4 vertices, a and b must share at least three neighbors. This means $(G - a, b)^s = K_{3,3}$. By Lemma 2.9, $G - b$ must be one of the graphs in Figure 10. By our assumption, b is adjacent to a, x, y , and z , with one other neighbor from the set $\{w_1, w_2, w_3, v_2, v_3\}$. In the case where $G - b$ looks like Figure 10 (left) we see that $G - v_1, w_1$ is planar. For the case of the right graph in the figure, observe that $G - v_1, x$ is planar. Hence if a and b have no degree-3 neighbors, then G is 2-apex. Therefore G must be H_{11} . \square

Lemma 6.4. *Let G be an $(11, 21)$ MMN2A graph with degree sequence $(3^3, 4^7, 5)$. Then G is N'_{11} .*

Proof. Let us begin by assuming that the degree-5 vertex b is adjacent to some vertex of degree 3. Then $G^* = (G - b)^s$ has degree sequence $(3^6, 4^3)$ and is therefore the

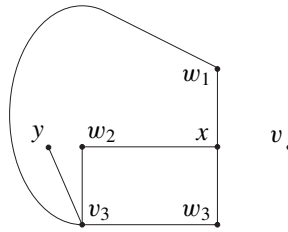


Figure 11. Remove v_1 and v_2 from $K_{4,4} - e$.

P_9 graph of Figure 6. Note that b is not adjacent to x , y , or z , since going from G to G^* did not change their degree. However, observing the graphs we obtain when removing x , y , or z , by Lemma 2.4 we see that b needs a path to all of them that does not utilize any of their neighbors in G^* . This is clearly impossible since there is at most one subdivision from G^* to $G - b$. Hence for all $v \in N(b)$, we have $\deg(v) = 4$.

Then $G - b$ must have the degree sequence $(3^8, 4^2)$. If the vertices of degree 4 in $G - b$ are not adjacent, then if v is one of those, $(G - b, v)^s$ has eight edges and is therefore planar, which is a contradiction. So choose $a \in V(G - b)$ such that $\deg(a) = 4$. Then if G is N2A, $(G - a, b)^s$ is $K_{3,3}$. When we add a back in, by Lemma 2.9, there are two cases, shown in Figure 10. However, for Figure 10 (right), we notice that b is not adjacent to v_1 since it can only be adjacent to vertices of degree 3 in $G - b$. This means that it is not near v_1 , which is required by Lemma 2.4. So $G - b$ is isomorphic to the graph illustrated in Figure 10 (left). As above, since b must be near v_1 , it must be adjacent to x . Now, $G - v_1, w_1$ will be planar unless $N(b)$ includes either $\{v_2, v_3\}$ or $\{w_2, w_3\}$. We will argue that it must be the latter. Suppose instead that $\{x, v_2, v_3\}$ is in $N(b)$ and $\{w_2, w_3\}$ is not. In particular, if $w_2 \notin N(b)$, then $G - v_3, w_3$ is planar, a contradiction. Similarly, if $w_3 \notin N(b)$, then $G - v_2, w_2$ gives a contradiction. This shows that it is not possible that $\{w_2, w_3\} \not\subset N(b)$, and so we can assume $\{w_2, w_3\} \subset N(b)$. Now $G - v_2, w_2$ is planar unless b is adjacent to y and $G - v_3, w_3$ shows z is adjacent to b as well, which means G is N'_{11} . \square

Lemma 6.5. *Let G be an $(11, 21)$ MMN2A graph with degree sequence $(3^2, 4^9)$. Then G is N_{11} .*

Proof. First assume that there exists a $v \in V(G)$ such that $\deg(v) = 4$ and the two vertices of degree 3 are neighbors of v . Then $(G - v)^s$ has degree sequence $(3^2, 4^6)$ and is the Petersen family graph $K_{4,4} - e$, illustrated in Figure 8. Thus $G - v$ is a subdivision of $K_{4,4} - e$. Note that in G , vertex v is adjacent to both x and y . The graph obtained from $K_{4,4} - e$ when we remove v_1 and v_2 is illustrated in Figure 11. Since v is adjacent to both x and y and the graph $G - v, v_1, v_2$ can be obtained from Figure 11 by only two subdivisions (the other neighbors of v), we see that $G - v_1, v_2$ is planar.



Figure 12. There are two or four edges between V_3 and V_4 .

We can now assume that the two degree-3 vertices of G have no common degree-4 neighbors. Let a be a degree-4 vertex that has a degree-3 neighbor. Then $G^* = (G - a)^s$ has degree sequence $(3^4, 4^5)$.

If G^* is not a simple graph, then, since it must be NA, by [Theorem 1.4](#), it is a Petersen family graph with an edge doubled. This means the Petersen family graph is P_9 ([Figure 6](#)), the only one of order 9. The doubled edge is between two degree-3 vertices in that figure. Using symmetry, we can assume it's v_1w_2 that's doubled. In G one of these edges is subdivided to give a degree-3 vertex whose neighbors are $a, v_1,$ and w_2 . None of these three are adjacent to the other degree-3 vertex, which is therefore w_1 or v_2 . By symmetry, we can assume w_1 is the other vertex of degree 3. In other words, G is formed from P_9 by adding a vertex b adjacent to v_1 and w_2 , and a vertex a with $N(a) = \{b, v_2, v_3, w_3\}$. Then $G - v_1, w_1$ is planar, a contradiction. So we can assume G^* is a simple graph and $G - a$ differs from it only by subdivision of an edge.

Notice first that if G^* has a degree-4 vertex v that has three or more degree-3 neighbors, then $(G^* - v)^s$ has at most nine edges and five vertices and is planar. We claim that there is a degree-4 vertex in G^* that has two neighbors of degree 3. Suppose not and let V_3 denote the set of degree-3 vertices of G^* and V_4 those of degree 4. As the degree sums in the two parts are even, there are an even number of edges between V_3 and V_4 . If there were six or more, then, by the pigeonhole principle, one of the degree-4 vertices would have two degree-3 neighbors, which is what we are trying to establish. If there were no edges in between, $G^* = K_4 \sqcup K_5$ would be apex, a contradiction. So there are two or four edges between V_3 and V_4 . (See [Figure 12](#).) In either case, removing a degree-4 vertex that has a degree-3 neighbor will result in a planar graph.

So, let $b \in V(G^*)$ be a degree-4 vertex with two degree-3 neighbors. Moreover, a and b have a common neighbor, as otherwise b has two degree-3 neighbors in G . Now, $G^* - b$ will be formed by subdividing two edges of a $(6, 10)$ graph G' having degree sequence $(3^4, 4^2)$. Since our assumption implies that G' is nonplanar, G' is one of the two graphs obtained by adding an edge to $K_{3,3}$ (see [Figure 13](#)).

Assume that G' is the $K_{3,3} + \bar{e}$ shown in [Figure 13](#) (left). Since G was a simple graph, there is at least one subdivision on one of the paired edges. This means b is adjacent to the vertex resulting from that subdivision. Notice that $G - v_3, w_3$ is

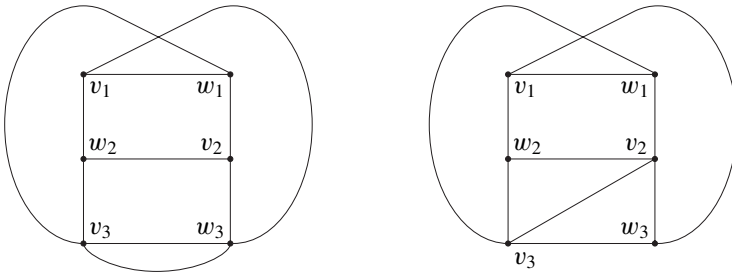


Figure 13. Two nonplanar (6, 10) graphs.

essentially a subdivision of the 4-cycle $v_1 w_1 v_2 w_2$ along with two more vertices that are not adjacent to one another. This graph is planar unless a and b are near the same edge, which is incident to either v_3 or w_3 in G' . On the other hand, by Lemma 2.4, b must have independent paths to each of the branch vertices of G' and this cannot happen if it is near two different edges adjacent to v_3 or w_3 . In other words, a and b are adjacent to the same edge, which is one of the pair between v_3 and w_3 .

Next, suppose a and b are adjacent to the same edge in the pair, but attached to the edge at two different vertices formed by subdividing that edge twice. By Lemma 2.4, b must have independent paths to each of the branch vertices of G' . Now, b is adjacent to two vertices formed by subdivisions of G' as well as two degree-3 vertices in that graph. This means that, in addition to one of the $v_3 w_3$ -edges, b is near an edge between two other vertices, say $v_1 w_1$. This gives b paths to four of the branch vertices and shows that the other two vertices, v_2 and w_2 , are the remaining neighbors of b .

Recall that $G - a, b$ is obtained from G' by exactly three edge subdivisions. If a and b do not share a vertex on the $v_3 w_3$ edge, it must be the vertex resulting from subdividing $v_1 w_1$ that is common. But this means there is no way to attach a to G' so that it will have independent paths to all the branch vertices. So far, we have subdivisions that show a is near a $v_3 w_3$ edge and $v_1 w_1$. The remaining two neighbors would have to be v_2 and w_2 . However, these vertices then have degree 5 in G , contradicting its $(3^2, 4^9)$ degree sequence.

We conclude that a and b attach at the same vertex of one of the paired edges of G' . Then as above, we can assume that b is near the edge $v_1 w_1$ and adjacent to v_2 and w_2 . Then those two vertices have degree 4 and are not adjacent to a . As there remains a single subdivision of G' , it must be on the edge $v_2 w_2$. So, a is near that edge, which forces a to be adjacent to v_1 and w_1 . This graph is N_{11} .

Now assume that G' is the simple graph $K_{3,3} + e$ illustrated in Figure 13 (right). The graph $G' - v_3$, shows us that both a and b are near w_1, w_2 , and w_3 . Similarly, $G' - w_3$ shows us that they are near v_3 and v_2 . Recall that b is adjacent to two of the degree-3 vertices of G' as well as two vertices formed by subdividing edges of G' .

Suppose b is adjacent to v_1 in $G - a$. Then b is adjacent to one of the w_i for $i \in \{1, 2, 3\}$, and by symmetry, we may assume w_1 . Since b is also near the other four vertices in G' , we may assume b 's other neighbors are vertices resulting from subdivisions of the edges w_2v_2 and v_3w_3 . Since a and b share at least one neighbor, we may assume (without loss of generality) that a is adjacent to the same vertex formed by subdividing w_3v_3 of G' .

There must be an additional subdivision of G' giving a neighbor of a . Since $\Delta(G) = 4$, the remaining two neighbors of a are drawn from $\{w_2, w_3\}$ and the vertex on v_2w_2 resulting from its subdivision. Suppose a is adjacent to w_2 and w_3 . As it must also be near v_2 and w_1 , it is also adjacent to a vertex formed by a subdivision of the edge v_2w_1 in G' . However, in this case v_2 has two neighbors of degree 3, a possibility ruled out at the beginning of the proof. This shows that, if a and b share exactly one neighbor, then b is not adjacent to v_1 . A similar argument starting with a instead of b shows that a is also not adjacent to v_1 , at least in the case where a and b share exactly one neighbor.

On the other hand, if we assume that a shares two neighbors with b , we can continue our search for a contradiction to the assertion that b is adjacent to v_1 . In this case, the common neighbors are the two vertices formed by subdividing v_2w_2 and v_3w_3 and a is adjacent to exactly one of w_2 and w_3 , say w_3 . Now, a must be near w_1 but if it is adjacent to a vertex formed by the subdivision of v_1w_1 or v_3w_1 , we again have the case of a degree-4 vertex with two degree-3 neighbors (v_1 and v_3 respectively). So it must be that a is adjacent to a vertex resulting from subdivision of the edge w_1v_2 . In this case, let x denote the common neighbor of a and b that is also a neighbor of v_3 and w_3 . Then $G - x, w_3$ is planar. This shows that b is not adjacent to v_1 .

So we know that b is not adjacent to v_1 in G' . Then without loss of generality it is adjacent to w_2 and w_3 . So, a is adjacent to w_1 or v_1 . If a is adjacent to v_1 , then a shares two neighbors with b . In other words, the vertices created by subdivisions in going from G' to $G - a, b$ that are neighbors of b are also neighbors of a . Since both a and b are near w_1 , suppose they are adjacent to a vertex resulting from subdivision of the edge v_1w_1 . Then since a is near w_2, w_3, v_2 , and v_3 , we may assume a is adjacent to vertices resulting from subdivisions of the edges w_2v_2 and v_3w_3 and that b is adjacent to one of these. However, in either case G has a degree-4 vertex with two degree-3 neighbors (v_3 and v_2 respectively).

Suppose instead that a and b are adjacent to a vertex produced by a subdivision of the edge v_2w_1 . (The symmetric case using the edge v_3w_1 will be similar.) Since a is near v_3 , it must be adjacent to a vertex formed by subdivision of the edge w_2v_3 or w_3v_3 (the other two options will not allow a to be near both w_2 and w_3). Without loss of generality it is w_3v_3 . Moreover, this forces b to share this neighbor as otherwise v_3 will have two degree-3 neighbors in G . The final neighbor of a makes a near w_2

but cannot lie on v_1w_2 or v_3w_2 lest we again have a vertex of degree 4 with two degree-3 neighbors. So a is adjacent to a vertex on the w_2v_2 edge. This is again N_{11} .

Finally, assume that neither a nor b is adjacent to v_1 in G' , that b is adjacent to w_2 and w_3 , and that a is adjacent to w_1 . The degree-3 vertices in G are then v_1 and the one adjacent to a formed by a subdivision of an edge in G' . Then the two subdivision vertices adjacent to b must also be adjacent to a . Since b is near w_1 , assume first that b is adjacent to a subdivision on the edge v_1w_1 in G' . Then the only way to make b near both v_2 and v_3 is by making it adjacent to a vertex formed by subdividing that edge. As a is also adjacent to that vertex, there is no way to make a near both w_2 and w_3 . So without loss of generality b (hence a) must be adjacent to a subdivision vertex on the edge v_2w_1 (as the symmetric case where a and b are adjacent to v_3w_1 is similar). Notice now that since a is near both w_2 and w_3 , either w_2 or w_3 will share a degree-3 neighbor with a . However, since they are both also neighbors of v_1 , we know that G will have a degree-4 vertex with two degree-3 neighbors and cannot be 2-apex. \square

Proposition 6.6. *If G is (11, 21) MMN2A, then G is in the Heawood family.*

Proof. Assume that G is an (11, 21) MMN2A graph. As we did in the previous cases, we may assume that the maximum vertex degree of G is 6 or less. Further, if G has more than one vertex of degree 6, then G is not MMN2A, since it must be the case that one of the degree-6 vertices has a degree-3 neighbor and removing such a vertex leaves one with a graph that simplifies to a graph that has no more than 14 edges, hence is not NA by [Theorem 1.4](#). This leaves us with the following degree sequences to consider: $(3^7, 5^3, 6)$, $(3^6, 4^2, 5^2, 6)$, $(3^5, 4^4, 5, 6)$, $(3^4, 4^6, 6)$, $(3^6, 4, 5^4)$, $(3^5, 4^3, 5^3)$, $(3^4, 4^5, 5^2)$, $(3^3, 4^7, 5)$, and $(3^2, 4^9)$.

We can throw out the first three sequences, since it is clear that the degree-6 vertex must have a neighbor of degree 3 and we find ourselves in the same situation as we were in at the beginning of this proof. Five of the remaining six sequences do in fact lead to an MMN2A graph and are treated in the five lemmas above.

This leaves only the degree sequence $(3^6, 4, 5^4)$. Suppose G is an MMN2A graph with this degree sequence. Each degree-5 vertex v has at most one degree-3 neighbor as otherwise $G - v$ simplifies to a graph of at most 14 edges and is not NA by [Theorem 1.4](#). This implies that the vertices of degrees 4 and 5, when considered separately, induce a K_5 subgraph, with four of the vertices having other neighbors in G . Choose $a, b \in V(G)$ such that $\deg(a) = \deg(b) = 5$, and consider $G - a, b$. Observe that the induced K_5 subgraph becomes a K_3 subgraph when a and b are removed and only two of its three vertices have neighbors in the rest of $G - a, b$. This means $(G - a, b)^s$ has nine edges, of which two are a double edge between the two remaining K_3 vertices. This graph is planar, which is a contradiction. Therefore there is no (11, 21) MMN2A graph G with degree sequence $(3^6, 4, 5^4)$. \square

7. 10-vertex graphs

We prove that a (10, 21) MMN2A graph is in the Heawood family. This is a corollary of the following proposition, originally proved in [Barsotti and Mattman 2013].

Proposition 7.1. *Let G be a graph with either $|V(G)| \leq 8$ or else $|V(G)| \leq 10$ and $|E(G)| \leq 21$. If G is N2A and a $Y\nabla$ move takes G to G' , then G' is also N2A.*

Proof. Since a graph of 20 or fewer edges is 2-apex [Mattman 2011], the only N2A graph with $|G| \leq 7$ is K_7 , which has no degree-3 vertices. So, the proposition is vacuously true for graphs of order 7 or less.

Suppose G is N2A with $|G| = 8$. As discussed in [Mattman 2011], G must be IK and we refer to the classification of such graphs due independently to Campbell et al. [2008] and Blain et al. [2007]. There are 23 IK graphs on eight vertices, but only four have a vertex of degree 3. In each case, a $Y\nabla$ move on that vertex results in K_7 , which is also N2A.

Again, graphs of size 20 or smaller are 2-apex. So, we can assume $\|G\| = 21$ and $|G| \geq 9$. If G is of order 9 and N2A, then, by [Mattman 2011, Proposition 1.6], G is a Heawood family graph (possibly with the addition of one or two isolated vertices). A $Y\nabla$ move results in the Heawood family graph H_8 or $K_7 \sqcup K_1$, both of which are N2A.

This leaves the case where $|G| = 10$. Assume G is a (10, 21) N2A graph that admits a $Y\nabla$ move to G' . For a contradiction, suppose G' is 2-apex with vertices a and b so that $G' - a, b$ is planar. Let v_0 be the degree-3 vertex in G at the center of the $Y\nabla$ move and v_1, v_2, v_3 the vertices of the resultant triangle in G' . Since G is N2A, it must be that $\{v_1, v_2, v_3\}$ is disjoint from $\{a, b\}$. Fix a planar representation of $G' - a, b$. The triangle $v_1 v_2 v_3$ divides the plane into two regions. Let H_1 be the induced subgraph on the vertices interior to the triangle and H_2 that of the vertices exterior. Then $|H_1| + |H_2| = 4$. Since G is N2A, there is an obstruction to converting the planar representation of $G' - a, b$ into a planar representation of $G - a, b$. This means that both H_1 and H_2 contain vertices adjacent to each of the triangle vertices $\{v_1, v_2, v_3\}$. In particular, H_1 and H_2 each have at least one vertex.

Suppose $|H_1| = |H_2| = 2$. The graph $G - b, v_1$ is nonplanar, but its subgraph $G - a, b, v_1$ is essentially a subgraph of $G' - a, b$ (with the addition of a degree-2 vertex v_0 on the edge $v_2 v_3$) and we will use the same planar representation for $G - a, b, v_1$ that we have for $G' - a, b$.

Since $G - b, v_1$ is not planar, there's an obstruction to placing a in the same plane. If we imagine putting a outside of a disk in the plane that covers $G - a, b, v_1$, we see that there is some vertex w in an H_i that is *hidden* from a . That is, although there's an edge $aw \in E(G)$, there is no path from a to w in the plane that avoids $G - b, v_1$. It follows that there's a cycle in $G - b, v_1$ with w in the interior and a on the exterior of the cycle.

Without loss of generality, the hidden vertex w is in $V(H_1) = \{c_1, d_1\}$, say $w = c_1$. This means we can assume that $c_1v_2d_1v_3$ is a 4-cycle in G , which, in the planar embedding of $G' - a, b$, is arranged with c_1 interior to the cycle $v_2d_1v_3$. However, since $G' - a, b$ is planar, this means c_1 is also hidden from v_1 and c_1v_1 is not an edge of the graph.

A similar argument using $G - b, v_2$ allows us to deduce a 4-cycle $c_2v_1d_2v_3$ using the vertices c_2 and d_2 of H_2 while showing $c_2v_2 \notin E(G)$. However, it follows that $G - b, v_3$ is planar, a contradiction.

So, we can assume $|H_1| = 3$, while H_2 consists of the vertex c_2 with $\{v_1, v_2, v_3\} \subset N(c_2)$. Suppose H_1 also has a vertex, c_1 , that is adjacent to all three triangle vertices. As $G - b, v_1$ is nonplanar, there's a vertex d_1 of H_1 that is hidden from a such that $c_1v_2d_1v_3$ is a cycle in G and $d_1v_1 \notin E(G)$. Similarly, $G - b, v_2$ shows that $c_1v_1e_1v_3$ is in G and e_1v_2 is not, e_1 being the third vertex of H_1 . Now, $G - b, v_3$ will be planar unless $d_1e_1 \in E(G)$. However, in that case, contracting d_1e_1 shows that $G' - a, b$ has a $K_{3,3}$ minor and is nonplanar, a contradiction.

In fact, the argument just given shows that there must be such a vertex $c_1 \in V(H_1)$ adjacent to all triangle vertices. That is, for $G - b, v_1$ to be nonplanar requires $x_1, x_2 \in V(H_1)$ so that $x_1v_2x_2v_3$ is a cycle, while $G - b, v_2$ gives vertices y_1, y_2 that form a cycle $y_1v_1y_2v_3$. Since $|H_1| = 3$, there are i, j so that $x_i = y_j$ and that vertex is adjacent to all v_i with $i = 1, 2, 3$.

We've shown that assuming G' is 2-apex leads to a contradiction. Thus, the proposition also holds in the case $|G| = 10$, which completes the proof. \square

Corollary 7.2. *If G is a (10, 21) MMN2A graph, then G is in the Heawood family.*

Proof. Suppose G is (10, 21) MMN2A. Recall that $\delta(G) \geq 3$ as otherwise a vertex deletion or edge contraction on a small-degree vertex gives a proper minor that is also N2A.

In [Mattman 2011], we showed that a graph of order 9 is MMN2A if and only if it is in the Heawood family. So, if G has a degree-3 vertex, then apply a $Y\nabla$ move at that vertex to get a graph G' . Then, by Proposition 7.1 and the classification of MMN2A graphs of order 9, G' is Heawood, whence G is too. So, we can assume $\delta(G) \geq 4$, which means the degree sequence of G is either $(4^8, 5^2)$ or $(4^9, 6)$.

Suppose there are vertices a and b such that $\|G - a, b\| = 11$. Then at least one of a and b has degree 5 or 6. Since $\delta(G) = 4$, we have that $\delta(G - a, b) \geq 2$ and $G - a, b$ is one of the graphs of Figure 14. In all three cases, both a and b must be adjacent to both v_3 and w_3 . For if, for example, a and v_3 are not adjacent, then $G - b, w_3$ would be planar. But, if a and b are adjacent to both, then v_3 and w_3 also have degree 5 in G , which contradicts the two given degree sequences for G . We conclude there is no choice of a and b such that $\|G - a, b\| = 11$.

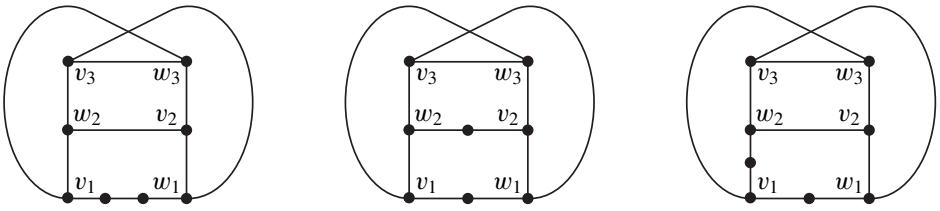


Figure 14. The three nonplanar $(8,11)$ graphs of minimum degree at least 2.

This means G must have degree sequence $(4^8, 5^2)$ with the two vertices of degree 5 adjacent and $G - a, b$ an $(8, 12)$ graph. There are two cases depending on whether or not a and b have a common neighbor in G . Suppose first that c is adjacent to both a and b . In $G - a, b$, vertex c will have degree 2 and we can contract an edge on c to arrive at either a $(7, 11)$ graph or else a multigraph with a doubled edge. Removing the extra edge if needed, let H denote the resulting $(7, 11)$ or $(7, 10)$ graph.

If H is $(7, 10)$, it is one of the two graphs of Figure 15. In the case of the graph on the left, the doubled edge must be that incident on the degree-1 vertex as $\delta(G - a, b) \geq 2$. But then the vertex labeled v_1 in the figure will have degree 5 in $G - a, b$, contradicting our assumption that a and b were the only vertices of degree greater than 4. So, we can assume H is the graph to the right in the figure. Up to symmetry, the doubled edge of H is either uv_1, v_1w_2 , or v_2w_2 . We'll examine the first case; the others are similar. Doubling uv_1 and adding back c leaves v_1 of degree 4 in $G - a, b$. Then $G - a, b, v_1$ simplifies to $K_{3,3} - v_1$. Since w_1, w_2 , and w_3 all have degree 3 in $G - a, b$, they each have exactly one of a and b as a neighbor in G . Suppose a is adjacent to w_2 . Then $G - a, v_1$ is planar, contradicting G being N2A. For the other two choices of edge doubling, one can again delete a resulting degree-4 vertex along with a or b to achieve a planar graph. So H being $(7, 10)$ leads to a contradiction.

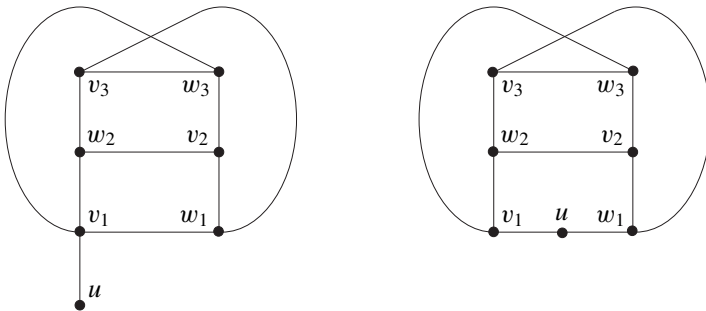


Figure 15. The two nonplanar $(7,10)$ graphs of minimum degree at least 1.

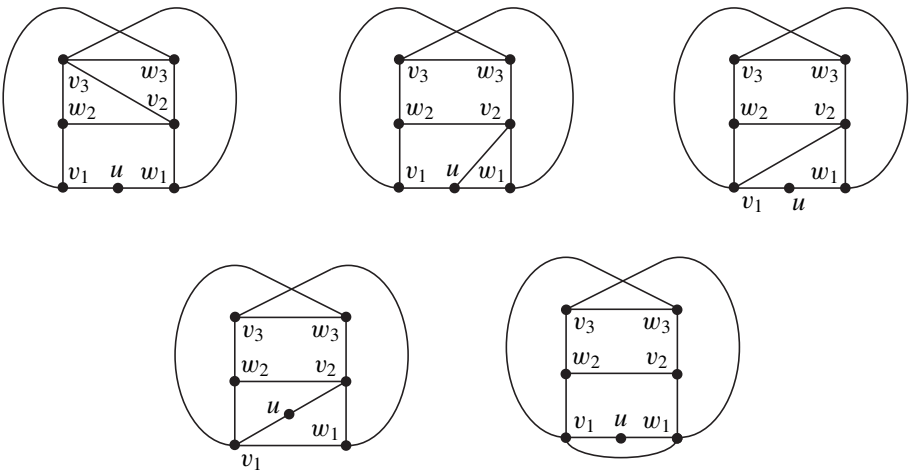


Figure 16. The five nonplanar (7,11) graphs of minimum degree at least 2.

If H is (7, 11), then $\delta(H) = \delta(G - a, b) \geq 2$ and H is one of the five graphs of Figure 16. Here we use a similar approach. Deleting one of the degree-4 vertices of H , call it x , results in a graph $G - a, b, x$ that simplifies to $K_{3,3} - v_1$. Since each of the degree-3 vertices of H is adjacent to exactly one of a and b , there will be an appropriate choice from those two, say a , such that $G - a, x$ is planar, which is a contradiction. So, H being (7, 11) is not possible and we conclude that there is no such vertex c that is adjacent to both a and b .

This means that $G - a, b$ is a nonplanar cubic graph (i.e., 3-regular) on eight vertices. There are two such graphs, shown in Figure 17. If $G - a, b$ is the graph to the left in Figure 17, note that the vertex labeled v is adjacent to exactly one of a and b , say a . Then $G - a, w$ is planar.

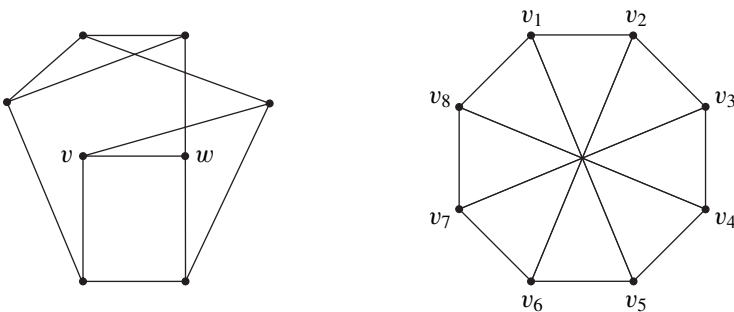


Figure 17. The two nonplanar cubic graphs of order 8.

Finally, assume that $G - a, b$ is the graph to the right in [Figure 17](#). Note that each vertex of $G - a, b$ is adjacent to exactly one of a and b in G . If a and b are adjacent to alternate vertices in the 8-cycle (for example if $\{v_1, v_3, v_5, v_7\} \subset N(a)$ and $\{v_2, v_4, v_6, v_8\} \subset N(b)$), we obtain graph 20 of [Figure 1](#), a Heawood family graph. If not, then we must have two consecutive vertices, say v_1 and v_2 , that share the same neighbor in $\{a, b\}$, say a . That is, we can assume $av_1, av_2 \in E(G)$. Then $G - a, v_3$ is planar, contradicting G being N2A.

In summary, if G of order 10 is N2A with $\delta(G) > 3$, it must be graph 20 of the Heawood family. □

Acknowledgments

This research was supported in part by a Provost's Research and Creativity Award and a Faculty Development Award from California State University, Chico. We thank the referee for a very close reading of an earlier version of the paper that resulted in many substantial improvements.

References

- [Barsotti and Mattman 2013] J. Barsotti and T. W. Mattman, "Intrinsically knotted graphs with 21 edges", preprint, 2013. [arXiv 1303.6911](#)
- [Blain et al. 2007] P. Blain, G. Bowlin, T. Fleming, J. Foisy, J. Hendricks, and J. Lacombe, "Some results on intrinsically knotted graphs", *J. Knot Theory Ramifications* **16**:6 (2007), 749–760. [MR 2341314](#) [Zbl 1162.57003](#)
- [Campbell et al. 2008] J. Campbell, T. W. Mattman, R. Ottman, J. Pyzer, M. Rodrigues, and S. Williams, "Intrinsic knotting and linking of almost complete graphs", *Kobe J. Math.* **25**:1-2 (2008), 39–58. [MR 2509265](#) [Zbl 1197.57005](#)
- [Conway and Gordon 1983] J. H. Conway and C. M. Gordon, "Knots and links in spatial graphs", *J. Graph Theory* **7**:4 (1983), 445–453. [MR 722061](#) [Zbl 0524.05028](#)
- [Ding and Dziobak 2016] G. Ding and S. Dziobak, "Excluded-minor characterization of apex-outerplanar graphs", *Graphs Combin.* **32**:2 (2016), 583–627.
- [Foisy 2002] J. Foisy, "Intrinsically knotted graphs", *J. Graph Theory* **39**:3 (2002), 178–187. [MR 1883594](#) [Zbl 1176.05022](#)
- [Goldberg et al. 2014] N. Goldberg, T. W. Mattman, and R. Naimi, "Many, many more intrinsically knotted graphs", *Algebr. Geom. Topol.* **14**:3 (2014), 1801–1823. [MR 3212585](#) [Zbl 1292.05091](#)
- [Hanaki et al. 2011] R. Hanaki, R. Nikkuni, K. Taniyama, and A. Yamazaki, "On intrinsically knotted or completely 3-linked graphs", *Pacific J. Math.* **252**:2 (2011), 407–425. [MR 2860431](#) [Zbl 1232.57003](#)
- [Johnson et al. 2010] B. Johnson, M. E. Kidwell, and T. S. Michael, "Intrinsically knotted graphs have at least 21 edges", *J. Knot Theory Ramifications* **19**:11 (2010), 1423–1429. [MR 2746195](#) [Zbl 1232.57004](#)
- [Kohara and Suzuki 1992] T. Kohara and S. Suzuki, "Some remarks on knots and links in spatial graphs", pp. 435–445 in *Knots 90* (Osaka, 1990), edited by A. Kawachi, de Gruyter, Berlin, 1992. [MR 1177440](#) [Zbl 0771.57002](#)

- [Lee et al. 2015] M. J. Lee, H. J. Kim, H. J. Lee, and S. Oh, “Exactly fourteen intrinsically knotted graphs have 21 edges”, *Algebr. Geom. Topol.* **15**:6 (2015), 3305–3322. MR MR3450762 Zbl 06535762
- [Mattman 2011] T. W. Mattman, “Graphs of 20 edges are 2-apex, hence unknotted”, *Algebr. Geom. Topol.* **11**:2 (2011), 691–718. MR 2782541 Zbl 1216.05017
- [Meringer 1999] M. Meringer, “Fast generation of regular graphs and construction of cages”, *J. Graph Theory* **30**:2 (1999), 137–146. MR 1665972 Zbl 0918.05062
- [Morris 2008] C. Morris, *A classification of all connected graphs on seven, eight, and nine vertices with respect to the property of intrinsic knotting*, master’s thesis, California State University, Chico, 2008, available at <http://www.csuchico.edu/~tmattman/CMTthesis.pdf>.
- [Ozawa and Tsutsumi 2007] M. Ozawa and Y. Tsutsumi, “Primitive spatial graphs and graph minors”, *Rev. Mat. Complut.* **20**:2 (2007), 391–406. MR 2351115 Zbl 1142.57004
- [Pierce 2014] M. Pierce, *Searching for and classifying the finite set of minor-minimal non-apex graphs*, honor’s thesis, California State University, Chico, 2014, available at <http://www.csuchico.edu/~tmattman/mpthesis.pdf>.
- [Robertson and Seymour 2004] N. Robertson and P. D. Seymour, “Graph minors, XX: Wagner’s conjecture”, *J. Combin. Theory Ser. B* **92**:2 (2004), 325–357. MR 2099147 Zbl 1061.05088
- [Robertson et al. 1995] N. Robertson, P. Seymour, and R. Thomas, “Sachs’ linkless embedding conjecture”, *J. Combin. Theory Ser. B* **64**:2 (1995), 185–227. MR 1339849 Zbl 0832.05032
- [Wagner 1937] K. Wagner, “Über eine Eigenschaft der ebenen Komplexe”, *Math. Ann.* **114**:1 (1937), 570–590. MR 1513158 Zbl 0017.19005

Received: 2015-01-12

Revised: 2015-06-23

Accepted: 2015-08-17

jbarsott@ucsc.edu

*Department of Mathematics, University of California,
Santa Cruz, CA 95064, United States*

tmattman@csuchico.edu

*Department of Mathematics and Statistics, California State
University, Chico, CA 95929-0525, United States*

Mathematical modeling of a surface morphological instability of a thin monocrystal film in a strong electric field

Aaron Wingo, Selahittin Cinar, Kurt Woods and Mikhail Khenner

(Communicated by Natalia Hritonenko)

A partial differential equation (PDE)-based model combining the effects of surface electromigration and substrate wetting is developed for the analysis of the morphological instability of a monocrystalline metal film in a high temperature environment typical to operational conditions of microelectronic interconnects and nanoscale devices. The model accounts for the anisotropies of the atomic mobility and surface energy. The goal is to describe and understand the time-evolution of the shape of the film surface. The formulation of a nonlinear parabolic PDE problem for the height function $h(x, t)$ of the film in the electric field is presented, followed by the results of the linear stability analysis of a planar surface. Computations of a fully nonlinear evolution equation are presented and discussed.

1. Introduction

The drift of ionized adsorbed atoms (adatoms) on a metal or semiconductor crystal surface due to their interaction with the “electron wind” is termed *surface electromigration*. The “wind” force on adatoms is the effect of a high-density direct current through the bulk of a crystal, which also heats up the surface—thus increasing the adatoms’ own kinetic energy. It is this combination that makes adatoms drift. Surface electromigration was studied theoretically in connection to the grain-boundary grooving in polycrystalline films [Averbuch et al. 2001; Maroudas 1995], the kinetic instabilities of crystal steps [Chang et al. 2006; Debierre et al. 2007; Stoyanov 1997], morphological stability of thin films [Dobbs and Krug 1994; Krug and Schimschak 1997; Barakat et al. 2012; Khenner 2013], and recently, as a way to fabricate nanometer-sized gaps in metallic films—suitable for testing of the conductive properties of single molecules and control of their functionalities [Barnes et al. 2010; Bolotin et al. 2007; Block et al. 2006]. Although the phenomenon of electromigration has been known for over 100 years, it became of practical interest

MSC2010: 35R37, 35Q74, 37N15, 65Z05, 74H55.

Keywords: nonlinear evolution PDEs, electromigration, surface diffusion, morphology, stability.

in 1966 when the first integrated circuits became commercially available. It is considered a key factor in determining the reliability of integrated circuits.

As we just mentioned, one recent technological application of electromigration is the fabrication of the nanoscale contacts (gaps) that are manufactured from the thin Ag films wetting the Si substrate [Barnes et al. 2010; Bolotin et al. 2007; Block et al. 2006]. The gap between contacts can be cyclically opened and closed. To open the contact, a strong electric current is applied at a low temperature of the film (~ 80 K), which enables the surface mass flow of adatoms across the narrow bridge, thus connecting the anode and the cathode, until the bridge breaks. To close the contact, the natural surface diffusion of adatoms across the gap is enabled by heating the film to the room temperature, all the while keeping the electric current.

Another example, more relevant to the present study, is a faceting of the initially planar surface of a crystalline thin film upon passing the current along the substrate. This way the so-called quantum wires can be fabricated [Dai et al. 2014]. Since the cross-section of a quantum wire is only a few nanometers, it possesses very special electronic properties, which makes it desirable for integration into nanoscale devices.

Here we further develop the PDE-based mathematical model of the film-surface morphological instability and evolution driven by the electromigration [Khenner 2013]. The very special feature of the presented model is that it accounts both for the wetting and the surface energy anisotropy effects. The surface morphological instability and evolution in a thin film system where the wetting, anisotropy, and electromigration are active have not been addressed theoretically, although PDE-based models of wetting and anisotropy [Davis et al. 2004; Gill and Wang 2008; Khenner 2008a; 2008b; Khenner et al. 2011], wetting and electromigration [Khenner 2013], and electromigration and anisotropy [Barakat et al. 2012] have been published.

The wetting effect emerges due to the existence of the attractive force between the adatoms and the substrate atoms; this force is nonnegligible because of the very small thickness of the film ($h \sim 10$ nm). The surface-energy anisotropy effect emerges due to the crystal nature of the film surface. The combination of the two effects results in a complicated nonlinear evolution PDE. We use the approach of [Khenner 2013] to build and analyze the model with the added anisotropy effect; first, the governing PDE is derived, and then we analytically obtain the stability regions of the planar surface in the space of the physical parameters and, for the values of the parameters such that the planar surface is unstable, compute the evolution of the small, one-wavelength surface perturbation on a periodic domain. The typical evolution scenarios, such as the evolution to a steady-state or the lateral surface drift, are presented.

2. Problem statement

We assume a simple one-dimensional geometry, where the surface is an open curve (without overhangs) in the xz -plane, described by a function $z = h(x)$. Since the

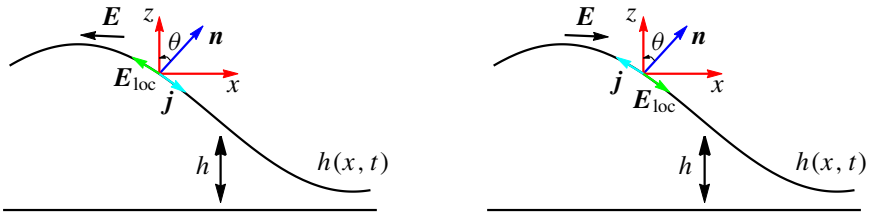


Figure 1. Sketch of the film surface $h(x, t)$ in the horizontal, constant electric field \mathbf{E} . Here $\mathbf{E}_{\text{loc}} = \mathbf{E} \cos \theta$ is the projection of \mathbf{E} on the surface. The surface atomic flux \mathbf{j} is in the direction opposite to \mathbf{E}_{loc} .

curve deforms with time, h is also a function of time t ; that is, $z = h(x, t)$. As is common in physics literature, the curve $h(x, t)$ is termed the *film surface* in the following, despite being a one-dimensional object.

Following [Khenner 2013], we focus on the case of the horizontal electric field (directed along the substrate and the initially planar film surface $h(x, 0) = \text{const.}$). As was stated in the [Introduction](#), we will incorporate the effects of substrate wetting by the film [Khenner 2008a], the anisotropy of the diffusional mobility $M(\theta)$ [Krug and Schimschak 1997], and weak anisotropy of the surface energy $\gamma(\theta)$ [Liu and Metiu 1993], where θ is the angle that the unit normal \mathbf{n} to the surface makes with the vertical coordinate axis z . From the mathematical standpoint, these effects will manifest in our model through various linear and nonlinear terms in the parabolic PDE for $h(x, t)$. The physicomathematical framework in which our model is firmly rooted has been established, beginning in 1960s, through the efforts of many prominent materials scientists, physicists, and mathematicians [Mullins 1963; Cahn et al. 1992; Cahn and Taylor 1994; Di Carlo et al. 1992; Dobbs and Krug 1994; Liu and Metiu 1993; Davis et al. 2004]. The mathematics of the accounting for the relevant physical effects are summarized below, and the physical foundations, as well as further mathematical detail, can be found in the cited papers. To illustrate the effects of the electromigration on film morphology, [Figure 1](#) depicts two directions of the electric field. In [Figure 1](#) (left), the electric current forces the adatoms downhill (from the crest to the trough); thus the surface becomes more planar with time. In [Figure 1](#) (right), the field in the opposite direction forces the adatoms uphill (from the trough to the crest) and thus the surface becomes less planar. This is the instability mechanism that we are investigating in this paper.

The dimensionless PDE governing the evolution of $h(x, t)$ has the form

$$h_t = B[M(h_x)(1 + h_x^2)^{-1/2}\mu_x]_x + A[M(h_x)(1 + h_x^2)^{-1/2}]_x. \quad (1)$$

In (1), $B > 0$ is the effective diffusivity of adatoms and $A \geq 0$ is the strength of the electric field. The first term on the right-hand side stems from the natural diffusion of adatoms (in the absence of the electric current) on a heated crystal surface. The

meaning of this term is that relocation of adatoms through diffusion changes the shape of the surface. It was first derived by Mullins [1963] in what is now considered the classical work. Similarly, the second term stems from the forced diffusion (drift) of adatoms caused by the electromigration force. It was derived in several papers, including [Dobbs and Krug 1994; Krug and Schimschak 1997; Khennner 2013]. The surface chemical potential $\mu(x, t)$ entering the Mullins term contains, in our model, the contributions from wetting (through the dependence of the surface energy γ on the film thickness h ; see (3) below) and anisotropy. The expression for $\mu(x, t)$ reads

$$\mu = (\gamma + \gamma_{\theta\theta})\kappa + (\gamma_h - h_x \gamma_{h\theta}) \cos \theta, \quad \cos \theta = (1 + h_x^2)^{-1/2}, \quad \kappa = -h_{xx}(1 + h_x^2)^{-3/2}. \quad (2)$$

Here κ is the curvature, and $\gamma(h, \theta)$ is the weakly anisotropic film-surface energy (tension):

$$\gamma(h, \theta) = 1 + \epsilon_\gamma \cos 4\theta + (G - 1 - \epsilon_\gamma \cos 4\theta)e^{-h}, \quad \theta = \arctan h_x, \quad (3)$$

where $G > 1$ and $0 \leq \epsilon_\gamma < \frac{1}{15}$ are the parameters; G is the ratio of the (dimensional) substrate energy to the (dimensional) surface energy, and ϵ_γ is the strength of the anisotropy. The interval for ϵ_γ implies that the “stiffness” $\gamma + \gamma_{\theta\theta}$ is larger than 0 for all θ in (2) when $\gamma = \gamma(\theta) = 1 + \epsilon_\gamma \cos 4\theta$ (the four-fold anisotropy typical for most semiconductor and metal crystals). This implies a negative effective “diffusivity” α_1 in the linearized PDE (1) (for $A = 0$): $h_t = \alpha_1 h_{xxxx}$, where $\alpha_1 < 0$. Such a linear PDE is well-posed; i.e., it is forward parabolic. If $\epsilon_\gamma \geq \frac{1}{15}$ (strong anisotropy, typical at comparatively low temperatures), then the PDE is backward parabolic for some θ -intervals and the regularization is required; usually the curvature-squared term is added to $\gamma(\theta)$ [Di Carlo et al. 1992], which raises the PDE order from the fourth to the sixth. In the presence of the electric current, the crystal temperature is high due to Joule heating, which justifies the restriction of the consideration to mild anisotropy. The choice $G > 1$ means that only wetting films are considered; i.e., the substrate energy is larger than the surface energy. Thus dewetting, meaning the substrate exposure, may occur only through the application of the external force, such as the electromigration.

The form of (3) results from the consideration of the conventional “two-layer” model for the film energy; for the discussion of that model see, for instance, [Davis et al. 2004] and the references therein. The parameters and their typical range of values are displayed in Table 1. Notice that the classical Mullins model assumes $\gamma = \text{const.}$ (the isotropic case without the wetting effect); thus the chemical potential reduces to $\mu = \gamma\kappa$. The form of the wetting potential contribution to the surface energy, $(\gamma_h - h_x \gamma_{h\theta}) \cos \theta$, is well established and is taken from [Davis et al. 2004].

In reference to the electromigration term in (1),

$$M(h_x) = \frac{1 + \beta \cos^2(N(\arctan h_x + \phi))}{1 + \beta \cos^2(N\phi)}, \quad \text{where } \beta, N, \phi = \text{const.}, \quad (4)$$

Physical parameters	Typical values	Range	Physical meaning
B	8	fixed	effective adatoms diffusivity
$M(0)$	1	fixed	adatoms mobility on the horiz. surface
h_0	3	$0 \leq h_0 \leq 20$	initial height of the film (same for all x)
A	72	$10 \leq A \leq 1000$	strength of the electric field
G	2	$1 < G \leq 100$	ratio of the substrate energy to the surface energy
$M'(0)$	-3	$-10 \leq M'(0) \leq 0$	derivative of adatoms' mobility on the horiz. surface

Table 1. Values of the dimensionless physical parameters. The definitions of these parameters in terms of the dimensional quantities can be found in [Khenner 2013; 2008b; Khenner et al. 2011]. The typical values in the second column result from the substitutions in these expressions of the published standard values of the dimensional parameters [Mullins 1963; Maroudas 1995; Dobbs and Krug 1994; Krug and Schimschak 1997; Davis et al. 2004; Liu and Metiu 1993], which have been measured in the experiments.

is the anisotropic diffusional mobility (notice that the denominator of the fraction is a constant value for the given nonnegative parameters β , N and ϕ); here β is the anisotropy strength, N is the number of crystallographic symmetry axes and ϕ is the angle between a symmetry direction and the average surface orientation. In this paper, β varies (resulting in a variation of $M'(0)$; see Table 1), $N = 4$, and $\phi = \pi/16$. Notice that $M(0) = 1$ for any β , N , ϕ . Equation (4) is taken from [Krug and Schimschak 1997].

Next, we begin by linearizing $M(h_x)$ about $h_x = 0$; i.e., we write $M(h_x) = M(0) + M'(0)h_x$, where $M(0)$ and $M'(0)$ will be later calculated from (4) for given β , N and ϕ (see [Khenner 2013]). Then

$$\frac{\partial M(h_x)}{\partial x} = \frac{\partial M(h_x)}{\partial h_x} h_{xx} = M'(0)h_{xx}, \quad (5)$$

and (1) now reads

$$h_t = BM'(0)h_{xx}(1+h_x^2)^{-1/2}\mu_x + B(M(0) + M'(0)h_x)\left[(1+h_x^2)^{-1/2}\mu_x\right]_x + AM'(0)h_{xx}(1+h_x^2)^{-1/2} + A(M(0) + M'(0)h_x)\left[(1+h_x^2)^{-1/2}\right]_x. \quad (6)$$

In order to compute μ_x in (6), we first calculate $\gamma_{\theta\theta}$, γ_h , $\gamma_{h\theta}$ using (3). Then we substitute these expressions in (2), use the trigonometric identities

$$\cos 4\theta = 8(\cos^4\theta - \cos^2\theta) + 1, \quad \sin 4\theta = 4 \sin \theta \cos \theta (2 \cos^2\theta - 1),$$

(where $\cos \theta = (1 + h_x^2)^{-1/2}$ and $\sin \theta = h_x(1 + h_x^2)^{-1/2}$) and obtain $\mu(x, t)$ in terms of h_x , h_x^2 , h_{xx} , etc. We then substitute $\mu(x, t)$ into (6) and the remaining differentiations are performed.

Finally, we employ the small-slope approximation. The spatial derivatives are replaced as $\partial/\partial x^k \rightarrow \epsilon^k \partial/\partial x^k$, the coefficients of the powers of ϵ are collected, and all but the coefficients of ϵ^k , $k = 1, 2, 3, 4$, are set to zero. Then, ϵ is set equal to 1. This results in the fourth-order, nonlinear PDE for $h(x, t)$:

$$\begin{aligned} h_t = & BM(0)(15\epsilon_\gamma - 1 + (1 - G - 15\epsilon_\gamma)e^{-h})h_{xxxx} \\ & + (AM'(0) - BM(0)(1 - G + \epsilon_\gamma)e^{-h})h_{xx} \\ & - Ah_{xx}(M(0)h_x + \frac{3}{2}h_x^2M'(0)) + Be^{-h}F, \end{aligned} \quad (7)$$

where

$$\begin{aligned} F = & M'(0)h_x^3(1 - G + \epsilon_\gamma) - 2M'(0)h_xh_{xx}(1 - G + \epsilon_\gamma) - M(0)h_x^4(1 - G - 7\epsilon_\gamma) \\ & + M(0)h_x^2(1 - G + \epsilon_\gamma) + 5M(0)h_x^2h_{xx}(1 - G - \frac{51}{5}\epsilon_\gamma) \\ & - 3M(0)h_xh_{xxx}(1 - G - 15\epsilon_\gamma) - 2M(0)h_{xx}^2(1 - G - 15\epsilon_\gamma). \end{aligned} \quad (8)$$

The first and second lines of (7) are composed of the linear contributions, while all terms in the third line are nonlinear (i.e., they are proportional to the products of the spatial derivatives of h). The terms in the first line emerge due to the natural diffusions of adatoms, mediated by a surface/substrate-interaction force, on a heated crystal surface with anisotropic surface energy. In the second line, the linear term that is proportional to B is also due to the natural diffusion mediated by the wetting effect, while another linear term there that is proportional to A is due to the electromigration drift of adatoms. In the third line, the two terms that are proportional to A also are due to the electromigration. Finally, the last contribution in the third line, $Be^{-h}F$, is the nonlinearity produced by the substrate wetting effect. This contribution, as well as the linear terms that are proportional to e^{-h} in the first and second lines of (7), drop out in the limit of a thick film, $h \rightarrow \infty$, where the film surface/substrate-interaction force vanishes. When $\epsilon_\gamma = 0$, equations (7) and (8) are reduced to [Khennner 2013, (15)]. In the following, it is important that the coefficients of the linear terms are negative, due to negativity of $M'(0)$ and weak anisotropy, $0 < \epsilon_\gamma < \frac{1}{15}$.

2.1. Example: analysis of a linear second-order PDE. Equation (7) is a well-posed, fourth-order, nonlinear parabolic PDE. The prototype *linear* fourth-order parabolic PDE is

$$h_t = \alpha_1 h_{xxxx} + \alpha_2 h_{xx}, \quad \alpha_1, \alpha_2 < 0. \quad (9)$$

This equation has the trivial solution $h(x, t) = h_0 = \text{const}$. In the physical context, this solution corresponds to a constant-height film for all values of x and t , that is, a film with a planar stationary surface. We call such a solution an equilibrium surface. The key issue is whether the equilibrium is stable or unstable with respect to small perturbations $\xi(x, t)$. This can be settled by substituting $h = h_0 + \xi(x, t)$ and then assuming $\xi(x, t)$ is a single Fourier mode: $\xi(x, t) = \xi_0 e^{\omega t} e^{ikx}$, where ξ_0 is the amplitude, $\omega(k)$ is the growth rate, and k is the wavenumber. (The wavelength, or

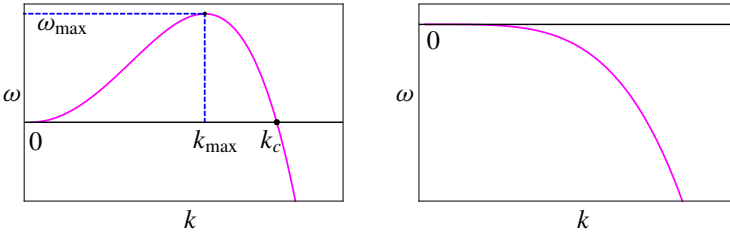


Figure 2. Two cases of the typical growth rate $\omega(k)$. Left: long-wave instability. Right: stability, $\omega(k) < 0$ for all k .

the spatial period, is $\lambda = 2\pi/k$.) Then one obtains the expression for the perturbation growth rate as a function of the wavenumber, the so-called dispersion relation

$$\omega(k) = \alpha_1 k^4 - \alpha_2 k^2. \quad (10)$$

For small k , the second term is dominant in this expression. For large k , it is the first term. Since $\alpha_2 < 0$, perturbations with small wavenumbers (large wavelengths) grow ($\omega(k) > 0$); because $\alpha_1 < 0$, perturbations with large wavenumbers (small wavelengths) decay ($\omega(k) < 0$). This is reflected in the shape of the curve $\omega(k)$ (see Figure 2 (left)), and correspondingly the instability is termed the *long-wavelength instability*. All perturbations with wavenumbers in the interval $0 < k < k_c$ grow, and all perturbations with wavenumbers greater than k_c decay; k_c is termed the *instability cut-off wavenumber*. The surface is unstable with respect to long-wavelength perturbations, and it is stable with respect to small-wavelength perturbations. In practice, the perturbation (induced, for instance, by a thermal noise) is not a single Fourier mode. However, most perturbations can be represented by a superposition of Fourier modes. Thus some modes grow and some decay. Among the unstable modes, there is a mode with the largest growth rate, ω_{\max} . This *most dangerous* mode will dominate over other modes shortly after the surface is destabilized, resulting in a surface deformation of the form $h(x, t) = h_0 + \xi_0 e^{\omega_{\max} t} \cos k_{\max} x$. Here, k_{\max} is the wavenumber for which $\omega = \omega_{\max}$, i.e., the maximum of $\omega(k)$ on the interval $0 \leq k \leq k_c$. In other words, k_{\max} is the positive solution of $d\omega/dk = 0$.

It is easy to show that for (10), we have $k_{\max} = k_c/\sqrt{2}$. First, we set the right-hand side of (10) to zero and solve for k :

$$\omega(k) = -\alpha_2 k^2 + \alpha_1 k^4 = 0 \implies k^2(-\alpha_2 + \alpha_1 k^2) = 0 \implies k = 0 \text{ or } k = \pm\sqrt{\alpha_2/\alpha_1}.$$

Since we need a positive solution, $k_c = \sqrt{\alpha_2/\alpha_1}$. To determine k_{\max} , we solve $d\omega/dk = 0$ for k ; that is,

$$-2\alpha_2 k + 4\alpha_1 k^3 = 0 \implies 2k(-\alpha_2 + 2\alpha_1 k^2) = 0 \implies k = 0 \text{ or } k = \pm\sqrt{\alpha_2/2\alpha_1}.$$

Again we take the positive solution. Thus, $k_{\max} = \sqrt{\alpha_2/2\alpha_1} = k_c/\sqrt{2}$.

Notice that the parameter α_1 in (9) cannot be positive. Otherwise, the short-wavelength perturbations will grow, which is not physically permissible, since in this case the surface is always unstable — such perturbations would be always present in the spectrum. However, an instability is not universal, and most material surfaces remain planar. Mathematically, (9) in the case of $\alpha_1 > 0$ is ill-posed; despite its higher order, it is similar to the (ill-posed) backward heat equation $h_t = -h_{xx}$. However, the parameter α_2 may be positive for some physical parameters' values. Then $\omega(k)$ is negative for all k (Figure 2 (right)), meaning that all perturbations decay and the surface restores its initial planar shape.

Equations such as (7) are nonlinear; thus the exponential growth of the most dangerous mode will not continue forever. Nonlinear terms in the equation will dampen growth, which usually results in a stationary, nontrivial solution which has the spatial form resembling the large-amplitude cosine curve. Determination of the stability of (7) and the form of the stationary solution will be discussed next.

3. Linear stability analysis

The dynamics of the film surface are governed by the nonlinear PDE (7). Toward our goal of determining stability of the surface with respect to small perturbations, we notice that (7) has the equilibrium solution $h = h_0 = \text{const.}$, and we linearize about this solution along the lines described above for (9). First, using the general small perturbation $\xi(x, t)$, we substitute $h = h_0 + \xi(x, t)$ in (7) and retain only the linear terms in ξ . Then, we substitute $\xi = \xi_0 e^{\omega t} e^{ikx}$, calculate the partial derivatives and divide out the factor $\xi_0 e^{\omega t}$. This results in the dispersion relation

$$\omega(k) = -BM(0)((G - 1 + 15\epsilon_\gamma)e^{-h_0} + 1 - 15\epsilon_\gamma)k^4 - (BM(0)(G - 1 - \epsilon_\gamma)e^{-h_0} + AM'(0))k^2. \quad (11)$$

3.1. Analysis of the dispersion relation (11). In this section we determine how the physical parameters of the problem affect the surface instability.

As we explained in Section 2, if $\omega(k) < 0$ for all k , then the surface is stable with respect to the perturbations of any wavenumber (Figure 2 (right)). When this condition does not hold, the surface is long-wave unstable (Figure 2 (left)). The degree of the instability is measured by the width of the domain under the dispersion curve $\omega(k)$. That is, the larger the cut-off wavenumber k_c is, the stronger the instability.

We notice that the dispersion relation (11) has the form of (10) and thus we identify

$$\begin{aligned} \alpha_1 &= -BM(0)((G - 1 + 15\epsilon_\gamma)e^{-h_0} + 1 - 15\epsilon_\gamma), \\ \alpha_2 &= BM(0)(G - 1 - \epsilon_\gamma)e^{-h_0} + AM'(0). \end{aligned}$$

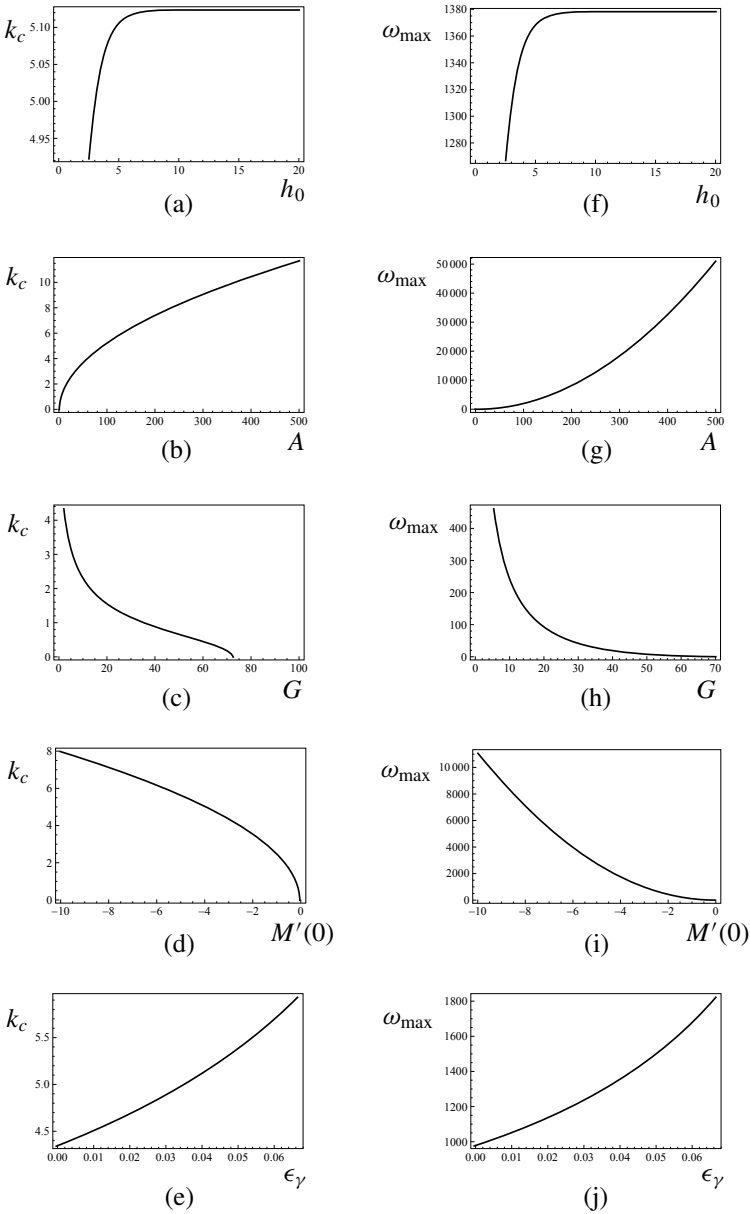


Figure 3. Characterization of the film linear stability. (a) k_c vs. h_0 ; (b) k_c vs. A ; (c) k_c vs. G ; (d) k_c vs. $M'(0)$; (e) k_c vs. ϵ_γ . In (f)–(j), ω_{\max} is plotted vs. the same variables. In each panel, all parameters except the single one that is varied are fixed to the typical values from the second column of Table 1. In (a)–(j), ϵ_γ is chosen equal to zero (isotropic evolution). The same strategy with regard to parameters is followed in Figures 4 and 5.

Then the expressions for k_c , k_{\max} and ω_{\max} are

$$k_c = \sqrt{\frac{Ae^{h_0}M'(0) + BM(0)(G - 1 - \epsilon_\gamma)}{BM(0)(1 - G - 15\epsilon_\gamma + (15\epsilon_\gamma - 1)e^{h_0})}}, \quad k_{\max} = \frac{k_c}{\sqrt{2}}, \quad (12)$$

$$\omega_{\max} = \frac{1}{4} \frac{(Ae^{h_0}M'(0) + BM(0)(G - 1) - BM(0)\epsilon_\gamma)^2}{BM(0)e^{h_0}(G - 1 + 15\epsilon_\gamma + (1 - 15\epsilon_\gamma)e^{h_0})}.$$

The wavelength of the most dangerous perturbation is $\lambda_{\max} = 2\pi/k_{\max}$.

The expressions for k_c , k_{\max} and ω_{\max} include new contributions due to the surface energy anisotropy (the terms proportional to ϵ_γ).

The film stability *decreases* with increasing h_0 , and this trend saturates around $h_0 = 5.12$ nm (Figure 3(a)). This is because the film wets the substrate and thus the attractive, cohesive force between the adatoms and the substrate atoms is stronger for thinner films (smaller h_0). Increasing the electric field strength A also makes the film less stable, but increasing G makes it more stable, since the substrate energy provides a stabilizing effect (Figures 3(b)–(c)). The stability of the film *decreases* with increasing $|M'(0)|$ and ϵ_γ (Figures 3(d)–(e)). The results in the panels (a)–(d) and (f)–(i) were obtained also in [Khennner 2013]; here these results are recomputed from (7). The results in the panels (e) and (j) are new; they stem from the new feature of the extended model: the accounting for the mild anisotropy of the film-surface energy.

4. Numerical solution of equation (7)

Using the information from the previous section on how the physical parameters affect the surface stability, in this section we compute the full nonlinear PDE (7) by implementing the method of lines (MOL) [Verwer and Sanz-Serna 1984; Schiesser 1994] in *Mathematica* [Wolfram 2016]. The MOL is a technique for solving partial differential equations by discretizing in all but one dimension, and then integrating the semidiscrete problem as a system of ordinary differential equations (ODEs). A significant advantage of the method is that it allows us to use the sophisticated general-purpose software [Hairer and Wanner 1999; Brown et al. 1989] that has been developed for numerically integrating large systems of ODEs. For the parabolic initial-boundary-value problems, the MOL typically is very efficient and accurate. Sophisticated adaptive MOL methods were also developed for some hyperbolic equations [Saucez et al. 2001].

The initial condition is the perturbation of $h = h_0$, and the boundary conditions are periodic:

$$h(x, 0) = h_0 + \delta \cos(k_{\max}x), \quad h(0, t) = h(\lambda_{\max}, t), \quad h'(0, t) = h'(\lambda_{\max}, t), \quad (13)$$

where δ is a small amplitude (we take $\delta = 0.01$). Periodic boundary conditions are used since the goal is to compute the evolution of a finite section of a periodic,

laterally unbounded surface. Notice that $h(x, 0)$ is the most dangerous (fastest growing) unstable perturbation according to the linear stability analysis in [Section 3](#).

Evolution of the perturbation is computed until the steady-state solution emerges. The steady-state solution is either a stationary or a traveling wave of wavelength λ_{\max} and constant amplitude. The amplitude and wave speed are studied as a function of the parameters A , h_0 , G , $M'(0)$, and ϵ_γ . It is important to emphasize here that the traveling wave is a nonlinear effect (which was overlooked in [\[Khenner 2013\]](#)); indeed, the perturbation growth rate $\omega(k)$ is real-valued (see [\(11\)](#)), indicating the absence of a linear traveling wave. In computations, the profile started to shift laterally only when the amplitude of a cosine-like perturbation became fairly large, indicating that nonlinearities in [\(7\)](#) are responsible. Another important observation is that for thick films, $h \gg 1$, the wave speed is zero; thus the traveling wave solution is caused by the nonlinear effect of substrate wetting, which is described by the term $Be^{-h}F$ in [\(7\)](#). The lateral drift of surface perturbations has been noted previously in surface electromigration problems; for instance, the drift is the hallmark of [\[Krug and Schimschak 1997\]](#), where it is caused by the nonlocality of the electric field.

Two sets of simulations are conducted, at different film heights: $h_0 = 3$ and $h_0 = 10$; when one parameter is varied, other parameters are fixed to their typical values in [Table 1](#). In addition, the height is also varied with all other parameters fixed. The results are displayed in [Figures 4 and 5](#).

[Figures 4\(a\)–\(b\)](#) show the effect of varying the film height on the amplitude and the wave speed. Both graphs show the sharp decreases and then the amplitude levels off, while the wave speed vanishes at $h_0 \approx 7$. The decrease of the wave amplitude and speed is expected, since the wetting potential and the corresponding driving force decay exponentially as the film thickness increases.

As seen in [Figures 4\(e\) and 5\(e\)](#), as the electric field parameter A is increased, the amplitude is decreased and then it levels off. As the graph is the same, we conclude that the dependence of the amplitude on A is unaffected (or is affected very weakly) by the height changes and by the traveling surface wave. Changing A affects the wave speed in a more complicated manner. As [Figure 5\(f\)](#) shows, increasing the strength of the field initially dampened the wave speed, but with further increase of the field, the wave speed also increases. The latter behavior is expected, since the strong field implies a fast adatoms drift. We will be looking into the reason for the initial wave speed decrease.

[Figures 4\(c\) and 5\(a\)](#) show that increasing the anisotropy strength ϵ_γ makes the amplitude smaller. The wave speed, however, increases with the increase of ϵ_γ , as shown in [Figure 5\(b\)](#). The amplitude variation is primarily affected by ϵ_γ and the simultaneous height changes make little difference.

In [Figure 5\(g\)](#), the amplitude increases only very little as the ratio of the energies G increases. (At $h_0 = 10$, the amplitude value stays constant (≈ 0.65) as G is

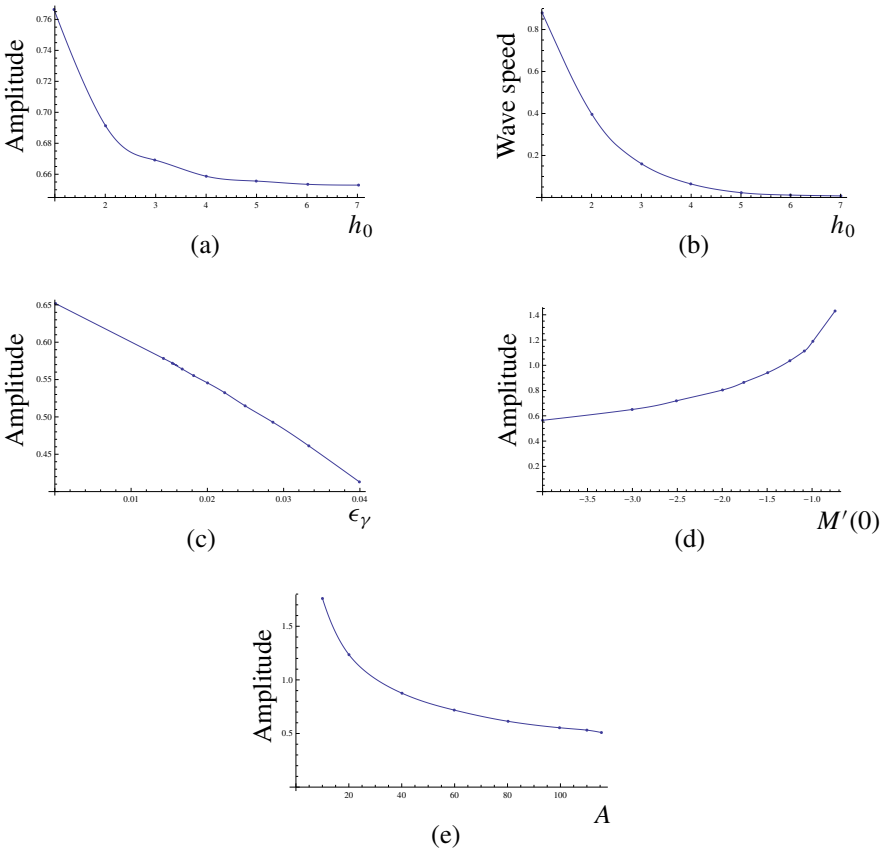


Figure 4. Graphs (a)–(b) show the effects of varying h_0 on the traveling wave amplitude and speed. Graphs (c)–(e) show the effects of varying ϵ_γ , $M'(0)$ and A on the stationary wave amplitude; here $h_0 = 10$, i.e., the film is thick. As is displayed in (b), for thick films the wave speed is zero.

varied; thus this graph is not shown in Figure 4.) In Figure 5(h), the wave speed increases as G increases.

Increasing the absolute value of the diffusional mobility derivative, $|M'(0)|$, results in the decrease of the amplitude and wave speed, as shown in Figures 4(d), 5(c) and 5(d). Once again we notice that the height variation does not seem to affect the trends that $M'(0)$ places on these characteristics.

Notice from the graphs of the amplitude in Figures 4 and 5 that the amplitude never reaches the value of h_0 , that is, 3 or 10. This means that the film’s local height is not zero, and therefore the film does not dewet the substrate. In other words, the film continuously covers the substrate at all times—the substrate is

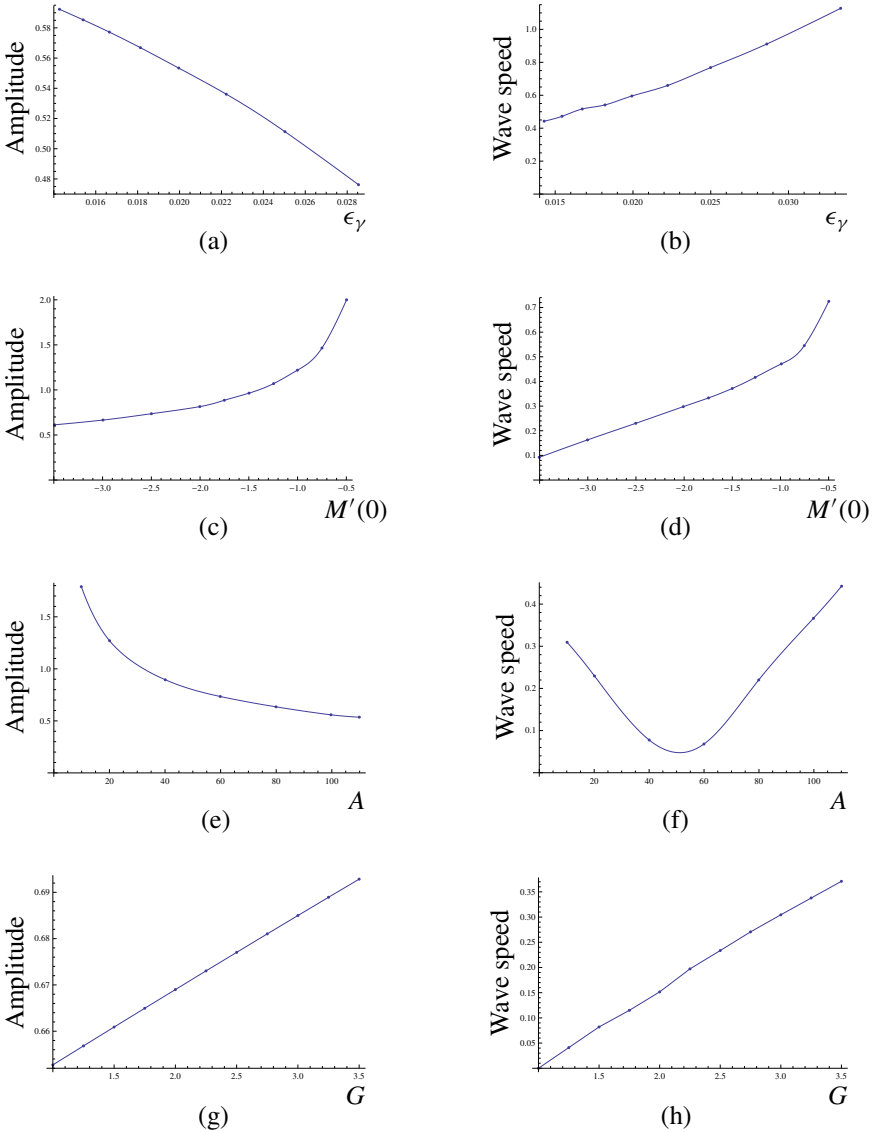


Figure 5. The effects of the parameters’ variation on the traveling wave amplitude and speed (the first and second column, respectively), for the thin film ($h_0 = 3$).

not exposed, despite the application of the electromigration. This can be expected, since the electric field is applied along the substrate, rather than across it. In the latter situation the film is more likely to dewet [Khenner 2013].

These results give insights into the complicated nonlinear dynamics of a film surface. Importantly, even though increasing A , ϵ_γ and $|M'(0)|$ results in the

decrease of the amplitude, the initial state of $\delta = 0.01$ is never reached. Thus at all times $t > 0$ the surface is always more deformed than it is initially.

The amplitude trends shown in Figures 4(a),(d),(e) and Figures 5(c),(e),(g) confirm the results obtained previously in [Khenner 2013]. The results for the amplitude and wave speed in other panels of these figures are new. As we noted above, the traveling wave solution was overlooked in [Khenner 2013] and thus the dependence of the traveling wave speed and amplitude on parameters, including the new parameter, i.e., the strength of the anisotropy ϵ_γ , was not computed there.

5. Conclusions

We performed the analysis of the partial differential equation model of the surface morphological evolution affected by electromigration, assuming a wetting solid film with the mildly anisotropic surface energy.

The linear stability analysis shows that the stability of the base planar state of the surface decreases with the increasing film thickness h_0 , the electric field strength A , the derivative of the diffusional mobility $|M'(0)|$ and the anisotropy strength ϵ_γ . The stability increases with increasing the ratio G of the substrate energy to the film energy, or equivalently, increasing the strength of the intermolecular attractive force between the adatoms and the atoms of the substrate.

We used the method of lines to numerically solve the fully nonlinear PDE. This way we found two outcomes of the surface evolution: the stationary wave relief for thick films, and the traveling surface wave (the surface drift) for thin films. We illustrated how all other physical parameters affect the amplitude of either wave (stationary or traveling), as well as the wave speed of the traveling wave. Our results also hint that there is no combination of the physically admissible parameters' values for which the film dewets the substrate.

Our numerical studies are on the periodic one-wavelength domain $x \in [0, \lambda_{\max}]$, and we used the cosine curve with the small initial amplitude to perturb the (constant) initial height. Future work will be focused on computing the evolution of a small random perturbation on the large periodic domain comprising many wavelengths. In this setup, the coarsening of the initial perturbation can be studied and predictions can be made about the pattern formation on the surface.

Acknowledgments

Aaron Wingo would like to thank his family and friends for all their support, but especially Dr. Khenner for his patience and guidance.

References

[Averbuch et al. 2001] A. Averbuch, E. Glickman, M. Israeli, M. Khenner, and M. Nathan, “Level set modeling of transient electromigration grooving”, *Comp. Mater. Sci.* **20** (2001), 235–250.

- [Barakat et al. 2012] F. Barakat, K. Martens, and O. Pierre-Louis, “Nonlinear wavelength selection in surface faceting under electromigration”, *Phys. Rev. Lett.* **109**:5 (2012), Article ID #056101.
- [Barnes et al. 2010] C. H. W. Barnes, A. B. Dominguez, L. L. Felix, S. I. Khondaker, Y. Majima, T. Mitrelias, F. Sfigakis, and L. Valladares, “Controlled electroplating and electromigration in nickel electrodes for nanogap formation”, *Nanotechnology* **21** (2010), Article ID #445304.
- [Block et al. 2006] T. Block, G. Gardinowski, H. Pfnür, J. Schmeidel, and C. Tegenkamp, “Switchable nanometer contacts: ultrathin Ag nanostructures on Si(100)”, *Appl. Phys. Lett.* **89**:6 (2006), Article ID #063120.
- [Bolotin et al. 2007] K. I. Bolotin, F. Kuemmeth, D. C. Ralph, and T. Taychatanapat, “Imaging electromigration during the formation of break junctions”, *Nano Lett.* **7** (2007), 652–656.
- [Brown et al. 1989] P. N. Brown, G. D. Byrne, and A. C. Hindmarsh, “VODE: a variable-coefficient ODE solver”, *SIAM J. Sci. Stat. Comput.* **10**:5 (1989), 1038–1051. MR 1009555 Zbl 0677.65075
- [Cahn and Taylor 1994] J. W. Cahn and J. E. Taylor, “Linking anisotropic sharp and diffuse surface motion laws via gradient flows”, *J. Stat. Phys.* **77**:1–2 (1994), 183–197. MR 1300532 Zbl 0844.35044
- [Cahn et al. 1992] J. W. Cahn, C. A. Handwerker, and J. E. Taylor, “Geometric models of crystal growth”, *Acta Met. Mat.* **40** (1992), 1443–1474.
- [Chang et al. 2006] J. Chang, C. Misbah, and O. Pierre-Louis, “Birth and morphological evolution of step bunches under electromigration”, *Phys. Rev. Lett.* **96**:19 (2006), Article ID #195901.
- [Dai et al. 2014] S. Dai, R. Yu, J. Zhao, and J. Zhu, “Kinetic faceting of the low index W surfaces under electrical current”, *Surf. Sci.* **625** (2014), 10–15.
- [Davis et al. 2004] S. H. Davis, A. A. Golovin, M. S. Levine, and T. V. Savina, “Faceting instability in the presence of wetting interactions: a mechanism for the formation of quantum dots”, *Phys. Rev. B* **70**:23 (2004), Article ID #235342.
- [Debierre et al. 2007] J.-M. Debierre, M. Dufay, and T. Frisch, “Electromigration-induced step meandering on vicinal surfaces: nonlinear evolution equation”, *Phys. Rev. B* **75**:4 (2007), Article ID #045413.
- [Di Carlo et al. 1992] A. Di Carlo, M. E. Gurtin, and P. Podio-Guidugli, “A regularized equation for anisotropic motion-by-curvature”, *SIAM J. Appl. Math.* **52**:4 (1992), 1111–1119. MR 1174049 Zbl 0800.73021
- [Dobbs and Krug 1994] H. T. Dobbs and J. Krug, “Current-induced faceting of crystal surfaces”, *Phys. Rev. Lett.* **73**:14 (1994), 1947–1950.
- [Gill and Wang 2008] S. P. A. Gill and T. Wang, “On the existence of a critical perturbation amplitude for the Stranski–Krastanov transition”, *Surf. Sci.* **602** (2008), 3560–3568.
- [Hairer and Wanner 1999] E. Hairer and G. Wanner, “Stiff differential equations solved by Runge–Kutta methods”, *J. Comput. Appl. Math.* **111**:1–2 (1999), 93–111. MR 1730588 Zbl 0945.65080
- [Khenner 2008a] M. Khenner, “Comparative study of a solid film dewetting in an attractive substrate potentials with the exponential and the algebraic decay”, *Math. Model. Nat. Phenom.* **3**:5 (2008), 16–29. MR 2477312 Zbl 06544732
- [Khenner 2008b] M. Khenner, “Morphologies and kinetics of a dewetting ultrathin solid film”, *Phys. Rev. B* **77**:24 (2008), Article ID #245445.
- [Khenner 2013] M. Khenner, “Analysis of a combined influence of substrate wetting and surface electromigration on a thin film stability and dynamical morphologies”, *C. R. Physique* **14** (2013), 607–618.
- [Khenner et al. 2011] M. Khenner, M. Levine, and W. T. Tekalign, “Stability of a strongly anisotropic thin epitaxial film in a wetting interaction with elastic substrate”, *Eur. Phys. Lett.* **93**:2 (2011), Article ID #26001.

- [Krug and Schimschak 1997] J. Krug and M. Schimschak, “Surface electromigration as a moving boundary value problem”, *Phys. Rev. Lett.* **78** (1997), 278–281.
- [Liu and Metiu 1993] F. Liu and H. Metiu, “Dynamics of phase separation of crystal surfaces”, *Phys. Rev. B* **48**:9 (1993), Article ID #5808.
- [Maroudas 1995] D. Maroudas, “Dynamics of transgranular voids in metallic thin films under electromigration conditions”, *Appl. Phys. Lett.* **67** (1995), 798–800.
- [Mullins 1963] W. W. Mullins, “Solid surface morphologies governed by capillarity”, pp. 17 in *Metal surfaces: structure, energetics and kinetics*, American Society for Metals, Cleveland, OH, 1963.
- [Saucez et al. 2001] P. Saucez, W. E. Schiesser, and A. Vande Wouwer (editors), *Adaptive method of lines*, Chapman and Hall, Boca Raton, FL, 2001. MR 1851429 Zbl 0986.65083
- [Schiesser 1994] W. E. Schiesser, *Computational mathematics in engineering and applied science: ODEs, DAEs, and PDEs*, CRC, Boca Raton, FL, 1994. MR 1272015 Zbl 0857.65083
- [Stoyanov 1997] S. Stoyanov, “Current-induced step bunching at vicinal surfaces during crystal sublimation”, *Surf. Sci.* **370** (1997), 345–354.
- [Verwer and Sanz-Serna 1984] J. G. Verwer and J. M. Sanz-Serna, “Convergence of method of lines approximations to partial differential equations”, *Computing* **33**:3-4 (1984), 297–313. MR 773930 Zbl 0546.65064
- [Wolfram 2016] Wolfram, “Mathematica”, 2016, <http://www.wolfram.com/mathematica>.

Received: 2015-01-28 Revised: 2015-07-08 Accepted: 2015-07-31

- aaron_wingo@mymail.eku.edu Department of Mathematics and Statistics,
Eastern Kentucky University, 521 Lancaster Avenue,
Richmond, KY 40475, United States
- scinar@math.uh.edu Department of Mathematics, University of Houston,
4800 Calhoun Road, Houston, TX 77004, United States
- kurt.woods774@topper.wku.edu General Motors Corporation, Bowling Green, KY 42101,
United States
- mikhail.khenner@wku.edu Department of Mathematics, Western Kentucky University,
1906 College Heights Boulevard, Bowling Green, KY 42101,
United States

Jacobian varieties of Hurwitz curves with automorphism group $\mathrm{PSL}(2, q)$

Allison Fischer, Mouchen Liu and Jennifer Paulhus

(Communicated by Nigel Boston)

The size of the automorphism group of a compact Riemann surface of genus $g > 1$ is bounded by $84(g - 1)$. Curves with automorphism group of size equal to this bound are called Hurwitz curves. In many cases the automorphism group of these curves is the projective special linear group $\mathrm{PSL}(2, q)$. We present a decomposition of the Jacobian varieties for all curves of this type and prove that no such Jacobian variety is simple.

1. Introduction

Let X be a compact Riemann surface of genus g (henceforth called a “curve”), and G its automorphism group with identity element denoted id_G . A result of Wedderburn gives the decomposition of the group ring $\mathbb{Q}G$,

$$\mathbb{Q}G \cong \bigoplus_i M_{n_i}(\Delta_i),$$

where $M_{n_i}(\Delta_i)$ denotes $n_i \times n_i$ matrices with coefficients in a division ring Δ_i . It is possible to decompose the Jacobian variety, JX , of the curve X into abelian varieties, up to isogeny \sim , as

$$JX \sim \bigoplus_i (e_i(JX))^{n_i}, \tag{1}$$

where e_i are certain idempotents in $\mathrm{End}(JX) \otimes_{\mathbb{Z}} \mathbb{Q}$. More details about this decomposition may be found in [Paulhus 2008]. It is important to note here that this decomposition may not be the finest possible decomposition. Some of the abelian variety factors $e_i(JX)$ could decompose further.

Decomposable Jacobian varieties have applications to rank and torsion questions in number theory [Howe et al. 2000; Rubin and Silverberg 2001]. In genus 2,

MSC2010: 14H40, 14H37, 20G05.

Keywords: Jacobian varieties, Hurwitz curves, projective special linear group, representation theory.

the elliptic curve factors appearing in these decompositions have interesting arithmetic properties (see [Cardona 2004; Earle 2006; Magaard et al. 2009], among many others).

The dimension, as an abelian variety, of the factor $e_i(J_X)$ in (1) is $\frac{1}{2}\langle\psi_i, \chi\rangle$, where $\langle\psi_i, \chi\rangle$ denotes the inner product of ψ_i , the i -th irreducible \mathbb{Q} -character labeled according to the Wedderburn decomposition, with χ , a character we define below called the *Hurwitz character*. To define the character χ , we consider the covering from X to its quotient $Y = X/G$, a curve with genus denoted g_Y . Let $h_1, \dots, h_s \in G$ be the monodromy of this covering. For any subgroup H of G , define the character χ_H to be the trivial character of H induced to G , and 1_G to be the trivial character of G . In this paper H is a cyclic subgroup generated by one element of the monodromy, which we write as $\langle h_i \rangle$. Note that with this notation $\chi_{\langle \text{id}_G \rangle}$ is the character associated to the regular representation. Define the *Hurwitz character* as

$$\chi = 2 \cdot 1_G + 2(g_Y - 1)\chi_{\langle \text{id}_G \rangle} + \sum_{j=1}^s (\chi_{\langle \text{id}_G \rangle} - \chi_{\langle h_j \rangle}), \quad (2)$$

which is the character of the representation of G on $H_{\text{et}}^1(X, \mathbb{Q}_\ell)$ [Milne 1980, Chapter V, §2]. To determine the dimensions of factors of JX from (1), we must know the automorphism group of X , the irreducible \mathbb{Q} -characters for that particular group, and the monodromy of the covering $X \rightarrow Y$.

The upper bound on the size of the automorphism group of a curve of genus $g > 1$ is given by $84(g - 1)$. Curves whose automorphism groups attain this bound are called *Hurwitz curves* and the groups themselves are called *Hurwitz groups*. Hurwitz groups have a long history in the study of triangle groups, Riemann surfaces, and hyperbolic geometry. See [Conder 1990] for a nice survey of these groups and their significance.

For all Hurwitz curves, the quotient curve Y is the projective line, so $g_Y = 0$. Since the quotient curve has genus 0, the monodromy of the covering is a set of elements $\{h_1, \dots, h_s\}$ in G such that $h_1 \cdots h_s = \text{id}_G$ and the set of all h_i generates G . The monodromy for Hurwitz curves is always of type $(2, 3, 7)$, meaning it consists of an element of order 2, an element of order 3, and an element of order 7, denoted in this paper by h_2, h_3 , and h_7 , respectively. (Equivalently, a Hurwitz group is a finite, nontrivial quotient of the $(2, 3, 7)$ -triangle group.) For Hurwitz curves, (2) may be simplified to

$$\chi = 2 \cdot 1_G + \chi_{\langle \text{id}_G \rangle} - \chi_{\langle h_2 \rangle} - \chi_{\langle h_3 \rangle} - \chi_{\langle h_7 \rangle}. \quad (3)$$

Let $\text{PSL}(2, q)$ denote the projective special linear group with coefficients in the finite field of order q . In this paper we will use (1) to decompose the Jacobian

varieties of all Hurwitz curves with automorphism group $\mathrm{PSL}(2, q)$. This decomposition may be found in [Theorem 10](#) and, in particular, in [Corollary 9](#) we prove that the Jacobian variety of these curves is never simple.

While there is an infinite family of Hurwitz curves with automorphism group $\mathrm{PSL}(2, q)$ (as we will see immediately below), there are many Hurwitz curves with other automorphism groups. For example, the alternating group A_n is a Hurwitz group for all $n \geq 168$ as well as for many smaller n [[Conder 1990](#)]. It is likely that a similar analysis would yield results about the decomposition of the Jacobians of these families of curves too.

Macbeath determines for which q the group $\mathrm{PSL}(2, q)$ is a Hurwitz group.

Theorem 1 [[Macbeath 1969](#)]. *The group $\mathrm{PSL}(2, q)$ is a Hurwitz group if and only if*

- (i) $q = 7$,
- (ii) q is a prime and congruent to $\pm 1 \pmod{7}$, or
- (iii) $q = p^3$ for a prime $p \equiv \pm 2$ or $\pm 3 \pmod{7}$.

Note that in both cases (ii) and (iii), we have $q \equiv \pm 1 \pmod{7}$. Case (i) occurs for a Hurwitz curve of genus 3, and the Jacobian is known to decompose as $JX \sim E^3$, where E is an elliptic curve [[Kuwata 2005](#)]. In case (ii), when $q = 13$ (and $g = 14$), the technique above may be used to show that $JX \sim E^{14}$, again for E some elliptic curve. Case (iii) includes the special case where $q = 8$. This corresponds to a genus 7 curve sometimes called the *Macbeath curve*. It has long been known that $JX \sim E^7$ [[Wolfart 2002](#)].

For odd q , $\mathrm{PSL}(2, q)$ has a well understood and relatively straightforward character table. Additionally, the monodromy of the coverings is not hard to find as (3) only requires knowledge of the monodromy *up to conjugation*. It turns out that, as we show below in [Proposition 2](#), for almost all q satisfying [Theorem 1](#), $\mathrm{PSL}(2, q)$ has only one conjugacy class of elements of order 2, one of elements of order 3, and three conjugacy classes of elements of order 7. This then allows us to compute the inner product $\langle \psi_i, \chi \rangle$ in all such examples and prove very general results about the Jacobian decompositions of curves with these groups as automorphism groups. The few exceptional q are either discussed above or at the end of the paper in [Section 6](#).

We begin in [Section 2](#) by reviewing known results about $G = \mathrm{PSL}(2, q)$. In particular, in [Section 2.3](#) we determine the irreducible \mathbb{Q} -characters, a key piece in our determination of the dimension of the factors in the Jacobian decompositions. In [Section 3](#) we compute the Hurwitz character χ , and in [Section 4](#) we compute the inner products. Finally we put the pieces together and present the Jacobian decomposition in [Section 5](#).

Using a different set of idempotents in $\mathbb{Q}G$ and the fact that $\mathrm{PSL}(2, q)$ has a *partition* (a set of subsets of G whose pairwise intersection is the identity and

whose union is the whole group), Kani and Rosen [1989, Example 2] describe a decomposition of a power of the Jacobian variety of curves with such automorphisms. The factors are themselves Jacobians of quotients of the curve by p -Sylow subgroups or Cartan subgroups of G .

2. Properties of $\mathrm{PSL}(2, q)$

Here we collect the relevant information about the group $G = \mathrm{PSL}(2, q)$. More details may be found in [Karpilovsky 1994] and we follow the notation in that book. For the rest of the paper, assume q is odd, $q > 27$, and q satisfies case (ii) or case (iii) in Theorem 1. All cases not covered by this are discussed above, except for $q = 27$, which we cover in Section 6.

First, the size of $\mathrm{PSL}(2, q)$ is

$$\frac{1}{2}q(q+1)(q-1).$$

To describe the character table of $\mathrm{PSL}(2, q)$ we need several special elements of $\mathrm{SL}(2, q)$. Let α be a generator of the group of units of the finite field with q^2 elements, let $\beta = \alpha^{q+1}$, and define b as the element of $\mathrm{SL}(2, q)$ determined by the map $x \rightarrow \alpha^{q-1}x$ for $x \in \mathbb{F}_{q^2}$. Additionally define elements of $\mathrm{SL}(2, q)$

$$a = \begin{bmatrix} \beta & 0 \\ 0 & \beta^{-1} \end{bmatrix}, \quad c = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \text{and} \quad d = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix}.$$

The images of the elements $a, b, c,$ and d in the quotient $\mathrm{PSL}(2, q)$ are denoted as $\bar{a}, \bar{b}, \bar{c},$ and \bar{d} . The element \bar{a} has order $\frac{1}{2}(q-1)$, the element \bar{b} has order $\frac{1}{2}(q+1)$ and the elements \bar{c} and \bar{d} each have order q .

2.1. Conjugacy classes. To determine the monodromy of the covering, we need to understand the conjugacy classes of elements of orders 2, 3, and 7. The representatives of the conjugacy classes of $\mathrm{PSL}(2, q)$ are $\bar{1}, \bar{c}, \bar{d}, \bar{a}^n,$ and \bar{b}^m , where $1 \leq n, m \leq \frac{1}{4}(q-1)$ if $q \equiv 1 \pmod{4}$, while $1 \leq n \leq \frac{1}{4}(q-3)$ and $1 \leq m \leq \frac{1}{4}(q+1)$ if $q \equiv -1 \pmod{4}$. We will write the conjugacy class of an element $h \in G$ as $[h]$.

Conjugacy classes with a representative \bar{a}^n have size $q(q+1)$, and conjugacy classes with a representative \bar{b}^m have size $q(q-1)$, with the exception of the conjugacy class containing elements of order 2 which has order half that size, (or $\frac{1}{2}q(q-1)$) [Karpilovsky 1994]. We will see in the proof of Proposition 2 that the conjugacy class of elements of order 2 is $[\bar{a}^{(q-1)/4}]$ if $q \equiv 1 \pmod{4}$ and $[\bar{b}^{(q+1)/4}]$ if $q \equiv -1 \pmod{4}$.

It turns out that χ as defined in (3) is 0 outside of the conjugacy classes of elements of orders 1, 2, 3, and 7, as we will see in Section 3. So it will be sufficient to only study these conjugacy classes of $\mathrm{PSL}(2, q)$ since any other conjugacy class will not contribute to our goal of computing the inner product of χ with the irreducible \mathbb{Q} -characters. But how many such conjugacy classes are there?

Proposition 2. *If $G = \text{PSL}(2, q)$ for q odd, greater than 27, and satisfying case (ii) or case (iii) in Theorem 1, then G has three distinct conjugacy classes of elements of order 7, and one each of elements of orders 2 and 3.*

Proof. When q is as in the proposition, since elements of the conjugacy classes represented by \bar{c} and \bar{d} have order q , the elements of order 7 can only lie in conjugacy classes represented by some power of \bar{a} or \bar{b} . (For $q = 7$ this need not be true as \bar{c} and \bar{d} both have order $q = 7$.)

Recall for a finite group G , the order of g^k for any $g \in G$ and positive integer k is $o(g^k) = o(g) / \gcd(k, o(g))$. Thus, 7 must divide the order of \bar{a} or the order of \bar{b} but not both, else it divides $\frac{1}{2}(q + 1) - \frac{1}{2}(q - 1) = 1$. Thus the conjugacy class(es) of order 7 are either represented by some power(s) of \bar{a} or some power(s) of \bar{b} .

First consider the case where $q \equiv 1 \pmod{4}$. Suppose that the conjugacy classes of elements of order 7 are represented by powers of \bar{a} (so $q \equiv 1 \pmod{7}$). The number of conjugacy classes will be the number of i such that $7 = o(\bar{a}) / \gcd(o(\bar{a}), i)$, where $1 \leq i \leq \frac{1}{4}(q - 1)$. Since 7 divides the order of \bar{a} , we let $o(\bar{a}) = 7j$ for some positive integer j . Then the number of i such that $7 = 7j / \gcd(7j, i)$ is the number of i that satisfy $\gcd(7j, i) = j$ and $1 \leq i \leq \frac{7}{2}j$. Since $o(\bar{a}) = \frac{1}{2}(q - 1)$ and $q > 13$, there are always three of them: $i = j$ (or $\frac{1}{14}(q - 1)$), $i = 2j$ (or $\frac{1}{7}(q - 1)$), and $i = 3j$ (or $\frac{3}{14}(q - 1)$). Hence the elements of order 7 are in the conjugacy classes represented by $\bar{a}^{(q-1)/14}$, $\bar{a}^{(q-1)/7}$, and $\bar{a}^{3(q-1)/14}$. A similar argument works if these classes are represented by powers of \bar{b} (or $q \equiv -1 \pmod{7}$). The elements of order 7 are in the conjugacy classes represented by $\bar{b}^{(q+1)/14}$, $\bar{b}^{(q+1)/7}$, and $\bar{b}^{3(q+1)/14}$.

Now, when $q \equiv -1 \pmod{4}$, the argument is identical except the bounds on i change to $1 \leq i \leq \frac{1}{4}(q - 3)$ if $q \equiv 1 \pmod{7}$ and $1 \leq i \leq \frac{1}{4}(q - 1)$ if $q \equiv -1 \pmod{7}$. The rest of the argument does not change and so there are three conjugacy classes of elements of order 7, again defined as $\bar{a}^{(q-1)/14}$, $\bar{a}^{(q-1)/7}$, and $\bar{a}^{3(q-1)/14}$ if $q \equiv 1 \pmod{7}$ or $\bar{b}^{(q+1)/14}$, $\bar{b}^{(q+1)/7}$, and $\bar{b}^{3(q+1)/14}$ if $q \equiv -1 \pmod{7}$.

The cases with orders 2 and 3 follow similarly. When $q \equiv 1 \pmod{4}$, the elements of order 2 are in the conjugacy class $[\bar{a}^{(q-1)/4}]$; when $q \equiv -1 \pmod{4}$, the elements of order 2 are in the conjugacy class $[\bar{b}^{(q+1)/4}]$. For elements of order 3, the conjugacy class is $[\bar{a}^{(q-1)/6}]$ if $q \equiv 1 \pmod{3}$ and $[\bar{b}^{(q+1)/6}]$ if $q \equiv -1 \pmod{3}$. (If $q = 27$ there are two conjugacy classes of elements of order 3. See Section 6 for this special case.) □

2.2. Character tables. Let ε be a primitive $(q - 1)$ -th root of unity and let δ be a primitive $(q + 1)$ -th root of unity, where $\varepsilon_{kn} = \varepsilon^{2kn} + \varepsilon^{-2kn}$ and $\delta_{tm} = -(\delta^{2tm} + \delta^{-2tm})$.

When $q \equiv 1 \pmod{4}$, the character table of $G = \text{PSL}(2, q)$ is given in Table 1 for $1 \leq m, n, t \leq \frac{1}{4}(q - 1)$ and $1 \leq k \leq \frac{1}{4}(q - 5)$ [Karpilovsky 1994, Theorem 8.9].

When $q \equiv -1 \pmod{4}$, the character table of $G = \text{PSL}(2, q)$ is given in Table 2 for $1 \leq n, k, t \leq \frac{1}{4}(q - 3)$ and $1 \leq m \leq \frac{1}{4}(q + 1)$ [Karpilovsky 1994, Theorem 8.11].

	$[\bar{1}]$	$[\bar{a}^n]$	$[\bar{b}^m]$	$[\bar{c}]$	$[\bar{d}]$
1_G	1	1	1	1	1
λ	q	1	-1	0	0
μ_k	$q+1$	ε_{kn}	0	1	1
θ_t	$q-1$	0	δ_{tm}	-1	-1
χ_1	$\frac{1}{2}(q+1)$	$(-1)^n$	0	$\frac{1}{2}(1+\sqrt{q})$	$\frac{1}{2}(1-\sqrt{q})$
χ_2	$\frac{1}{2}(q+1)$	$(-1)^n$	0	$\frac{1}{2}(1-\sqrt{q})$	$\frac{1}{2}(1+\sqrt{q})$

Table 1. The character table of $G = \text{PSL}(2, q)$ for $q \equiv 1 \pmod{4}$.

	$[\bar{1}]$	$[\bar{a}^n]$	$[\bar{b}^m]$	$[\bar{c}]$	$[\bar{d}]$
1_G	1	1	1	1	1
λ	q	1	-1	0	0
μ_k	$q+1$	ε_{kn}	0	1	1
θ_t	$q-1$	0	δ_{tm}	-1	-1
γ_1	$\frac{1}{2}(q-1)$	0	$(-1)^{m+1}$	$\frac{1}{2}(-1+\sqrt{-q})$	$\frac{1}{2}(-1-\sqrt{-q})$
γ_2	$\frac{1}{2}(q-1)$	0	$(-1)^{m+1}$	$\frac{1}{2}(-1-\sqrt{-q})$	$\frac{1}{2}(-1+\sqrt{-q})$

Table 2. The character table of $G = \text{PSL}(2, q)$ for $q \equiv -1 \pmod{4}$.

2.3. Irreducible \mathbb{Q} -characters. The character tables above give the irreducible \mathbb{C} -characters of $\text{PSL}(2, q)$ but we need \mathbb{Q} -characters to compute the dimensions of the factors of the Jacobian decompositions. Since all irreducible \mathbb{C} -characters of $\text{PSL}(2, q)$ have Schur index 1 [Janusz 1974], it is sufficient to find the Galois conjugates of all \mathbb{C} -characters.

The characters 1_G and λ are already \mathbb{Q} -characters, and it is clear that $\chi_1 + \chi_2$ and $\gamma_1 + \gamma_2$ are \mathbb{Q} -characters as their noninteger entries are Galois conjugates. This leaves the μ_k and θ_t characters.

Proposition 3. (a) Let r be a divisor of $\frac{1}{2}(q - 1)$ and define the set

$$M_r = \begin{cases} \{\mu_i \mid 1 \leq i \leq \frac{1}{4}(q - 5) \text{ and } \gcd(i, \frac{1}{2}(q - 1)) = r\} & \text{if } q \equiv 1 \pmod{4}, \\ \{\mu_i \mid 1 \leq i \leq \frac{1}{4}(q - 3) \text{ and } \gcd(i, \frac{1}{2}(q - 1)) = r\} & \text{if } q \equiv -1 \pmod{4}. \end{cases}$$

The sum of the characters in each M_r is an irreducible \mathbb{Q} -character of $\text{PSL}(2, q)$.

(b) Let s be a divisor of $\frac{1}{2}(q + 1)$ and define the set

$$\Theta_s = \begin{cases} \{\theta_i \mid 1 \leq i \leq \frac{1}{4}(q - 1) \text{ and } \gcd(i, \frac{1}{2}(q - 1)) = s\} & \text{if } q \equiv 1 \pmod{4}, \\ \{\theta_i \mid 1 \leq i \leq \frac{1}{4}(q - 3) \text{ and } \gcd(i, \frac{1}{2}(q - 1)) = s\} & \text{if } q \equiv -1 \pmod{4}. \end{cases}$$

The sum of the characters in each Θ_s is an irreducible \mathbb{Q} -character of $\text{PSL}(2, q)$.

	$[\bar{a}]$	$[\bar{a}^2]$...	$[\bar{a}^{(q-1)/4}]$
μ_1	$\rho + \rho^{-1}$	$\rho^2 + \rho^{-2}$...	$\rho^{(q-1)/4} + \rho^{-(q-1)/4}$
μ_2	$\rho^2 + \rho^{-2}$	$\rho^4 + \rho^{-4}$...	$\rho^{(q-1)/2} + \rho^{-(q-1)/2}$
\vdots	\vdots	\vdots	\ddots	\vdots
$\mu_{(q-5)/4}$	$\rho^{(q-5)/4} + \rho^{-(q-5)/4}$	$\rho^{(q-5)/2} + \rho^{-(q-5)/2}$...	$\rho^{(q-5)(q-1)/16} + \rho^{-(q-5)(q-1)/16}$

Table 3. Values of μ_k on conjugacy classes of elements \bar{a}^n when $q \equiv 1 \pmod 4$.

Proof. We prove (a) below. The argument for (b) is almost identical. Since the only nonrational values of the μ_k characters are their values on the $[\bar{a}^n]$, we only need to consider the values on these conjugacy classes. For simplicity of notation, we define ρ to be ε^2 , so ρ is a primitive $\frac{1}{2}(q-1)$ -th root of unity. Then the values of the μ_k on the conjugacy classes $[\bar{a}^n]$ in the case where $q \equiv 1 \pmod 4$ are given in Table 3. (For $q \equiv -1 \pmod 4$, replace $\frac{1}{4}(q-5)$ in the last row with $\frac{1}{4}(q-3)$ and change the exponent in the last column from $\frac{1}{4}(q-1)$ to $\frac{1}{4}(q-3)$.)

Fix a particular μ_k with $\gcd(k, \frac{1}{2}(q-1)) = r$. The Galois orbit is completely determined by $\mu_k([\bar{a}])$ since the values of μ_k on the conjugacy classes with representative powers of \bar{a} are sums of powers of the summand of $\mu_k([\bar{a}])$ (as seen in Table 3). So it is enough to find the Galois conjugates of $\mu_k([\bar{a}])$. Now $\mu_k([\bar{a}]) = \rho^k + \rho^{-k}$, where ρ^k is a primitive $\frac{1}{2r}(q-1)$ -th root of unity. The Galois conjugates of this will be sums of the other primitive $\frac{1}{2r}(q-1)$ -th roots of unity. By a simple order argument, we determine that ρ^i is a primitive $(\frac{1}{2}(q-1) / \gcd(i, \frac{1}{2}(q-1)))$ -th root of unity. So the other primitive $\frac{1}{2r}(q-1)$ -th roots of unity appear for exactly those μ_i such that $\gcd(i, \frac{1}{2}(q-1)) = r$. So the irreducible \mathbb{Q} -character associated with μ_k will be the sum of μ_k with the other characters μ_i such that $\gcd(i, \frac{1}{2}(q-1)) = \gcd(k, \frac{1}{2}(q-1)) = r$. \square

Example. We demonstrate the previous proposition with an example. Consider $q = 29 \equiv 1 \pmod 4$. Here $\frac{1}{2}(q-1) = 14$, $\frac{1}{2}(q+1) = 15$, $\frac{1}{4}(q-5) = 6$, and $\frac{1}{4}(q-1) = 7$ and so there are 6 μ_k characters and 7 θ_t characters. The only divisors of $\frac{1}{2}(q-1)$ less than 6 are 1 and 2. From Proposition 3(a) we have two distinct sets

$$M_1 = \{\mu_i \mid \gcd(i, 14) = 1\} = \{\mu_1, \mu_3, \mu_5\},$$

$$M_2 = \{\mu_i \mid \gcd(i, 14) = 2\} = \{\mu_2, \mu_4, \mu_6\}.$$

The divisors of $\frac{1}{2}(q+1)$ less than 7 are 1, 3, and 5, so from Proposition 3(b) there are three distinct sets

$$\Theta_1 = \{\theta_i \mid \gcd(i, 15) = 1\} = \{\theta_1, \theta_2, \theta_4, \theta_7\},$$

$$\Theta_3 = \{\theta_i \mid \gcd(i, 15) = 3\} = \{\theta_3, \theta_6\},$$

$$\Theta_5 = \{\theta_i \mid \gcd(i, 15) = 5\} = \{\theta_5\}.$$

Therefore when $q = 29$, there are two irreducible \mathbb{Q} -characters of degree $q + 1$ ($\mu_1 + \mu_3 + \mu_5$ and $\mu_2 + \mu_4 + \mu_6$) and three irreducible \mathbb{Q} -characters of degree $q - 1$ ($\theta_1 + \theta_2 + \theta_4 + \theta_7$, $\theta_3 + \theta_6$, and θ_5).

We also need the values of the irreducible \mathbb{Q} -characters from [Proposition 3](#) for the inner product computation of the dimensions of the factors in [\(1\)](#). In the rest of the paper, for any character μ_k , we denote by r the $\gcd(k, \frac{1}{2}(q - 1))$, and for any character θ_t , we denote by s the $\gcd(t, \frac{1}{2}(q + 1))$. Thus M_r from [Proposition 3\(a\)](#) will contain the character μ_k and Θ_s from [Proposition 3\(b\)](#) will contain θ_t . The value of the characters in [Proposition 3](#) will be the value of μ_k (or θ_t) times the number of irreducible \mathbb{C} -characters in the set M_r (or Θ_s). The size of M_r is half the number of i such that $\gcd(i, \frac{1}{2}(q - 1)) = r$, or half the number of i such that $\gcd(i, \frac{1}{2r}(q - 1)) = 1$. This is $\frac{1}{2}\phi(\frac{1}{2r}(q - 1))$, where $\phi(x)$ is the Euler phi function. Similarly, the size of Θ_s is equal to $\frac{1}{2}\phi(\frac{1}{2s}(q + 1))$. Additionally for our computations, we will only need the values of the characters on conjugacy classes of orders 1, 2, 3, and 7, as it turns out that the Hurwitz character χ is 0 outside these conjugacy classes. This means the inner product we use to compute the dimension of the factors of the Jacobian will not be impacted by the values outside of these conjugacy classes. Again, see [Section 3](#) and [\(5\)](#).

Determining the value of each μ_k or θ_t on the relevant conjugacy classes boils down to whether elements of that order are powers of \bar{a} or \bar{b} . The next three propositions give the values of these characters on conjugacy classes of elements of orders 2, 3, and 7, respectively.

Proposition 4. *Consider the conjugacy class of elements of order 2 in $\mathrm{PSL}(2, q)$ for q satisfying the conditions in [Proposition 2](#).*

(a) *When $q \equiv 1 \pmod{4}$, the irreducible \mathbb{Q} -characters from [Proposition 3\(a\)](#) evaluate to $(-1)^k \phi(\frac{1}{2r}(q - 1))$, while the irreducible \mathbb{Q} -characters from [Proposition 3\(b\)](#) evaluate to 0.*

(b) *When $q \equiv -1 \pmod{4}$, the irreducible \mathbb{Q} -characters from [Proposition 3\(a\)](#) evaluate to 0, while the irreducible \mathbb{Q} -characters from [Proposition 3\(b\)](#) evaluate to $(-1)^{t+1} \phi(\frac{1}{2s}(q + 1))$.*

Proof. (a) As we saw in the proof of [Proposition 2](#), the conjugacy class of elements of order 2 is represented by a power of either \bar{a} or \bar{b} , depending on whether $q \equiv \pm 1 \pmod{4}$. In the first case, it is $[\bar{a}^{(q-1)/4}]$. Consider the value of one μ_k on this conjugacy class:

$$\varepsilon_{k(q-1)/4} = \varepsilon^{k(q-1)/2} + \varepsilon^{-k(q-1)/2}.$$

Since ε is a primitive $(q-1)$ -th root of unity, $\varepsilon^{(q-1)/2}$ is a primitive second root of unity, i.e., -1 . Thus $\varepsilon_{k(q-1)/4} = (-1)^k + (-1)^{-k}$. When k is odd, this value

is -2 , and when k is even, this value is 2 . Combining this value with the number of characters in the set M_r yields the value of

$$(-1)^k \phi\left(\frac{q-1}{2r}\right).$$

The \mathbb{Q} -characters which are sums of the characters in Θ_s (as in Proposition 3(b)) are 0 on this class in this case. From the character table for this case, it is clear that each θ_t has a value of 0 on any conjugacy class of the form $[\bar{a}^n]$ and hence the sum of such characters also has a value of 0.

(b) When $q \equiv -1 \pmod{4}$, the conjugacy class is represented by $\bar{b}^{(q+1)/4}$ and so the \mathbb{Q} -characters in Proposition 3(a) are 0 on that class since each μ_k evaluates to 0. A similar argument as for $q \equiv 1 \pmod{4}$ gives that θ_t will be 2 when t is odd and -2 when t is even. Then the irreducible \mathbb{Q} -characters in Proposition 3(b) evaluate to this value multiplied by the size of Θ_s . This gives

$$(-1)^{t+1} \phi\left(\frac{q+1}{2s}\right). \quad \square$$

Proposition 5. *Consider the conjugacy class of elements of order 3 in $\text{PSL}(2, q)$ for q satisfying the conditions in Proposition 2.*

(a) *When $q \equiv 1 \pmod{3}$ the irreducible \mathbb{Q} -characters in Proposition 3(a) evaluate to*

$$\begin{cases} \phi\left(\frac{1}{2r}(q-1)\right) & \text{if } k \equiv 0 \pmod{3}, \\ -\frac{1}{2}\phi\left(\frac{1}{2r}(q-1)\right) & \text{otherwise,} \end{cases}$$

while the irreducible \mathbb{Q} -characters from Proposition 3(b) evaluate to 0.

(b) *When $q \equiv -1 \pmod{3}$, the characters described in Proposition 3(a) evaluate to 0, while the irreducible \mathbb{Q} -characters in Proposition 3(b) evaluate to*

$$\begin{cases} -\phi\left(\frac{1}{2s}(q+1)\right) & \text{if } t \equiv 0 \pmod{3}, \\ \frac{1}{2}\phi\left(\frac{1}{2s}(q+1)\right) & \text{otherwise.} \end{cases}$$

Proof. As was discussed in the proof of Proposition 2, the conjugacy class of elements of order 3 is represented by $\bar{a}^{(q-1)/6}$ or $\bar{b}^{(q+1)/6}$.

(a) Consider the value of μ_k :

$$\varepsilon_{k(q-1)/6} = \varepsilon^{k(q-1)/3} + \varepsilon^{-k(q-1)/3}.$$

Since ε is a primitive $(q-1)$ -th root of unity, $\varepsilon^{(q-1)/3}$ is a third root of unity, which we call ω . Thus, $\varepsilon_{k(q-1)/4} = \omega^k + \omega^{-k}$. When $3 \mid k$, this is 2 and when $3 \nmid k$, this is $\varepsilon_{k(q-1)/4} = \omega + \omega^2 = -1$. This value, together with the size of M_r gives the value of the irreducible \mathbb{Q} -characters in Proposition 3(a) on elements of order 3. Since each θ_t evaluates to 0 on the conjugacy classes represented by powers of \bar{a} , the irreducible \mathbb{Q} -characters from Proposition 3(b) also evaluate to 0.

(b) A similar argument may be used when $q \equiv -1 \pmod 3$ (or the elements of order 3 are in the conjugacy class represented by $\bar{b}^{(q+1)/6}$). \square

Proposition 6. *Consider the conjugacy classes of elements of order 7 in $\text{PSL}(2, q)$ for q satisfying the conditions in Proposition 2.*

(a) *When $q \equiv 1 \pmod 7$, the characters in Proposition 3(b) evaluate to 0, while the irreducible \mathbb{Q} -characters from Proposition 3(a) evaluate to*

$$\begin{cases} \phi\left(\frac{1}{2r}(q-1)\right) & \text{if } k \equiv 0 \pmod 7, \\ -\frac{1}{2}\phi\left(\frac{1}{2r}(q-1)\right) & \text{otherwise.} \end{cases}$$

(b) *When $q \equiv -1 \pmod 7$, the irreducible \mathbb{Q} -characters in Proposition 3(a) evaluate to 0, while the irreducible \mathbb{Q} -characters from Proposition 3(b) evaluate to*

$$\begin{cases} -\phi\left(\frac{1}{2s}(q+1)\right) & \text{if } t \equiv 0 \pmod 7, \\ \frac{1}{2}\phi\left(\frac{1}{2s}(q+1)\right) & \text{otherwise.} \end{cases}$$

Proof. From the proof of Proposition 2 we know that the three conjugacy classes of order 7 are represented by $\bar{a}^{(q-1)/14}$, $\bar{a}^{(q-1)/7}$, and $\bar{a}^{3(q-1)/14}$ or $\bar{b}^{(q+1)/14}$, $\bar{b}^{(q+1)/7}$, and $\bar{b}^{3(q+1)/14}$.

(a) If $q \equiv 1 \pmod 7$ (equivalently the conjugacy classes of elements of order 7 are represented by powers of \bar{a}) then μ_k evaluates to $\zeta^k + \zeta^{-k}$ on these conjugacy classes, where ζ is a primitive 7th root of unity. If $7 \mid k$, then $\zeta^k + \zeta^{-k}$ is 2 and if $7 \nmid k$, then $\zeta^k + \zeta^{-k}$ is -1 . Combining this with the size of the set M_r or Θ_s gives the result.

(b) A similar argument follows for $q \equiv -1 \pmod 7$ except we are considering conjugacy classes represented by powers of \bar{b} . \square

3. Computation of the Hurwitz character

Recall from (3) that in order to compute χ , we need to determine $\chi_{\langle \text{id}_G \rangle}$, $\chi_{\langle h_2 \rangle}$, $\chi_{\langle h_3 \rangle}$, and $\chi_{\langle h_7 \rangle}$. Let H be a subgroup of G . By the definition of χ_H , the induced character of the trivial character of H is

$$\chi_H(g) = \frac{1}{H} \sum_{x \in G} \chi^o(xgx^{-1}), \quad \text{where } \chi^o(g) = \begin{cases} 1 & \text{if } g \in H, \\ 0 & \text{if } g \notin H. \end{cases}$$

Note that $\chi_{\langle \text{id}_G \rangle}$ is just the regular representation

$$\chi_{\langle \text{id}_G \rangle}(g) = \begin{cases} |G| & \text{if } g = \text{id}_G, \\ 0 & \text{if } g \neq \text{id}_G. \end{cases}$$

To compute the remaining three characters, we need several facts from Section 2.1 and a lemma, which is an immediate consequence of the orbit-stabilizer theorem considering the group action of conjugation.

Lemma 7. *Let G be a group and $g, h \in G$ with g not the identity. The number of $x \in G$ such that $xgx^{-1} = h$ is the size of the centralizer of h if $g \in [h]$ and 0 otherwise.*

Consider $\chi_{\langle h_2 \rangle}$. We know

$$\chi_{\langle h_2 \rangle}(g) = \frac{1}{2} \sum_{x \in G} \chi^o(xgx^{-1}). \tag{4}$$

For each $g \in G$, we must determine the number of $x \in G$ such that $xgx^{-1} = \text{id}_G$ or h_2 , since $\langle h_2 \rangle = \{\text{id}_G, h_2\}$. The case of $xgx^{-1} = \text{id}_G$ follows from the fact that, for any group G and $g \in G$ not the identity, there is no $x \in G$ so that $xgx^{-1} = \text{id}_G$. Thus the number of $x \in G$ such that $xgx^{-1} = \text{id}_G$ or h_2 is the size of G when g is the identity and 0 otherwise. For $\chi_{\langle h_2 \rangle}(g)$ when $g \neq \text{id}_G$, if $g \notin [h_2]$ then this number is 0, else we must determine the number of $x \in G$ so that $xgx^{-1} = h_2$. By Lemma 7, this is the size of the centralizer of h_2 . Recall that under the action of conjugation, orbits are conjugacy classes. By the orbit-stabilizer theorem, $|C_G(h_2)| = |G|/|[h_2]|$. For h_2 of order 2, we have $|[h_2]| = \frac{1}{2}q(q+1)$ when $q \equiv 1 \pmod{4}$, and $|[h_2]| = \frac{1}{2}q(q-1)$ when $q \equiv -1 \pmod{4}$, hence $|C_G(h_2)| = q-1$ if $q \equiv 1 \pmod{4}$ and $|C_G(h_2)| = q+1$ if $q \equiv -1 \pmod{4}$. Plugging these values into (4) gives

$$\chi_{\langle h_2 \rangle}(g) = \begin{cases} \frac{1}{2}|G| & \text{if } g = \text{id}_G, \\ \frac{1}{2}(q-1) & \text{if } g \in [h_2] \text{ and } q \equiv 1 \pmod{4}, \\ \frac{1}{2}(q+1) & \text{if } g \in [h_2] \text{ and } q \equiv -1 \pmod{4}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, we calculate $\chi_{\langle h_3 \rangle}$. As before, for each $g \in G$, we need to find the number of $x \in G$ so that $xgx^{-1} \in \langle h_3 \rangle = \{1, h_3, h_3^2\}$, and the formula in this case is

$$\chi_{\langle h_3 \rangle}(g) = \frac{1}{3} \sum_{x \in G} \chi^o(xgx^{-1}).$$

When $g = \text{id}_G$, we have $\chi_{\langle h_3 \rangle}(\text{id}_G) = \frac{1}{3}|G|$. Else by Lemma 7 and the fact that $h_3^2 \in [h_3]$, we have $\chi_{\langle h_3 \rangle}(g) = \frac{2}{3}|C_G(h_3)|$ if $g \in [h_3]$ and 0 otherwise. From Section 2.1 we know $|[h_3]| = q(q-1)$ if $3 \mid \frac{1}{2}(q+1)$ and $|[h_3]| = q(q+1)$ if $3 \mid \frac{1}{2}(q-1)$. Then $|C_G(h_3)| = \frac{1}{2}(q+1)$ if $3 \mid \frac{1}{2}(q+1)$ and $|C_G(h_3)| = \frac{1}{2}(q-1)$ if $3 \mid \frac{1}{2}(q-1)$, and

$$\chi_{\langle h_3 \rangle}(g) = \begin{cases} \frac{1}{3}|G| & \text{if } g = \text{id}_G, \\ \frac{1}{3}(q-1) & \text{if } g \in [h_3] \text{ and } q \equiv 1 \pmod{3}, \\ \frac{1}{3}(q+1) & \text{if } g \in [h_3] \text{ and } q \equiv 2 \pmod{3}, \\ 0 & \text{otherwise.} \end{cases}$$

$q \pmod{84}$	Value for elements of order			
	1	2	3	7
± 1	$\frac{1}{42} G $	$-\frac{1}{2}(q \mp 1)$	$-\frac{1}{3}(q \mp 1)$	$-\frac{1}{7}(q \mp 1)$
± 13	$\frac{1}{42} G $	$-\frac{1}{2}(q \mp 1)$	$-\frac{1}{3}(q \mp 1)$	$-\frac{1}{7}(q \pm 1)$
± 29	$\frac{1}{42} G $	$-\frac{1}{2}(q \mp 1)$	$-\frac{1}{3}(q \pm 1)$	$-\frac{1}{7}(q \mp 1)$
± 43	$\frac{1}{42} G $	$-\frac{1}{2}(q \pm 1)$	$-\frac{1}{3}(q \mp 1)$	$-\frac{1}{7}(q \mp 1)$

Table 4. Values of χ' on conjugacy classes of elements of orders 1, 2, 3, and 7.

For $\chi_{\langle h_7 \rangle}$, as with the proof of the other characters, $\chi_{\langle h_7 \rangle}(\text{id}_G) = \frac{1}{7}|G|$. To compute the value on other elements, observe that for any g of order 7, we know that g and g^{-1} are in the same conjugacy class [Karpilovsky 1994, Corollary 8.3] but g , g^2 , and g^3 are all in distinct conjugacy classes. Combining Lemma 7 with this information gives us that $\chi_{\langle h_7 \rangle}(g) = \frac{2}{7}|C_G(h_7)|$, and we know the sizes of the conjugacy classes by Section 2.1. Putting all this information together, the value of $\chi_{\langle h_7 \rangle}(g)$ is

$$\chi_{\langle h_7 \rangle}(g) = \begin{cases} \frac{1}{7}|G| & \text{if } g = \text{id}_G, \\ \frac{1}{7}(q-1) & \text{if } g \in [h_7] \text{ and } q \equiv 1 \pmod{7}, \\ \frac{1}{7}(q+1) & \text{if } g \in [h_7] \text{ and } q \equiv -1 \pmod{7}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the values of χ are invariant under the three conjugacy classes of elements of order 7. This means we do not have to find in which conjugacy class of elements of order 7 the monodromy exists in order to compute (3) (i.e., we do not have to explicitly find h_7 , we can just use the formula above for any element of order 7).

We will use χ to calculate inner products with irreducible \mathbb{Q} -characters to find the dimensions of the factors in the Jacobian variety decomposition. To simplify later calculations, we rewrite χ as $\chi = 2 \cdot 1_G + \chi'$, where $\chi' = \chi_{\langle 1_G \rangle} - \chi_{\langle h_2 \rangle} - \chi_{\langle h_3 \rangle} - \chi_{\langle h_7 \rangle}$. Then, the inner product of an irreducible \mathbb{Q} -character ψ_i and χ will be $\langle \psi_i, \chi \rangle = 2\langle \psi_i, 1_G \rangle + \langle \psi_i, \chi' \rangle$. But since ψ_i and 1_G are orthogonal when $\psi_i \neq 1_G$, we have that $\langle \psi_i, \chi \rangle$ is simply $\langle \psi_i, \chi' \rangle$ in all cases except for the trivial character.

Table 4 gives the values of χ' on the conjugacy classes of elements of orders 1, 2, 3, and 7, computed by combining all the data in this section. Additionally, $\chi'(g) = 0$ if g is not in one of these conjugacy classes.

4. Inner product computations

Our next goal is to use our computation of χ' in Section 3 and the irreducible \mathbb{Q} -characters in Section 2.3 to compute the inner products $\langle \psi_i, \chi' \rangle$. Consider

$\langle \psi_i, \chi' \rangle$, where ψ_i is an irreducible \mathbb{Q} -character of $\mathrm{PSL}(2, q)$. The formula for the inner product is

$$\langle \psi_i, \chi' \rangle = \frac{1}{|G|} \sum_{g \in G} \psi_i(g) \chi'(g^{-1}).$$

Since g and g^{-1} are in the same conjugacy class and χ' is 0 for all elements that are not of order 1, 2, 3, or 7, we have the formula

$$\begin{aligned} \langle \psi_i, \chi' \rangle = \frac{1}{|G|} & (\psi_i(\mathrm{id}_G) \chi'(\mathrm{id}_G) + |[h_2]| \psi_i(h_2) \chi'(h_2) \\ & + |[h_3]| \psi_i(h_3) \chi'(h_3) + 3|[h_7]| \psi_i(h_7) \chi'(h_7)). \end{aligned}$$

In Section 3 we saw that

$$|[h_2]| \chi'(h_2) = -\frac{1}{2}|G|, \quad |[h_3]| \chi'(h_3) = -\frac{2}{3}|G|, \quad 3|[h_7]| \chi'(h_7) = -\frac{6}{7}|G|.$$

The formula for the inner product reduces to

$$\langle \psi_i, \chi' \rangle = \frac{1}{42} \psi_i(\mathrm{id}_G) - \frac{1}{2} \psi_i(h_2) - \frac{2}{3} \psi_i(h_3) - \frac{6}{7} \psi_i(h_7). \tag{5}$$

Since the values of the irreducible \mathbb{Q} -characters are based on whether the conjugacy classes of elements of orders 2, 3 and 7 are represented by \bar{a} or \bar{b} (which depends on the residue of q modulo 3, 4, or 7), the values of these characters, and the subsequent inner products, will depend on what q is modulo $3 \cdot 4 \cdot 7 = 84$.

4.1. Trivial character. Recall that χ is the Hurwitz character, and $\chi = 2 \cdot 1_G + \chi'$.

Proposition 8. $\langle 1_G, \chi \rangle = 0.$

Proof. By the calculation of χ , we have that $\langle 1_G, \chi \rangle = 2\langle 1_G, 1_G \rangle + \langle 1_G, \chi' \rangle$ and $\langle 1_G, 1_G \rangle = 1$. Consider $\langle 1_G, \chi' \rangle$. We use (5) to get

$$\langle 1_G, \chi' \rangle = \frac{1}{42} \cdot 1_G(1) - \frac{1}{2} \cdot 1_G(h_2) - \frac{2}{3} \cdot 1_G(h_3) - \frac{6}{7} \cdot 1_G(h_7) = \frac{1}{42} - \frac{1}{2} - \frac{2}{3} - \frac{6}{7} = -2.$$

Thus, $\langle 1_G, \chi' \rangle = 2 - 2 = 0.$ □

All other irreducible \mathbb{Q} -characters of $\mathrm{PSL}(2, q)$ have degree greater than 1. Hence by (1), where the n_i correspond to the degree of the i -th irreducible \mathbb{Q} -character, the decomposition of JX must have more than one factor.

Corollary 9. *No Hurwitz curve with automorphism group $\mathrm{PSL}(2, q)$ has a simple Jacobian variety.*

4.2. Character of degree q . Recall λ is the character of degree q . We again apply (5). Since the value of λ is either 1 or -1 depending on whether the element is in a conjugacy class represented by powers of \bar{a} or \bar{b} , we get that $\langle \lambda, \chi' \rangle = \frac{1}{42}(q - u)$, where u is given in Table 5 and the positive u -values correspond to positive $q \bmod 84$ values and the same holds for the negative values.

$q \pmod{84}$	Value of u
± 1	± 85
± 13	± 13
± 29	± 29
± 43	± 43

Table 5. Values for u in $\langle \lambda, \chi' \rangle$.

4.3. Characters of degree $\frac{1}{2}(q \pm 1)$. For $q \equiv 1 \pmod{4}$, this irreducible \mathbb{Q} -character is $\chi_1 + \chi_2$ and evaluates to $q + 1$ on the identity, $2(-1)^n$ on the conjugacy classes $[\bar{a}^n]$, and 0 on the conjugacy classes $[\bar{b}^m]$. Furthermore, the conjugacy class of elements of order 2 will always be in the set of conjugacy classes $[\bar{a}^n]$. We use (5) again, which becomes

$$\langle \chi_1 + \chi_2, \chi' \rangle = \frac{q + 1}{42} - \frac{(\chi_1 + \chi_2)(h_2)}{2} - \frac{2(\chi_1 + \chi_2)(h_3)}{3} - \frac{6(\chi_1 + \chi_2)(h_7)}{7}.$$

Determining these values depends on whether $q \equiv \pm 1 \pmod{3}$ and whether $q \equiv \pm 1 \pmod{7}$ (as we have discussed above, this distinguishes the cases where the elements of orders 3 and 7 are in conjugacy classes represented by powers of \bar{a} or \bar{b}). But additionally we need to determine if n is even or odd to determine the sign of $\chi_1 + \chi_2$. Recall n is given by $\frac{1}{6}(q - 1)$ for elements of order 3 and $\frac{1}{14}(q - 1)$ for elements of order 7. This requires us to consider values modulo $3 \cdot 4 \cdot 7 \cdot 2 = 168$.

Similar arguments will give us the values for $\gamma_1 + \gamma_2$ when $q \equiv -1 \pmod{4}$. In all cases, the inner product is given by $\frac{1}{42}(q - v)$, where v is given in Table 6. In the table, the positive values of $q \pmod{168}$ correspond to the positive v -values and the same holds for the negative values.

4.4. Characters of degree $q \pm 1$. The computations for the inner products of χ' with sums of μ_k or θ_t are similar. We recall the values of these \mathbb{Q} -characters

$q \pmod{168}$	Values of v
± 1	± 169
± 13	± 13
± 29	± 29
± 41	± 41
± 43	± 43
± 85	± 85
± 97	± 97
± 113	± 113

Table 6. Values for v in $\langle \chi_1 + \chi_2, \chi' \rangle$ or $\langle \gamma_1 + \gamma_2, \chi' \rangle$.

	$q \equiv 1 \pmod 4$	$q \equiv -1 \pmod 4$
μ_k	$z + (f - 1) \cdot 84$	$z - f \cdot 84$
θ_t	$z + f \cdot 84$	$z - (f - 1) \cdot 84$

Table 7. Values of w for the inner products of χ' with characters of degree $q \pm 1$.

on the conjugacy classes of orders 1, 2, 3, and 7 from Section 2.3. The values depend on whether the conjugacy classes are powers of \bar{a} or \bar{b} . To describe the value in all cases, we define two additional values. For $r = \gcd(k, \frac{1}{2}(q - 1))$ and $s = \gcd(t, \frac{1}{2}(q + 1))$, define f to be the number of 2, 3, and 7 which divide r (or s). Also define z to be the least residue of q modulo 84. Then the inner product with irreducible \mathbb{Q} -characters from Proposition 3(a) is

$$\frac{1}{2}\phi\left(\frac{q - 1}{2r}\right) \frac{q - w}{42},$$

and the inner product with irreducible \mathbb{Q} -characters from Proposition 3(b) is

$$\frac{1}{2}\phi\left(\frac{q + 1}{2s}\right) \frac{q - w}{42},$$

where w is given in Table 7.

Example. Continuing from the example in Section 2.3, let $q = 29$, so z also is 29. When $r = 1$ (or $s = 1$ or 5) then $f = 0$ and when $r = 2$ (or $s = 3$) we have $f = 1$. In this case (since $q = z$) if $f = 1$, then the value of the inner product on the corresponding irreducible \mathbb{Q} -character which is the sum of characters in M_r ($r = 2$) will be 0 and if $f = 0$, the value on the inner product of the corresponding irreducible \mathbb{Q} -character which is the sum of characters in Θ_s ($s = 1$ or 5) will also be 0. This just leaves two nonzero values to compute ($r = 1$ and $s = 3$),

$$\langle \mu_1 + \mu_3 + \mu_5, \chi' \rangle = \frac{1}{2}\phi\left(\frac{28}{2}\right) \cdot \left(\frac{29+55}{42}\right) = \frac{6}{2} \cdot 2 = 6$$

and

$$\langle \theta_3 + \theta_6, \chi' \rangle = \frac{1}{2}\phi\left(\frac{30}{6}\right) \cdot \left(\frac{29+55}{42}\right) = \frac{4}{2} \cdot 2 = 4.$$

5. Decomposition of Jacobian varieties

As described in the introduction, Jacobian varieties may be factored into the direct product of abelian varieties as in (1). The dimension of the factors is half of the inner product computed in Section 4. Collecting the information in the previous section we get the following result.

Theorem 10. *Let X be a Hurwitz curve with full automorphism group $\text{PSL}(2, q)$, where q is odd and $q > 27$. Let u, v , and w be as given in Tables 5, 6, and 7, respectively.*

When $q \equiv 1 \pmod 4$, the Jacobian variety of X is isogenous to

$$A^q \oplus B^{(q+1)/2} \oplus \prod_{\substack{r \mid (q-1)/2 \\ r < (q-5)/4}} C_r^{q+1} \oplus \prod_{\substack{s \mid (q+1)/2 \\ s < (q-1)/4}} D_s^{q-1},$$

and when $q \equiv -1 \pmod 4$, the Jacobian variety of X is isogenous to

$$A^q \oplus B^{(q-1)/2} \oplus \prod_{\substack{r \mid (q-1)/2 \\ r < (q-3)/4}} C_r^{q+1} \oplus \prod_{\substack{s \mid (q+1)/2 \\ s < (q-3)/4}} D_s^{q-1},$$

where the factors in the decomposition are abelian varieties and

- *A has dimension $\frac{1}{84}(q - u)$,*
- *B has dimension $\frac{1}{84}(q - v)$,*
- *each C_r has dimension $\frac{1}{168}\phi\left(\frac{1}{2r}(q - 1)\right) \cdot (q - w)$,*
- *and each D_s has dimension $\frac{1}{168}\phi\left(\frac{1}{2s}(q + 1)\right) \cdot (q - w)$.*

As mentioned in the introduction, the decomposition technique does not guarantee that the factors are indecomposable. Also, when determining w , note that the product indexed by r corresponds to inner products of characters which are sums of μ_k characters, and the product indexed by s corresponds to inner products of characters which are sums of the θ_t characters.

6. Special case

In the special case when $q = 27 = 3^3$, there are still three conjugacy classes of elements of order 7 and one of elements of order 2; however, there are now two conjugacy classes of elements of order 3. When we apply the decomposition technique to this special case we find

$$JX \sim E_1^{13} \times A_3^{26} \times E_2^{27},$$

where the E_i are elliptic curves and A_3 is a dimension-3 abelian variety. These factors correspond to nonzero inner products of χ with the character $\gamma_1 + \gamma_2$, a sum of θ_t , and λ , respectively.

Acknowledgments

The authors thank Grinnell College for providing generous summer funding through the Mentored Advanced Project program. They also thank the anonymous referee for suggestions which improve the exposition and clarity of the paper.

References

- [Cardona 2004] G. Cardona, “ \mathbb{Q} -curves and abelian varieties of GL_2 -type from dihedral genus 2 curves”, pp. 45–52 in *Modular curves and abelian varieties*, edited by J. Cremona et al., Progr. Math. **224**, Birkhäuser, Basel, 2004. [MR 2058641](#) [Zbl 1080.11045](#)
- [Conder 1990] M. Conder, “Hurwitz groups: a brief survey”, *Bull. Amer. Math. Soc. (N.S.)* **23**:2 (1990), 359–370. [MR 1041434](#) [Zbl 0716.20015](#)
- [Earle 2006] C. J. Earle, “The genus two Jacobians that are isomorphic to a product of elliptic curves”, pp. 27–36 in *The geometry of Riemann surfaces and abelian varieties*, edited by J. M. Muñoz Porras et al., Contemp. Math. **397**, Amer. Math. Soc., Providence, RI, 2006. [MR 2217995](#) [Zbl 1099.14017](#)
- [Howe et al. 2000] E. W. Howe, F. Leprévost, and B. Poonen, “Large torsion subgroups of split Jacobians of curves of genus two or three”, *Forum Math.* **12**:3 (2000), 315–364. [MR 1748483](#) [Zbl 0983.11037](#)
- [Janusz 1974] G. J. Janusz, “Simple components of $Q[SL(2, q)]$ ”, *Comm. Algebra* **1** (1974), 1–22. [MR 0344323](#) [Zbl 0281.20003](#)
- [Kani and Rosen 1989] E. Kani and M. Rosen, “Idempotent relations and factors of Jacobians”, *Math. Ann.* **284**:2 (1989), 307–327. [MR 1000113](#) [Zbl 0652.14011](#)
- [Karpilovsky 1994] G. Karpilovsky, *Group representations*, vol. 3, North-Holland Mathematics Studies **180**, North-Holland Publishing Co., Amsterdam, 1994. [MR 1280715](#) [Zbl 0804.20001](#)
- [Kuwata 2005] M. Kuwata, “Quadratic twists of an elliptic curve and maps from a hyperelliptic curve”, *Math. J. Okayama Univ.* **47** (2005), 85–97. [MR 2198864](#) [Zbl 1161.11353](#)
- [Macbeath 1969] A. M. Macbeath, “Generators of the linear fractional groups”, pp. 14–32 in *Number Theory* (Houston, Tex., 1967), vol. XII, Proc. Sympos. Pure Math., Amer. Math. Soc., Providence, R.I., 1969. [MR 0262379](#) [Zbl 0192.35703](#)
- [Magaard et al. 2009] K. Magaard, T. Shaska, and H. Völklein, “Genus 2 curves that admit a degree 5 map to an elliptic curve”, *Forum Math.* **21**:3 (2009), 547–566. [MR 2526800](#) [Zbl 1174.14025](#)
- [Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, 1980. [MR 559531](#) [Zbl 0433.14012](#)
- [Paulhus 2008] J. Paulhus, “Decomposing Jacobians of curves with extra automorphisms”, *Acta Arith.* **132**:3 (2008), 231–244. [MR 2403651](#) [Zbl 1142.14017](#)
- [Rubin and Silverberg 2001] K. Rubin and A. Silverberg, “Rank frequencies for quadratic twists of elliptic curves”, *Experiment. Math.* **10**:4 (2001), 559–569. [MR 1881757](#) [Zbl 1035.11025](#)
- [Wolfart 2002] J. Wolfart, “Regular dessins, endomorphisms of Jacobians, and transcendence”, pp. 107–120 in *A panorama of number theory or the view from Baker’s garden* (Zürich, 1999), edited by G. Wüstholz, Cambridge Univ. Press, 2002. [MR 1975447](#) [Zbl 1042.14016](#)

Received: 2015-02-05

Revised: 2015-07-08

Accepted: 2015-07-20

fischera@grinnell.edu*Department of Mathematics and Statistics, Grinnell College, Grinnell, IA 50112, United States*liumouch@grinnell.edu*Department of Mathematics and Statistics, Grinnell College, Grinnell, IA 50112, United States*paulhusj@grinnell.edu*Department of Mathematics and Statistics, Grinnell College, Grinnell, IA 50112, United States*

Avoiding approximate repetitions with respect to the longest common subsequence distance

Serina Camungol and Narad Rampersad

(Communicated by Joshua Cooper)

Ochem, Rampersad, and Shallit gave various examples of infinite words avoiding what they called approximate repetitions. An approximate repetition is a factor of the form xx' , where x and x' are close to being identical. In their work, they measured the similarity of x and x' using either the Hamming distance or the edit distance. In this paper, we show the existence of words avoiding approximate repetitions, where the measure of similarity between adjacent factors is based on the length of the longest common subsequence. Our principal technique is the so-called “entropy compression” method, which has its origins in Moser and Tardos’s algorithmic version of the Lovász local lemma.

1. Introduction

A now classical result of Thue [1906] showed the existence of an infinite word over a 3-letter alphabet avoiding *squares*, that is, factors of the form xx . Ochem, Rampersad, and Shallit [Ochem et al. 2008] generalized the work of Thue by constructing infinite words over a finite alphabet that avoid factors of the form xx' , where x and x' are close to being identical. In most of their work, the closeness of x and x' was measured using the Hamming distance; they also have some results where the edit distance was used instead. Here, we measure the closeness of two words based on the length of their longest common subsequence.

The most common metrics used to measure the distance between strings are the edit distance, the Hamming distance, and the longest common subsequence metric. The edit distance is the most general: it is defined as the smallest number of single-letter insertions, deletions, and substitutions needed to transform one string into the other. The other two distances can be viewed as restricted versions of the edit distance: the Hamming distance (between strings of the same length) is the edit distance where only the substitution operation is permitted; the longest common subsequence metric allows only insertions and deletions.

MSC2010: 68R15.

Keywords: approximate repetition, longest common subsequence, entropy compression.

The study of the longest common subsequence of two (or several) sequences has a lengthy history (which, at least initially, was motivated by the biological problem of comparing long protein or genomic sequences). For example, Chvátal and Sankoff [1975] explored the following question: given two random sequences of length n over a k -letter alphabet, what is the expected length of their longest common subsequence? Questions concerning longest common subsequences in words continue to be studied to this day (see the recent preprint [Bukh and Zhou 2016], for example).

Ochem, Rampersad, and Shallit [2008] previously studied the avoidability of approximate squares with respect to Hamming distance and edit distance. Using the longest common subsequence metric has not yet been done, so it is the aim of this paper to consider the avoidability of approximate squares with respect to this measure of distance.

Our main result is nonconstructive—indeed it seems to be quite difficult to find explicit constructions for words avoiding the kinds of repetitions we consider here—and is based on the so-called “entropy compression” method, which originates from Moser and Tardos’s algorithmic version [2010] of the Lovász local lemma. This method has recently been applied very successfully in combinatorics on words, for instance in [Grytczuk et al. 2013; 2011]. Ochem and Pinlou [2014] also recently resolved a longstanding conjecture of Cassaigne using this method (this was also accomplished independently by Blanchet-Sadri and Woodhouse [2013] using a different method).

2. Measuring similarity

The definitions given in this section are essentially those of Ochem et al. except that they are based on the longest common subsequence distance rather than the Hamming distance.

For words x and x' , let $\text{lcs}(x, x')$ denote the length of a longest common subsequence of x and x' . For example,

$$\text{lcs}(0120, 1220) = 3.$$

Given two words x, x' of the same length, we define their *similarity*, $s(x, x')$, by

$$s(x, x') := \frac{\text{lcs}(x, x')}{|x|}.$$

For example,

$$s(20120121, 02102012) = \frac{3}{4}.$$

The *similarity coefficient* $\text{sc}(z)$ of a finite word z is defined to be

$$\text{sc}(z) := \max\{s(x, x') : xx' \text{ a subword of } z \text{ and } |x| = |x'|\}.$$

If $sc(z) = \alpha$, we say that z is α -similar. If z is an infinite word, then its similarity coefficient is defined by

$$sc(z) := \sup\{s(x, x') : xx' \text{ a subword of } z \text{ and } |x| = |x'|\}.$$

Again, if $sc(z) = \alpha$ then we say that z is α -similar.

3. Infinite words with low similarity

Our main result is the following:

Theorem 1. *Let $0 < \alpha < 1$ and let $k > 16^{1/\alpha}$ be an integer. Then there exists an infinite word z over an alphabet of size k such that $sc(z) \leq \alpha$.*

To prove this, we follow the method of Grytczuk, Kozik, and Witkowski [Grytczuk et al. 2011]. We begin by defining a randomized algorithm which attempts to construct a word S of length n with similarity coefficient at most α by a sort of backtracking procedure. Let s_i denote the i -th element of S .

Input : n, k, α

- 1: $S = \emptyset, i = 1$
- 2: **while** $i \leq n$ **do**
- 3: randomly choose $a \in \{1, \dots, k\}$ and set $s_i = a$
- 4: **if** $sc(s_1s_2 \dots s_i) \leq \alpha$ **then** set i to $i + 1$
- 5: **else** $s_1s_2 \dots s_i$ is β -similar, $\beta > \alpha$, and contains a subword xx' such that $|x| = |x'| = \ell$, $\ell \leq i/2$ and $s(x, x') = \beta$, say $x = s_{t+1}s_{t+2} \dots s_{t+\ell}$ and $x' = s_{t+\ell+1}s_{t+\ell+2} \dots s_{t+2\ell}$, where $t + 2\ell = i$.
- 6: **for** $t + \ell + 1 \leq j \leq t + 2\ell$ **do**
- 7: delete s_j
- 8: **end for**
- 9: set $i = t + \ell + 1$
- 10: **end if**
- 11: **end while**

Algorithm 1. Choosing a sequence with similarity coefficient at most α .

The algorithm generates consecutive terms of a sequence S by choosing symbols at random (uniformly and independently). Every time a β -similar subword xx' is created, where $\beta > \alpha$, the algorithm erases x' , to ensure that the β -similar subword is deleted. Note that in line 6 the subword xx' must occur as a suffix of $s_1s_2 \dots s_i$ (i.e., $t + 2\ell = i$), since if it occurred elsewhere it would have been detected at an earlier stage of the algorithm and its second half would have been deleted.

It is easy to see that the algorithm terminates after a word of length n with similarity coefficient at most α has been produced. The general idea is to prove the algorithm cannot continue forever with all possible evaluations of the random inputs.

Fix a real number α . We will show that for every positive integer n there exists a word of length n with similarity coefficient at most α . The existence of an infinite word with the same property then follows by a standard compactness argument.

Let n be a positive integer, and suppose for the sake of contradiction that every possible execution of the algorithm fails to produce a sequence of length n . We are going to count the possible executions of the algorithm in two ways.

Suppose the algorithm runs for M steps. By “step” we mean appending a letter to the sequence S (which only happens in line 3). Let a_1, a_2, \dots, a_M be the sequence of values chosen randomly and independently in the first M steps of the algorithm. Each a_j , $1 \leq j \leq M$, can take k different values; thus there are k^M such sequences.

The second way of counting involves analyzing the behavior of the algorithm. For a fixed evaluation of the first M random choices of the algorithm we define a 4-tuple (R, X, Y, S) , called a *log*, whose elements consist of the following:

- A route R in the upper right quadrant of the Cartesian plane, going from coordinate $(0, 0)$ to coordinate $(2M, 0)$, with possible moves $(1, 1)$ and $(1, -1)$, which never goes below the axis $y = 0$.
- A sequence X over $\{1, \dots, k\} \cup \{*\}$ whose elements correspond to the peaks on the route R , where a *peak* is defined as a move $(1, 1)$ followed immediately by a move $(1, -1)$.
- A sequence Y over $\{0, *\}$ whose elements also correspond to the peaks in R .
- A sequence S over $\{1, \dots, k\}$ produced after M steps of the algorithm.

The values of R , X , Y , and S are determined as follows. Each time the algorithm appends a letter to the sequence S , we append a move $(1, 1)$ to the route R and every time an s_i is deleted we append $(1, -1)$. Every down-step $(1, -1)$ corresponds to an up-step $(1, 1)$ so we never reach below the y -axis. At the end of computations we add to the route R one down-step for each element of S which was not deleted at any point in the algorithm, bringing us to the point $(2M, 0)$. If a β -similar word is created, say xx' , we concatenate to X the word obtained from x' by replacing the elements of the longest common subsequence of x and x' with the symbol star $(*)$. We also concatenate to Y the word obtained from x by replacing the elements of the longest common subsequence of x and x' with the star symbol and setting all other positions equal to zero. At the end of computations we pad X and Y with enough stars so that $|X| = |Y| = M$. Lastly, S is the sequence produced by [Algorithm 1](#) after making M random selections from $\{1, \dots, k\}$.

Example 2. For example, let us choose $\alpha = \frac{37}{50}$. Then $\lceil 16^{\frac{50}{37}} \rceil = 43$ and we have alphabet $\{1, \dots, 43\}$ and $\log \{R = \emptyset, X = \emptyset, Y = \emptyset, S = \emptyset\}$. Suppose we create the word 12324541465 after 11 steps of the algorithm. Each of our steps avoids creating a β -similar word, so at each step we append $(1, 1)$ to R and the randomly selected letter to S . Thus we have

$$\{R = (1, 1)^{11}, X = \emptyset, Y = \emptyset, S = 12324541465\}.$$

Suppose in the 12th step of the algorithm we append 4 to S ; then our log becomes

$$\{R = (1, 1)^{12}, X = \emptyset, Y = \emptyset, S = 123245414654\}.$$

Observe that the factor $xx' = 45414654$ is $\frac{3}{4}$ -similar, where $x = 4541$, $x' = 4654$ and the longest common subsequence of x and x' is 454. As $\frac{3}{4} > \frac{37}{50}$, we replace the longest common subsequence elements of x and x' with stars and we append $*6**$ to X and $***0$ to Y . We then delete x' and append to R a $(1, -1)$ for each deleted element. This results in the log

$$\{R = (1, 1)^{12}(1, -1)^4, X = *6**, Y = ***0, S = 12324541\}.$$

Lemma 3. Every log corresponding to an execution of the algorithm uniquely determines the sequence a_1, a_2, \dots, a_M of the first M values chosen randomly and independently in this execution of Algorithm 1.

Proof. Let us fix some $\log (R, X, Y, S)$. Before we decode a_1, a_2, \dots, a_M , we do some preparatory analysis. We construct a sequence $D = (d_1, d_2, \dots, d_p)$, corresponding to the lengths of consecutive down-steps, $(1, -1)$, of R . Let $N = d_1 + d_2 + \dots + d_p$. Next we delete the last $M - N$ stars from X and partition the resulting sequence into blocks of lengths d_1, d_2, \dots, d_p . Let X' be the sequence of these blocks; i.e.,

$$X' = (x_1x_2 \dots x_{d_1}, x_{d_1+1}x_{d_1+2} \dots x_{d_1+d_2}, \dots, x_{N-d_p+1}x_{N-d_p+2} \dots x_N).$$

We do the same for the sequence Y , obtaining a sequence of blocks

$$Y' = (y_1y_2 \dots y_{d_1}, y_{d_1+1}y_{d_1+2} \dots y_{d_1+d_2}, \dots, y_{N-d_p+1}y_{N-d_p+2} \dots y_N).$$

Next we use information from route R to determine which s_i , $1 \leq i \leq n$, were not deleted at each step of Algorithm 1 and to find the coordinates of the blocks which were deleted at line 6 of the algorithm. Notice that appending some letter from $\{1, \dots, k\}$ to S corresponds to some up-step $(1, 1)$ on the route R , while deleting an s_i corresponds to some down-step $(1, -1)$ on the route R . We analyze the route R , starting from the point $(0, 0)$ to the point $(2M, 0)$. Assume the first peak occurs between the j -th and $(j+1)$ -th step. As this is the first time that we erase elements, we know that s_1, \dots, s_j are the only nondeleted elements at this point. From the

number of down-steps on R we deduce the length of the deleted block, say there are d_1 down-steps, and remember that for this peak we deleted $s_{j-d_1+1}, s_{j-d_1+2}, \dots, s_j$. Now again each up-step on R denotes appending some value of $\{1, \dots, k\}$ to S . Continuing on in this manner, we are able to determine exactly which position was set last as we reach the next peak. From this information it is easy to determine which positions were deleted as a result of erasing the repetition. We repeat these operations until we reach the end of the route R .

After these preparatory measures we are ready to decode a_1, a_2, \dots, a_M . We consider the sequence R in reverse order, from the point $(2M, 0)$ to the point $(0, 0)$, modifying the sequences X' and Y' from the preparatory step and the final sequence S . We use information encoded in S, X' and Y' , as well as knowledge from the preparatory step.

As we process the elements of R in reverse order, suppose we encounter an up-step. Note that each up-step corresponds to some a_j . In the preparatory analysis we determined the indices of elements a_j in S , so each time there is an up-step of R we assign to a_j a value from the appropriate s_i (where i was determined in the preparatory step), and delete s_i .

Now we suppose that we encounter a down-step of R (or rather, a block of down-steps of R). At the end of R there is some number of down-steps corresponding to the last nondeleted elements of S (the elements added at the end of computations); we skip these elements and move on. The first block of down-steps that follows an up-step has length d_p and corresponds to the last element of X' , say X'_N , as well as the last element of Y' , say Y'_N . Let $s_i, s_{i+1}, \dots, s_{i+d_p-1}$ be the elements of S that were deleted at each down-step in this block of down-steps. We reconstruct the values of these elements by using the information from $s_{i-d_p}, s_{i-d_p+1}, \dots, s_{i-1}, Y'_N$, and X'_N .

Together, the elements of $s_{i-d_p}, s_{i-d_p+1}, \dots, s_{i-1}$ that correspond to the star elements of Y'_N form the longest common subsequence of $s_{i-d_p}, s_{i-d_p+1}, \dots, s_{i-1}$ and $s_i, s_{i+1}, \dots, s_{i+d_p-1}$; call this sequence LCS . The values of $s_i, s_{i+1}, \dots, s_{i+d_p-1}$ are obtained by replacing the stars in X'_N with the elements of LCS . We add these elements to the end of S and repeat the process. Continuing in this manner, we are able to reconstruct all deleted blocks, and therefore the entire sequence a_1, a_2, \dots, a_M . \square

We have just shown that there is an injective mapping between the set of all sequences of randomly chosen values during the execution of the algorithm and the set of all logs. Consequently, the number of different logs is always greater than or equal to the number of possible sequences $a_1, a_2, a_3, \dots, a_M$. We now derive an upper bound for the number of possible logs.

The number of possible routes R , of length $2M$ and possible moves $(1, 1)$ and $(1, -1)$, in the upper right quadrant of the Cartesian plane is the M -th Catalan number C_M .

To count X we first note that $|X| = M$ and that each deleted factor x' has (strictly) more than $\alpha|x'|$ star positions, so it follows that X has more than αM star positions. Let j be the number of stars in X . There are k choices for the $M - j$ nonstar positions in X , so there are $\binom{M}{j}k^{M-j}$ possibilities for X . Now if X has j positions with stars, then so does Y , and the remaining positions in Y are 0's. Thus, there are $\binom{M}{j}$ possibilities for Y , and hence $\binom{M}{j}^2 k^{M-j}$ possibilities for the pair (X, Y) . Summing over all j , we conclude that there are

$$\sum_{j=\lceil \alpha M \rceil}^M \binom{M}{j}^2 k^{M-j}$$

possibilities for the pair (X, Y) .

The sequence S consists of at most n elements of value between 1 and k , so there are $(k^{n+1} - 1)/(k - 1)$ possible sequences S .

Multiplying these individual bounds together brings us to the conclusion that the number of possible logs is at most

$$\frac{k^{n+1} - 1}{k - 1} C_M \sum_{j=\lceil \alpha M \rceil}^M \binom{M}{j}^2 k^{M-j}.$$

Comparing with the number k^M of possible choices for the sequence a_1, \dots, a_M we get the inequality

$$k^M \leq \frac{k^{n+1} - 1}{k - 1} C_M \sum_{j=\lceil \alpha M \rceil}^M \binom{M}{j}^2 k^{M-j}.$$

Asymptotically, the Catalan numbers C_M satisfy $C_M \sim 4^M/(M\sqrt{\pi M})$, and $\binom{M}{j} < 2^M$, which implies that

$$k^M \ll \frac{k^{n+1} - 1}{k - 1} \frac{4^M}{M\sqrt{\pi M}} \sum_{j=\lceil \alpha M \rceil}^M (2^M)^2 k^{M-j}.$$

Simplifying we get that

$$\begin{aligned} k^M &\ll \frac{k^{n+1} - 1}{k - 1} \frac{4^M}{M\sqrt{\pi M}} 4^M \sum_{j=\lceil \alpha M \rceil}^M k^{M-j} \\ &= \frac{k^{n+1} - 1}{k - 1} \frac{16^M}{M\sqrt{\pi M}} \sum_{j=0}^{M-\lceil \alpha M \rceil} k^j \\ &= \frac{16^M}{M\sqrt{\pi M}} \frac{(k^{n+1} - 1)(k^{M-\lceil \alpha M \rceil+1} - 1)}{(k - 1)^2} \leq k^{n+2} \frac{16^M}{M\sqrt{\pi M}} \frac{k^{M(1-\alpha)}}{(k - 1)^2}. \end{aligned}$$

It is easy to verify that when $k > 16^{1/\alpha}$, the last expression in the above calculation is $o(k^M)$, which is a contradiction. This contradiction implies that for some specific choices of a_1, a_2, \dots [Algorithm 1](#) stops (i.e., produces a word of length n with similarity coefficient at most α). This completes the proof of [Theorem 1](#).

4. Similarity coefficients for small alphabets

Almost certainly, the bound of $16^{1/\alpha}$ for the size of the alphabet needed to obtain an infinite word with similarity coefficient at most α is far larger than the true optimal alphabet size. For example, for $\alpha = 0.9$ we get an alphabet size of 22, which is surely much larger than necessary. In this section we investigate the following question: given an alphabet Σ of size k , what is the smallest similarity coefficient possible over all infinite words over Σ ? Implementing an algorithm similar to that of [Section 3](#) allows us to get an idea of which values of α , where $0 < \alpha < 1$, are avoidable and unavoidable. Given a similarity coefficient α to avoid, a length n , and an alphabet size k , the algorithm starts at 0 and appends letters until a word of length n with similarity coefficient less than α is obtained. If a factor with similarity coefficient at least α is created, the last appended letter is deleted. If appending no other letter avoids α , the algorithm deletes yet another letter, and so on and so forth. The algorithm continues until a word of length n is produced. If no word of length n avoids α , the algorithm returns the longest word avoiding α . If, on the other hand, the algorithm produces words with similarity coefficient less than α for longer and longer values of n , then we take this as evidence that there exists an infinite word over a k -letter alphabet with similarity coefficient less than α . We performed this computation for various alphabet sizes, and the results can be found in [Table 1](#).

For each lower bound reported in the table, we are certain that there does not exist an infinite word with this similarity coefficient. However, the upper bounds are only conjectural: the backtracking algorithm described above produces long words with similarity coefficient less than the stated bound, but we have no conclusive proof that an infinite word exists.

alphabet size	similarity coefficient
3	$0.888 < \alpha < 0.901$
4	$0.690 < \alpha < 0.760$
5	$0.590 < \alpha < 0.700$
6	$0.500 < \alpha < 0.650$
7	$0.450 < \alpha < 0.650$
8	$0.400 < \alpha < 0.570$

Table 1. Results of the backtracking algorithm. (Upper bounds are conjectural.)

alphabet size	similarity coefficient	prefix length	factor length
3	-	2401	500
4	$11/12$	912	500
5	$16/19$	9261	399
6	$10/13$	9261	312
7	-	5000	218
8	$12/15$	5000	445

Table 2. Results of computer calculations on Moulin Ollagnier’s words.

In fact, we cannot produce a single explicit construction (with proof) of an infinite word with similarity coefficient less than 1. However, computer calculations suggest that the so-called *Dejean words* seem to have fairly low similarity (though not nearly as low as the values given in Table 1). We now report the results of our computer calculations on the words constructed by Moulin Ollagnier [1992] in order to verify Dejean’s Conjecture for small alphabet sizes. For each alphabet size $k = 3, \dots, 11$, Ollagnier constructed an infinite word over a k -letter alphabet. Each such word verified a conjecture of Dejean [1972] concerning the repetitions avoidable on a k -letter alphabet. See [Moulin Ollagnier 1992] for the precise nature of the construction as well as the details of Dejean’s conjecture. In Table 2, we report the largest similarity coefficient found among all factors of Moulin Ollagnier’s words, up to a certain length. In the table, “prefix length” is the length of the prefix of the infinite word that we examined, “factor length” is the maximum length of the factors of this prefix that we examined, and a “-” signifies a continuous increase in similarity coefficient as the lengths of the factors increase.

Two natural problems suggest themselves:

- (1) Determine the similarity coefficients of Moulin Ollagnier’s words.
- (2) For each alphabet size k , determine the least similarity coefficient among all infinite words over a k -letter alphabet.

The second question is likely quite difficult. Even an answer just for the 3-letter alphabet would be nice to have.

Acknowledgments

Rampersad is supported by an NSERC Discovery Grant.

References

[Blanchet-Sadri and Woodhouse 2013] F. Blanchet-Sadri and B. Woodhouse, “Strict bounds for pattern avoidance”, *Theoret. Comput. Sci.* **506** (2013), 17–28. MR Zbl

- [Bukh and Zhou 2016] B. Bukh and L. Zhou, “Twins in words and long common subsequences in permutations”, *Israel J. Math.* (online publication April 2016).
- [Chvátal and Sankoff 1975] V. Chvátal and D. Sankoff, “Longest common subsequences of two random sequences”, *J. Appl. Probability* **12** (1975), 306–315. [MR](#) [Zbl](#)
- [Dejean 1972] F. Dejean, “Sur un théorème de Thue”, *J. Combinatorial Theory Ser. A* **13** (1972), 90–99. [MR](#) [Zbl](#)
- [Grytczuk et al. 2011] J. Grytczuk, J. Kozik, and M. Witkowski, “Nonrepetitive sequences on arithmetic progressions”, *Electron. J. Combin.* **18**:1 (2011), #P209. [MR](#)
- [Grytczuk et al. 2013] J. Grytczuk, J. Kozik, and P. Micek, “New approach to nonrepetitive sequences”, *Random Structures Algorithms* **42**:2 (2013), 214–225. [MR](#) [Zbl](#)
- [Moser and Tardos 2010] R. A. Moser and G. Tardos, “A constructive proof of the general Lovász local lemma”, *J. ACM* **57**:2 (2010), Art. 11. [MR](#)
- [Moulin Ollagnier 1992] J. Moulin Ollagnier, “Proof of Dejean’s conjecture for alphabets with 5, 6, 7, 8, 9, 10 and 11 letters”, *Theoret. Comput. Sci.* **95**:2 (1992), 187–205. [MR](#) [Zbl](#)
- [Ochem and Pinlou 2014] P. Ochem and A. Pinlou, “Application of entropy compression in pattern avoidance”, *Electron. J. Combin.* **21**:2 (2014), #P2.7. [MR](#)
- [Ochem et al. 2008] P. Ochem, N. Rampersad, and J. Shallit, “Avoiding approximate squares”, *Internat. J. Found. Comput. Sci.* **19**:3 (2008), 633–648. [MR](#) [Zbl](#)
- [Thue 1906] A. Thue, *Über unendliche Zeichenreihen*, Videnskabs-Selskabets Skrifter Math.-Naturv. Klasse **1906** No. 7, Jacob Dybwad for the Fridtjof Nansens Fond, Kristiania, 1906. [JFM](#)

Received: 2015-03-20

Revised: 2015-09-06

Accepted: 2015-09-17

serina.camungol@gmail.com

*Department of Mathematics and Statistics, University of
Winnipeg, 515 Portage Ave., Winnipeg MB R3B 2E9, Canada*

narad.rampersad@gmail.com

*Department of Mathematics and Statistics, University of
Winnipeg, 515 Portage Ave., Winnipeg MB R3B 2E9, Canada*

Prime vertex labelings of several families of graphs

Nathan Diefenderfer, Dana C. Ernst, Michael G. Hastings,
Levi N. Heath, Hannah Prawzinsky, Briahna Preston,
Jeff Rushall, Emily White and Alyssa Whittemore

(Communicated by Joseph A. Gallian)

A simple and connected n -vertex graph has a prime vertex labeling if the vertices can be injectively labeled with the integers $1, 2, 3, \dots, n$ such that adjacent vertices have relatively prime labels. We will present previously unknown prime vertex labelings for new families of graphs, including cycle pendant stars, cycle chains, prisms, and generalized books.

1. Introduction

The focus of this paper is prime vertex labelings, wherein adjacent vertices of simple, connected graphs are assigned integer labels that are relatively prime. Currently, the two most prominent open conjectures involving prime vertex labelings are:

- All tree graphs have a prime vertex labeling (Entringer–Tout conjecture).
- All unicyclic graphs have a prime vertex labeling [Seoud and Youssef 1999].

While we will address one infinite family that is unicyclic, our primary concern will be nonunicyclic graphs.

A *graph* G is a set of vertices, $V(G)$, together with a set of edges, $E(G)$, connecting some subset, possibly empty, of the vertices. If $u, v \in V(G)$ are connected by an edge, we say u and v are *adjacent*. The *degree* of a vertex u is the number of edges incident to u . A *subgraph* H of a graph G is a graph whose vertex set is a subset of that of G , and whose adjacency relation is a subset of that of G restricted to this subset.

We will restrict our attention to graphs that are simple (i.e., graphs that do not have multiple edges between pairs of vertices nor edges that connect a vertex to itself) and connected (i.e., graphs that do not consist of two or more disjoint “pieces”). For the remainder of this paper, all graphs are assumed to be simple and connected.

MSC2010: 05C78.

Keywords: graph labeling, prime vertex labeling, prime graphs.

This research was supported by the National Science Foundation grant #DMS-1148695 through the Center for Undergraduate Research in Mathematics (CURM).

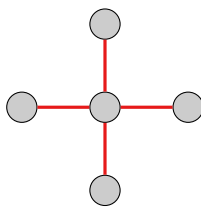


Figure 1. The star S_4 .

Next, we define a few important families of graphs. For $n \geq 2$, an n -*path* (or simply *path*), denoted P_n , is a connected graph consisting of two vertices of degree 1 and $n - 2$ vertices of degree 2. For $n \geq 3$, an n -*cycle* (or simply *cycle*), denoted C_n , is a connected graph consisting of n vertices, each of degree 2. Note that both P_n and C_n have n vertices, while P_n has $n - 1$ edges and C_n has n edges. An n -*star* (or simply *star*), denoted S_n , is a graph consisting of one vertex of degree n , called the *center*, and n vertices of degree 1. Note that S_n consists of $n + 1$ vertices and n edges. The star S_4 is shown in Figure 1.

A simple and connected n -vertex graph is said to have a *prime vertex labeling* if the vertices can be injectively labeled with the integers $1, 2, 3, \dots, n$ such that adjacent vertices have relatively prime labels. For brevity, if a graph has a prime vertex labeling, we will say that the graph is *prime*. The many familiar families of graphs that are known to be prime include paths, cycles, and stars.

In this paper, we will present previously unknown prime vertex labelings for several infinite families of graphs, including cycle pendant stars (Section 2), cycle chains (Section 3), prisms (Section 4), and generalized books (Section 5). Finally, some conjectures and potential future work will be described in Section 6.

2. Cycle pendant stars

The focus of this section is on a type of unicyclic graph (i.e., a graph containing exactly one cycle as a subgraph), which was inspired by Seoud and Youssef's conjecture [1999] that all unicyclic graphs are prime. Instead of attempting to prove the conjecture outright, which we anticipate would require advanced machinery from linear algebra, most authors concentrate on finding prime labelings for specific families of unicyclic graphs. In particular, this was our endeavor in [Diefenderfer et al. 2015].

Every vertex lying on the cycle of a unicyclic graph will be referred to as a *cycle vertex*. In a unicyclic graph, a *spur* is an edge with exactly one vertex on the cycle. The noncycle vertex of a spur is called a *spur vertex*. For example, the graph shown in Figure 2 is a unicyclic graph with five spurs. In this case, the vertices labeled by c_1, c_2, c_3 , and c_4 are cycle vertices, while the vertices labeled by p_1, p_2, p_3, p_4 , and p_5 are spur vertices.

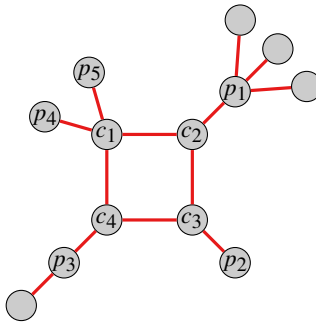


Figure 2. Example of a unicyclic graph with five spurs.

Seoud and Youssef [1999] investigated cycles with identical complete binary trees attached to each cycle vertex. These unicyclic graphs can also be viewed as cycles with various levels of the star S_2 attached to one another. This observation led to the following generalization, which features a single “level” of stars, but with differing star sizes. A *cycle pendant star*, denoted $C_n \star P_2 \star S_m$, is the graph that results from attaching the path P_2 to each vertex of C_n followed by attaching the star S_m at its center to each spur vertex. For example, the graph $C_5 \star P_2 \star S_6$ is shown in Figure 3. Note that our \star notation is not a construction typically found in the literature and refers to “selectively gluing” copies of one graph to another.

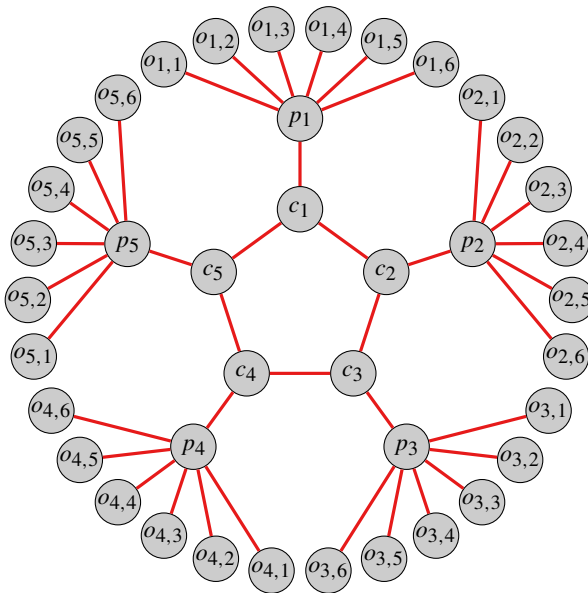


Figure 3. The graph $C_5 \star P_2 \star S_6$.

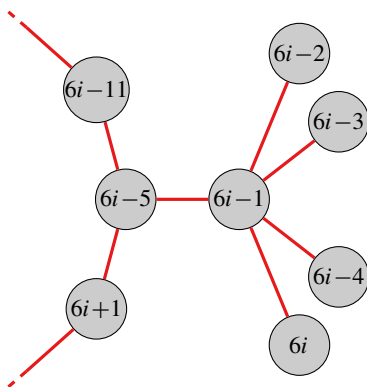


Figure 4. The generalized prime vertex labeling of $C_n \star P_2 \star S_4$.

Theorem 2.1. All $C_n \star P_2 \star S_m$ with $0 \leq m \leq 8$ are prime.

Proof. The cases involving $m = 0, 1, 2, 3$ correspond to familiar graphs with known prime vertex labelings. Seoud and Youssef [1999] showed that pendant graphs ($m = 0$), double pendant graphs ($m = 1$), and graphs with identical complete binary trees attached to the spur vertices of a pendant graph are prime (which includes the case $m = 2$). Diefenderfer et al. [2015] showed that $C_n \star P_2 \star S_3$ has a prime vertex labeling.

Now, consider $C_n \star P_2 \star S_m$, with $4 \leq m \leq 8$. Let c_1, c_2, \dots, c_n denote the consecutive cycle vertices of C_n , let p_i denote the spur vertex adjacent to c_i , and let $o_{i,k}$, with $1 \leq k \leq m$, denote the outer vertices adjacent to p_i . For example, see the identification of vertices depicted in Figure 3. For each case, it is straightforward and routine to verify that all adjacent vertices have relatively prime labels. However, for the reader’s benefit, we will describe the case involving $m = 6$ in detail. Similar reasoning is required for each of the remaining cases.

For the $m = 4$ case, the labeling function $f : V \rightarrow \{1, 2, \dots, 6n\}$ is given by

$$\begin{aligned}
 f(c_i) &= 6i - 5 && \text{if } 1 \leq i \leq n, \\
 f(p_i) &= 6i - 1 && \text{if } 1 \leq i \leq n, \\
 f(o_{i,1}) &= 6i - 2 && \text{if } 1 \leq i \leq n, \\
 f(o_{i,2}) &= 6i - 3 && \text{if } 1 \leq i \leq n, \\
 f(o_{i,3}) &= 6i - 4 && \text{if } 1 \leq i \leq n, \\
 f(o_{i,4}) &= 6i && \text{if } 1 \leq i \leq n.
 \end{aligned}$$

Figure 4 depicts the generalized labeling function in this case for $C_n \star P_2 \star S_4$. The prime vertex labeling of $C_4 \star P_2 \star S_4$ using this labeling appears in Figure 5.

For the $m = 5$ case, the labeling function $f : V \rightarrow \{1, 2, \dots, 7n\}$ is given by

$$\begin{aligned}
 f(c_i) &= 7i - 6 \quad \text{if } 1 \leq i \leq n, \\
 f(p_i) &= \begin{cases} 7i - 2 & \text{if } i \equiv_6 1, 3, \\ 7i - 3 & \text{if } i \equiv_6 2, 4, \\ 7i - 4 & \text{if } i \equiv_6 5, \\ 7i - 5 & \text{if } i \equiv_6 0, i \not\equiv_{30} 0, \\ 7i - 1 & \text{if } i \equiv_{30} 0, \end{cases} \\
 f(o_{i,1}) &= \begin{cases} 7i - 5 & \text{if } i \not\equiv_6 0 \text{ or } i \equiv_{30} 0, \\ 7i - 4 & \text{if } i \equiv_6 0, i \not\equiv_{30} 0, \end{cases} \\
 f(o_{i,2}) &= \begin{cases} 7i - 4 & \text{if } i \not\equiv_6 0, 5 \text{ or } i \equiv_{30} 0, \\ 7i - 3 & \text{if } i \equiv_6 5 \text{ or } i \equiv_6 0, i \not\equiv_{30} 0, \end{cases} \\
 f(o_{i,3}) &= \begin{cases} 7i - 3 & \text{if } i \not\equiv_6 0, 2, 4, 5 \text{ or } i \equiv_{30} 0, \\ 7i - 2 & \text{if } i \not\equiv_6 1, 3, \end{cases} \\
 f(o_{i,4}) &= \begin{cases} 7i - 2 & \text{if } i \equiv_{30} 0, \\ 7i - 1 & \text{if } i \not\equiv_{30} 0, \end{cases} \\
 f(o_{i,5}) &= 7i \quad \text{if } 1 \leq i \leq n.
 \end{aligned}$$

As an example, the prime vertex labeling of $C_3 \star P_2 \star S_5$ using this labeling appears in [Figure 6](#).

For $m = 6$, the labeling function $f : V \rightarrow \{1, 2, \dots, 8n\}$ is given by

$$\begin{aligned}
 f(c_i) &= 8i - 7 \quad \text{if } 1 \leq i \leq n, \\
 f(p_i) &= \begin{cases} 8i - 3 & \text{if } i \not\equiv_3 0, \\ 8i - 5 & \text{if } i \equiv_3 0, i \not\equiv_{15} 0, \\ 8i - 1 & \text{if } i \equiv_{15} 0, \end{cases} \\
 f(o_{i,1}) &= 8i - 6 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,2}) &= \begin{cases} 8i - 5 & \text{if } i \not\equiv_3 0 \text{ or } i \equiv_{15} 0, \\ 8i - 3 & \text{if } i \equiv_3 0, i \not\equiv_{15} 0, \end{cases} \\
 f(o_{i,3}) &= 8i - 4 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,4}) &= \begin{cases} 8i - 3 & \text{if } i \equiv_{15} 0, \\ 8i - 1 & \text{if } i \not\equiv_{15} 0, \end{cases} \\
 f(o_{i,5}) &= 8i - 2 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,6}) &= 8i \quad \text{if } 1 \leq i \leq n.
 \end{aligned}$$

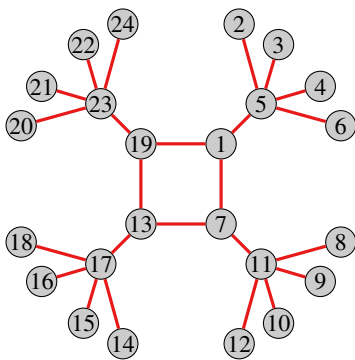


Figure 5. A prime vertex labeling of $C_4 \star P_2 \star S_4$.

As an example, the prime vertex labeling of $C_5 \star P_2 \star S_6$ using this labeling appears in Figure 7. When $m = 6$, the cycle vertex labels will be consecutive integers congruent to 1 modulo 8. So, each pair of adjacent labels are odd integers that differ by 8 and hence are relatively prime. The remaining adjacencies to consider involve three cases. These cases refer to the three possible labelings for the spur vertices and within each case there are seven adjacencies to check.

Case 1. Assume $i \not\equiv_3 0$, so that $f(p_i) = 8i - 3$, and refer to Figure 8 for the corresponding labeling. Then

- (1) $8i - 7$ and $8i - 3$ are odd integers that differ by 4;
- (2) $8i - 3$ and $8i - 6$ are not multiples of 3 and differ by 3;
- (3) $8i - 3$ and $8i - 5$ are consecutive odd integers;
- (4) $8i - 3$ and $8i - 4$ are consecutive integers;
- (5) $8i - 3$ and $8i - 2$ are consecutive integers;
- (6) $8i - 3$ and $8i - 1$ are consecutive odd integers;
- (7) $8i - 3$ and $8i$ are not multiples of 3 and differ by 3.

Case 2. Next, assume that $i \equiv_3 0$ and $i \not\equiv_{15} 0$, so that $f(p_i) = 8i - 5$, and refer to Figure 9 for the corresponding labeling. Then

- (1) $8i - 7$ and $8i - 5$ are consecutive odd integers;
- (2) $8i - 5$ and $8i - 6$ are consecutive integers;
- (3) $8i - 5$ and $8i - 4$ are consecutive integers;
- (4) $8i - 5$ and $8i - 3$ are consecutive odd integers;
- (5) $8i - 5$ and $8i - 2$ are not multiples of 3 and differ by 3;
- (6) $8i - 5$ and $8i - 1$ are odd integers that differ by 4;
- (7) $8i - 5$ and $8i$ are not multiples of 5 and differ by 5.

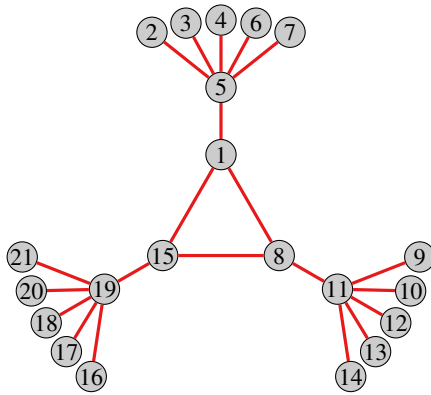


Figure 6. A prime vertex labeling of $C_3 \star P_2 \star S_5$.

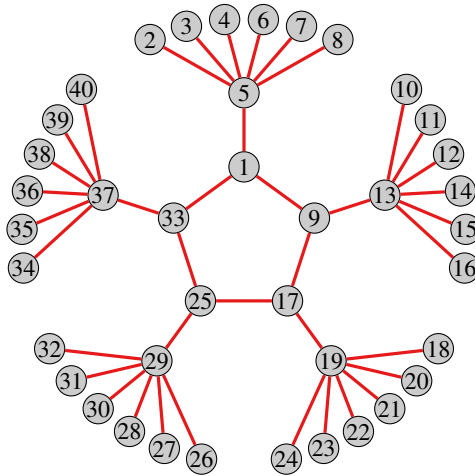


Figure 7. A prime vertex labeling of $C_5 \star P_2 \star S_6$.

Case 3. Lastly, assume $i \equiv_{15} 0$, so that $f(p_i) = 8i - 1$, and refer to [Figure 10](#) for the corresponding labeling. Then

- (1) $8i - 7$ and $8i - 1$ are odd integers that are not multiples of 3 and differ by 6;
- (2) $8i - 1$ and $8i - 6$ are not multiples of 5 and differ by 5;
- (3) $8i - 1$ and $8i - 5$ are odd integers that differ by 4;
- (4) $8i - 1$ and $8i - 4$ are not multiples of 3 and differ by 3;
- (5) $8i - 1$ and $8i - 3$ are consecutive odd integers;
- (6) $8i - 1$ and $8i - 2$ are consecutive integers;
- (7) $8i - 1$ and $8i$ are consecutive integers.

In all cases, adjacent labels are relatively prime, showing that $C_n \star P_2 \star S_6$ is prime.

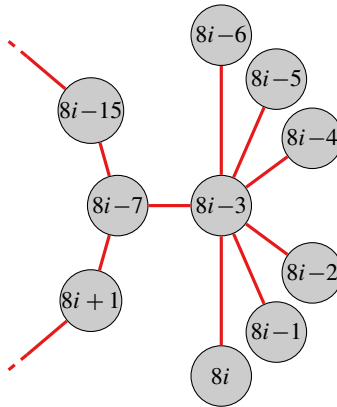


Figure 8. The generalized prime vertex labeling of $C_n \star P_2 \star S_6$ for $i \not\equiv_3 0$ (Case 1).

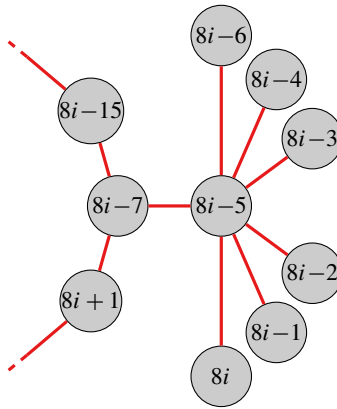


Figure 9. The generalized prime vertex labeling of $C_n \star P_2 \star S_6$ for $i \equiv_3 0$ and $i \not\equiv_{15} 0$ (Case 2).

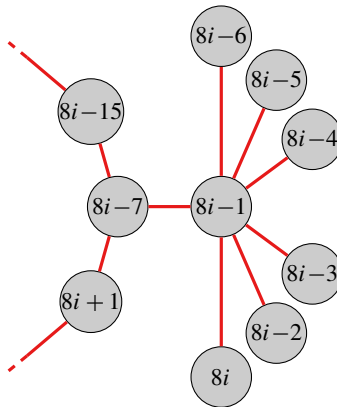


Figure 10. The generalized prime vertex labeling of $C_n \star P_2 \star S_6$ for $i \equiv_{15} 0$ (Case 3).

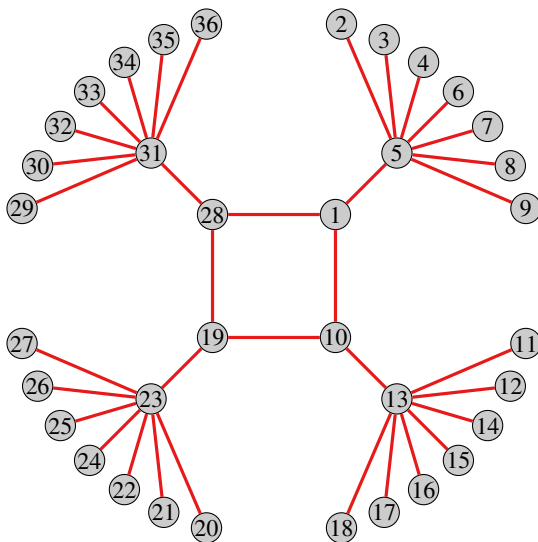


Figure 11. A prime vertex labeling of $C_4 \star P_2 \star S_7$.

Next, for $m = 7$, the labeling function $f : V \rightarrow \{1, 2, \dots, 9n\}$ is given by

$$\begin{aligned}
 f(c_i) &= 9i - 8 \quad \text{if } 1 \leq i \leq n, \\
 f(p_i) &= \begin{cases} 9i - 4 & \text{if } i \equiv_2 1, \\ 9i - 5 & \text{if } i \equiv_2 0, i \not\equiv_{10} 0, \\ 9i - 7 & \text{if } i \equiv_{10} 0, i \not\equiv_{70} 0, \\ 9i - 1 & \text{if } i \equiv_{70} 0, \end{cases} \\
 f(o_{i,1}) &= \begin{cases} 9i - 7 & \text{if } i \not\equiv_{10} 0 \text{ or } i \equiv_{70} 0, \\ 9i - 5 & \text{if } i \equiv_{10} 0, i \not\equiv_{70} 0, \end{cases} \\
 f(o_{i,2}) &= 9i - 6 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,3}) &= \begin{cases} 9i - 5 & \text{if } i \equiv_2 1 \text{ or } i \equiv_{70} 0, \\ 9i - 4 & \text{if } i \equiv_2 0, i \not\equiv_{70} 0, \end{cases} \\
 f(o_{i,4}) &= \begin{cases} 9i - 4 & \text{if } i \equiv_{70} 0, \\ 9i - 1 & \text{if } i \not\equiv_{70} 0, \end{cases} \\
 f(o_{i,5}) &= 9i - 3 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,6}) &= 9i - 2 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,7}) &= 9i \quad \text{if } 1 \leq i \leq n.
 \end{aligned}$$

Figure 11 shows the prime vertex labeling of $C_4 \star P_2 \star S_7$ using this labeling.

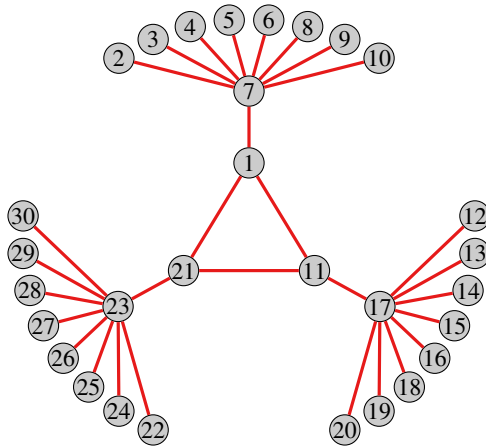


Figure 12. A prime vertex labeling of $C_3 \star P_2 \star S_8$.

Finally, for the $m = 8$ case, the labeling function $f : V \rightarrow \{1, 2, \dots, 10n\}$ is given by

$$\begin{aligned}
 f(c_i) &= 10i - 9 \quad \text{if } 1 \leq i \leq n, \\
 f(p_i) &= \begin{cases} 10i - 3 & \text{if } i \not\equiv_3 0, \\ 10i - 7 & \text{if } i \equiv_3 0, i \not\equiv_{21} 0, \\ 10i - 1 & \text{if } i \equiv_{21} 0, \end{cases} \\
 f(o_{i,1}) &= 10i - 8 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,2}) &= \begin{cases} 10i - 7 & \text{if } i \equiv_{21} 0 \text{ or } i \equiv_3 1, 2, \\ 10i - 3 & \text{if } i \equiv_3 0, i \not\equiv_{21} 0, \end{cases} \\
 f(o_{i,3}) &= 10i - 6 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,4}) &= 10i - 5 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,5}) &= 10i - 4 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,6}) &= \begin{cases} 10i - 3 & \text{if } i \equiv_{21} 0, \\ 10i - 1 & \text{if } i \not\equiv_{21} 0, \end{cases} \\
 f(o_{i,7}) &= 10i - 2 \quad \text{if } 1 \leq i \leq n, \\
 f(o_{i,8}) &= 10i \quad \text{if } 1 \leq i \leq n.
 \end{aligned}$$

The prime vertex labeling of $C_3 \star P_2 \star S_8$ using this labeling appears in [Figure 12](#). \square

Given the relative simplicity of the labelings for the graphs in [Theorem 2.1](#), it might be surprising that determining prime vertex labelings for $C_n \star P_2 \star S_m$ with $m \geq 9$ appears to be more difficult. Consistent with Seoud and Youssef’s conjecture, we expect that all cycle pendant stars are prime.

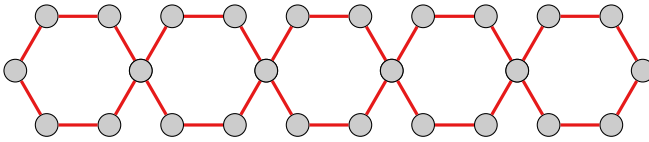


Figure 13. An example of the cycle chain C_6^5 .

3. Cycle chains

The “gluing together” of identical cycles appears in various guises in the literature. But the construction of chains of cycles, with adjacent cycles sharing a single common vertex, is not prevalent. For this reason, we require the following definition, where we assume that n is even. The graph C_n^2 results from attaching two n -cycles together at a single shared vertex. Continuing in this manner, we define C_n^3 by attaching a third n -cycle to one of the n -cycles of C_n^2 in a similar fashion so that the cycle containing two shared vertices consists of two identical $(n/2)$ -paths. Recursively, the graph C_n^m consists of a “chain” of m consecutive n -cycles. We refer to each of the graphs in this family as a *cycle chain*. For example, the cycle chain C_6^5 consisting of five consecutive 6-cycles is shown in Figure 13.

In what follows, we will show that each C_n^m is prime for $n = 4, 6, 8$ and all m . The labeling functions involved are all similar and relatively simple, which is a consequence of the vertex identification employed on each family. We also show that one of these labeling schemes generalizes in an obvious way, providing a prime vertex labeling for a specific family of cycle chains associated with Mersenne primes.

Theorem 3.1. *All C_4^m are prime.*

Proof. The vertices of C_4^m are identified as follows. First, the vertices of C_1 are identified clockwise as $c_{1,1}, c_{1,2}, c_{1,3}, c_{1,4}$, where $c_{1,4}$ is the vertex also belonging to C_2 . The remaining vertex identifications are based on the parity of i .

If i is even, the vertices $c_{i,1}, c_{i,2}, c_{i,3}, c_{i,4}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} . If i is odd and greater than 1, the vertices $c_{i,1}$ and $c_{i,2}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} , and the vertices $c_{i,3}$ and $c_{i,4}$ are identified counterclockwise from the vertex that is counterclockwise adjacent to the common vertex of C_i and C_{i-1} . Note that in both cases, $c_{i,4}$ is the common vertex of C_i and C_{i+1} .

The graph C_4^m can now be prime labeled via

$$f(c_{i,k}) = \begin{cases} k + 1 & \text{if } i = 1, 1 \leq k \leq 4, \\ 3i + k - 1 & \text{if } 2 \leq i \leq m, 1 \leq k \leq 3, \text{ except } i = m \text{ and } k = 3, \\ 1 & \text{if } i = m, k = 3. \end{cases}$$

It is simple to verify that all adjacent vertices receive relatively prime labels. □

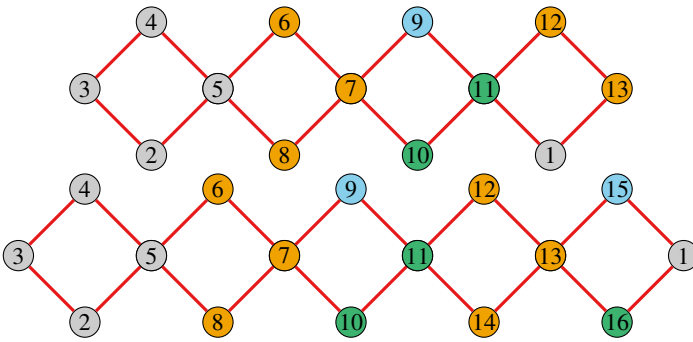


Figure 14. Prime vertex labelings of C_4^4 and C_4^5 .

In Figure 14, we see two labeled examples of C_4^m , one for each of the possible locations of the integer 1, the last vertex label to be assigned.

Theorem 3.2. All C_6^m are prime.

Proof. The vertices of C_6^m are identified as follows. First, the vertices of C_1 are identified counterclockwise as $c_{1,1}, c_{1,2}, \dots, c_{1,6}$, where $c_{1,1}$ is the vertex also belonging to C_2 . The remaining vertex identifications are based on the congruence class modulo 3 to which i belongs.

If $i \equiv 0, 2$, the vertices $c_{i,1}$ and $c_{i,2}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} , and the vertices $c_{i,3}, c_{i,4}, c_{i,5}$ are identified counterclockwise from the vertex that is counterclockwise adjacent to the common vertex of C_i and C_{i-1} . If $i \equiv 1$, the vertices $c_{i,1}, c_{i,2}, \dots, c_{i,5}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} . Note that in both cases, $c_{i,5}$ is the common vertex of C_i and C_{i+1} .

The graph of C_6^m is then labeled using the function given by

$$f(c_{i,k}) = \begin{cases} k & \text{if } i = 1, 1 \leq k \leq 6, \\ 5i + k - 4 & \text{if } i > 1, 1 \leq k \leq 5. \end{cases}$$

Again, it is straightforward to verify that all adjacent vertices receive relatively prime labels, and the result follows. □

A prime vertex labeling of C_6^5 using Theorem 3.2 appears in Figure 15.

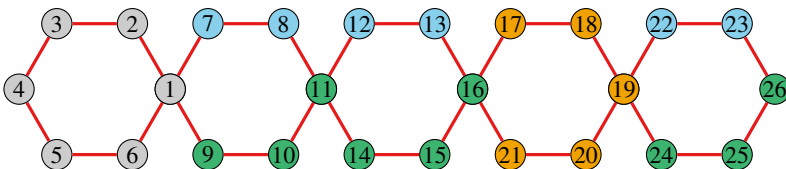


Figure 15. A prime vertex labeling of C_6^5 .

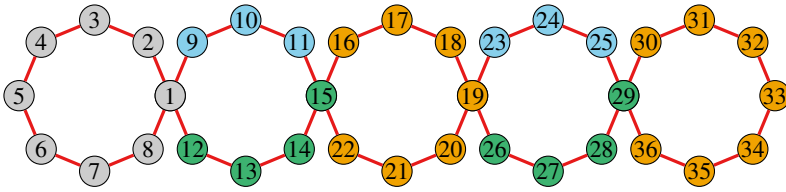


Figure 16. A prime vertex labeling of C_8^5 .

Theorem 3.3. All C_8^m are prime.

Proof. The vertices of C_8^m are identified as follows. The vertices of C_1 are identified counterclockwise as $c_{1,1}, c_{1,2}, \dots, c_{1,8}$, where $c_{1,1}$ is the vertex also belonging to C_2 . The remaining vertex identifications are based on the parity of i .

If i is even, the vertices $c_{i,1}, c_{i,2}, c_{i,3}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} , and the vertices $c_{i,4}, c_{i,5}, c_{i,6}, c_{i,7}$ are identified counterclockwise from the vertex that is counterclockwise adjacent to the common vertex of C_i and C_{i-1} . If i is odd and greater than 1, the vertices $c_{i,1}, c_{i,2}, \dots, c_{i,7}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} . Note that in both cases, $c_{i,4}$ is the common vertex of C_i and C_{i+1} .

The graph of C_8^m is then labeled using the function given by

$$f(c_{i,k}) = \begin{cases} k & \text{if } i = 1, 1 \leq k \leq 8, \\ 7i + k - 6 & \text{if } i > 1, 1 \leq k \leq 7. \end{cases}$$

It is simple to verify that all adjacent vertices receive relatively prime labels. \square

An example of a prime vertex labeling of C_8^5 using the labeling in [Theorem 3.3](#) appears in [Figure 16](#).

For any positive integer k , a Mersenne number is an integer of the form $M_k = 2^k - 1$. If $M_k = 2^k - 1$ is a prime number, then M_k is called a *Mersenne prime*. The first few Mersenne primes are $M_2 = 2^2 - 1 = 3$, $M_3 = 2^3 - 1 = 7$, and $M_5 = 2^5 - 1 = 31$. There are 49 known Mersenne primes.

Theorem 3.4. Let $k \in \mathbb{N}$, $k \geq 3$, and let $n = 2^k$. If $2^k - 1$ is a Mersenne prime, then C_n^m has a prime vertex labeling.

Proof. Both the vertex identification and the labeling function follow from what appeared in the proof of [Theorem 3.3](#). It should be noted that these are different from the machinery used to show that C_4^m are prime. The vertices of C_1 are identified clockwise as $c_{1,1}, c_{1,2}, \dots, c_{1,2^k}$, where $c_{1,1}$ is the vertex also belonging to C_2 . The remaining vertex identifications are based on the parity of i .

If i is even, the vertices $c_{i,1}, c_{i,2}, \dots, c_{i,2^{k-1}-1}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} , and

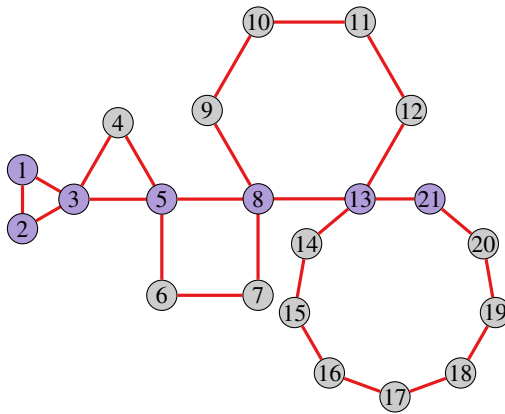


Figure 17. A prime vertex labeling of the Fibonacci chain C_F^5 .

the vertices $c_{i,2^{k-1}}, c_{i,2^{k-1}+1}, \dots, c_{i,2^k-1}$ are identified counterclockwise from the vertex that is counterclockwise adjacent to the common vertex of C_i and C_{i-1} . If i is odd and greater than 1, the vertices $c_{i,1}, c_{i,2}, \dots, c_{i,2^k-1}$ are identified clockwise from the vertex that is clockwise adjacent to the common vertex of C_i and C_{i-1} . Note that $c_{i,2^k-1}$ is the common vertex of C_i and C_{i+1} in both cases.

The graph C_n^m can be prime labeled using the function given by

$$f(c_{i,k}) = \begin{cases} k & \text{if } i = 1, 1 \leq k \leq n, \\ (n-1)i + k - (n-2) & \text{if } i > 1, 1 \leq k \leq n-1. \end{cases}$$

It is again relatively straightforward to verify that all adjacent vertices receive relatively prime labels. □

Another historically significant sequence of integers that can be used to generate prime chains of cycles is the sequence of Fibonacci numbers. However, in this case, we will attach cycles of increasing size determined by the terms of the Fibonacci sequence. Recall that the Fibonacci numbers are defined by the recurrence relation $F_i = F_{i-1} + F_{i-2}$, where $F_1 = 1$ and $F_2 = 1$. The first several Fibonacci numbers are 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144. To construct our graph, begin with a single path consisting of $m + 2$ vertices. Starting from one end of the path, identify the vertices as p_1, p_2, \dots, p_{m+2} . First, add an edge between p_1 and p_3 . Next, for $i \geq 3$, build out a cycle by adding an additional F_i edges between p_i and p_{i+1} . The resulting graph consists of m consecutive cycles, where the first cycle will consist of three vertices and, for $j \geq 2$, the j -th cycle will consist of $F_{j+1} + 1$ vertices. We denote this graph by C_F^m and call it a *Fibonacci cycle chain*.

The well-known fact that consecutive Fibonacci numbers are relatively prime leads directly to the following result.

Theorem 3.5. *All Fibonacci cycle chains C_F^m are prime.*

Proof. To label a Fibonacci cycle chain, begin by consecutively labeling the path initially used to construct the graph with the Fibonacci numbers starting with the second 1 of the sequence. The remaining vertices can be labeled using the complement of the Fibonacci sequence in the obvious way. \square

Figure 17 shows a Fibonacci cycle chain with a prime vertex labeling. Note that the colored vertices are those that are assigned Fibonacci numbers as labels, which form a path within the graph.

4. Prisms

A *prism graph* is a graph of the form $C_n \times P_2$, which consists of an inner and an outer n -cycle connected with spurs. Prajapati and Gajjar [2014] showed that if $n + 1$ is prime, then $C_n \times P_2$ has a prime vertex labeling. Moreover, they also showed that if $n \geq 3$ is odd, then $C_n \times P_2$ does not have a prime vertex labeling. In this section, we will prove that if $n - 1$ is prime, then $C_n \times P_2$ has a prime vertex labeling. The remaining cases involving prism graphs are currently open.

When n is even, the initial strategy for labeling prism graphs is to divide the set $\{1, 2, \dots, 2n\}$ into two subsets, namely $\{1, 2, \dots, n\}$ and $\{n + 1, n + 2, \dots, 2n\}$. We then attempt to label the inner-cycle vertices clockwise using the consecutive integers from 1 to n and label the outer-cycle vertices with consecutive integers from $n + 1$ to $2n$ in the same direction so that the vertex labeled $n + 1$ is adjacent to the vertex labeled 2. In this case, the difference between the labels assigned to adjacent pairs of inner and outer vertices will be $n - 1$ except between 1 and $2n$. Since $n - 1$ is prime, most pairs of labels for inner and outer vertices will be guaranteed to be relatively prime. However, we have a problem with the labels $n - 1$ and $2(n - 1)$ that the labeling in the next theorem will address by swapping the labels 1 and n with $n - 1$ and $2n$, respectively. It should be noted that our technique does not generalize to arbitrary $C_n \times P_2$ (for n even).

Theorem 4.1. *If $n - 1$ is a prime number and $n \geq 4$, then $C_n \times P_2$ is prime.*

Proof. Denote the vertices on the inner cycle by $c_{1,1}, c_{1,2}, \dots, c_{1,n}$ and their corresponding vertices on the outer cycle by $c_{2,1}, c_{2,2}, \dots, c_{2,n}$. The labeling function $f : V \rightarrow \{1, 2, \dots, 2n\}$ is given by

$$f(c_{1,i}) = \begin{cases} i & \text{if } i = 2, 3, \dots, n - 2, \\ n - 1 & \text{if } i = 1, \\ 1 & \text{if } i = n - 1, \\ 2n & \text{if } i = n, \end{cases}$$

and

$$f(c_{2,i}) = \begin{cases} i + n - 1 & \text{if } i = 2, 3, \dots, n, \\ n & \text{if } i = 1. \end{cases}$$

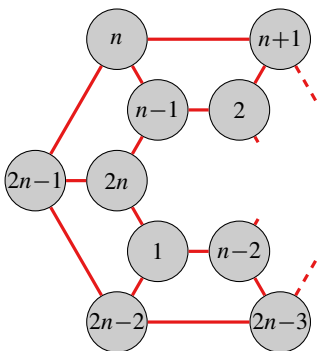


Figure 18. The generalized labeling for the prism $C_n \times P_2$ when $n - 1$ is prime.

Recall that consecutive integers are relatively prime, which forces most pairs of adjacent inner-cycle vertices and most pairs of adjacent outer-cycle vertices to receive relatively prime labels. To verify the relative primeness of the labels on the remaining pairs of adjacent vertices, first note that, for $i = 2, 3, \dots, n - 2$, we have

$$(f(c_{1,i}), f(c_{2,i})) = (i, i + n - 1) = 1.$$

Next, for $i = 1$, we see that

$$(f(c_{1,1}), f(c_{2,1})) = (n - 1, n) = 1.$$

Finally, observe that

$$(f(c_{1,1}), f(c_{1,2})) = (n - 1, 2) = 1,$$

$$(f(c_{1,1}), f(c_{1,n})) = (n - 1, 2n) = 1,$$

$$(f(c_{1,1}), f(c_{1,2})) = (n - 1, 2) = 1,$$

$$(f(c_{1,n-2}), f(c_{2,n-2})) = (n - 2, 2n - 3) = 1.$$

In **Figure 18**, we see a portion of the labeling of $C_n \times P_2$ ($n - 1$ prime). Note that vertices absent from the figure are labeled clockwise using consecutive integers. Thus, $C_n \times P_2$ has a prime vertex labeling when $n - 1$ is a prime number. \square

As an example, **Figure 19** depicts the labeling of **Theorem 4.1** for $C_6 \times P_2$.

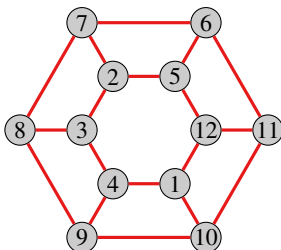


Figure 19. A prime vertex labeling of $C_6 \times P_2$.

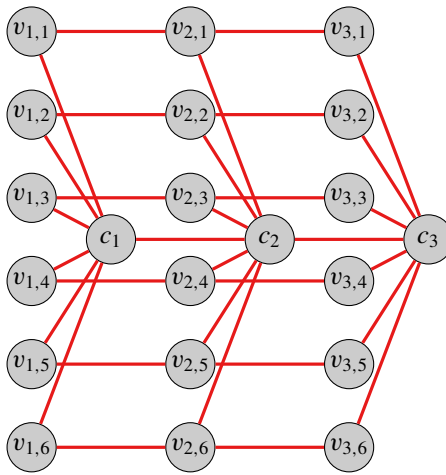


Figure 20. The generalized book $S_6 \times P_3$.

5. Generalized books

A *book* is a graph of the form $S_n \times P_2$, where S_n is the star with n spur vertices and P_2 is the path with two vertices. Using a simple parity argument, Seoud and Youssef [1999] showed that all books have a prime vertex labeling. In this section, we extend their work by providing a prime vertex labeling for some *generalized books*, which are graphs of the form $S_n \times P_m$. Observe that $S_n \times P_m$ looks like $m - 1$ books glued together. The generalized book $S_6 \times P_3$ is shown in Figure 20.

Theorem 5.1. All $S_n \times P_m$ with $3 \leq m \leq 7$ are prime.

Proof. We will handle each of the cases involving $3 \leq m \leq 7$ separately. For each $3 \leq m \leq 7$, let c_1, c_3, \dots, c_m denote the vertices on the path through the center of each star S_n . Next, let $v_{i,1}, v_{i,2}, \dots, v_{i,n}$ denote the vertices of degree 1 on the i -th star so that $v_{i,k}$ in the i -th star is adjacent to $v_{i+1,k}$ in the $(i+1)$ -th star. For an example, see the identification of vertices depicted in Figure 20.

First, we consider the generalized book $S_n \times P_3$. We define our labeling function $f : V \rightarrow \{1, 2, \dots, 3n + 3\}$ as follows. Let $f(c_j) = j$. For k odd, define $f(v_{i,k}) = 3k - i + 4$. For k even, let $f(v_{1,k}) = 3k + 2$, $f(v_{2,k}) = 3k + 3$, and $f(v_{3,k}) = 3k + 1$. Using this labeling, one can quickly see that $S_n \times P_3$ has a prime vertex labeling. For an example, see the labeling of $S_6 \times P_3$ given in Figure 21.

For $S_n \times P_4$, we define our labeling function $f : V \rightarrow \{1, 2, \dots, 4n + 4\}$ as follows. Let $f(c_j) = j$. For $k \equiv 1 \pmod 3$, let $f(v_{1,k}) = 4k + 2$, $f(v_{2,k}) = 4k + 3$, $f(v_{3,k}) = 4k + 4$, and $f(v_{4,k}) = 4k + 1$. For $k \not\equiv 1 \pmod 3$, define $f(v_{i,k}) = 4k - i + 5$. It then follows that each $S_n \times P_4$ has a prime vertex labeling. An example of this labeling for $S_6 \times P_4$ appears in Figure 22.

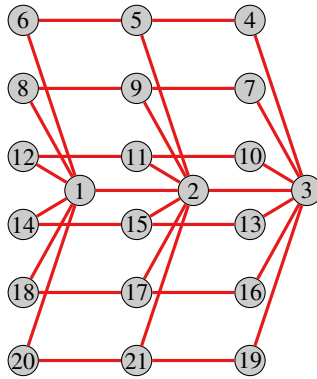


Figure 21. A prime vertex labeling of $S_6 \times P_3$.

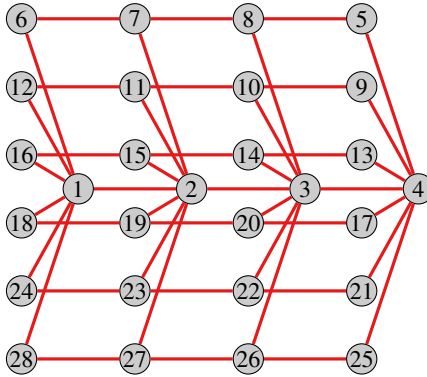


Figure 22. A prime vertex labeling of $S_6 \times P_4$.

Next, for $S_n \times P_5$, we define our labeling function $f : V \rightarrow \{1, 2, \dots, 5n + 5\}$ in the following manner. Let $f(c_j) = j$. The rest of the labeling function is given by

$$f(v_{i,k}) = \begin{cases} 11 - i & \text{if } k = 1, \\ 5k + 1 + i & \text{if } 1 \leq i \leq 4, k > 1, k - 1 \equiv_6 1, 5, \\ 5k + 1 & \text{if } i = 5, k > 1, k - 1 \equiv_6 1, 5, \\ 5k + 2 + i & \text{if } 1 \leq i \leq 3, k > 1, k - 1 \equiv_6 0, 2, \\ 5k + 6 - i & \text{if } i = 4, 5, k > 1, k - 1 \equiv_6 0, 2, \\ 5k + 5 - i & \text{if } 1 \leq i \leq 3, k > 1, k - 1 \equiv_6 3, \\ 5k + 5 & \text{if } i = 4, k > 1, k - 1 \equiv_6 3, \\ 5k + 1 & \text{if } i = 5, k > 1, k - 1 \equiv_6 3, \\ 5k + 6 - i & \text{if } k > 1, k - 1 \equiv_6 4. \end{cases}$$

It then follows that each $S_n \times P_5$ has a prime vertex labeling. For an example, see the labeling of $S_7 \times P_5$ given in [Figure 23](#).

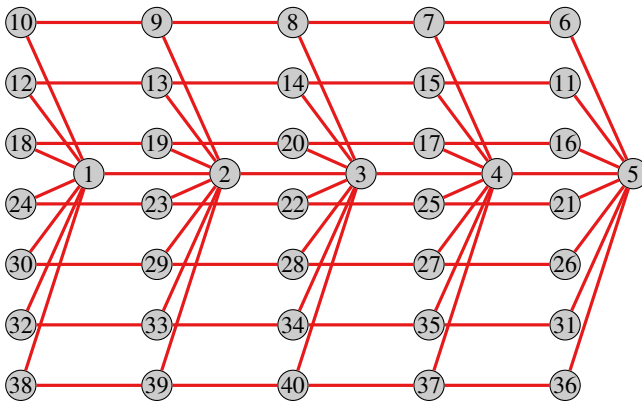


Figure 23. A prime vertex labeling of $S_7 \times P_5$.

For $S_n \times P_6$, we define our labeling function $f : V \rightarrow \{1, 2, \dots, 6n + 6\}$ as follows. Let $f(c_j) = j$ and let

$$f(v_{i,k}) = \begin{cases} 6(k+1) + 1 - i & \text{if } k = 1, 2, \\ 6k + 1 + i & \text{if } 1 \leq i \leq 5, k = 3, \\ 6k + 1 & \text{if } i = 6, k = 3, \\ 6k + 1 - i & \text{if } k > 3, k - 3 \equiv_5 1, 2, 3, 4, \\ 6k + 1 + i & \text{if } 1 \leq i \leq 5, k > 3, k - 3 \equiv_5 0, \\ 6k + 1 & \text{if } i = 6, k > 3, k - 3 \equiv_5 0. \end{cases}$$

Using this labeling, we see that each $S_n \times P_6$ has a prime vertex labeling. An example of this labeling for $S_8 \times P_6$ appears in [Figure 24](#).

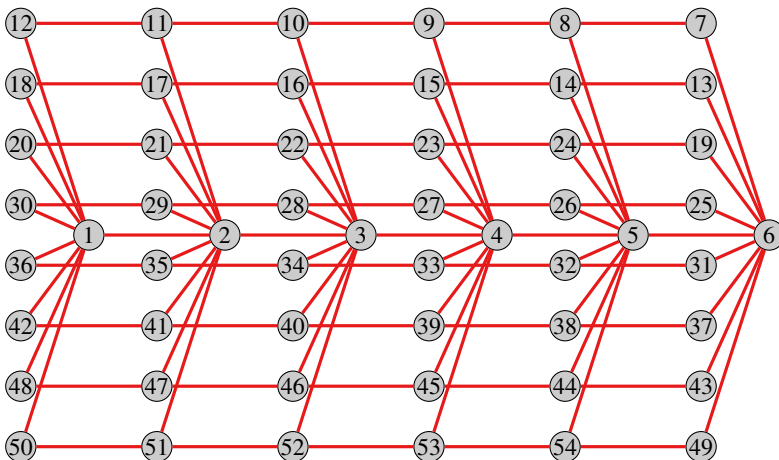


Figure 24. A prime vertex labeling of $S_8 \times P_6$.

Lastly, for $S_n \times P_7$, we define our labeling function $f : V \rightarrow \{1, 2, \dots, 7n+7\}$ as follows. Let $f(c_j) = j$. The remaining portion of the labeling function will involve 10 ordered “row” (i.e., set of corresponding positions in each star) permutation patterns, which we will denote by $A, B, C, D, E, F, G, H, I$, and J . Each ordered row permutation pattern takes the seven consecutive integers that will be used to label the vertices in the k -th row ($7k + 1, \dots, 7k + 7$) and assigns them the labels $w_1 = 7k + 1, w_2 = 7k + 2, \dots, w_7 = 7k + 7$. For instance, if permutation A is applied to the k -th row, then the first vertex in the k -th row is given the label $w_2 = 7k + 2$, the second vertex in the k -th row is given the label $w_3 = 7k + 3$, etc. Here are the 10 row permutations written in 1-line notation:

$$\begin{aligned}
 A &= [w_2, w_3, w_4, w_5, w_6, w_7, w_1], \\
 B &= [w_2, w_3, w_6, w_7, w_4, w_5, w_1], \\
 C &= [w_3, w_2, w_7, w_6, w_5, w_4, w_1], \\
 D &= [w_5, w_6, w_7, w_2, w_3, w_4, w_1], \\
 E &= [w_4, w_5, w_6, w_7, w_2, w_3, w_1], \\
 F &= [w_3, w_4, w_5, w_6, w_7, w_2, w_1], \\
 G &= [w_6, w_7, w_2, w_3, w_4, w_5, w_1], \\
 H &= [w_2, w_1, w_3, w_7, w_6, w_5, w_4], \\
 I &= [w_7, w_6, w_1, w_2, w_3, w_4, w_5], \\
 J &= [w_7, w_6, w_5, w_4, w_3, w_2, w_1].
 \end{aligned}$$

After assigning $f(c_j) = j$ to the center of each star, the labeling for $S_n \times P_7$ has the following 30-row repeating pattern:

$$C, E, J, A, J, A, D, E, J, A, F, G, C, H, J, A, J, A, I, E, F, E, J, A, D, E, J, A, J, A.$$

It is straightforward to check that this “block” of 30 rows provides a prime vertex labeling for $S_{30} \times P_7$. To label the next block of 30 rows in the generalized book, simply add 210 to each of the labels from the first block of 30 rows. Since any of the vertices in this second block can only be adjacent to the center of its star and any neighbors in its row, and since congruence classes modulo 2, 3, 4, 5, 6, and 7 are invariant under addition of 210, all adjacent vertices in this second block must still have relatively prime labels. Thus, this second block has a prime vertex labeling. It then follows that the entire generalized book has a prime vertex labeling. \square

We conjecture that all generalized books $S_n \times P_m$ have a prime vertex labeling; however, it appears that extending our results using the current approach becomes increasingly difficult as one increases the size of the path P_m .

6. Conclusion

The prime vertex labeling functions included in this paper for cycle pendant stars may be extended to somewhat larger sizes of stars. Specifically, we conjecture that similar processes will work for cycle pendant stars up to stars of size 15. The reasoning for this restriction on stars of size 15 is the following result of Pillai [1941].

Proposition 6.1. *When $k \geq 17$, we can find k consecutive integers such that no integer in the set is relatively prime to all other integers in the set.*

This implies that a new labeling scheme must be devised for finding prime vertex labelings of $C_n \star P_2 \star S_m$ for $m \geq 15$.

In Section 3, we constructed prime vertex labelings for cycle chains having cycles of sizes 4, 6, and 8, respectively. We conjecture that all cycle chains are prime. Using the fact that consecutive Fibonacci numbers are relatively prime, we constructed a prime graph. One can likely construct similar graphs using other well-known sequences that exhibit traits of relative primeness.

We conjecture that all prisms constructed from even-order cycles have prime vertex labelings. We also have preliminary results implying that some families of generalized prisms of the form $C_n \times P_m$ for $m > 2$ have prime vertex labelings.

In addition, we conjecture that every generalized book has a prime vertex labeling. But our pairwise matching approach would need to be replaced in order for this conjecture to be realized. For the interested reader, additional information can be found in Gallian's dynamic survey on graph labelings [1998].

References

- [Diefenderfer et al. 2015] N. Diefenderfer, M. G. Hastings, L. N. Heath, H. Prawzinsky, B. Preston, E. White, and A. Whittemore, "Prime vertex labelings of families of unicyclic graphs", *Rose-Hulman Undergrad. Math J.* **16**:1 (2015), 253–269. [MR 3367500](#)
- [Gallian 1998] J. A. Gallian, "A dynamic survey of graph labeling", *Electron. J. Combin.* Dynamic Surveys:6 (1998). 18th ed., December 7, 2015. [MR 1668059](#) [Zbl 0953.05067](#)
- [Pillai 1941] S. S. Pillai, "On m consecutive integers—III", *Proc. Indian Acad. Sci. (A)* **13**:6 (1941), 530–533. [MR 0004824](#) [Zbl 0063.06228](#)
- [Prajapati and Gajjar 2014] U. M. Prajapati and S. J. Gajjar, "Some results on prime labeling", *Open J. Discret. Math.* **4**:3 (2014), 60–66.
- [Seoud and Youssef 1999] M. A. Seoud and M. Z. Youssef, "On prime labeling of graphs", *Congr. Numer.* **141** (1999), 203–215. [MR 1745237](#) [Zbl 0968.05068](#)

Received: 2015-03-21

Revised: 2015-08-12

Accepted: 2015-08-17

ned32@nau.edu

Northern Arizona University, Flagstaff, AZ 86011,
United States

dana.ernst@nau.edu

Northern Arizona University, Flagstaff, AZ 86001,
United States

mgh64@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
lnh57@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
hpp3@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
bep38@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
jeffrey.rushall@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
ekw49@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>
anw99@nau.edu	<i>Northern Arizona University, Flagstaff, AZ 86011, United States</i>

Presentations of Roger and Yang's Kauffman bracket arc algebra

Martin Bobb, Dylan Peifer, Stephen Kennedy and Helen Wong

(Communicated by Colin Adams)

The Jones polynomial for knots and links was a breakthrough discovery in the early 1980s. Since then, it's been generalized in many ways; in particular, by considering knots and links which live in thickened surfaces and by allowing arcs between punctures or marked points on the boundary of the surface. One such generalization was recently introduced by Roger and Yang and has connections with hyperbolic geometry. We provide generators and relations for Roger and Yang's Kauffman bracket arc algebra of the torus with one puncture and the sphere with three or fewer punctures.

Roger and Yang's Kauffman bracket arc algebra is a generalization of the well-known Kauffman bracket skein algebra of a surface, whose definition in [Turaev 1988; Przytycki 1991] is based on Kauffman's skein theoretic description of the Jones polynomial for knots and links [Jones 1985; Kauffman 1987]. Later, the skein algebra of a hyperbolic surface was interpreted as a quantization of the surface's Teichmüller space from hyperbolic geometry [Turaev 1991; Bullock et al. 1999; Przytycki and Sikora 2000]. Interest thus grew for the skein algebra, as a construction important in the Jones polynomial skein theory but also deeply related to Teichmüller theory.

Following this body of work on the skein algebra, Roger and Yang introduced a "skein algebra of arcs" to be a skein theory version of Penner's decorated Teichmüller space. Penner [1987] defined the decorated Teichmüller space as an alternate way to describe the hyperbolic structures of a surface using lengths of both simple closed curves and arcs between punctures on the surface (each decorated with a choice of horoball). Roger and Yang defined their arc algebra as a quantization of Penner's decorated Teichmüller space, roughly in the same way that the skein algebra is a quantization of the usual Teichmüller space. The arc algebra includes both simple closed curves and arcs between punctures on the surface. In addition to the two

MSC2010: primary 57M27; secondary 57M50.

Keywords: Kauffman bracket skein algebra, Kauffman bracket arc algebra.

Wong was supported in part by NSF Grant DMS-1105692 .

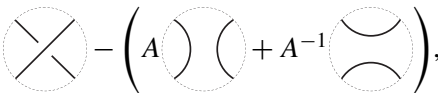
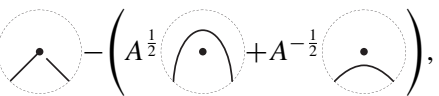
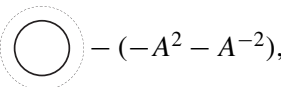
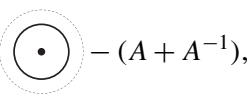
usual bracket skein relations for framed links, there are two extra relations for arcs and loops near the punctures.

Understanding the algebraic structure of Roger and Yang’s arc algebra is an important first step to exploring its role as an intermediary between quantum topology and hyperbolic geometry. Here, we seek finite presentations of the arc algebra for some simple surfaces, namely for spheres with three or fewer punctures and for tori with one or no punctures. In the companion paper [Bobb et al. 2016], we show that the arc algebra is finitely generated. Our work is inspired by analogous statements for the skein algebra found in [Bullock and Przytycki 2000; Bullock 1999].

1. The Kauffman bracket arc algebra

Let $F_{g,n}$ denote a compact, orientable surface of genus g with n points p_1, p_2, \dots, p_n removed. The points removed are the *punctures*. Let A be an indeterminate, with formal square roots $A^{\frac{1}{2}}$ and $A^{-\frac{1}{2}}$. In addition, let there be an indeterminate v_i associated to each puncture p_i , and let $R_n = \mathbb{Z}[A^{\pm\frac{1}{2}}][v_1^{\pm 1}, v_2^{\pm 1}, \dots, v_n^{\pm 1}]$ denote the ring of Laurent polynomials in the commuting variables $A^{\frac{1}{2}}$ and v_1, \dots, v_n .

A framed curve in the thickened surface $F_{g,n} \times [0, 1]$ is the union of framed knots and framed arcs that go from puncture to puncture. (See [Roger and Yang 2014] for a precise definition.) Let $\mathcal{G}(F_{g,n})$ be the R_n -module freely generated by the framed curves in $F_{g,n} \times [0, 1]$, up to isotopy, and let $\mathcal{K}(F_{g,n})$ be the submodule generated by terms of the following four forms:

- (1) skein relation: 
- (2) puncture-skein relation on i -th puncture: v_i 
- (3) framing relation: 
- (4) puncture-framing relation: 

where the diagrams in each form are assumed to be identical outside of the small balls depicted. Let $\mathcal{A}(F_{g,n})$ denote the quotient $\mathcal{G}(F_{g,n})/\mathcal{K}(F_{g,n})$.

There is a natural stacking operation for framed curves in the thickened surface $F_{g,n} \times [0, 1]$ which extends to $\mathcal{A}(F_{g,n})$. That is, if $[L_1], [L_2] \in \mathcal{A}(F_{g,n})$ are respectively represented by framed curves L_1, L_2 in $F_{g,n} \times [0, 1]$, the product

$$[L_1] * [L_2] = [L'_1 \cup L'_2] \in \mathcal{A}(F_{g,n})$$

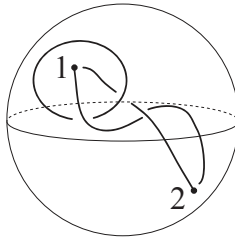


Figure 1. A framed curve with three components on $F_{0,2}$.

is represented by the union of the framed curve $L'_1 \subset F_{g,n} \times [0, \frac{1}{2}]$ (obtained by rescaling L_1 in $F_{g,n} \times [0, 1]$) and the framed curve $L'_2 \subset F_{g,n} \times [\frac{1}{2}, 1]$ (obtained by rescaling L_2 in $F_{g,n} \times [0, 1]$). This stacking operation makes $\mathcal{A}(F_{g,n})$ into an algebra, called the *Kauffman bracket arc algebra* of the surface $F_{g,n}$.

Diagrams in this paper represent projections of framed curves onto $F_{g,n}$ with over- and under-crossing information depicted by breaks in the projection at double-points in the projection, and where the framing of curves is vertical, at right angles to the plane of the paper. Although more than one component of a framed curve can end at any puncture, they must do so at different heights. Diagrams will indicate the order in height of the crossings as necessary. **Figure 1** shows a framed curve consisting of three components (two framed arcs and a framed knot) in a sphere with two punctures. No further labeling at punctures is necessary in **Figure 1** since arcs intersect each puncture only twice.

Figure 2 shows a product of two framed curves on a twice-punctured torus. The product can be simplified by using a Reidemeister 2 move followed by relation (2) (puncture-skein relation) to “pull off” a pair of strands that meet at a puncture and

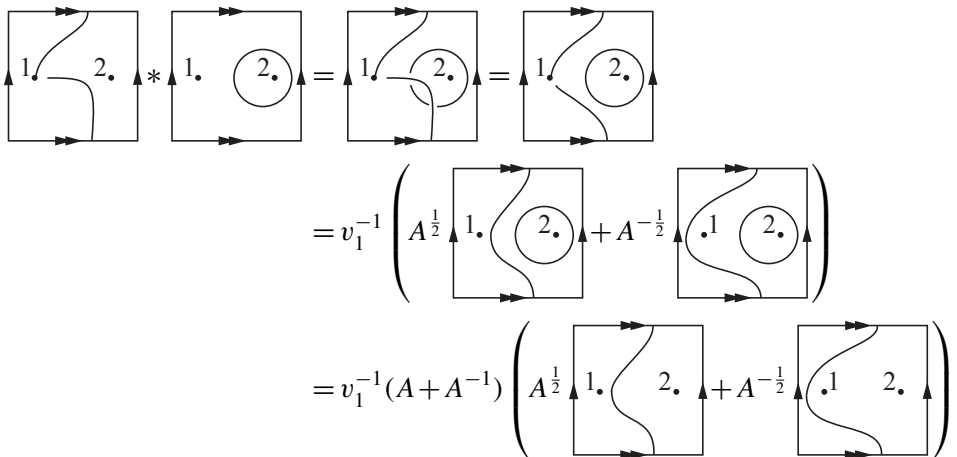


Figure 2. Rewriting a framed curve in $\mathcal{A}(F_{1,2})$.

relation (4) (puncture-framing relation) to “remove” trivial components enclosing a puncture.

The Kauffman bracket skein algebra $\mathcal{S}(F_{g,n})$ defined by Turaev [1988] and Przytycki [1991] is closely related to the arc algebra $\mathcal{A}(F_{g,n})$. Recall that the skein algebra $\mathcal{S}(F_{g,n})$ can be constructed by considering the quotient $\mathcal{G}_0(F_{g,n})/\mathcal{K}_0(F_{g,n})$, where $\mathcal{G}_0(F_{g,n})$ is the $\mathbb{Z}[A, A^{-1}]$ -module generated by the framed links in the thickened surface $F_{g,n} \times [0, 1]$ and $\mathcal{K}_0(F_{g,n})$ is the submodule generated by only relation (1) (the skein relation) and relation (3) (the framing relation) from above. Again, multiplication is induced by the stacking of framed links in the thickened surface. Compared with the skein algebra $\mathcal{S}(F_{g,n})$, the definition of the arc algebra $\mathcal{A}(F_{g,n})$ differs in two ways: in the choice of a larger ring and in the inclusion of two extra relations.

Lemma 1.1. *There exists a well-defined nontrivial algebra homomorphism*

$$\psi : \mathcal{S}(F_{g,n}) \rightarrow \mathcal{A}(F_{g,n})$$

so that $\psi([K]) = [K]$ for a framed link K in $F_{g,n}$.

Proof. Consider the map $i : \mathcal{G}_0(F_{g,n}) \rightarrow R_n \otimes \mathcal{G}_0(F_{g,n})$ with $i(x) = 1 \otimes x$ that changes the scalars, and the map $j : R_n \otimes \mathcal{G}_0(F_{g,n}) \rightarrow \mathcal{G}(F_{g,n})$ with $j(p \otimes x) = p \cdot x$ that includes the framed links into the framed curves. Let $\hat{\psi} = j \circ i$. Notice that $\hat{\psi}(\mathcal{K}_0(F_{g,n})) \subseteq \mathcal{K}(F_{g,n})$. Thus $\hat{\psi} : \mathcal{G}_0(F_{g,n}) \rightarrow \mathcal{G}(F_{g,n})$ descends to a map $\psi : \mathcal{G}_0(F_{g,n})/\mathcal{K}_0(F_{g,n}) \rightarrow \mathcal{G}(F_{g,n})/\mathcal{K}(F_{g,n})$. □

We are interested in the image of ψ . In certain small cases, it generates $\mathcal{A}(F_{g,n})$, a fact we will exploit later on page 697.

2. Generators and relations for the arc algebra

A general strategy for finding generating sets for $\mathcal{A}(F_{g,n})$ is to rewrite framed curves using ones with fewer crossings. We say that a framed knot is a *simple knot* if it allows a projection onto $F_{g,n} \times \{0\}$ without any crossings and it does not bound a disk containing one or no punctures. A framed arc is a *simple arc* if its endpoints are at distinct punctures and it allows a projection without any crossings. A *simple curve* is either a simple knot or a simple arc.

Lemma 2.1. *If a set of elements $\{x_1, x_2, \dots, x_m\}$ generates the simple curves then it generates all of $\mathcal{A}(F_{g,n})$.*

Proof. Suppose we have a basis element $[L] \in \mathcal{A}(F_{g,n})$ represented by a framed curve $L \subset F_{g,n} \times [0, 1]$. By application of the skein relation and the puncture-skein relation, $[L]$ may be written as a linear combination of skeins represented by framed curves each of which has no crossings and intersects a puncture at most once. In particular, the connected components of each framed curve can be isotoped to be at

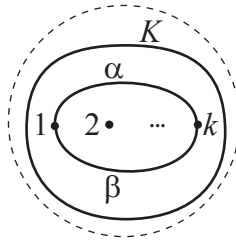


Figure 3. A neighborhood of K in $F_{0,n}$.

different heights, so $[L]$ is a linear combination of products of simple knots, simple arcs, and possibly some loops that bound disks containing one or no punctures. Those latter loops may be removed by application of the framing and puncture relations. Thus $[L]$ is a linear combination of simple curves. Since $\{x_1, x_2, \dots, x_m\}$ generate the simple curves, $[L]$ is also in the set generated by $\{x_1, x_2, \dots, x_m\}$. \square

Remark. Observe that if $(A^2 - 1)$ is invertible, then the puncture-skein relation implies that

$$\text{Diagram} = v_i A^{\frac{1}{2}} (A^2 - 1)^{-1} \left(- \text{Diagram} + A \text{Diagram} \right)$$

and

$$\text{Diagram} = v_i A^{-\frac{1}{2}} (A^{-2} - 1)^{-1} \left(- \text{Diagram} + A^{-1} \text{Diagram} \right).$$

So when $A^2 - 1$ is invertible, if a set of elements generates only the simple arcs, then it generates all of $\mathcal{A}(F_{g,n})$ by [Lemma 2.1](#). However, in the following examples, we will work under the most general set-up, and we will *not* assume that $A^2 - 1$ is invertible.

Arc algebra of punctured spheres. We begin by a refinement of [Lemma 2.1](#) in the case of punctured spheres, $F_{0,n}$ with $n \geq 2$.

Proposition 2.2. *If a set of elements $\{x_1, x_2, \dots, x_n\}$ of the arc algebra $\mathcal{A}(F_{0,n})$ generates the simple arcs, then it generates the entire algebra.*

Proof. By [Lemma 2.1](#), it suffices to show that any simple knot can be rewritten in terms of simple arcs. Given a simple knot $K \subseteq F_{0,n} \times [0, 1]$, notice that it has a projection which separates $F_{0,n}$ into two punctured disks. Let D be the punctured disk bounded by K with the smaller number of punctures, say p_1, \dots, p_k .

Since K is a simple knot, $k \geq 2$. There exist two disjoint simple arcs from p_1 to p_k such that the union of their projections onto $F_{0,n}$ encloses the remaining punctures p_2, \dots, p_{k-1} . Let α and β be the skeins represented by these two arcs, respectively. See [Figure 3](#).

Consider the product $\alpha * \beta \in \mathcal{A}(F_{0,n})$ and apply the puncture-skein relation twice:

$$\alpha * \beta = \text{Diagram} = \frac{1}{v_1 v_k} \left(A \cdot \text{Diagram}_1 + \text{Diagram}_2 + \text{Diagram}_3 + A^{-1} \cdot \text{Diagram}_4 \right).$$

The diagram shows the product of two arcs α and β on a disk with punctures $1, 2, \dots, k$. The result is a linear combination of four terms, each enclosed in a dashed circle. The first term is A times a disk with two arcs and punctures $1, 2, \dots, k$. The second term is a disk with one arc and punctures $1, 2, \dots, k$. The third term is a disk with a knot K and punctures $1, 2, \dots, k$. The fourth term is A^{-1} times a disk with two arcs and punctures $1, 2, \dots, k$.

Thus $[K]$ can be rewritten as a linear combination involving the product of two simple arcs (α and β) and three knots bounding disks with strictly fewer punctures. Notice also that a knot bounding a disk with one or no punctures can be removed using the puncture relation or the framing relation, respectively. Thus by induction, we are done. \square

Sphere with two punctures.

Theorem 2.3. $\mathcal{A}(F_{0,2}) = R_2\langle \alpha \mid \alpha^2 = -v_1^{-1}v_2^{-1}(A - A^{-1})^2 \rangle$, where α is represented by a simple arc between the two punctures of $F_{0,2}$.

Proof. On $F_{0,2}$, any simple arc must start at one puncture and end at the other without intersecting itself. Up to isotopy, there is only one such arc, and let α be the skein represented by that arc. By Proposition 2.2, α generates the algebra. Note that in the arc algebra,

$$\begin{aligned} \alpha^2 &= \text{Diagram} = \frac{1}{v_1 v_2} \left(A \cdot \text{Diagram}_1 + 2 \cdot \text{Diagram}_2 + A^{-1} \cdot \text{Diagram}_3 \right) \\ &= v_1^{-1} v_2^{-1} \left((A + A^{-1})(A + A^{-1}) + 2(-A^2 - A^{-2}) \right) \\ &= -v_1^{-1} v_2^{-1} (A - A^{-1})^2. \end{aligned}$$

The diagram shows α^2 as a disk with two arcs and two punctures. The result is a linear combination of three terms, each enclosed in a dashed circle. The first term is A times a disk with two arcs and two punctures. The second term is 2 times a disk with two arcs and two punctures. The third term is A^{-1} times a disk with two arcs and two punctures.

In particular, this shows that α^2 is not linearly independent from 1 and α . As α is the only generator, this is the only relation of $\mathcal{A}(F_{0,2})$. \square

Sphere with three punctures. To determine the arc algebra $\mathcal{A}(F_{0,3})$, we will need the following lemma from algebra.

Lemma 2.4. Let A and B be R -algebras. Suppose x_1, x_2, \dots, x_n are elements of the algebra A and ρ is some algebra homomorphism of A to the algebra B . If the elements $\rho(x_1), \rho(x_2), \dots, \rho(x_n)$ are linearly independent in B , then x_1, x_2, \dots, x_n are linearly independent in A .

Proof. We prove the contrapositive. Suppose x_1, x_2, \dots, x_n are linearly dependent in A . Then there exist coefficients $k_1, k_2, \dots, k_n \in R$ so that

$$k_1x_1 + k_2x_2 + \dots + k_nx_n = 0$$

and at least one k_i is nonzero. Since ρ is an R -algebra homomorphism,

$$k_1\rho(x_1) + k_2\rho(x_2) + \dots + k_n\rho(x_n) = 0.$$

So $\rho(x_1), \rho(x_2), \dots, \rho(x_n)$ are linearly dependent in B as well. □

Theorem 2.5.

$$A(F_{0,3}) = R_3\langle \alpha_1, \alpha_2, \alpha_3 \mid \alpha_i\alpha_{i+1} = \alpha_{i+1}\alpha_i = v_{i+2}^{-1}\delta\alpha_{i+2}, v_{i+1}v_{i+2}\alpha_i^2 = \delta^2 \rangle,$$

where α_i is represented by the simple arc connecting the punctures p_{i+1} to p_{i+2} in the thrice-punctured sphere $F_{0,3}$, with $i = 1, 2, 3$ and indices interpreted modulo 3, and where $\delta = (A^{\frac{1}{2}} + A^{-\frac{1}{2}})$.

Proof. Since the only simple arcs in $F_{0,3}$ are those connecting distinct punctures, it follows that α_1, α_2 , and α_3 generate the arc algebra $K[F_{0,3}]$. Observe that

$$\begin{aligned} \alpha_i^2 &= \text{Diagram: a dashed circle with punctures } i, i+1, i+2. \text{ Arcs connect } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2. \\ &= v_{i+1}^{-1}v_{i+2}^{-1} \left(A \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i+2 \right) + \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i+1 \\ &\quad + \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i \right) + A^{-1} \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i+2 \\ &= v_{i+1}^{-1}v_{i+2}^{-1} (A(A + A^{-1}) + (-A^2 - A^{-2}) + (A + A^{-1}) + A^{-1}(A + A^{-1})) \\ &= v_{i+1}^{-1}v_{i+2}^{-1} (A^{\frac{1}{2}} + A^{-\frac{1}{2}})^2 \\ &= v_{i+1}^{-1}v_{i+2}^{-1} \delta^2. \end{aligned}$$

Also,

$$\begin{aligned} \alpha_i * \alpha_{i+1} &= \text{Diagram: a dashed circle with punctures } i, i+1, i+2. \text{ Arcs connect } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2. \\ &= v_{i+2}^{-1} \left(A^{\frac{1}{2}} \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i+1 \right) + A^{-\frac{1}{2}} \text{Diagram: } i \text{ to } i+1 \text{ and } i+1 \text{ to } i+2 \text{ with a loop at } i+2 \\ &= v_{i+2}^{-1} (A^{\frac{1}{2}} + A^{-\frac{1}{2}}) \alpha_{i+2} \\ &= v_{i+2}^{-1} \delta \alpha_{i+2} \end{aligned}$$

and similarly

$$\begin{aligned}
 \alpha_{i+1} * \alpha_i &= \text{Diagram} = v_{i+2}^{-1} \left(A^{\frac{1}{2}} \text{Diagram} + A^{-\frac{1}{2}} \text{Diagram} \right) \\
 &= v_{i+2}^{-1} (A^{\frac{1}{2}} + A^{-\frac{1}{2}}) \alpha_{i+2} \\
 &= v_{i+2}^{-1} \delta \alpha_{i+2}.
 \end{aligned}$$

We next show that these are the only relations. Notice that the relations above imply that any product $\alpha_i * \alpha_j$ can be rewritten as either a scalar multiple when $i = j$ or as a multiple of the remaining α_k for $k \neq i, j$. Thus any word in $\alpha_1, \alpha_2,$ and α_3 can be rewritten in terms of a scalar multiple of one or zero generators. So any other relation among the generators $\alpha_1, \alpha_2,$ and α_3 can be expressed in the form

$$k_0 + k_1 \alpha_1 + k_2 \alpha_2 + k_3 \alpha_3 = 0,$$

where $k_i \in R_3$. We will show that $1, \alpha_1, \alpha_2, \alpha_3$ are linearly independent, so that the $k_i = 0$.

Recall that a left regular representation of a group is the linear representation provided by multiplication of group elements on the left. Based on the similarity of the algebra elements $1, \alpha_1, \alpha_2, \alpha_3$ from $\mathcal{A}(F_{0,3})$ with the group elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$, we define a left regular representation ρ for $\mathcal{A}(F_{0,3})$ by

$$\begin{aligned}
 \rho(1) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & \rho(\alpha_1) &= \begin{bmatrix} 0 & v_2^{-1} v_3^{-1} \delta^2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & v_2^{-1} \delta \\ 0 & 0 & v_3^{-1} \delta & 0 \end{bmatrix}, \\
 \rho(\alpha_2) &= \begin{bmatrix} 0 & 0 & v_1^{-1} v_3^{-1} \delta^2 & 0 \\ 0 & 0 & 0 & v_1^{-1} \delta \\ 1 & 0 & 0 & 0 \\ 0 & v_3^{-1} \delta & 0 & 0 \end{bmatrix}, & \rho(\alpha_3) &= \begin{bmatrix} 0 & 0 & 0 & v_1^{-1} v_2^{-1} \delta^2 \\ 0 & 0 & v_1^{-1} \delta & 0 \\ 0 & v_2^{-1} \delta & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.
 \end{aligned}$$

Note that the coefficients from each column are exactly those given by the equations describing left multiplication by α_i . In particular, they are the coefficients in the equations $\alpha_i * 1 = \alpha_i, \alpha_i * \alpha_i = v_{i+1}^{-1} v_{i+2}^{-1} \delta^2$, and $\alpha_i * \alpha_{i+1} = \alpha_{i+1} * \alpha_i = v_{i+2}^{-1} \delta \alpha_{i+2}$ for all i . The matrices $\rho(1), \rho(\alpha_1), \rho(\alpha_2)$ and $\rho(\alpha_3)$ are clearly linearly independent, as can be determined by looking at their first columns. Thus by [Lemma 2.4](#), we have that $1, \alpha_1, \alpha_2, \alpha_3$ are linearly independent. Hence there are no more relations to be found in $K(F_{0,3})$. This also shows that $\{\alpha_1, \alpha_2, \alpha_3\}$ is a minimal set of generators. \square

Surfaces with zero or one punctures. Recall that $R = \mathbb{Z}[A, A^{-1}]$ is a subring of $R_n = \mathbb{Z}[A^{\frac{1}{2}}, A^{-\frac{1}{2}}][v_1^{\pm 1}, v_2^{\pm 1}, \dots, v_n^{\pm 1}]$, and from [Lemma 1.1](#), there exists an algebra homomorphism ψ , which maps the R -algebra $\mathcal{S}(F_{g,n})$ to the R_n -algebra $\mathcal{A}(F_{g,n})$.

First observe that when $n=0$, the relations in $\mathcal{S}(F_{g,0})$ are exactly those in $\mathcal{A}(F_{g,0})$. That is, $R_0 \otimes \mathcal{K}_0(F_{g,0}) \cong \mathcal{K}(F_{g,0})$. Moreover, the map ψ from the proof of [Lemma 1.1](#) is injective when $n=0$ and acts as the identity on simple knots. Since all simple curves are simple knots in this case and the image of ψ contains all simple knots, the image of ψ generates all of the arc algebra $\mathcal{A}(F_{g,0})$ by [Lemma 2.1](#). Thus $R_0 \otimes \mathcal{S}(F_{g,0}) \cong \mathcal{A}(F_{g,0})$, and any presentation of $\mathcal{S}(F_{g,0})$ provides a presentation of $\mathcal{A}(F_{g,0})$.

When $n=1$, again there are no simple arcs, so that the image of ψ generates all of the arc algebra $\mathcal{A}(F_{g,1})$. So any set generating $\mathcal{S}(F_{g,1})$ also generates $\mathcal{A}(F_{g,1})$. However, the map ψ is no longer injective. Specifically, the relations for the Kauffman skein algebra and the relations for the Kauffman arc algebra will differ; the puncture-framing relation from the Kauffman arc algebra is not a relation in the Kauffman skein algebra. However, this is the only difference. Hence $R_1 \otimes \mathcal{S}(F_{g,1}) / \mathcal{K}_{\text{pfr}}(F_{g,1}) \cong \mathcal{A}(F_{g,1})$, where $\mathcal{K}_{\text{pfr}}(F_{g,1})$ is the submodule generated by only the puncture-framing relation. In summary, the generators of $\mathcal{A}(F_{g,1})$ are generators of $\mathcal{S}(F_{g,1})$, but the relations of $\mathcal{A}(F_{g,1})$ are relations of $\mathcal{S}(F_{g,1})$ along with one corresponding to the puncture-framing relation.

Torus with zero or one punctures. As an example, let us examine the cases of the closed torus and the torus with one puncture. From [\[Bullock and Przytycki 2000\]](#), we have that the Kauffman skein algebras $\mathcal{S}(F_{1,0})$ and $\mathcal{S}(F_{1,1})$ are both generated as $\mathbb{Z}[A, A^{-1}]$ -modules by three simple closed curves $\gamma_1, \gamma_2, \gamma_3$ such that γ_1 and γ_2 intersect once and γ_3 is one of two curves that meet both γ_1 and γ_2 once. Moreover, if ∂ represents a small loop around the puncture of $F_{1,1}$, then

$$\partial = A\gamma_1\gamma_2\gamma_3 - A^2\gamma_1^2 - A^{-2}\gamma_2^2 - A^2\gamma_3^2 + A^2 + A^{-2}. \quad (1)$$

In the skein algebra $\mathcal{S}(F_{1,0})$, we have $\partial = -A^2 - A^{-2}$. Up to a change in scalars from R to R_0 , a presentation of the arc algebra $\mathcal{A}(F_{1,0})$ is the same as the presentation of the skein algebra $\mathcal{S}(F_{1,0})$. That is,

$$\begin{aligned} \mathcal{A}(F_{1,0}) = R_0 \langle \gamma_1, \gamma_2, \gamma_3 \mid A\gamma_i\gamma_{i+1} - A^{-1}\gamma_{i+1}\gamma_i = (A^2 - A^{-2})\gamma_{i+1} \text{ and} \\ -A^2 - A^{-2} = A\gamma_1\gamma_2\gamma_3 - A^2\gamma_1^2 - A^{-2}\gamma_2^2 - A^2\gamma_3^2 + A^2 + A^{-2} \rangle, \end{aligned}$$

where the indices are interpreted modulo 3. On the other hand, in the once-punctured torus, we have $\partial = A + A^{-1}$ in the arc algebra. Thus

$$\begin{aligned} \mathcal{A}(F_{1,1}) = R_1 \langle \gamma_1, \gamma_2, \gamma_3 \mid A\gamma_i\gamma_{i+1} - A^{-1}\gamma_{i+1}\gamma_i = (A^2 - A^{-2})\gamma_{i+1} \text{ and} \\ A + A^{-1} = A\gamma_1\gamma_2\gamma_3 - A^2\gamma_1^2 - A^{-2}\gamma_2^2 - A^2\gamma_3^2 + A^2 + A^{-2} \rangle. \end{aligned}$$

Acknowledgements

The authors would like to thank Eric Egge and Francis Bonahon for helpful discussions, and the Carleton College Mathematics Department for their support and encouragement throughout this research.

References

- [Bobb et al. 2016] M. Bobb, S. Kennedy, D. Peifer, and H. Wong, “Roger and Yang’s Kauffman bracket arc algebra is finitely generated”, *J. Knot Theory Ramifications* **26** (online publication April 2016).
- [Bullock 1999] D. Bullock, “A finite set of generators for the Kauffman bracket skein algebra”, *Math. Z.* **231**:1 (1999), 91–101. MR 1696758 Zbl 0932.57016
- [Bullock and Przytycki 2000] D. Bullock and J. H. Przytycki, “Multiplicative structure of Kauffman bracket skein module quantizations”, *Proc. Amer. Math. Soc.* **128**:3 (2000), 923–931. MR 1625701 Zbl 0971.57021
- [Bullock et al. 1999] D. Bullock, C. Frohman, and J. Kania-Bartoszyńska, “Understanding the Kauffman bracket skein module”, *J. Knot Theory Ramifications* **8**:3 (1999), 265–277. MR 1691437 Zbl 0932.57015
- [Jones 1985] V. F. R. Jones, “A polynomial invariant for knots via von Neumann algebras”, *Bull. Amer. Math. Soc. (N.S.)* **12**:1 (1985), 103–111. MR 766964 Zbl 0564.57006
- [Kauffman 1987] L. H. Kauffman, “State models and the Jones polynomial”, *Topology* **26**:3 (1987), 395–407. MR 899057 Zbl 0622.57004
- [Penner 1987] R. C. Penner, “The decorated Teichmüller space of punctured surfaces”, *Comm. Math. Phys.* **113**:2 (1987), 299–339. MR 919235 Zbl 0642.32012
- [Przytycki 1991] J. H. Przytycki, “Skein modules of 3-manifolds”, *Bull. Polish Acad. Sci. Math.* **39**:1–2 (1991), 91–100. MR 1194712 Zbl 0762.57013
- [Przytycki and Sikora 2000] J. H. Przytycki and A. S. Sikora, “On skein algebras and $Sl_2(\mathbb{C})$ -character varieties”, *Topology* **39**:1 (2000), 115–148. MR 1710996 Zbl 0958.57011
- [Roger and Yang 2014] J. Roger and T. Yang, “The skein algebra of arcs and links and the decorated Teichmüller space”, *J. Differential Geom.* **96**:1 (2014), 95–140. MR 3161387 Zbl 1290.53080
- [Turaev 1988] V. G. Turaev, “The Conway and Kauffman modules of a solid torus”, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **167**:Issled. Topol. 6 (1988), 79–89, 190. In Russian; translated in *J. Soviet Math.* **52**:1 (1990), 2799–2805. MR 964255 Zbl 0673.57004
- [Turaev 1991] V. G. Turaev, “Skein quantization of Poisson algebras of loops on surfaces”, *Ann. Sci. École Norm. Sup. (4)* **24**:6 (1991), 635–704. MR 1142906 Zbl 0758.57011

Received: 2015-05-17

Accepted: 2015-07-31

mbobb@math.utexas.edu

Department of Mathematics, University of Texas at Austin,
Austin, TX 78712, United States

djp282@cornell.edu

Department of Mathematics, Cornell University,
Ithaca, NY 14853, United States

skennedy@carleton.edu

Department of Mathematics, Carleton College,
Northfield, MN 55057, United States

hwong@carleton.edu

Department of Mathematics, Carleton College,
Northfield, MN 55057, United States

Arranging kings k -dependently on hexagonal chessboards

Robert Doughty, Jessica Gonda, Adriana Morales, Berkeley Reiswig,
Josiah Reiswig, Katherine Slyman and Daniel Pritikin

(Communicated by Arthur T. Benjamin)

Tessellate the plane into rows of hexagons. Consider a subset of $2n$ rows of these hexagons, each row containing $2n$ hexagons, forming a rhombus-shaped chessboard of $4n^2$ spaces. Two kings placed on the board are said to “attack” each other if their spaces share a side or corner. Placing kings in alternating spaces of every other row results in an arrangement where no two of the n^2 kings are attacking each other. According to our specific distance metric, n^2 is in fact the largest number of kings that can be placed on such a board with no two kings attacking one another, for a maximum “density” of $\frac{1}{4}$. We consider a generalization of this maximum density problem, instead requiring that no king attacks more than k other kings for $0 \leq k \leq 12$. For instance when $k = 2$ the density is at most $\frac{1}{3}$. For each k we give constructive lower bounds on the density, and use systems of inequalities and discharging arguments to yield upper bounds, where the bounds match in most cases.

1. Introduction

Consider the task of arranging as many king pieces as possible on a standard 8×8 chessboard so that no two squares containing kings share a side or corner. Note that the board partitions into sixteen 2×2 patches, and at most one king can reside in any patch. Yet placing a king in the upper left corner of each patch satisfies our requirements. So, in an optimal placement we have sixteen kings occupying $\frac{1}{4}$ of the board. We can generalize this problem as follows. Consider whole numbers k and n . What is the maximum number of kings that can be placed on an $n \times n$ board such that each king-occupied space shares at most k edges and corners with other king-occupied spaces? Note that in the previous example, $k = 0$ and $n = 8$. In fact, in [Ionascu et al. 2008] the following question is investigated:

MSC2010: 90C05, 90C27.

Keywords: k -dependence, combinatorial chessboard, optimization, discharging, linear programming.

Given a whole number $k \leq 8$ (8 being the maximum number of squares a king can attack), what is the maximum number s of kings that can be placed on an $m \times n$ board so that no king attacks more than k other kings? When m and n are large, how large can the density $s/(mn)$ be?

Similar problems have also been studied on alternate boards. For instance, [Bode et al. 2003] studies a similar problem on triangular boards with $k = 0$, and [Bode and Harborth 2003] looks at a similar problem for knights on both triangular and hexagonal boards with $k = 0$. Other interesting articles on combinatorial chessboard problems include [Fricke et al. 1995; Haynes et al. 1998; Hedetniemi et al. 1998; Watkins 2004]. In this paper, we consider a board in which the spaces are regular hexagons arranged into an $n \times n$ rhombus as in Figure 1. As in Władysław Gliški's hexagonal chess [Bodlaender 1996], a king occupying a hexagon will be said to attack the 12 other hexagons "nearest" its hexagon (as shown in the left half of Figure 2). In particular, we study the following:

Given a whole number $k \leq 12$, what is the maximum number s of kings that can be placed on an $n \times n$ hexagonal board so that no king attacks more than k other kings? When n is large, how large can the density s/n^2 be?

For most values of k , we find tight bounds on the optimal density s/n^2 when n is large. For those values of k where our bounds are not tight, the gaps between our upper and lower bounds are reasonably close, and we conjecture that a limiting density exists.

2. Notation and terminology

We establish some definitions and notation since we are not using the normal $n \times n$ chessboard. Consider a tiling of the plane by regular hexagons, each hexagon having two vertical sides. We call a finite collection of hexagons a *hexagonal board* and call each hexagon in that tiling a *space*. For convenience we consider only hexagonal boards B_n where the hexagons form an $n \times n$ rhombus in the plane as in Figure 1, where $n \geq 5$. We label the spaces on our board (as in Figure 1)

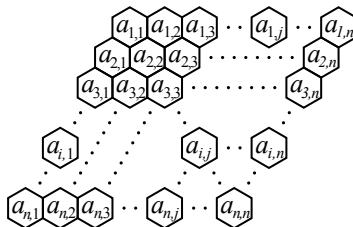


Figure 1. An $n \times n$ hexagonal board in the shape of a rhombus.

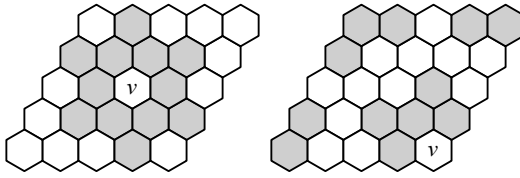


Figure 2. Some examples of $\mathcal{R}(v)$ shaded on a 5×5 board.

from top to bottom, left to right, as $a_{i,j}$, where i denotes the row the space is in and j denotes the diagonal column the space is in. We make this assumption since any finite set of spaces in the tiling will be contained in a suitably large rhombus of hexagons. We use scripted capital letters (such as \mathcal{A}) to name subsets of a board \mathcal{B}_n and nonscripted letters (such as a or A) to label individual spaces of a board \mathcal{B}_n . To avoid distinguishing between a space and a king occupying that space, we introduce the following terminology. When a subset \mathcal{A} of \mathcal{B}_n is specified, we refer to each space $v \in \mathcal{B}_n$ as being a *king* if $v \in \mathcal{A}$. We define the *realm* (or *neighborhood*) of a space $v \in \mathcal{B}_n$, denoted $\mathcal{R}(v)$, to be the set of spaces a king in space v attacks. In particular, we define the spaces in $\mathcal{R}(v)$ to create a “wrap-around” property on the board as follows: Given a space $a_{i,j} \in \mathcal{B}_n$ we define $g, g', g'', g''', h, h', h'', h''' \in (1, 2, \dots, n)$ as

$$\begin{aligned} g &\equiv i - 2 \pmod{n}, & h &\equiv j - 2 \pmod{n}, \\ g' &\equiv i - 1 \pmod{n}, & h' &\equiv j - 1 \pmod{n}, \\ g'' &\equiv i + 1 \pmod{n}, & h'' &\equiv j + 1 \pmod{n}, \\ g''' &\equiv i + 2 \pmod{n}, & h''' &\equiv j + 2 \pmod{n}. \end{aligned}$$

Finally we define $\mathcal{R}(a_{i,j})$ to be the set of spaces

$$\{a_{g,h'}, a_{g',h}, a_{g',h'}, a_{g',j}, a_{g',h''}, a_{i,h'}, a_{i,h''}, a_{g'',h'}, a_{g'',j}, a_{g'',h''}, a_{g''',h''}, a_{g''',h''}\}.$$

Note that if the board is less than 4×4 , the realm of a king overlaps itself. To simplify matters, the theorems in this paper only address $n \times n$ boards where $n \geq 4$, so a king’s realm always contains 12 spaces. (This is because realm spaces that are otherwise prevented by the boundary of the board are instead extended past the boundary and associated with the opposite side. Additionally when $n \geq 4$ it is impossible for the realm to overlap itself.) As an example,

$$\mathcal{R}(a_{1,1}) = \{a_{(n-1),n}, a_{n,(n-1)}, a_{n,n}, a_{n,1}, a_{n,2}, a_{1,n}, a_{1,2}, a_{2,n}, a_{2,1}, a_{2,2}, a_{2,3}, a_{3,2}\}.$$

A *placement of kings* on a hexagonal board \mathcal{B}_n is a subset \mathcal{A} of \mathcal{B}_n , the members of which we call *kings*. If u, v are kings, u is said to *attack* v if $v \in \mathcal{R}(u)$. A placement of kings \mathcal{A} is *k-dependent* if $\mathcal{R}(v)$ contains at most k kings for all $v \in \mathcal{A}$. We denote the collection of all k -dependent placements of kings on \mathcal{B}_n by $\mathcal{B}_n(k)$.

Our notation of a k -dependent arrangement of kings relates to the following graph-theoretic terminology. For a graph $G = (V, E)$ and a vertex $v \in V$, the set of vertices directly joined by an edge to v is denoted by $N(v)$. A subset \mathcal{A} of V is called *independent* if $N(a) \cap \mathcal{A} = \emptyset$ for each vertex $a \in \mathcal{A}$, and more generally, \mathcal{A} is called *k -dependent* (for a specified constant k) if $|N(a) \cap \mathcal{A}| \leq k$ for each $a \in \mathcal{A}$. The *k -dependence number* of $G = (V, E)$ is the maximum cardinality among k -dependent subsets of V .

The *maximum k -count* on \mathcal{B}_n , denoted by $\text{MNK}(k, n)$, is the maximum size of a k -dependent arrangement or placement of kings on \mathcal{B}_n . That is,

$$\text{MNK}(k, n) = \max\{|\mathcal{A}| : \mathcal{A} \in \mathcal{B}_n(k)\}$$

when $n \geq 4$. We let $\mu(k)$ denote the least upper bound of the density of kings in k -dependent placements. That is,

$$\mu(k) = \sup\left\{\frac{|\mathcal{A}|}{n^2} : \mathcal{A} \in \mathcal{B}_n(k), n \geq 4\right\}.$$

Alternatively, $\mu(k) = \sup\{\text{MNK}(k, n)/n^2 : n \geq 4\}$.

3. Initial upper bounds

Following tradition, we refer to the following technique as *linear programming*, even though we are not optimizing some objective function subject to constraints in the standard sense. For each k , we produce a simple system of linear inequalities valid for all $\mathcal{A} \in \mathcal{B}_n(k)$. Summing these inequalities will yield upper bounds for $\mu(k)$. Let $T(k)$ be the maximum number of kings in $\mathcal{R}(v)$ for $v \in \mathcal{A}^c = \mathcal{B}_n \setminus \mathcal{A}$ over all $\mathcal{A} \in \mathcal{B}_n(k)$ for $n \geq 4$. That is, if v is a non-king with respect to some k -dependent placement, then the largest number of kings possible in $\mathcal{R}(v)$ is $T(k)$. Given a placement $\mathcal{A} \subseteq \mathcal{B}_n$ the *indicator function* of \mathcal{A} over \mathcal{B}_n is

$$\chi_{\mathcal{A}}(w) = \begin{cases} 1 & \text{if } w \in \mathcal{A}, \\ 0 & \text{if } w \in \mathcal{A}^c. \end{cases}$$

Theorem 3.1. *Given $n \geq 4$, for all $\mathcal{A} \in \mathcal{B}_n(k)$ we have*

$$|\mathcal{A}| \leq \frac{T(k)}{T(k) - k + 12} n^2.$$

Proof. Let $\mathcal{A} \in \mathcal{B}_n(k)$. Then for each $v \in \mathcal{B}_n$, we have that $\sum_{w \in \mathcal{R}(v)} \chi_{\mathcal{A}}(w)$ is the number of kings that are in the realm of v . We consider the inequality

$$(T(k) - k)\chi_{\mathcal{A}}(v) + \sum_{w \in \mathcal{R}(v)} \chi_{\mathcal{A}}(w) \leq T(k). \tag{1}$$

If $\chi_{\mathcal{A}}(v) = 0$, then (1) holds by the definition of $T(k)$. Suppose $\chi_{\mathcal{A}}(v) = 1$. Then $\sum_{w \in \mathcal{R}(v)} \chi_{\mathcal{A}}(w) \leq k$ because \mathcal{A} is a k -dependent placement of kings. Therefore, since $(T(k) - k) \cdot 1 + k \leq T(k)$, inequality (1) holds whether or not $\chi_{\mathcal{A}}(v) = 0$. Now, we sum the inequalities of form (1), over all choices of $v \in \mathcal{B}_n$, and simplify the result:

$$\begin{aligned} \sum_{v \in \mathcal{B}_n} \left((T(k) - k) \chi_{\mathcal{A}}(v) + \sum_{w \in \mathcal{R}(v)} \chi_{\mathcal{A}}(w) \right) &\leq \sum_{v \in \mathcal{B}_n} T(k), \\ \left((T(k) - k) \sum_{v \in \mathcal{B}_n} \chi_{\mathcal{A}}(v) \right) + 12 \sum_{v \in \mathcal{B}_n} \chi_{\mathcal{A}}(v) &\leq T(k) n^2, \\ (T(k) - k + 12) \sum_{v \in \mathcal{B}_n} \chi_{\mathcal{A}}(v) &\leq T(k) n^2. \end{aligned}$$

Hence,

$$|\mathcal{A}| \leq \frac{T(k)}{T(k) - k + 12} n^2. \quad \square$$

Therefore, Theorem 3.1 establishes that

$$\frac{|\mathcal{A}|}{n^2} \leq \frac{T(k)}{T(k) - k + 12}$$

for all $k \in \{0, 1, 2, \dots, 12\}$ and $n \geq 4$.

Finding T -values. For each $k \in \{0, 1, 2, \dots, 12\}$, Theorem 3.1 gives us an upper bound for $|\mathcal{A}|/n^2$, the fraction of \mathcal{B}_n that can be occupied by a k -dependent set of kings. Since these upper bounds depend upon k and $T(k)$, we must find the exact values of $T(k)$ for each separate choice of k . We refer to values of $T(k)$ as T -values.

The following illustrates our process for finding the T -values for each k . To see why $T(0) = 4$, suppose $\mathcal{A} \in \mathcal{B}_n(0)$ for some $n \geq 4$ and consider some $v \in \mathcal{A}^c$. We label the spaces in $\mathcal{R}(v)$ as in Figure 3. Partition $\mathcal{R}(v)$ into the sets $\{A, B, C\}$, $\{D, E, F\}$, $\{G, H, I\}$, and $\{J, K, L\}$. Since \mathcal{A} is 0-dependent, the maximum number of kings in each of $\{A, B, C\}$, $\{D, E, F\}$, $\{G, H, I\}$, and $\{J, K, L\}$ is 1. Therefore, $T(0) \leq 4$. Also, since $\{A, D, G, J\}$ is a 0-dependent placement, we have $T(0) \geq 4$. Therefore $T(0) = 4$.

Figure 4 demonstrates lower bounds for all remaining T -values. The reader can check by hand that $T(k)$ equals the lower bound given below each picture.

Therefore, Theorem 3.1 establishes upper bounds for $\mu(k)$ as seen below:

k	0	1	2	3	4	5	6	7	8	9	10	11	12
Upper bound for $\mu(k)$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{3}{8}$	$\frac{8}{17}$	$\frac{5}{9}$	$\frac{10}{17}$	$\frac{2}{3}$	$\frac{12}{17}$	$\frac{3}{4}$	$\frac{4}{5}$	$\frac{6}{7}$	$\frac{12}{13}$	1

A board consisting entirely of kings would be 12-dependent. Therefore, $\mu(12) = 1$.

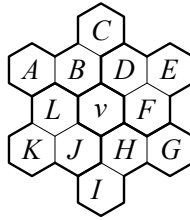


Figure 3. $\mathcal{R}(v)$ partitioned into four subsets.

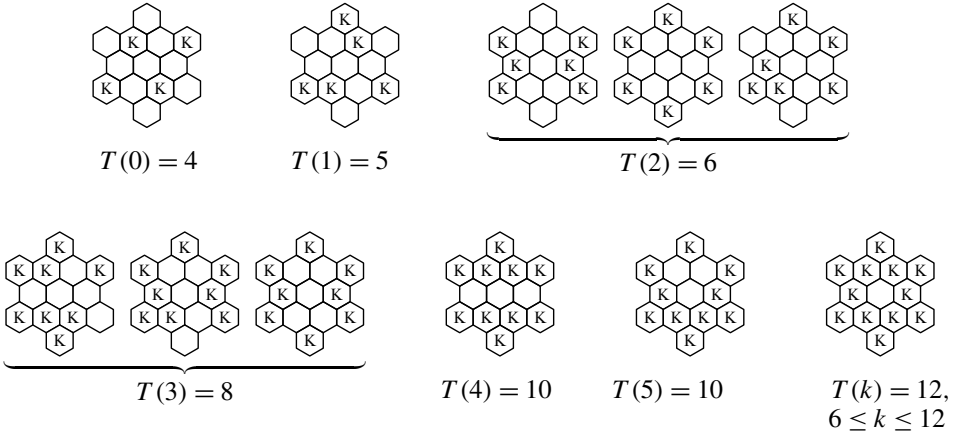


Figure 4. Lower bounds for the T -values for $0 \leq k \leq 12$.

4. Lower bounds

Ideally, we wish to find matching upper and lower bounds for $\mu(k)$ for each k , so as to determine $\mu(k)$ exactly. In this section we give constructive lower bounds for each such $\mu(k)$. To establish lower bounds for $\mu(k)$ we create patterns via “puzzle pieces”. We use them to construct arbitrarily large k -dependent placements with calculable density. To construct these placements, we “stamp” the puzzle piece

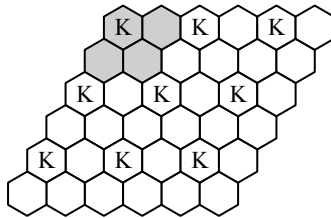


Figure 5. The shaded region represents the puzzle piece $P(0)$, which has been stamped nine times to create a 0-dependent placement on a 6×6 board.

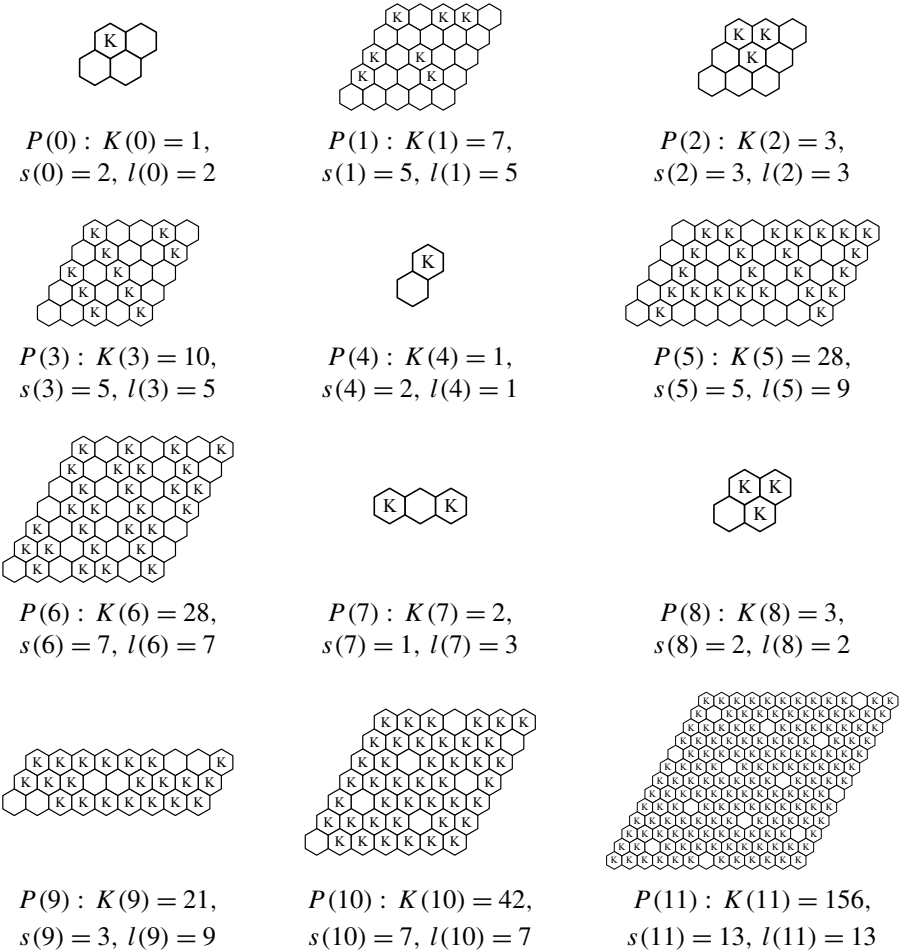


Figure 6. Puzzle pieces $P(k)$ that result in a k -dependent placement.

(using translated copies) as many times as needed to fill large $n \times n$ boards. For example, in Figure 5, the shaded region represents a 2×2 puzzle piece, which we stamped a total of nine times to create a 6×6 placement. This results in a 0-dependent placement. We call this method of obtaining a k -dependent placement the *stamping method*. For a given k , we define $P(k)$ (as in Figure 6) to be a particular $s \times l$ puzzle piece which results in a k -dependent placement. Additionally, for a given $P(k)$, we define $K(k)$ to be the number of kings in $P(k)$. We denote the number of rows in $P(k)$ by $s(k)$ and the number of diagonal columns in $P(k)$ by $l(k)$.

Theorem 4.1. For a given k , puzzle piece $P(k)$, and $n \geq \max\{s, l, 4\}$, we have

$$\mu(k) \geq \frac{K(k)}{s(k)l(k)}.$$

Proof. Let $n = q_1s(k) + r_1 = q_2l(k) + r_2$ for integers $q_1, q_2, r_1, r_2 \geq 0$, where $r_1 < s(k)$ and $r_2 < l(k)$. Consider a particular k and its corresponding puzzle piece $P(k)$ placed in the top-left of the $n \times n$ board. Let \mathcal{Z} be the set of kings in the $n \times n$ board at this stage. We define

$$\mathcal{Z}' = \{a_{i+cs, j+dl} : a_{i,j} \in \mathcal{Z}, c \in \{0, 1, 2, \dots, q_1 - 1\}, d \in \{0, 1, 2, \dots, q_2 - 1\}\}.$$

Therefore, \mathcal{Z}' is the placement of kings arising from a specific stamp and the process of stamping it in \mathcal{B}_n .

Since q_1q_2 is the number of copies of the puzzle piece on the $n \times n$ board, $|\mathcal{Z}'| = K(k)q_1q_2$. Thus, since $\text{MNK}(k, n)$ is the maximum number of kings on a k -dependent $n \times n$ board, we have $\text{MNK}(k, n) \geq K(k)q_1q_2$ for $n \geq \max\{s, l\}$. Also, note that $q_1 = (n - r_1)/s(k)$ and $q_2 = (n - r_2)/l(k)$. So it follows that

$$q_1q_2 = \frac{(n - r_1)(n - r_2)}{s(k)l(k)}.$$

Thus,

$$\frac{\text{MNK}(k, n)}{n^2} \geq \frac{K(k)(n - r_1)(n - r_2)}{s(k)l(k)n^2}.$$

This implies

$$\mu(k) \geq \sup \left\{ \frac{K(k)(n - r_1)(n - r_2)}{s(k)l(k)n^2} : n \geq \max\{s, l, 4\} \right\} = \frac{K(k)}{s(k)l(k)}. \quad \square$$

Our choices of $P(k)$ for $0 \leq k \leq 12$ are shown in [Figure 6](#). Each of these is crafted carefully to optimize the maximum proportion of kings, yielding the best lower bound we can manage. The following is a table of the lower bounds we constructed for $\mu(k)$ using the stamping method:

k	0	1	2	3	4	5	6	7	8	9	10	11	12
Lower bound for $\mu(k)$	$\frac{1}{4}$	$\frac{7}{25}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{4}{7}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{7}{9}$	$\frac{6}{7}$	$\frac{12}{13}$	1

5. Tightening the upper bounds

Although the upper and lower bounds for $\mu(k)$ were matched for some k using the methods in [Sections 3 and 4](#), we were unable to match others. In this section we use two additional methods to attempt to bring down the upper bound for $\mu(k)$ to match the lower bound.

Taxation. We now use a standard discharging method, which we refer to as *taxation*, to improve the upper bound for $\mu(k)$ when $k \in \{1, 2, 3, 4\}$. In this method, we call any non-king space a *pawn*. So, for a placement of kings \mathcal{A} , the set of pawns is \mathcal{A}^c . To understand taxation, consider the following scenario. Initially, each pawn starts with X dollars. Then each pawn pays all its money to the kings adjacent

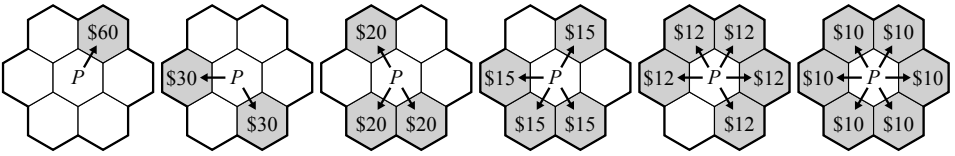


Figure 7. Examples of a taxation rule.

to it, following taxation rules. For example, each pawn could start with \$60 and distribute it evenly among all the kings adjacent to it. Some examples of a pawn following this taxation rule can be seen in Figure 7, where the shaded spaces are kings and the white spaces are pawns.

An $\$X$ taxation rule for $\mathcal{A} \in \mathcal{B}_n(k)$ is a function $f : \mathcal{A}^c \times \mathcal{A} \rightarrow [0, X]$ such that $\sum_{K \in \mathcal{A}} f(P, K) = X$ for every $P \in \mathcal{A}^c$. Note that in an $\$X$ taxation rule, $\sum_{(P, K) \in \mathcal{A}^c \times \mathcal{A}} f(P, K) = X|\mathcal{A}^c|$. When f is understood, we denote the value of funds a king K receives by $F(K) = \sum_{P \in \mathcal{A}^c} f(P, K)$. In an $\$X$ taxation rule, we know that the total amount of taxes paid, and also received, is $\$X|\mathcal{A}^c|$. So if each king receives a known minimum amount of money, $\$Y$, we can calculate an upper bound for $|\mathcal{A}|$.

Theorem 5.1. Consider a placement of kings $\mathcal{A} \in \mathcal{B}_n(k)$, where $n \geq 4$. If, following an $\$X$ taxation rule, $F(K) \geq Y$ for all $K \in \mathcal{A}$, then

$$|\mathcal{A}| \leq \frac{X}{X + Y} n^2.$$

Proof. Consider a placement $\mathcal{A} \in \mathcal{B}_n(k)$. The number of kings is $|\mathcal{A}|$, and the number of pawns is $n^2 - |\mathcal{A}|$. Suppose for some $\$X$ taxation rule f , we have $F(K) \geq Y$ for each king $K \in \mathcal{A}$. Comparing the amounts paid by pawns and received by kings, we see that

$$Y|\mathcal{A}| \leq X(n^2 - |\mathcal{A}|) \implies (X + Y)|\mathcal{A}| \leq Xn^2 \implies |\mathcal{A}| \leq \frac{X}{X + Y} n^2. \quad \square$$

For a given k , a general $\$X$ taxation plan (tax plan) $G(\mathcal{A})$ is a function G that assigns an $\$X$ taxation rule f to each $\mathcal{A} \in \mathcal{B}_n(k)$. Given such a G , let

$$\mathcal{Y}_G = \min\{F(K) : K \in \mathcal{A} \text{ and } \mathcal{A} \in \mathcal{B}_n(k)\}.$$

Corollary 5.2. Suppose that $\mathcal{Y}_G \geq Y$ for some $\$X$ taxation rule G . Then

$$\mu(k) \leq \frac{X}{X + Y}.$$

Proof. Suppose $\mathcal{Y}_G \geq Y$ for some $\$X$ taxation rule G . For all $\mathcal{A} \in \mathcal{B}_n(k)$,

$$|\mathcal{A}| \leq \frac{X}{X + \mathcal{Y}_G} n^2 \leq \frac{X}{X + Y} n^2,$$

where the first inequality follows from [Theorem 5.1](#). So, $(X/(X+Y))n^2$ is an upper bound on $|\mathcal{A}|$ for all $\mathcal{A} \in \mathcal{B}_n(k)$. Thus, $X/(X+Y)$ is an upper bound for $\mu(k)$. \square

For each $k \in \{1, 2, 3, 4\}$, we choose a convenient X and Y and show that there exists an $\$X$ tax plan G such that $\mathcal{Y}_G \geq Y$. Thus, by [Corollary 5.2](#), $\mu(k) \leq X/(X+Y)$. To achieve this, it is helpful to partition each placement \mathcal{A} into parts and analyze how much funding each part receives collectively. We define a *cluster* in a placement of kings \mathcal{A} to be a nonempty set $\mathcal{K} \subseteq \mathcal{A}$ such that for each $v, w \in \mathcal{K}$, there exists a sequence $v = v_1, v_2, \dots, v_s = w$, where v_i and v_{i+1} are adjacent for each $i \in \{1, 2, \dots, s-1\}$. When a tax plan is understood, we denote the funds a cluster $\mathcal{K} \subseteq \mathcal{A}$ receives by $F(\mathcal{K}) = \sum_{K \in \mathcal{K}} F(K)$. Since each k -dependent placement of kings is partitioned into clusters in our arguments, it suffices to show that $\$Y|\mathcal{K}| \leq F(\mathcal{K})$ for each cluster $\mathcal{K} \in \mathcal{B}_n(k)$ we consider.

Corollary 5.3. *Consider an $\$X$ tax plan for a given k and $n \geq 4$. Suppose for each $\mathcal{A} \in \mathcal{B}_n(k)$ that \mathcal{A} partitions into clusters where $\$Y|\mathcal{K}| \leq F(\mathcal{K})$ for each cluster \mathcal{K} in that partition. Then $\mu(k) \leq X/(X+Y)$.*

In all of our tax plans, any pawn with no adjacent kings is assumed to divide its money equally between all the kings in \mathcal{A} . Since we are looking for a maximum number of kings on $\mathcal{B}_n(k)$ for some k with $n \geq 4$, we need not worry about the case $\mathcal{A} = \emptyset$, so there will always be at least one king in any relevant arrangement. In the remainder of this subsection the following notation is used in the diagrams. A *white* hexagon denotes a pawn. A *shaded* hexagon denotes a king. We omit proofs of the following four propositions, as they took many pages of case analysis.

Tax plan for $k = 1$. We show that $\mu(1) \leq \frac{7}{25}$ using a taxation argument. We employ a tax plan where each pawn starts with $\$14$ and distributes it evenly among kings adjacent to it unless the pawn and king arrangement is one shown in [Figure 8](#). When the arrangement is as in [Figure 8\(a\)](#), the pawn P_1 pays $\$6$ to K_1 and $\$8$ to K_2 . When the arrangement is as in [Figure 8\(b\)–\(c\)](#) the pawn P_2 pays $\$8$ to K_1 and $\$6$ to K_2 . When the arrangement is as in [Figure 8\(d\)–\(e\)](#) the pawn P_3 pays $\$8$ to K_1 and $\$6$ to K_2 . Note that the orientation of these figures is arbitrary, with respect to rotation and reflection.

Proposition 5.4. *Every king in a 1-dependent arrangement on \mathcal{B}_n receives at least $\$36$ using the aforementioned tax plan.*

By [Proposition 5.4](#), $\mu(1) \leq \frac{7}{25}$.

Tax plan for $k = 2$. We show that $\mu(2) \leq \frac{1}{3}$ using a taxation argument. We employ a tax plan where each pawn starts with $\$6$ and distributes it evenly among kings adjacent to it unless the pawn and king arrangement is as shown in [Figure 9](#). When the arrangement is as in [Figure 9](#), the pawn P_1 pays $\$3$ to K_1 and $\$1.50$ each to K_2 and K_3 .

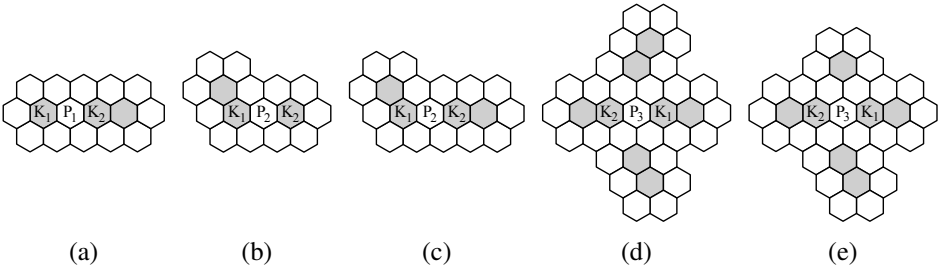


Figure 8. Pawn and king arrangements for the tax plan for $k = 1$.

Proposition 5.5. *Every king in a 2-dependent arrangement on \mathcal{B}_n receives at least \$12 using the aforementioned tax plan.*

By Proposition 5.5, $\mu(2) \leq \frac{1}{3}$.

Tax plan for $k = 3$. We show that $\mu(3) \leq \frac{2}{5}$ using a taxation argument. We employ a tax plan where each pawn starts with \$4 and distributes it evenly among kings adjacent to it unless the pawn and king arrangement is as shown in Figure 9. When the arrangement is as in Figure 9, the pawn P_1 pays \$2 to K_3 and \$1 each to K_1 and K_2 .

Proposition 5.6. *Every king in a 3-dependent arrangement on \mathcal{B}_n receives at least \$6 using the aforementioned tax plan.*

By Proposition 5.6, $\mu(3) \leq \frac{2}{5}$.

Tax plan for $k = 4$. We show that $\mu(4) \leq \frac{1}{2}$ using a taxation argument. We employ a tax plan where each pawn starts with \$60 and distributes it evenly amongst all kings adjacent to it.

Proposition 5.7. *Every king in a 4-dependent arrangement on \mathcal{B}_n receives at least \$60 using the aforementioned tax plan.*

By Proposition 5.7, $\mu(4) \leq \frac{1}{2}$.

Further linear programming. We improve our bounds using a more general form of linear programming we call *weighting patterns*. In this method, we seek more complicated linear inequality constraints that are valid for all k -dependent sets.

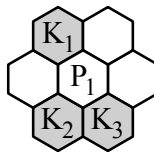


Figure 9. The special case for $k = 2$ and $k = 3$.

Given $n \geq 4$ and a *weighting function* $\varpi : \mathcal{B}_n \rightarrow [0, \infty)$, not everywhere zero, let $W(\varpi)$ denote the *total weight*, $\sum_{v \in \mathcal{B}_n} \varpi(v)$, of ϖ . Given k and ϖ on \mathcal{B}_n , let $M_k(n, \varpi)$ denote the maximum value of $\sum_{v \in \mathcal{B}_n} \varpi(v) \chi_{\mathcal{A}}(v)$ over all $\mathcal{A} \in \mathcal{B}_n(k)$. So, for any such ϖ and any $\mathcal{A} \in \mathcal{B}_n(k)$, we always have

$$\sum_{v \in \mathcal{B}_n} \varpi(v) \chi_{\mathcal{A}}(v) \leq M_k(n, \varpi).$$

To define a weighting function ϖ we refer to a figure with mapped values in their corresponding spaces. For example, if we defined ϖ by Figure 10(a), then $\varpi(a_{5,4}) = 10740$. Given a weighting function ϖ , we define the *shifted function* $\varpi_{x,y} : \mathcal{B}_n \rightarrow [0, \infty)$ by $\varpi_{x,y}(a_{i,j}) = \varpi(a_{(i+x) \bmod n, (j+y) \bmod n})$. Note that $\varpi_{0,0} = \varpi$.

Theorem 5.8. *Consider any weighting function ϖ on \mathcal{B}_n . Then*

$$\mu(k) \leq \frac{M_k(n, \varpi)}{W(\varpi)}$$

whenever $n \geq n_0 \geq 4$, where n_0 is the minimum value such that the weighting pattern ϖ fits within an $n_0 \times n_0$ rhombus.

Proof. Given $n \geq n_0 \geq 4$, let $\mathcal{A} \in \mathcal{B}_n(k)$, and let $\Gamma = \{0, 1, 2, \dots, n-1\}$. For any $x, y \in \Gamma$, we know that $\varpi_{x,y}$ is also a weighting function on \mathcal{B}_n . Thus we have

$$\sum_{a_{i,j} \in \mathcal{B}_n} \varpi_{x,y}(a_{i,j}) \chi_{\mathcal{A}}(a_{i,j}) \leq M_k(n, \varpi).$$

Therefore

$$\begin{aligned} \sum_{x,y \in \Gamma} \left(\sum_{a_{i,j} \in \mathcal{B}_n} \varpi_{x,y}(a_{i,j}) \chi_{\mathcal{A}}(a_{i,j}) \right) &\leq \sum_{x,y \in \Gamma} M_k(n, \varpi), \\ \sum_{a_{i,j} \in \mathcal{B}_n} \left(\chi_{\mathcal{A}}(a_{i,j}) \left(\sum_{x,y \in \Gamma} \varpi_{x,y}(a_{i,j}) \right) \right) &\leq n^2 M_k(n, \varpi), \\ \sum_{a_{i,j} \in \mathcal{B}_n} \left(\chi_{\mathcal{A}}(a_{i,j}) \left(\sum_{x,y \in \Gamma} \varpi(a_{(i+x) \bmod n, (j+y) \bmod n}) \right) \right) &\leq n^2 M_k(n, \varpi), \\ \sum_{a_{i,j} \in \mathcal{B}_n} (\chi_{\mathcal{A}}(a_{i,j}) W(\varpi)) &\leq n^2 M_k(n, \varpi), \end{aligned}$$

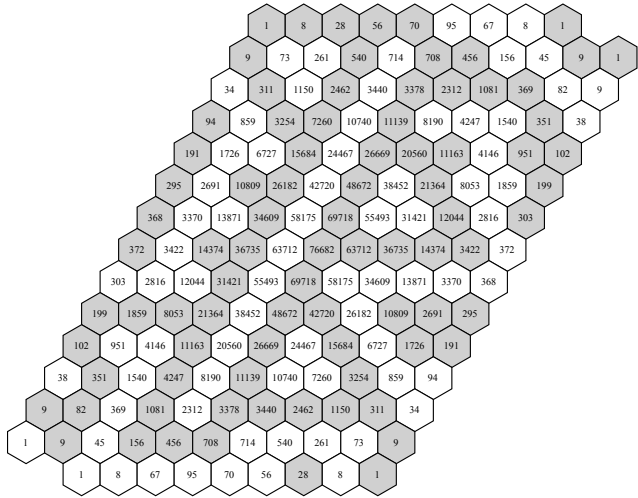
and thus

$$|\mathcal{A}| \cdot W(\varpi) \leq n^2 M_k(n, \varpi) \implies \frac{|\mathcal{A}|}{n^2} \leq \frac{M_k(n, \varpi)}{W(\varpi)}.$$

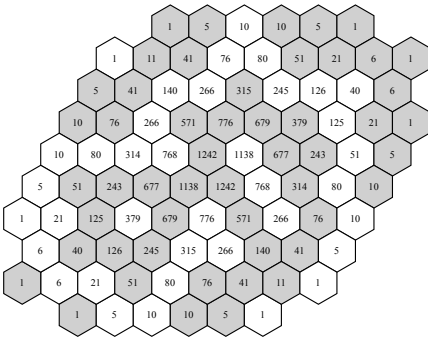
Since $n \geq 4$ and $\mathcal{A} \in \mathcal{B}_n(k)$ were arbitrary, we have

$$\mu(k) \leq \frac{M_k(n, \varpi)}{W(\varpi)}.$$

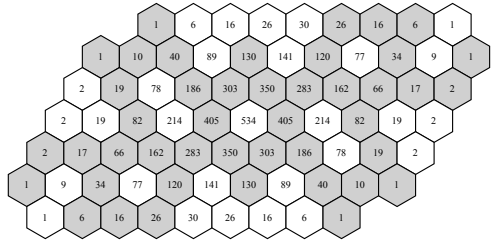
□



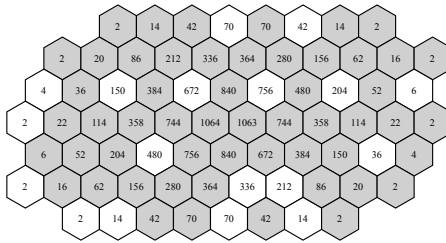
(a) $k = 5$



(b) $k = 6$



(c) $k = 7$



(d) $k = 9$

Figure 10. Weighting functions ϖ for $k \in \{5, 6, 7, 9\}$.

Using [Theorem 5.8](#), we improve our upper bounds for some $\mu(k)$ values by conveniently choosing the weighting patterns in [Figure 10](#) to define ϖ for $k \in \{5, 6, 7, 9\}$. The shaded hexagons represent a k -dependent placement of kings resulting in the largest $M_k(n, \varpi)$.

Applying four weighting functions and employing a tedious case analysis yields the following upper bounds for $\mu(k)$:

k	5	6	7	9
Upper bound for $\mu(k)$	$\frac{902416}{1636944}$	$\frac{11099}{17872}$	$\frac{4520}{6474}$	$\frac{12287}{15359}$

6. Conclusion

When the upper and lower bounds for $\mu(k)$ match, we know

$$\lim_{n \rightarrow \infty} \left(\frac{\text{MNK}(k, n)}{n^2} \right) = \mu(k).$$

In the cases where the upper and lower bounds for $\mu(k)$ do not match, it remains to determine the values of $\lim_{n \rightarrow \infty} (\text{MNK}(k, n)/n^2)$, if indeed these limits exist. In summary, combining results from various methods, the best results we found concerning the values of $\mu(k)$ are

$$\begin{aligned} \mu(0) &= \frac{1}{4}, & \mu(1) &= \frac{7}{25}, & \mu(2) &= \frac{1}{3}, & \mu(3) &= \frac{2}{5}, & \mu(4) &= \frac{1}{2}, \\ \frac{8}{15} &\leq \mu(5) \leq \frac{902416}{1636944}, & \frac{3}{5} &\leq \mu(6) \leq \frac{11099}{17872}, & \frac{2}{3} &\leq \mu(7) \leq \frac{4520}{6474}, & \mu(8) &= \frac{3}{4}, \\ \frac{7}{9} &\leq \mu(9) \leq \frac{12287}{15359}, & \mu(10) &= \frac{6}{7}, & \mu(11) &= \frac{12}{13}, & \mu(12) &= 1. \end{aligned}$$

Although our bounds for $\mu(k)$ are tight for most values of k , we were unable to match the bounds for $k \in \{5, 6, 7, 9\}$. The tightness of these bounds can perhaps be improved through computer programming and other mathematical methods. For example, one could try a “reverse taxation” process to lower the upper bound for $\mu(9)$. In such a process the kings would be given money to distribute to the pawns, potentially limiting the number of cases in the case analysis. Additionally, a computer could be used to test larger weighting patterns, although we have reason to believe that it is unlikely they will lead to a tight bound. Finally, by computer search or by hand, one could search for patterns which raise the lower bound.

To create a new problem, one can investigate $\mu(k)$ on similar boards with a change in the realm of a king. Moreover, one can consider boards consisting of other shapes, or in higher dimensions. For example, one could examine the k -dependence of kings on an $n \times n$ rhombus tiled with equilateral triangles, a parallelepiped whose surface is tiled with equilateral triangles, or a parallelepiped internally tiled with regular tetrahedra. A similar type of problem one might investigate is the domination number as in [Haynes et al. 1998] of an $n \times n$ hexagonal board on a torus using our king’s realm. One could also examine the domination number of a board with respect to different realms or of a different board altogether. We encourage the investigation of these problems, because they seem to us ideal problems for upper-level undergraduates with backgrounds or interests in discrete mathematics.

Acknowledgements

The last-named author (Pritikin) mentored the remaining authors in this research, for which they are grateful. All authors extend their thanks to the SUMSRI program at Miami University for funding the research.

References

- [Bode and Harborth 2003] J.-P. Bode and H. Harborth, “Independence for knights on hexagon and triangle boards”, *Discrete Math.* **272**:1 (2003), 27–35. [MR 2019197](#) [Zbl 1029.05111](#)
- [Bode et al. 2003] J.-P. Bode, H. Harborth, and M. Harborth, “King independence on triangle boards”, *Discrete Math.* **266**:1–3 (2003), 101–107. [MR 1991709](#) [Zbl 1015.05067](#)
- [Bodlaender 1996] H. Bodlaender, “Glinksy’s hexagonal chess”, website, 1996, available at <http://www.chessvariants.com/hexagonal.dir/hexagonal.html>.
- [Fricke et al. 1995] G. H. Fricke, S. M. Hedetniemi, S. T. Hedetniemi, A. A. McRae, C. K. Wallis, M. S. Jacobson, H. W. Martin, and W. D. Weakley, “Combinatorial problems on chessboards: a brief survey”, pp. 507–528 in *Graph theory, combinatorics, and algorithms, I* (Kalamazoo, MI, 1992), edited by Y. Alavi and A. Schwenk, Wiley, New York, 1995. [MR 1405836](#) [Zbl 0843.05061](#)
- [Haynes et al. 1998] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, *Fundamentals of domination in graphs*, Monographs and Textbooks in Pure and Applied Mathematics **208**, Marcel Dekker, New York, 1998. [MR 1605684](#) [Zbl 0890.05002](#)
- [Hedetniemi et al. 1998] S. M. Hedetniemi, S. T. Hedetniemi, and R. Reynolds, “Combinatorial problems on chessboards, II”, pp. 133–162 in *Domination in graphs*, edited by T. W. Haynes et al., Monogr. Textbooks Pure Appl. Math. **209**, Dekker, New York, 1998. [MR 1605691](#) [Zbl 0888.05036](#)
- [Ionascu et al. 2008] E. J. Ionascu, D. Pritikin, and S. E. Wright, “ k -dependence and domination in kings graphs”, *Amer. Math. Monthly* **115**:9 (2008), 820–836. [MR 2463293](#) [Zbl 1203.05119](#)
- [Watkins 2004] J. J. Watkins, *Across the board: the mathematics of chessboard problems*, Princeton University Press, 2004. [MR 2041306](#) [Zbl 1108.00007](#)

Received: 2015-07-22

Revised: 2015-07-31

Accepted: 2015-09-17

doughtrl@miamioh.edu

Miami University, Oxford, OH 45162, United States

jlgl31@zips.uakron.edu

The University of Akron, Akron, OH 44304, United States

adriana.morales@upr.edu

University of Puerto Rico, San Juan, 00931, Puerto Rico

berkeleyreiswig@gmail.com

Anderson University, Anderson, SC 29621, United States

jreiswig@math.sc.edu

University of South Carolina, Columbia, SC 29208, United States

kat.slyman@gmail.com

Wake Forest University, Winston-Salem, NC 27106, United States

pritikd@miamioh.edu

Department of Mathematics, Miami University, 123 Bachelor Hall, 301 S. Patterson Ave., Oxford, OH 45056, United States

Gonality of random graphs

Andrew Deveau, David Jensen, Jenna Kainic and Dan Mitropolsky

(Communicated by Ravi Vakil)

The gonality of a graph is a discrete analogue of the similarly named geometric invariant of algebraic curves. Motivated by recent progress in Brill–Noether theory for graphs, we study the gonality of random graphs. In particular, we show that the gonality of a random graph is asymptotic to the number of vertices.

1. Introduction

In the moduli space of curves, the locus of Brill–Noether general curves is a dense open subset [Griffiths and Harris 1980]. In the moduli space of tropical curves, however, the Brill–Noether general locus is open [Lim et al. 2012; Len 2014] and nonempty [Cools et al. 2012], but it is not dense [Jensen 2014]. A natural question, therefore, is *how likely is it that a graph is Brill–Noether general?*

In this paper, we approach this question by studying the gonality of Erdős–Rényi random graphs. Recall that an Erdős–Rényi random graph $G(n, p)$ is obtained by fixing n vertices and, for each pair of vertices, introducing an edge between them with probability p . We will often refer to such graphs simply as random graphs, where the probability distribution is understood to be that of Erdős–Rényi. It is common to define the probability p as a function of n , and to consider the expected value of combinatorial invariants as n increases. Throughout, we use \mathbb{P} and \mathbb{E} to denote the probability and expected value, respectively.

The current article is a natural follow-up to other recent work on the divisor theory of random graphs. Most notably, Lorenzini [2008] asked about the distribution of divisor class groups of random graphs, and in [Clancy et al. 2015b] it is conjectured that they are distributed according to a variation of the Cohen–Lenstra heuristics. This conjecture is proved in [Wood 2014], expanding on the preliminary work of [Clancy et al. 2015a].

Before stating our main result, we briefly recall the basic theory of divisors on graphs. For a more detailed account, see [Baker and Norine 2007] and [Baker 2008]. A *divisor* on a simple graph G is an element of the free abelian group on the

MSC2010: 05C80, 14H51, 14T05.

Keywords: random graphs, gonality, chip-firing, Brill–Noether theory.

vertices of G , and a divisor $D = \sum_{v \in V(G)} a_v v$ is said to be *effective* if $a_v \geq 0$ for all v . Given a divisor $D = \sum_{v \in V(G)} a_v v$ and a vertex v' , we may *fire* v' to obtain a new divisor $D' = \sum_{v \in V(G)} b_v v$, where

$$b_v = \begin{cases} a_v - \text{val}(v) & \text{if } v = v', \\ a_v + 1 & \text{if } v \text{ is adjacent to } v', \\ a_v & \text{otherwise,} \end{cases}$$

where $\text{val}(v)$ is the valence of the vertex v . Two divisors are *equivalent* if one can be obtained from the other by firing a sequence of vertices, and we say that a divisor D has positive rank if $D - v$ is equivalent to an effective divisor for all vertices v in G . The *gonality* $\text{gon}(G)$ is the smallest degree of a divisor with positive rank. Our main result is the following.

Theorem 1.1. *Let $p(n) = c(n)/n$, and suppose that $\log(n) \ll c(n) \ll n$. Then*

$$\mathbb{E}(\text{gon}(G(n, p))) \sim n.$$

[Theorem 1.1](#) essentially says that the expected gonality of a random graph is as high as possible. We note, however, that the gonality of random graphs nevertheless falls short of that for a general curve, as the genus of a random graph is asymptotically $\frac{1}{2}c(n)n$, and if $c(n)$ is unbounded, this grows faster than n . From this perspective, it may be more natural to study the gonality of random *regular* graphs, as the genus of such graphs grows in proportion to the number of vertices. The case of 3-regular graphs would be particularly interesting, as such graphs correspond to top-dimensional strata of the moduli space of tropical curves.

Although [Theorem 1.1](#) follows directly from the earlier work of de Bruyn and Gijswijt [\[2014\]](#) and Wang et al. [\[2011\]](#), it appears to be unknown to experts in tropical Brill–Noether theory. At the time of writing, we became aware of simultaneous work by Amini and Kool [\[2016\]](#), in which they use an improvement on the spectral methods of [\[Cornelissen et al. 2015\]](#) to show that the gonality of a random graph is bounded above and below by constant multiples of n . Our results are essentially a tightening of these bounds, so that both upper and lower bounds are asymptotic to n , which indeed is conjectured in [\[Amini and Kool 2016, Section 5.2\]](#). Their techniques apply additionally to metric graphs, which we do not discuss here, and to the case of random regular graphs, which they show to have gonality bounded above and below by constant multiples of n as well.

Also of note is the bound that we provide on the error term $n - \mathbb{E}(\text{gon}(G(n, p)))$ (see [Theorem 3.3](#)). In the future, it would be interesting to explore with what precision we can bound this term.

A more complete study of the Brill–Noether theory of random graphs would involve divisors of rank greater than one. A natural generalization of the current

line of inquiry would be to study the Clifford index of random graphs, defined as

$$\text{Cliff}(G) := \min_{D \in \text{Jac}(G)} \{ \deg(D) - 2r(D) \mid r(D) > 0 \text{ and } r(K_G - D) > 0 \}.$$

Note that if the minimum in this expression is obtained by a divisor of rank one, then $\text{Cliff}(G) = \text{gon}(G) - 2$. The Clifford index of an algebraic curve C is known to always be either $\text{gon}(C) - 2$ or $\text{gon}(C) - 3$ [Coppens and Martens 1991]. The corresponding statement remains open for graphs, but if true, it would imply that the Clifford index of a random graph is asymptotic to the number of vertices as well.

2. A lower bound

In this section, we obtain a lower bound on the expected gonality of a random graph. The first step is to identify a lower bound for the gonality of an arbitrary graph. This is done in [de Bruyn and Gijswijt 2014], where it is shown that the *treewidth* of a graph is a lower bound for the gonality.

Definition 2.1. A *tree decomposition* of a graph G is a tree T whose nodes are subsets of the vertices of G , satisfying the following properties:

- (1) Each vertex of G is contained in at least one node of T .
- (2) If two nodes of T both contain a given vertex v , then all nodes of the tree in the unique path between these two nodes must contain v as well.
- (3) If two vertices v and w are adjacent in G , then there is a node of T that contains both v and w .

The *width* of a tree decomposition is one less than the number of vertices in its largest node. The *treewidth* $\text{tw}(G)$ of a graph G is the minimum width among all possible tree decompositions of G .

Proposition 2.2 [de Bruyn and Gijswijt 2014]. *Let G be a simple connected graph. Then*

$$\text{gon}(G) \geq \text{tw}(G).$$

Although we will not use it, we note the following simple consequence.

Corollary 2.2.1. *For a simple connected graph G ,*

$$\text{gon}(G) \geq \min\{\text{val}(v) \mid v \in V(G)\}.$$

Proof. The result follows immediately from Proposition 2.2 and the fact that $\text{tw}(G) \geq \min\{\text{val}(v) \mid v \in V(G)\}$ (see [Bodlaender and Koster 2011]). \square

The treewidth of random graphs has been studied extensively in [Wang et al. 2011; Gao 2012].

Lemma 2.3 [Wang et al. 2011]. *Let $p(n) = c(n)/n$, and suppose that $c(n) \ll n$ is unbounded. Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{tw}(G(n, p)) \geq n - o(n)) = 1.$$

Theorem 2.4. *Let $p(n) = c(n)/n$, and suppose that $\log(n) \ll c(n) \ll n$. Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{gon}(G(n, p)) \geq n - o(n)) = 1.$$

Proof. A random graph is always simple, and by a well-known result of Erdős and Rényi [1959], the assumption $c(n) \gg \log(n)$ implies that such a graph is connected with probability approaching 1. It follows that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(\text{gon}(G(n, p)) \geq n - o(n) \text{ and } G(n, p) \text{ is connected}) \\ = \lim_{n \rightarrow \infty} \mathbb{P}(\text{gon}(G(n, p)) \geq n - o(n)). \end{aligned}$$

By Proposition 2.2, if $G(n, p)$ is connected, then $\text{gon}(G(n, p)) \geq \text{tw}(G(n, p))$, and it follows that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{gon}(G(n, p)) \geq n - o(n)) \geq \lim_{n \rightarrow \infty} \mathbb{P}(\text{tw}(G(n, p)) \geq n - o(n)) = 1,$$

where the final equality follows from Lemma 2.3. □

3. An upper bound

In this section, we obtain an upper bound on the expected gonality of a random graph. Together with the results of the previous section, this will imply that the gonality of a random graph is asymptotically equal to the number of vertices. We note that the number of vertices n is a very simple upper bound for the gonality of a graph and, together with Theorem 2.4, this would be enough to establish the main theorem. We actually go a bit further and obtain a bound on the expected value of $n - \text{gon}(G(n, p))$. In the future, it would be interesting to explore this with higher precision.

Recall that an *independent set* in a graph is a set of vertices, no pair of which are connected by an edge. The *independence number* $\alpha(G)$ of a graph G is defined to be the maximal size of an independent set.

Proposition 3.1. *If G is a simple connected graph with n vertices, then $\text{gon}(G) \leq n - \alpha(G)$.*

Proof. Let I be a maximal independent set, and let D be the sum of the vertices in the complement of I . We will show that D has positive rank. If $v \notin I$, then $D - v$ is effective by definition. On the other hand, if $v \in I$, then since all of the neighbors of v are not in I and the graph is simple, by firing all of the vertices other than v we obtain an effective divisor equivalent to D with at least one chip on v . It follows that D has rank at least one, hence $\text{gon}(G) \leq \text{deg}(D) = n - \alpha(G)$. □

Note that gonality $n - 1$ is achieved by the complete graph K_n , so this bound is sharp. Note further that the complete graph is the only simple graph with n vertices whose gonality is $n - 1$.

The expected independence number of a random graph has been studied in [Frieze 1990].

Lemma 3.2 [Frieze 1990]. *Let $p(n) = c(n)/n$, and suppose that $c(n) \ll n$ is unbounded. For any $\epsilon > 0$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \alpha(G(n, p)) - \frac{2}{p(n)} (\log c(n) - \log \log c(n) - \log 2 + 1) \right| \leq \frac{\epsilon}{p(n)} \right) = 1.$$

From this, we can conclude the following.

Theorem 3.3. *Let $p(n) = c(n)/n$, and suppose that $\log(n) \ll c(n) \ll n$. Then*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\text{gon}(G(n, p)) \leq n - \frac{2}{p(n)} (\log c(n) - \log \log c(n) - \log 2 + 1) \right) = 1.$$

Proof. By Lemma 3.2, for any $\epsilon > 0$, we have

$$\alpha(G(n, p)) > \frac{2}{p(n)} (\log c(n) - \log \log c(n) - \log 2 + 1 - \epsilon)$$

with probability 1 as n approaches infinity. By Proposition 3.1, the number $n - \alpha(G(n, p))$ is an upper bound for the gonality of $G(n, p)$. \square

Proof of Theorem 1.1. Again, the assumption that $c(n) \gg \log(n)$ implies that the random graph is connected with high probability. By Theorem 3.3, the gonality of a random graph is bounded above by $n - o(n)$. Similarly, by Theorem 2.4, the gonality of a random graph is bounded below by $n - o(n)$. It follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}(\text{gon}(G(n, p))) = \lim_{n \rightarrow \infty} \frac{n - o(n)}{n} = 1. \quad \square$$

Acknowledgements

This paper was written as part of the 2014 Summer Undergraduate Math Research at Yale (SUMRY) program. We would like to extend our thanks to everyone involved in the program, and in particular to Sam Payne, who suggested this project. We also thank Matt Kahle for a particularly fruitful discussion.

References

- [Amini and Kool 2016] O. Amini and J. Kool, “A spectral lower bound for the divisorial gonality of metric graphs”, *Int. Math. Res. Notices* **2016**:8 (2016), 2423–2450.
- [Baker 2008] M. Baker, “Specialization of linear systems from curves to graphs”, *Algebra Number Theory* **2**:6 (2008), 613–653. MR 2448666 Zbl 1162.14018
- [Baker and Norine 2007] M. Baker and S. Norine, “Riemann–Roch and Abel–Jacobi theory on a finite graph”, *Adv. Math.* **215**:2 (2007), 766–788. MR 2355607 Zbl 1124.05049

- [Bodlaender and Koster 2011] H. L. Bodlaender and A. M. C. A. Koster, “Treewidth computations, II: Lower bounds”, *Inform. and Comput.* **209**:7 (2011), 1103–1119. [MR 2829452](#) [Zbl 1220.68071](#)
- [de Bruyn and Gijswijt 2014] J. v. D. de Bruyn and D. Gijswijt, “Treewidth is a lower bound on graph gonality”, preprint, 2014. [arXiv 1407.7055](#)
- [Clancy et al. 2015a] J. Clancy, N. Kaplan, T. Leake, S. Payne, and M. Wood, “On a Cohen–Lenstra heuristic for Jacobians of random graphs”, *J. Algebraic Combin.* **42**:3 (2015), 701–723. [MR 3403177](#) [Zbl 1325.05146](#)
- [Clancy et al. 2015b] J. Clancy, T. Leake, and S. Payne, “A note on Jacobians, Tutte polynomials, and two-variable zeta functions of graphs”, *Exp. Math.* **24**:1 (2015), 1–7. [MR 3305035](#) [Zbl 1310.05121](#)
- [Cools et al. 2012] F. Cools, J. Draisma, S. Payne, and E. Robeva, “A tropical proof of the Brill–Noether theorem”, *Adv. Math.* **230**:2 (2012), 759–776. [MR 2914965](#) [Zbl 1325.14080](#)
- [Coppens and Martens 1991] M. Coppens and G. Martens, “Secant spaces and Clifford’s theorem”, *Compositio Math.* **78**:2 (1991), 193–212. [MR 1104787](#) [Zbl 0741.14035](#)
- [Cornelissen et al. 2015] G. Cornelissen, F. Kato, and J. Kool, “A combinatorial Li–Yau inequality and rational points on curves”, *Math. Ann.* **361**:1-2 (2015), 211–258. [MR 3302619](#) [Zbl 06399408](#)
- [Erdős and Rényi 1959] P. Erdős and A. Rényi, “On random graphs, I”, *Publ. Math. Debrecen* **6** (1959), 290–297. [MR 0120167](#) [Zbl 0092.15705](#)
- [Frieze 1990] A. M. Frieze, “On the independence number of random graphs”, *Discrete Math.* **81**:2 (1990), 171–175. [MR 1054975](#) [Zbl 0712.05052](#)
- [Gao 2012] Y. Gao, “Treewidth of Erdős–Rényi random graphs, random intersection graphs, and scale-free random graphs”, *Discrete Appl. Math.* **160**:4-5 (2012), 566–578. [MR 2876340](#) [Zbl 1239.05166](#)
- [Griffiths and Harris 1980] P. Griffiths and J. Harris, “On the variety of special linear systems on a general algebraic curve”, *Duke Math. J.* **47**:1 (1980), 233–272. [MR 563378](#) [Zbl 0446.14011](#)
- [Jensen 2014] D. Jensen, “The locus of Brill–Noether general graphs is not dense”, preprint, 2014. [arXiv 1405.6338](#)
- [Len 2014] Y. Len, “The Brill–Noether rank of a tropical curve”, *J. Algebraic Combin.* **40**:3 (2014), 841–860. [Zbl 06371201](#)
- [Lim et al. 2012] C. M. Lim, S. Payne, and N. Potashnik, “A note on Brill–Noether theory and rank-determining sets for metric graphs”, *Int. Math. Res. Not.* **2012**:23 (2012), 5484–5504. [MR 2999150](#) [Zbl 1328.14099](#)
- [Lorenzini 2008] D. Lorenzini, “Smith normal form and Laplacians”, *J. Combin. Theory Ser. B* **98**:6 (2008), 1271–1300. [MR 2462319](#) [Zbl 1175.05088](#)
- [Wang et al. 2011] C. Wang, T. Liu, P. Cui, and K. Xu, “A note on treewidth in random graphs”, pp. 491–499 in *Combinatorial optimization and applications*, edited by W. Wang et al., Lecture Notes in Comput. Sci. **6831**, Springer, Heidelberg, 2011. [MR 2878459](#) [Zbl 05938199](#)
- [Wood 2014] M. M. Wood, “The distribution of sandpile groups of random graphs”, preprint, 2014. [arXiv 1402.5149](#)

Received: 2015-07-22 Revised: 2015-08-28 Accepted: 2015-08-30

andrew.deveau@yale.edu *Yale University, New Haven, CT 06511, United States*

dave.h.jensen@gmail.com *Department of Mathematics, University of Kentucky, 719
Patterson Office Tower, Lexington, KY 40506, United States*

jenna.kainic@yale.edu *Yale University, New Haven, CT 06511, United States*

daniel.mitropolsky@yale.edu *Yale University, New Haven, CT 06511, United States*

Guidelines for Authors

Submissions in all mathematical areas are encouraged. All manuscripts accepted for publication in *Involve* are considered publishable in quality journals in their respective fields, and include a minimum of one-third student authorship. Submissions should include substantial faculty input; faculty co-authorship is strongly encouraged. Authors may submit manuscripts in PDF format on-line at the Submission page at the [Involve website](#).

Originality. Submission of a manuscript acknowledges that the manuscript is original and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *Involve* are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use L^AT_EX but submissions in other varieties of T_EX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibT_EX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

involve

2016

vol. 9

no. 4

Affine hyperbolic toral automorphisms	541
COLIN THOMSON AND DONNA K. MOLINEK	
Rings of invariants for the three-dimensional modular representations of elementary abelian p -groups of rank four	551
THÉO PIERRON AND R. JAMES SHANK	
Bootstrap techniques for measures of center for three-dimensional rotation data	583
L. KATIE WILL AND MELISSA A. BINGHAM	
Graphs on 21 edges that are not 2-apex	591
JAMISON BARSOTTI AND THOMAS W. MATTMAN	
Mathematical modeling of a surface morphological instability of a thin monocrystal film in a strong electric field	623
AARON WINGO, SELAHITTIN CINAR, KURT WOODS AND MIKHAIL KHENNER	
Jacobian varieties of Hurwitz curves with automorphism group $\mathrm{PSL}(2, q)$	639
ALLISON FISCHER, MOUCHEN LIU AND JENNIFER PAULHUS	
Avoiding approximate repetitions with respect to the longest common subsequence distance	657
SERINA CAMUNGOL AND NARAD RAMPERSAD	
Prime vertex labelings of several families of graphs	667
NATHAN DIEFENDERFER, DANA C. ERNST, MICHAEL G. HASTINGS, LEVI N. HEATH, HANNAH PRAWZINSKY, BRIAHNA PRESTON, JEFF RUSHALL, EMILY WHITE AND ALYSSA WHITTEMORE	
Presentations of Roger and Yang's Kauffman bracket arc algebra	689
MARTIN BOBB, DYLAN PEIFER, STEPHEN KENNEDY AND HELEN WONG	
Arranging kings k -dependently on hexagonal chessboards	699
ROBERT DOUGHTY, JESSICA GONDA, ADRIANA MORALES, BERKELEY REISWIG, JOSIAH REISWIG, KATHERINE SLYMAN AND DANIEL PRITIKIN	
Gonality of random graphs	715
ANDREW DEVEAU, DAVID JENSEN, JENNA KAINIC AND DAN MITROPOLSKY	